# Analiza și Prognoza Seriilor de Timp

## Capitolul 7: Cointegrare și Modele VECM

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

## Cuprins

## Obiective de învățare

**La finalul acestui capitol, veți fi capabili să:**

- **Cointegrare**: Înțelegeți conceptul și relațiile de echilibru pe termen lung
- **Regresia falsă**: Recunoașteți și evitați problema regresiei false (spurious regression)
- **Engle-Granger**: Aplicați metoda în doi pași pentru testarea cointegrării
- **Johansen**: Efectuați testul pentru cointegrare multiplă
- **VECM** (Vector Error Correction Model — Model Vectorial de Corecție a Erorilor): Estimați și interpretați aceste modele
- **Viteza de ajustare**: Analizați coeficienții $\alpha$ și vectorii de cointegrare $\beta$
- **Python**: Implementați analiza de cointegrare cu aplicații practice

# De ce contează cointegrarea?

## Provocarea

- **Nestaționaritate**: Multe serii economice/financiare sunt I(1) — PIB, prețuri acțiuni, cursuri valutare au rădăcini unitare
- **Regresia standard**: Cu variabile I(1) $\Rightarrow$ rezultate false; diferențierea pierde informația pe termen lung

## Soluția: Cointegrarea

- **Trend stochastic comun**: Unele serii nestaționare se mișcă împreună pe termen lung — această relație poate fi modelată.

## Premiul Nobel 2003

- Clive Granger a primit Premiul Nobel în Economie (împreună cu Robert Engle)
- Contribuția: "metode pentru analiza seriilor de timp economice cu tendințe comune"

## Aplicații practice

### Finanțe

- **Pairs Trading**: Tranzacționarea spread-ului între acțiuni cointegrate
- **Structura la termen**: rate ale dobânzii pe termen scurt și lung
- **Spot-Futures**: Prețurile spot și futures converg la maturitate

### Macroeconomie

- **Consum și Venit**: Ipoteza venitului permanent
- **Bani și Prețuri**: Teoria cantitativă a banilor
- **PPP** (Purchasing Power Parity — Paritatea Puterii de Cumpărare): Cursuri valutare și niveluri de prețuri

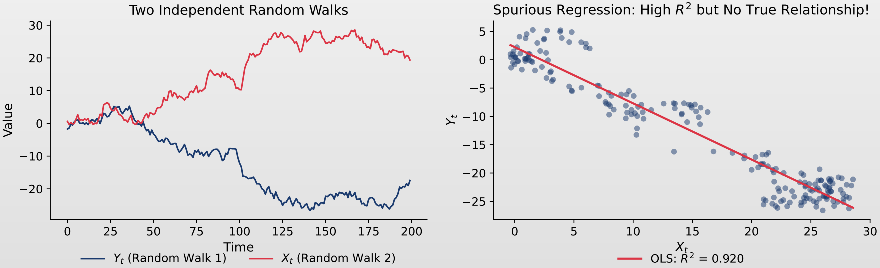
### Analiza politicilor

- **Politica Fiscală**: Cheltuieli guvernamentale și venituri fiscale
- **Politica Monetară**: Transmiterea ratelor dobânzii
- **Piața Muncii**: Salarii și productivitate

# Regresia falsă: exemplu vizual

## Atenție

- **Rezultat**: Două mersuri aleatorii complet independente prezintă $R^2$ ridicat doar din întâmplare; $R^2$ ridicat apare frecvent, iar $R^2 \to 1$ când $T \to \infty$ (Phillips, 1986)
- De aceea avem nevoie de analiza cointegrării




TSA_ch7_spurious_regression

# Problema regresiei false

## Granger & Newbold (1974)

- **Configurare**: Regresia unui mers aleatoriu pe un alt mers aleatoriu independent
  - $Y_t = \alpha + \beta X_t + u_t$, unde $Y_t$ și $X_t$ sunt procese I(1) independente

## Simptomele regresiei false

- **Coeficienți**: $R^2$ ridicat (adesea $> 0.9$) și statistici $t$ semnificative
  - Chiar dacă variabilele sunt complet necorelate.
- **Diagnostic**: Statistica Durbin-Watson foarte mică ($DW \approx 0$)
  - Reziduurile sunt nestaționare (au rădăcină unitară)

## Regulă practică (Granger)

- Dacă $R^2 > DW$, suspectați regresie falsă.

## Corelații false în lumea reală

### Explorarea datelor poate produce corelații fără sens

- Cu suficiente variabile și serii lungi de timp, apar tipare pur întâmplătoare

- Distanța dintre Neptun și Uranus $\leftrightarrow$ Prețul acțiunilor SAP SE (2002–2023)
- Utilizarea porumbului OMG în South Dakota $\leftrightarrow$ Căutări Google "i cant even" (2004–2023)
- Ratingurile serialului *Two and a Half Men* $\leftrightarrow$ Combustibil pentru avioane în Serbia (2006–2015)
- Popularitatea meme-ului "Its Wednesday my dudes" $\leftrightarrow$ Prețul acțiunilor Boeing (2006–2023)

### Lecție

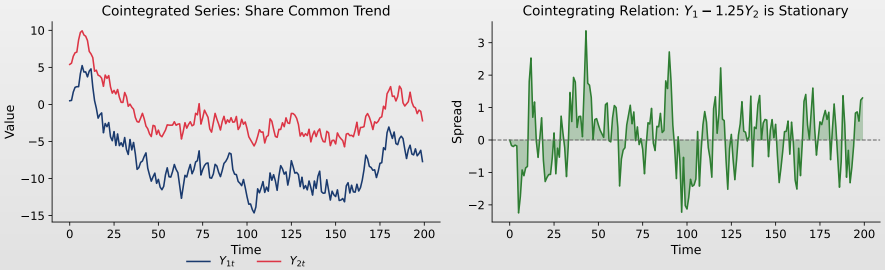
- Corelație ridicată $\neq$ cauzalitate
- Seriile nestaționare cu tendințe comune produc $R^2$ ridicat prin construcție
- Testați întotdeauna staționaritatea și cointegrarea înainte de a interpreta rezultatele regresiei

Explorați mai multe exemple: `tylervigen.com/spurious-correlations`

# Cointegrarea: exemplu vizual

## De reținut

- **Cointegrare**: Ambele serii sunt I(1) și evoluează împreună
- Combinația lor liniară (spread-ul) este staționară ⇒ aceasta este cointegrarea.




TSA_ch7_cointegrated_series

## Definiția cointegrării

### Definiție 1 (Cointegrare (Engle & Granger, 1987))

- **Definiție**: Variabilele $Y_{1t}, \ldots, Y_{kt}$ sunt **cointegrate de ordinul** $(d, b)$, notat $CI(d, b)$, dacă:
  1. Toate variabilele sunt integrate de ordinul $d$: $Y_{it} \sim I(d)$
  2. Există o combinație liniară $\beta' \mathrm{Y}_t$ care este integrată de ordinul $(d - b)$, unde $b > 0$

### Cazul cel mai comun: $CI(1, 1)$

- **Variabile**: Sunt $I(1)$ (au rădăcini unitare) $\Rightarrow$ combinația liniară este $I(0)$
- **Vectorul de cointegrare**: $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_k)' \Rightarrow$ definește echilibrul pe termen lung

### Observație

- **Unicitate**: Vectorul de cointegrare este unic doar până la înmulțire scalară. Se normalizează: $\beta_1 = 1$

## Intuiție: tendințe stochastice comune

### De ce apare cointegrarea?

- **Tendințe stochastice comune**: Variabilele cointegrate împart un trend comun
  - $Y_{1t} = \gamma_1 \tau_t + S_{1t}, \quad Y_{2t} = \gamma_2 \tau_t + S_{2t}$
  - $\tau_t$ este un mers aleatoriu comun și $S_{it}$ sunt componente staționare

### Combinația liniară elimină tendința

- $\gamma_2 Y_{1t} - \gamma_1 Y_{2t} = \gamma_2 S_{1t} - \gamma_1 S_{2t} \sim I(0)$

### Interpretare economică

- Cointegrarea reprezintă o **relație de echilibru pe termen lung**
  - Variabilele pot devia pe termen scurt
  - Dar sunt *trase înapoi* spre echilibru în timp
- Vectorul de cointegrare definește echilibrul

# Rangul de cointegrare

## Câte relații de cointegrare?

- **Condiție**: Pentru $k$ variabile $I(1) \Rightarrow$ maximum $r = k - 1$ relații de cointegrare
- **Cazuri posibile**:
  - $r = 0$: Nu există cointegrare (variabilele divergează)
  - $r = k$: Toate variabilele sunt de fapt $I(0)$ — inconsistent cu ipoteza $I(1)$

## Exemplu: 3 variabile

- **Rangul de cointegrare**:
  - $r = 0$: Nu există cointegrare
  - $r = 1$: O relație de cointegrare
  - $r = 2$: Două relații de cointegrare (doar 1 tendință comună)

## Observație

- **Relația**: Numărul de tendințe stochastice comune $= k - r$

## Metoda în doi pași Engle-Granger

### Pasul 1: Estimarea regresiei de cointegrare

- **Regresia OLS** (Ordinary Least Squares — Metoda Celor Mai Mici Pătrate): $Y_t = \alpha + \beta X_t + e_t$
- **Reziduuri**: $\hat{e}_t = Y_t - \hat{\alpha} - \hat{\beta} X_t$

### Pasul 2: Testarea staționarității reziduurilor

- **Testul ADF** (Augmented Dickey-Fuller): $\Delta \hat{e}_t = \rho \hat{e}_{t-1} + \sum_{j=1}^{p} \gamma_j \Delta \hat{e}_{t-j} + v_t$
- **Important**: Regresia auxiliară *nu* include constantă, deoarece reziduurile OLS au media zero prin construcție
- **Ipoteze**:
  - $H_0$: $\rho = 0$ (rădăcină unitară $\Rightarrow$ nu există cointegrare)
  - $H_1$: $\rho < 0$ (staționare $\Rightarrow$ există cointegrare)

### Important

- Folosiți **valorile critice Engle-Granger**, nu cele ADF standard! (mai negative deoarece reziduurile sunt estimate)

# Valorile critice Engle-Granger

## Valori critice pentru testul de cointegrare

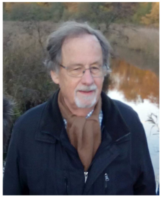
| Număr de Variabile | 1% | 5% | 10% |
|---|---|---|---|
| 2 | $-3.90$ | $-3.34$ | $-3.04$ |
| 3 | $-4.29$ | $-3.74$ | $-3.45$ |
| 4 | $-4.64$ | $-4.10$ | $-3.81$ |
| 5 | $-4.96$ | $-4.42$ | $-4.13$ |

- **Sursa**: Bazat pe estimările MacKinnon (1991), $T = 100$

## Limitările metodei Engle-Granger

- **Un singur vector**: Testează doar pentru o singură relație de cointegrare
  - Rezultatele depind de variabila aleasă ca dependentă
- **Eșantioane mici**: Bias pentru vectorul de cointegrare estimat
  - Nu se pot testa ipoteze asupra vectorului de cointegrare

## Portret de cercetător: Søren Johansen



*1939

W Wikipedia

### Biografie

- Statistician și econometrist danez, Profesor Emerit la Universitatea din Copenhaga
- Cunoscut pentru abordarea matematică riguroasă a econometriei
- Membru al Econometric Society; numeroase distincții în știința statistică

### Contribuții principale

- **Testul de cointegrare Johansen** (1988, 1991) — maxima verosimilitate pentru vectori multipli de cointegrare
- **Statisticile trace și eigenvalue maxim** pentru rangul de cointegrare
- **Estimarea VECM** — legătura între cointegrare și modele cu corecție a erorilor
- Cadrul standard pentru analiza multivariată a cointegrării

## Testul de cointegrare Johansen

### Avantaje față de Engle-Granger

- **Vectori multipli**: Testează pentru multiple relații de cointegrare
  - Permite testarea restricțiilor asupra vectorilor de cointegrare
- **Estimare MLE** (Maximum Likelihood Estimation — Estimare prin Maxima Verosimilitate): Mai eficientă
  - Nu necesită alegerea unei variabile dependente

### Punct de plecare: VAR (Vector Autoregression) în niveluri

- $Y_t = c + A_1 Y_{t-1} + A_2 Y_{t-2} + \cdots + A_p Y_{t-p} + \varepsilon_t$

### Următorul pas

- **Transformare**: Rescriem în forma VECM

## Derivare: De la VAR la VECM

### Punct de plecare: VAR(p) în niveluri

- $Y_t = A_1 Y_{t-1} + A_2 Y_{t-2} + \cdots + A_p Y_{t-p} + \varepsilon_t$

### Pasul 1: Scădem $Y_{t-1}$ din ambii membri

- **Transformare**:
  - $Y_t - Y_{t-1} = A_1 Y_{t-1} + A_2 Y_{t-2} + \cdots + A_p Y_{t-p} - Y_{t-1} + \varepsilon_t$
  - $\Delta Y_t = (A_1 - I) Y_{t-1} + A_2 Y_{t-2} + \cdots + A_p Y_{t-p} + \varepsilon_t$

### Obiectiv

- Rescriem astfel încât toți termenii să fie fie în **niveluri** ($Y_{t-1}$), fie în **diferențe** ($\Delta Y_{t-j}$)

# Derivare: De la VAR la VECM (cont.)

## Pasul 2: Adunăm și scădem termeni strategici

- **Manipulare**: Adunăm și scădem $A_2 Y_{t-1}$
- **Rezultat**: $\Delta Y_t = (A_1 + A_2 - I)Y_{t-1} - A_2 \Delta Y_{t-1} + A_3 Y_{t-3} + \cdots + \varepsilon_t$
- **Procedură**: Continuăm cu $A_3 Y_{t-1}$, etc.

## Pasul 3: Forma generală VECM

- **Rezultat**: $\Delta Y_t = \mathbf{\Pi} Y_{t-1} + \sum_{j=1}^{p-1} \mathbf{\Gamma}_j \Delta Y_{t-j} + \varepsilon_t$

## Matricele principale

- **Impact termen lung**: $\boxed{\mathbf{\Pi} = \sum_{i=1}^{p} A_i - I}$
- **Dinamică termen scurt**: $\boxed{\mathbf{\Gamma}_j = -\sum_{i=j+1}^{p} A_i, \quad j = 1, \ldots, p-1}$

## Derivare: Verificare cu VAR(2)

### Exemplu: VAR(2)

- **Punct de plecare**: $Y_t = A_1 Y_{t-1} + A_2 Y_{t-2} + \varepsilon_t$
  - Scădem $Y_{t-1}$: $\Delta Y_t = (A_1 - I)Y_{t-1} + A_2 Y_{t-2} + \varepsilon_t$
  - Adunăm și scădem $A_2 Y_{t-1}$: $\Delta Y_t = (A_1 + A_2 - I)Y_{t-1} + A_2(Y_{t-2} - Y_{t-1}) + \varepsilon_t$
- **Rezultat VECM**:
  - $\Delta Y_t = \underbrace{(A_1 + A_2 - I)}_{\boldsymbol{\Pi}} Y_{t-1} \underbrace{-A_2}_{\boldsymbol{\Gamma}_1} \Delta Y_{t-1} + \varepsilon_t$

### Verificare

- **VAR(2)**: $\boldsymbol{\Pi} = A_1 + A_2 - I$ și $\boldsymbol{\Gamma}_1 = -A_2$
  - Folosind formula: $\boldsymbol{\Gamma}_1 = -\sum_{i=2}^{2} A_i = -A_2 \quad \checkmark$

## Reprezentarea VECM

### Modelul vectorial de corecție a erorilor

- **Ecuația VECM**:
  - $\Delta Y_t = c + \mathbf{\Pi} Y_{t-1} + \sum_{j=1}^{p-1} \mathbf{\Gamma}_j \Delta Y_{t-j} + \varepsilon_t$
- **Componente**:
  - $\mathbf{\Pi} = A_1 + A_2 + \cdots + A_p - I$ (matricea impactului pe termen lung)
  - $\mathbf{\Gamma}_j = -(A_{j+1} + \cdots + A_p)$ (dinamica pe termen scurt)

### Observație: Rangul lui $\mathbf{\Pi}$

- **Rangul lui $\mathbf{\Pi}$** determină cointegrarea:
  - $\text{rank}(\mathbf{\Pi}) = 0$: Nu există cointegrare (VAR în diferențe)
  - $\text{rank}(\mathbf{\Pi}) = k$: Toate variabilele sunt $I(0)$ (VAR în niveluri)
  - $0 < \text{rank}(\mathbf{\Pi}) = r < k$: Cointegrare cu $r$ vectori de cointegrare

## Descompunerea lui $\mathbf{\Pi}$

### Când rank($\mathbf{\Pi}$) $= r < k$

- **Descompunere**: $\mathbf{\Pi} = \boldsymbol{\alpha}\boldsymbol{\beta}'$
  - $\boldsymbol{\beta}$: matricea $k \times r$ a vectorilor de cointegrare
  - $\boldsymbol{\alpha}$: matricea $k \times r$ a coeficienților de ajustare

### Interpretare

- **Termenul de corecție**: $\boldsymbol{\beta}'\mathrm{Y}_{t-1}$ = deviații de la echilibrul pe termen lung
- **Viteza de ajustare**: $\boldsymbol{\alpha}$ = cât de repede se revine la echilibru

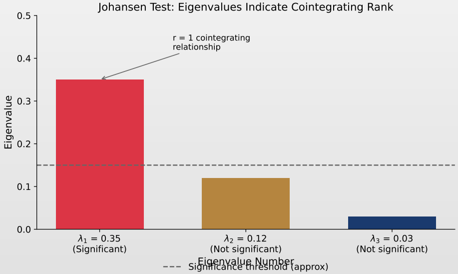
### Formula VECM

- **Ecuația**: $\Delta \mathrm{Y}_t = \mathrm{c} + \boldsymbol{\alpha}(\boldsymbol{\beta}'\mathrm{Y}_{t-1}) + \sum_{j=1}^{p-1} \mathbf{\Gamma}_j \Delta \mathrm{Y}_{t-j} + \varepsilon_t$

# Testul Johansen: interpretare vizuală

## Interpretare

- **Valorile proprii**: Cele semnificative (peste pragul critic) indică relații de cointegrare
- În acest exemplu, doar prima valoare proprie este semnificativă, sugerând $r = 1$




TSA_ch7_johansen_eigenvalues

## Statisticile testului Johansen

### Două statistici de Test

- **Bază**: Valorile proprii $\hat{\lambda}_1 > \hat{\lambda}_2 > \cdots > \hat{\lambda}_k$ ale unei anumite matrici
- **Testul Trace**: Testează $H_0$: rang $\leq r$ vs $H_1$: rang $> r$
  - $\lambda_{\text{trace}}(r) = -T\sum_{i=r+1}^{k} \ln(1-\hat{\lambda}_i)$
- **Testul Valorii Proprii Maxime**: Testează $H_0$: rang $= r$ vs $H_1$: rang $= r+1$
  - $\lambda_{\max}(r, r+1) = -T\ln(1-\hat{\lambda}_{r+1})$

### Observație

- **Valori critice**: Din Johansen & Juselius (1990)
- Depind de numărul de variabile $k$ și componentele deterministe (constante, trend)

## Procedura de testare

### Testare secvențială (testul Trace)

1. Testați $H_0$: $r = 0$ vs $H_1$: $r > 0$
   - Dacă nu se respinge: Nu există cointegrare. Stop.
   - Dacă se respinge: Cel puțin un vector de cointegrare. Continuăm.
2. Testați $H_0$: $r \leq 1$ vs $H_1$: $r > 1$
   - Dacă nu se respinge: $r = 1$. Stop.
   - Dacă se respinge: Cel puțin doi vectori de cointegrare. Continuăm.
3. Continuăm până când $H_0$ nu se respinge...

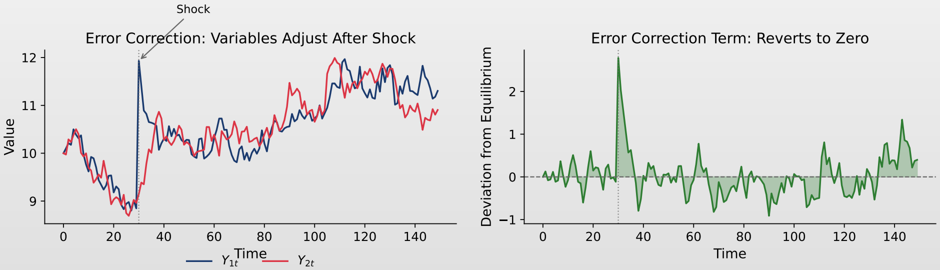
### Componentele deterministe

- **Alegere**: Specificația trebuie aleasă cu atenție
  - Fără constantă, fără trend (rar utilizat)
  - Constantă doar în relația de cointegrare
  - Constantă în ambele (cel mai comun)
  - Constantă + trend în relația de cointegrare
  - Constantă + trend în ambele

## Mecanismul de corecție a erorilor: vizualizare

### Interpretare

- **Corecția erorilor**: Când seriile deviază de la echilibru (zonele umbrite), mecanismul de ajustare le trage înapoi
- Deviațiile pozitive duc la ajustare în jos, deviațiile negative duc la ajustare în sus




TSA_ch7_error_correction

## Structura VECM

### Specificația completă VECM

- Pentru $k = 2$ variabile cu $r = 1$ relație de cointegrare:

$$\Delta Y_{1t} = c_1 + \alpha_1(Y_{1,t-1} - \beta Y_{2,t-1}) + \gamma_{11}\Delta Y_{1,t-1} + \gamma_{12}\Delta Y_{2,t-1} + \varepsilon_{1t}$$
$$\Delta Y_{2t} = c_2 + \alpha_2(Y_{1,t-1} - \beta Y_{2,t-1}) + \gamma_{21}\Delta Y_{1,t-1} + \gamma_{22}\Delta Y_{2,t-1} + \varepsilon_{2t}$$

### Componente

- **Corecția erorilor**: $(Y_{1,t-1} - \beta Y_{2,t-1})$ = deviație de la echilibru
  - $\alpha_1, \alpha_2$ = viteze de ajustare (ar trebui să aibă semne opuse)
- **Dinamica pe termen scurt**: $\gamma_{ij}$ = coeficienți lag-uri diferențiate
  - $\varepsilon_{it}$ = inovații

## Interpretarea coeficienților de ajustare

### Coeficienții $\alpha$

- **Relația de echilibru**: $Y_1 - \beta Y_2 = 0$
  - $\alpha_1 < 0$: $Y_1$ se ajustează în jos când este deasupra echilibrului
  - $\alpha_2 > 0$: $Y_2$ se ajustează în sus când $Y_1$ este deasupra echilibrului

### Exogenitate slabă

- Dacă $\alpha_i = 0$, variabila $Y_i$ **nu** răspunde la dezechilibru
  - $Y_i$ este **slab exogenă** pentru parametrii pe termen lung
  - Cealaltă variabilă face toată ajustarea
- Poate simplifica estimarea (abordare cu o singură ecuație)

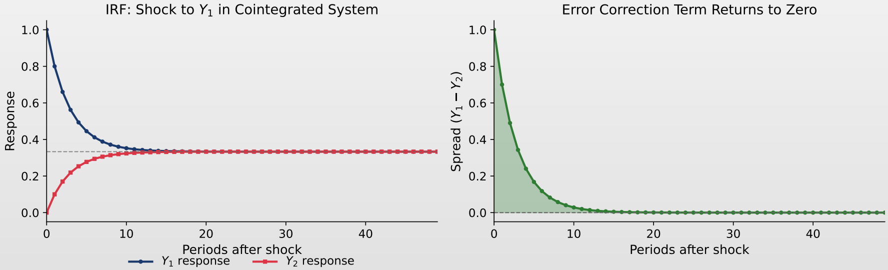
### Testare

- **Exogenitate slabă**: $H_0 : \alpha_i = 0$ folosind testul raportului de verosimilitate

# Funcțiile IRF (Impulse Response Functions) ale VECM

## Interpretare

- **Efecte permanente**: Într-un sistem cointegrat, șocurile au efecte permanente asupra nivelurilor
- Sistemul revine la echilibru ⇒ converg către o nouă valoare pe termen lung




TSA_ch7_vecm_irf

# VECM vs VAR în diferențe

## Când variabilele sunt cointegrate

| | VAR în Diferențe | VECM |
|---|---|---|
| Info pe termen lung | Pierdută | Păstrată |
| Dinamică pe termen scurt | Da | Da |
| Corecție a erorilor | Nu | Da |
| Prognoză | Slabă (termen lung) | Mai bună |
| Interpretare IRF | Doar termen scurt | Ambele |

## Teorema reprezentării Granger

- **Implicație**: Dacă variabilele sunt cointegrate, trebuie să existe o reprezentare de corecție a erorilor
- Ignorarea cointegrării = specificare greșită a modelului.

# Teorema Reprezentării Granger (1983)

TSA_ch7_granger_theorem

## Enunț formal

Fie $Y_t$ un vector $K \times 1$ de variabile $I(1)$. Dacă $Y_t$ este cointegrat cu rang de cointegrare $r$ $(0 < r < K)$, cu $\boldsymbol{\Pi} = \boldsymbol{\alpha\beta}'$, atunci:

1. **Există** o reprezentare VECM: $\Delta Y_t = \boldsymbol{\alpha\beta}' Y_{t-1} + \sum_{j=1}^{p-1} \boldsymbol{\Gamma}_j \Delta Y_{t-j} + \varepsilon_t$
2. **Există** o reprezentare MA (Wold): $\Delta Y_t = C(L)\varepsilon_t$, cu $C(1) \neq 0$, $\text{rang}(C(1)) = K - r$
3. **Reciproc**: Dacă $\Delta Y_t$ admite o reprezentare MA cu $\text{rang}(C(1)) = K - r < K$, atunci $Y_t$ este cointegrat CI(1,1) cu rang $r$

## Implicație fundamentală

Cointegrarea $\Leftrightarrow$ Corecția erorilor $\Leftrightarrow$ Trend comun. Cele trei reprezentări sunt **echivalente**.

# Demonstrație (schiță) — Teorema Reprezentării Granger

TSA ch7 granger theorem

## 1. Reprezentarea medie mobilă (MA)

Din VECM $\Delta Y_t = \alpha\beta' Y_{t-1} + \sum_{j=1}^{p-1} \Gamma_j \Delta Y_{t-j} + \varepsilon_t$ se obține prin inversare:

$$\Delta Y_t = C(L)\varepsilon_t, \quad \text{unde } C(1) = \beta_\perp(\alpha'_\perp(I_K - \sum_{i=1}^{p-1} \Gamma_i)\beta_\perp)^{-1}\alpha'_\perp$$

cu $\alpha_\perp, \beta_\perp$ complementele ortogonale ($\alpha'\alpha_\perp = 0$, $\beta'\beta_\perp = 0$), iar $\text{rang}(C(1)) = K - r$.

## 2. Reprezentarea trendurilor comune (Beveridge–Nelson)

Prin sumarea $\Delta Y_s$ de la $s = 1$ la $t$:

$$Y_t = C(1)\sum_{s=1}^{t} \varepsilon_s + C^*(L)\varepsilon_t + Y_0$$

- $C(1)\sum_{s=1}^{t} \varepsilon_s$: **trenduri stocastice** (mersuri aleatorii), sursa nestaționarității
- $C^*(L)\varepsilon_t$: **componentă staționară**
- Deoarece $\text{rang}(C(1)) = K - r$, există exact $K - r$ **trenduri comune**

## Consecință

$K$ variabile $I(1)$ cu rang de cointegrare $r \Rightarrow K - r$ trenduri stocastice comune. Cointegrarea *elimină* trenduri: $\beta' Y_t$ anulează componenta nestaționară deoarece $\beta' C(1) = 0$. *Ref. Johansen (1991), Teorema 4.2.*

## Flux de lucru practic

### Procedură pas cu pas

1. **Teste de Rădăcină Unitară**: Verificați că toate variabilele sunt $I(1)$
   - ADF, KPSS (Kwiatkowski-Phillips-Schmidt-Shin) pe niveluri și prime diferențe
2. **Selectarea numărului de lag-uri**: alegeți $p$ pentru VAR în niveluri
   - Folosiți AIC (Akaike), BIC (Bayesian Information Criterion) sau teste LR (Likelihood Ratio) secvențiale
3. **Testul de Cointegrare**: Teste trace/valoare proprie maximă Johansen
   - Determinați rangul de cointegrare $r$
4. **Estimați VECM**: Dacă $0 < r < k$
   - Estimați $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\boldsymbol{\Gamma}_j$
5. **Diagnostice**: Verificați reziduurile pentru autocorelație, normalitate
6. **Analiză**: IRF, FEVD (Forecast Error Variance Decomposition), teste de ipoteze

## Capcane frecvente

### Puncte de atenție

- **Rupturi structurale**: Pot cauza rădăcini unitare sau cointegrare false
- **Procese aproape de rădăcină unitară**: Testele au putere scăzută
- **Selectarea lag-urilor**:
  - **Prea multe lag-uri**: Supraparametrizare, pierdere de eficiență
  - **Prea puține lag-uri**: Autocorelație reziduală, estimări distorsionate
- **Specificație deterministă greșită**: Afectează valorile critice
- **Eșantioane mici**: Testul Johansen supradimensionat în eșantioane mici

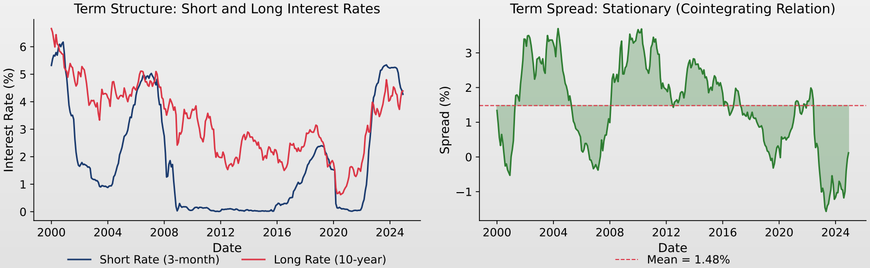
### Recomandare

- **Verificați întotdeauna**:
  - Diagnosticele reziduale (testul Portmanteau, normalitatea)
  - Stabilitatea relației de cointegrare estimate în timp
  - Sensibilitatea la lungimea lag-urilor și specificația deterministă

# Exemplu 1: structura la termen a ratelor dobânzii

## Ipoteza așteptărilor

- **Concluzie**: Ratele pe termen scurt și lung împart o tendință comună
- Spread-ul (prima de termen) este staționar ⇒ dovadă de cointegrare.

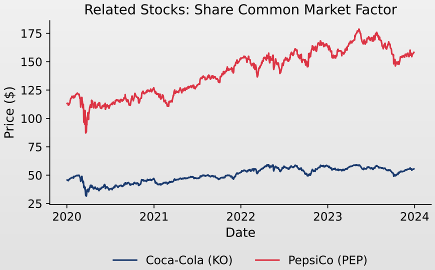


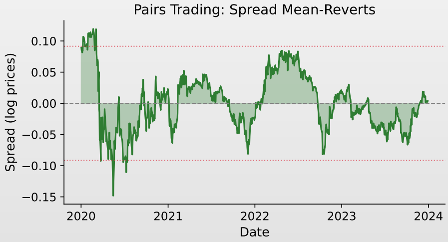

TSA_ch7_interest_rates_coint

# Exemplu 2: pairs trading în finanțe

## Strategie

- **Pairs trading**: Găsiți perechi de acțiuni cointegrate (ex., Coca-Cola & Pepsi)
- Când spread-ul deviază de la medie, tranzacționați așteptând revenirea la medie




TSA_ch7_pairs_trading

# Ratele dobânzii: teoria economică

## Ipoteza așteptărilor în structura la termen

- **Formula**: Rata pe termen lung ca medie a ratelor viitoare așteptate
  - $R_t^{(n)} = \frac{1}{n}\sum_{i=0}^{n-1} E_t[r_{t+i}] +$ prima de termen
- **Implicație**: Dacă prima de termen este constantă, $r_t$ și $R_t$ sunt cointegrate
  - Vectorul de cointegrare: $(1, -1)$

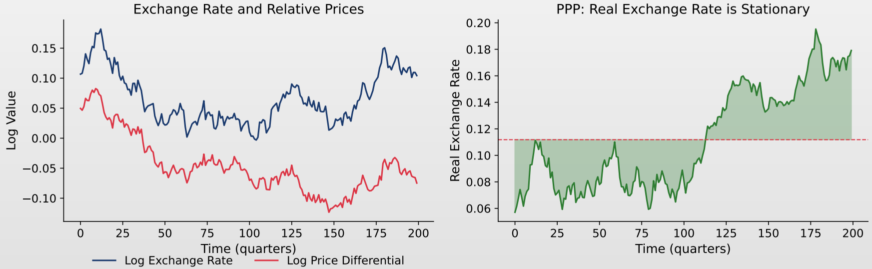
## Rezultate empirice

- **Teste de rădăcină unitară**: Ambele rate sunt $I(1)$
  - O relație de cointegrare (testul Johansen)
- **Vectorul de cointegrare**: $\approx (1, -1)$, spread-ul este staționar
  - Rata pe termen scurt se ajustează la dezechilibru (rata pe termen lung este slab exogenă)

# Exemplu 3: paritatea puterii de cumpărare (PPP)

## Teoria PPP

- **Formula**: $e_t = p_t - p_t^*$ (cursul de schimb logaritmic = diferențialul de preț)
- Cursul de schimb real ar trebui să fie staționar pe termen lung

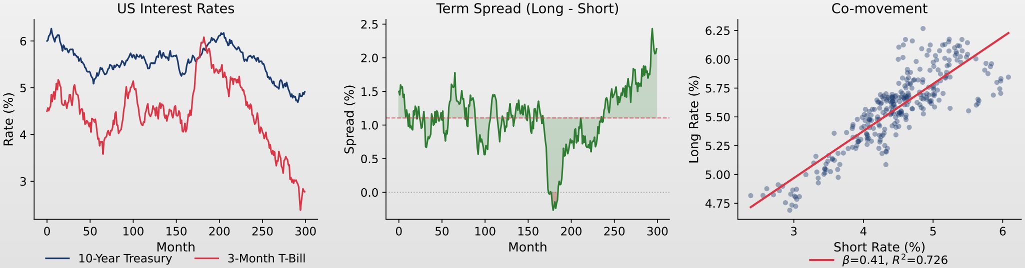



TSA_ch7_ppp_cointegration

# Studiu de caz: Cointegrarea ratelor dobânzii

## Date

- **Rate dobândă SUA**: Pe termen lung (10 ani) și scurt (3 luni)
- **Observație**: Ambele serii sunt I(1), dar spread-ul pare staționar

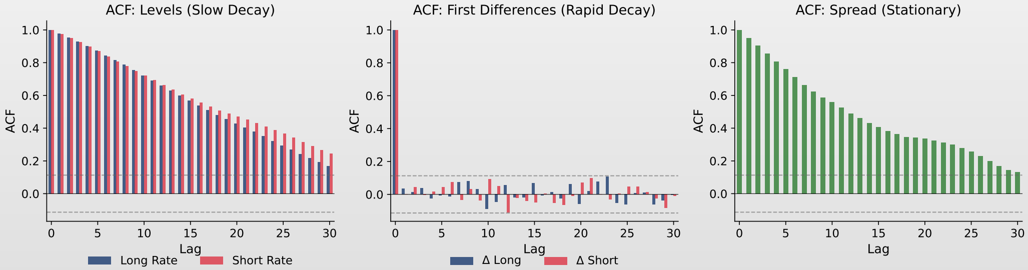



TSA_ch7_case_raw_data

# Pasul 1: Teste de rădăcină unitară

## Rezultate

- **ACF** (Autocorrelation Function) **niveluri**: Descreștere lentă ⇒ nestaționaritate; după diferențiere: scădere rapidă ⇒ I(1)
- **ACF spread**: Staționar ⇒ posibilă cointegrare.

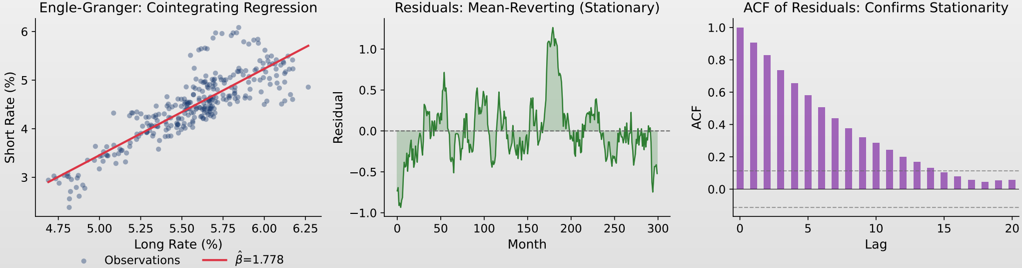



TSA_ch7_case_unit_root

# Pasul 2: Testul Engle-Granger de cointegrare

## Rezultate

- **Regresie**: Rata pe termen scurt $= \alpha + \beta \times$ Rata pe termen lung $+ \varepsilon_t$
- **Concluzie**: Seriile sunt cointegrate
- Există relație de echilibru pe termen lung

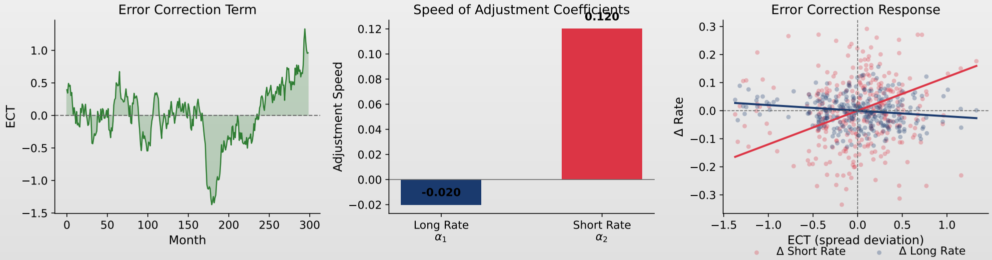



TSA_ch7_case_cointegration

# Pasul 3: Estimare VECM

## Model și ecuații stilizate

- **VECM(2)**: Rang de cointegrare $= 1$, vector $\approx (1, -1)$
- $\Delta r_t = 0.02 - 0.15(r_{t-1} - R_{t-1}) + \text{lag-uri} + \varepsilon_{1t}$; $\Delta R_t = 0.01 - 0.02(r_{t-1} - R_{t-1}) + \text{lag-uri} + \varepsilon_{2t}$
- **Interpretare**: Rata pe termen scurt se ajustează rapid ($\alpha_1 = -0.15$); rata pe termen lung aproape slab exogenă ($\alpha_2 \approx 0$)

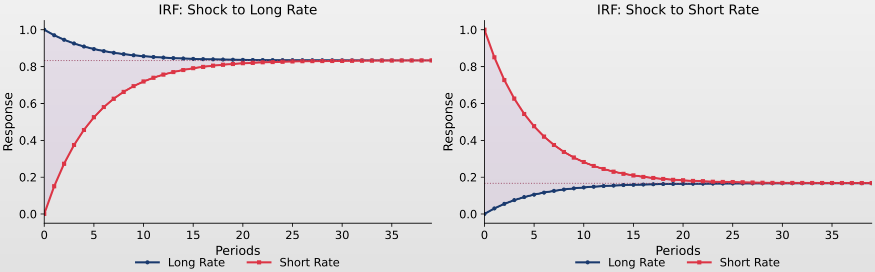



TSA_ch7_case_vecm

# Pasul 4: Funcții de răspuns la impuls

## Interpretare

- **Efecte permanente**: Șocuri în rata pe termen lung afectează persistent ambele rate
- **Cointegrare**: Efectele nu converg la zero ⇒ caracteristică a seriilor cointegrate

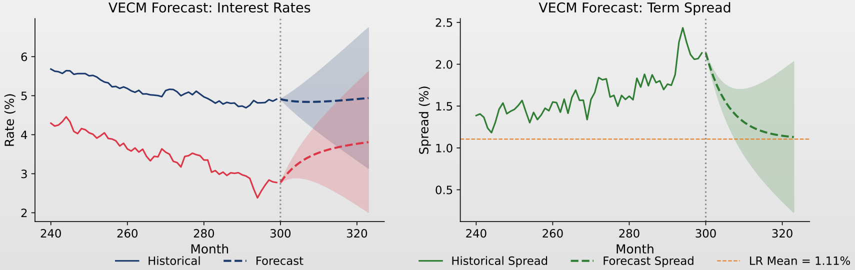



TSA_ch7_case_irf

# Pasul 5: Prognoza VECM

## Prognoză

- **Orizont**: 24 de luni pentru ambele rate simultan
- **Avantaj**: VECM menține relația de cointegrare în prognoză




TSA_ch7_case_forecast

# Exercițiu AI: Gândire critică

**Prompt de testat în ChatGPT / Claude / Copilot**

"Folosind yfinance, descarcă prețurile zilnice de închidere ale aurului (GC=F) și argintului (SI=F) din 2019-01-01 până în 2024-12-31 (aprox. 1.500 observații). Testează dacă fiecare serie este I(1), testează cointegrarea cu Engle-Granger și Johansen, și estimează un model VECM. Analizează parametrii de viteză de ajustare. Vreau cod Python complet."

**Exercițiu**:

1. Rulați prompt-ul într-un LLM (Large Language Model) la alegere și analizați critic răspunsul.
2. Verifică dacă fiecare serie este I(1) înainte de a testa cointegrarea?
3. Folosește atât Engle-Granger cât și Johansen? Care sunt avantajele fiecăruia?
4. Cum determină rangul de cointegrare? Statistica Trace vs max-eigenvalue?
5. Interpretează corect coeficienții $\alpha$ (de ajustare)?

- **Atenție**: Codul generat de AI poate rula fără erori și arăta profesional
- *Asta nu înseamnă că e corect*

# Concluzii principale

## Concepte principale

- **Cointegrare**: Variabile $I(1)$ cu combinație liniară staționară (trend stochastic comun)
- **Regresie falsă**: $R^2$ ridicat cu variabile $I(1)$ necorelate
- **Corecție a erorilor**: VECM $\Rightarrow$ VAR cu termeni de corecție pentru sisteme cointegrate

## Metode de testare

- **Engle-Granger**: Simplu, doi pași, un singur vector de cointegrare
- **Johansen**: Vectori multipli, bazat pe MLE (teste trace + max-eigen)

## De reținut

- **Putere scăzută**: Testele au putere scăzută în eșantioane mici
- **Teorie**: Teoria economică ar trebui să ghideze specificația
- **Validare**: Verificați întotdeauna validitatea modelului.

# Extensie: Cointegrare cu ruptură structurală

## Limitare a testelor standard

Testele Engle-Granger și Johansen presupun **stabilitatea** relației de cointegrare pe întreaga perioadă. Dacă relația se schimbă (e.g., criză financiară, schimbare de regim), aceste teste **pierd putere**.

## Testul Gregory-Hansen (1996)

- Testează $H_0$: **fără cointegrare** vs. $H_1$: cointegrare cu **o ruptură structurală** la un moment necunoscut $\tau$
- Trei variante:
  - **C**: schimbare de nivel (constantă)
  - **C/T**: schimbare de nivel și trend
  - **C/S**: schimbare de regim (toți coeficienții)
- **Procedură**: Se calculează statistica ADF pentru fiecare punct de ruptură posibil $\tau \in [0{,}15T,\ 0{,}85T]$; se raportează $\inf_\tau \text{ADF}(\tau)$
- Valorile critice: tabelate de Gregory & Hansen (1996)

## Aplicații tipice

Paritatea puterii de cumpărare (PPP) cu rupturi la schimbări de regim valutar, relația Fisher cu schimbări de politică monetară

# Cointegrare cu prag (Threshold Cointegration)

TSA ch7 threshold ecm

## Motivație

Modelele liniare ECM presupun **viteza de ajustare constantă** indiferent de mărimea sau semnul deviației de la echilibru. În realitate, ajustarea poate fi **asimetrică**: rapidă când deviația este mare, lentă sau inexistentă când este mică (costuri de tranzacție, rigidități).

## Modelul Threshold ECM (Balke & Fomby, 1997)

Fie $z_t = Y_t - \hat{\beta}' X_t$ eroarea de echilibru. Modelul cu prag:

$$\Delta Y_t = \rho_1 \, z_{t-1} \, \mathbb{I}(z_{t-1} \leq \gamma) + \rho_2 \, z_{t-1} \, \mathbb{I}(z_{t-1} > \gamma) + \varepsilon_t$$

- $\gamma$: **pragul** (estimat endogen); $\mathbb{I}(\cdot)$: funcția indicator
- **Regimul 1** ($z_{t-1} \leq \gamma$): ajustare lentă sau inactivă ($\rho_1 \approx 0$)
- **Regimul 2** ($z_{t-1} > \gamma$): ajustare rapidă ($\rho_2 < 0$, semnificativ)

## Testare și aplicații

- **Testul sup-LM** (Hansen & Seo, 2002): $H_0$: ECM liniar vs. $H_1$: threshold ECM
- Aplicații: transmiterea prețurilor pe piețele de mărfuri, legea prețului unic, asimetria pe piețele energetice

# Modelul cu tranziție lină (STECM)

TSA ch7 stecm

## Smooth Transition ECM (Kapetanios, Shin & Snell, 2006)

Generalizare a modelului cu prag prin **tranziție continuă** între regimuri:

$$\Delta Y_t = \rho\, z_{t-1}\, G(z_{t-d};\, \gamma,\, c) + \varepsilon_t$$

cu funcția de tranziție **logistică**:

$$G(z;\, \gamma,\, c) = \left[1 + \exp\left(-\gamma(z - c)\right)\right]^{-1}, \quad \gamma > 0$$

## Cazuri speciale și variante

- $\gamma \to 0$: $G \to 0{,}5$ constant $\Rightarrow$ ECM **liniar** (caz particular)
- $\gamma \to \infty$: $G \to \mathbb{I}(z > c)$ $\Rightarrow$ model cu **prag** (salt discret)
- **ESTAR** (Exponential STAR): $G(z) = 1 - \exp(-\gamma\, z^2)$ — ajustare **simetrică**: rapidă pentru deviații mari (ambele semne), lentă în apropierea echilibrului

## Testul KSS (Kapetanios, Shin & Snell, 2003)

Test de rădăcină unitară neliniară, pre-condiție pentru STECM. Testează $H_0$: rădăcină unitară vs. $H_1$: proces ESTAR staționar global. Bazat pe aproximare Taylor de ordinul I a funcției de tranziție.

# Cointegrare neliniară: sinteză

TSA_ch7_nonlinear_cointegration

## Comparație: ECM liniar vs. Threshold vs. STECM

| Criteriu | ECM liniar | Threshold ECM | STECM |
|---|---|---|---|
| Ajustare | simetrică | asimetrică | graduală |
| Tranziție | — | discretă (salt) | continuă (lină) |
| Nr. regimuri | 1 | 2 (sau mai multe) | continuum |
| Viteză ajustare | constantă $\rho$ | $\rho_1 \neq \rho_2$ | $\rho \cdot G(z)$ variabilă |
| Parametru cheie | $\rho$ | $\gamma$ (prag) | $\gamma$ (viteză tranziție) |

## Când să utilizați modele neliniare?

- **Pas 1**: Estimați ECM liniar; testați reziduurile cu **testul BDS** sau **RESET** pentru neliniaritate
- **Pas 2**: Dacă se respinge liniaritatea, estimați modelul cu prag (test sup-LM) sau STECM
- **Pas 3**: Comparați prin criterii informaționale (AIC, BIC) și putere predictivă out-of-sample

## Instrumente software

- **R**: pachetele `tsDyn` (TVECM, SETAR), `apt` (threshold cointegration), `nardl` (NARDL)
- **Python**: `statsmodels` (ECM liniar) + implementare proprie STECM; `arch` pentru teste neliniare

## Ce urmează?

### Extensii și subiecte conexe

- **VECM Structural**: Identificarea șocurilor structurale
- **Cointegrare cu prag**: Ajustare neliniară
- **Cointegrare de panel**: Secțiuni transversale multiple
- **Cointegrare fracționară**: Memorie lungă
- **Cointegrare variabilă în timp**: Schimbări de regim

- **Întrebări?**

# Formule principale – Rezumat

## Cointegrare

- **Definiție**: $Y_t - \beta X_t = u_t \sim I(0)$
- **Interpretare**: Echilibru pe termen lung

## Test Engle-Granger

- **Pas 1**: $Y_t = \alpha + \beta X_t + u_t$
- **Pas 2**: Test ADF pe $\hat{u}_t$
- **Notă**: Valori critice speciale

## Rang de cointegrare

- **Rangul** $r$: $0 \leq r \leq K - 1$ relații

## Model VECM

- **Ecuație**: $\Delta \mathbf{Y}_t = \mathbf{\Pi}\mathbf{Y}_{t-1} + \sum_{i=1}^{p-1} \mathbf{\Gamma}_i \Delta \mathbf{Y}_{t-i} + \varepsilon_t$
- **Factorizare**: $\mathbf{\Pi} = \boldsymbol{\alpha}\boldsymbol{\beta}'$

## Interpretare $\boldsymbol{\alpha}$ și $\boldsymbol{\beta}$

- $\boldsymbol{\beta}$: Vectori de cointegrare
- $\boldsymbol{\alpha}$: Viteza de ajustare

## Test Johansen

- **Trace**: $\lambda_{trace} = -T \sum_{i=r+1}^{K} \ln(1 - \hat{\lambda}_i)$
- **Max-Eigen**: $\lambda_{max} = -T \ln(1 - \hat{\lambda}_{r+1})$

# Întrebarea 1

## Întrebare

- Analizați cele două perechi de serii I(1) de mai jos. Care pereche este cointegrată?

## Variante de răspuns

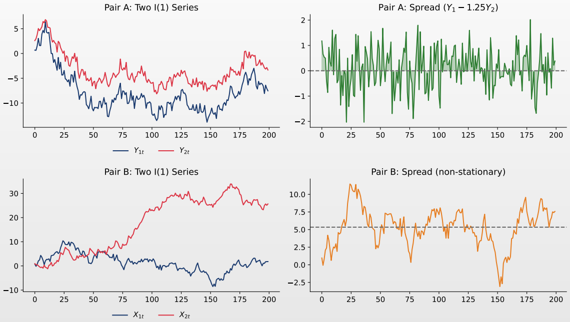
**(A)** Perechea A, deoarece seriile au aceeași tendință

**(B)** Perechea B, deoarece seriile sunt necorelate

**(C)** Perechea A, deoarece spread-ul lor este staționar

**(D)** Ambele perechi sunt cointegrate

# Întrebarea 1: Răspuns




**Răspuns: (C)**

- Cointegrare = combinație liniară staționară, nu doar corelație
- Spread-ul perechii B nu este staționar ⇒ nu sunt cointegrate

TSA_ch7_quiz1_cointegration_concept

# Întrebarea 2

## Întrebare

Ambele regresii de mai jos au $R^2$ ridicat. Cum puteți distinge o regresie falsă de una genuină?

## Variante de răspuns

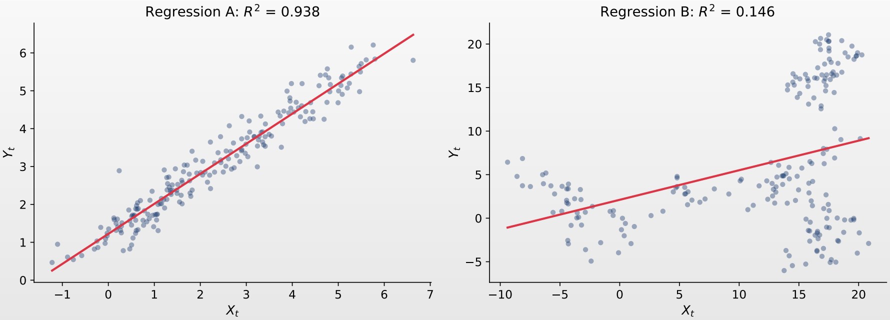
**(A)** Nu se poate distinge – ambele au $R^2$ ridicat

**(B)** Testând staționaritatea reziduurilor: reziduuri staționare = cointegrare genuină

**(C)** Verificând semnificația coeficientului $\beta$

**(D)** Comparând valorile lui $R^2$: mai mare = relație mai reală

# Întrebarea 2: Răspuns




## Răspuns: (B)

- Testul Engle-Granger: dacă reziduurile OLS sunt staționare (ADF), relația este genuină
- $R^2$ ridicat *nu* implică relație reală între variabile $I(1)$.

TSA_ch7_quiz2_spurious_regression

## Întrebarea 3

### Întrebare

În care scenariu rata pe termen lung este slab exogenă (nu se ajustează la dezechilibru)?

### Variante de răspuns

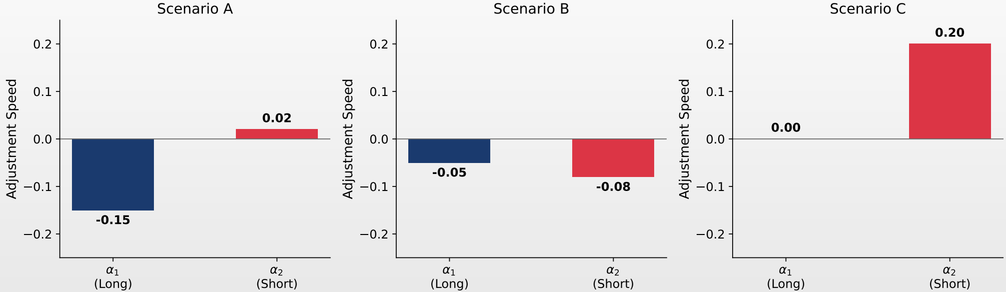
**(A)** Scenariul A: $\alpha_1 = -0.15$, $\alpha_2 = 0.02$

**(B)** Scenariul B: $\alpha_1 = -0.05$, $\alpha_2 = -0.08$

**(C)** Scenariul C: $\alpha_1 = 0.00$, $\alpha_2 = 0.20$

**(D)** Niciun scenariu – ambele variabile trebuie să se ajusteze

# Întrebarea 3: Răspuns




**Răspuns: (C)**

- $\alpha_1 = 0$: rata pe termen lung nu răspunde la dezechilibru (slab exogenă)
- Toată ajustarea este făcută de rata pe termen scurt ($\alpha_2 = 0.20$)

TSA_ch7_quiz3_error_correction

# Întrebarea 4

## Întrebare

- Având rezultatele testului Johansen Trace pentru $K = 3$ variabile, care este rangul de cointegrare?

## Variante de răspuns

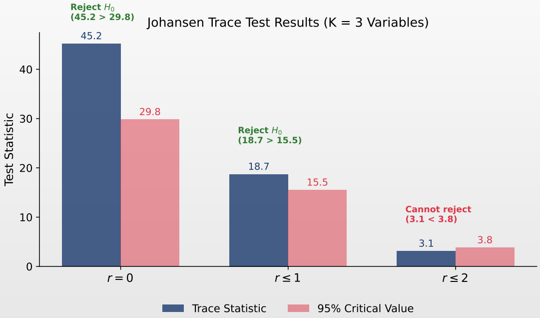
**(A)** $r = 0$ (nicio relație de cointegrare)

**(B)** $r = 1$ (o relație de cointegrare)

**(C)** $r = 2$ (două relații de cointegrare)

**(D)** $r = 3$ (sistem complet staționar)

# Întrebarea 4: Răspuns




## Răspuns: (C)

- Respingem $H_0 : r = 0$ (45.2 > 29.8) și $H_0 : r \leq 1$ (18.7 > 15.5)
- Nu respingem $H_0 : r \leq 2$ (3.1 < 3.8) $\Rightarrow$ rangul este $r = 2$

TSA_ch7_quiz4_johansen_rank

## Întrebarea 5

### Întrebare

- Care este avantajul principal al VECM față de VAR în diferențe pentru prognoză?

### Variante de răspuns

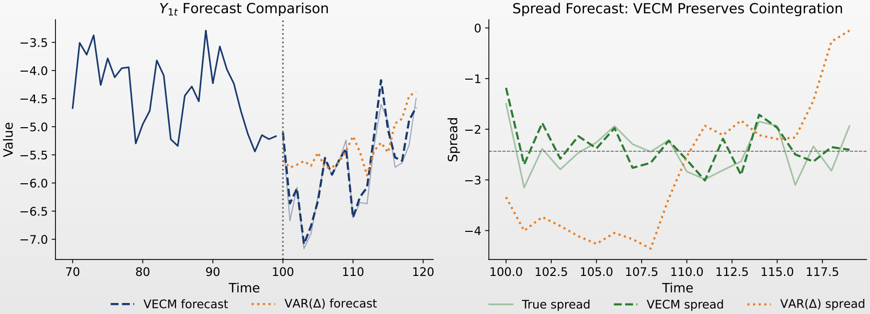
**(A)** VECM are mai puțini parametri de estimat

**(B)** VECM menține relația de cointegrare în prognoze pe termen lung

**(C)** VAR în diferențe nu poate produce prognoze

**(D)** Nu există avantaj – ambele sunt echivalente

# Întrebarea 5: Răspuns




## Răspuns: (B)

- VAR(Δ) pierde relația de nivel ⇒ spread-ul diverge
- VECM încorporează echilibrul pe termen lung ⇒ prognoza rămâne coerentă

TSA_ch7_quiz5_vecm_forecast

# Bibliografie I

## Lucrări fundamentale cointegrare


- Engle, R.F., & Granger, C.W.J. (1987). Co-Integration and Error Correction: Representation, Estimation, and Testing, *Econometrica*, 55(2), 251–276.
- Johansen, S. (1988). Statistical Analysis of Cointegration Vectors, *Journal of Economic Dynamics and Control*, 12(2-3), 231–254.
- Johansen, S. (1991). Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models, *Econometrica*, 59(6), 1551–1580.


## Manuale VECM și cointegrare


- Juselius, K. (2006). *The Cointegrated VAR Model: Methodology and Applications*, Oxford University Press.
- Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*, Springer.

# Bibliografie II

## Teste și aplicații


- Phillips, P.C.B., & Ouliaris, S. (1990). Asymptotic Properties of Residual Based Tests for Cointegration, *Econometrica*, 58(1), 165–193.
- Hamilton, J.D. (1994). *Time Series Analysis*, Princeton University Press.
- Banerjee, A., Dolado, J.J., Galbraith, J.W., & Hendry, D.F. (1993). *Co-Integration, Error-Correction, and the Econometric Analysis of Non-Stationary Data*, Oxford University Press.


## Resurse online și cod

- **Quantlet**: `https://quantlet.com` – Platformă de cod pentru metode cantitative
- **Quantinar**: `https://quantinar.com` – Platformă de învățare pentru metode cantitative
- **GitHub TSA**: `https://github.com/QuantLet/TSA/tree/main/TSA_ch7` – Cod Python pentru acest capitol

# Vă Mulțumim!

Întrebări?

Materialele cursului sunt disponibile la: https://danpele.github.io/Time-Series-Analysis/

Quantlet Quantinar