Χρονοσειρές - Μάθημα 4

System is a set of interacting or interdependent components forming an integrated whole. Fields that study the general properties of systems include systems theory, cybernetics, dynamical systems, thermodynamics and complex systems. They investigate the abstract properties of systems' matter and organization, looking for **concepts** and **principles** that are independent of domain, substance, type, or temporal scale.

A **stochastic process** is the counterpart to a **deterministic process** (or **deterministic system**). Instead of dealing with **only one possible way the process might develop over time** (as in the case, for example, of solutions of an ordinary differential equation), in a stochastic or random process **there is some indeterminacy described by probability distributions**. This means that even if the initial condition (or starting point) is known, there are many possibilities the process might go to, but some paths may be more probable and others less so.

A scientific **model** seeks to **represent** empirical objects, phenomena, and **physical processes** in a logical and objective way. All models are in simulacra, that is, simplified reflections of reality, but, despite their inherent falsity, they are nevertheless extremely useful. Complete and true representation may be impossible, but scientific debate often concerns **which is the better model for a given task**, e.g., which is the more accurate climate model for seasonal forecasting.

Models for time series data can have many forms and represent different stochastic processes.

πηγή: Wikipedia

Σχέση μεταξύ AR και MA διαδικασιών

Αυτοπαλινδρομούμενη διαδικασία τάξης p, AR(p)

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + Z_t \qquad Z_t \sim \mathrm{WN}(0, \sigma_Z^2)$$

$$(1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p) X_t = Z_t$$

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p$$

$$\phi(B) X_t = Z_t$$

⇓ AR(p) στάσιμη

$$X_t = [\phi(B)]^{-1} Z_t$$

⇓ $\psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \ldots$ τέτοιο ώστε $\phi(B)\psi(B) = 1$

$X_t = \psi(B) Z_t$ **MA(∞)**

Διαδικασία κινούμενου μέσου τάξης q, MA(q)

$$X_t = Z_t - \theta_1 Z_{t-1} - \theta_2 Z_{t-2} - \cdots - \theta_q Z_{t-q}$$

$$X_t = (1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q) Z_t$$

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q$$

$$X_t = \theta(B) Z_t$$

⇓ MA(q) αντιστρέψιμη

$$[\theta(B)]^{-1} X_t = Z_t$$

⇓ $\pi(B) = 1 - \pi_1 B - \pi_2 B^2 + \ldots$ τέτοιο ώστε $\pi(B)\theta(B) = 1$

$\pi(B) X_t = Z_t$ **AR(∞)**

AR(p) ↔ MA(∞)

AR(p) και MA(q) έχουν δυϊκή σχέση

MA(q) ↔ AR(∞)

Wold's decomposition (1)
every covariance-stationary time series can be written as an infinite moving average (MA(∞)) process of its innovation process.

Η αυτοσυσχέτιση και μερική αυτοσυσχέτιση των AR(p) και MA(q) έχουν επίσης δυϊκή σχέση

AR(p): ρ_τ φθίνει εκθετικά προς 0, $\phi_{\tau\tau}$ μηδενίζεται για $\tau > p$
MA(q): $\phi_{\tau\tau}$ φθίνει εκθετικά προς 0, ρ_τ μηδενίζεται για $\tau > q$

Μικτή διαδικασία ARMA(p,q)

Αυτοπαλινδρομούμενη διαδικασία τάξης p, AR(p)

Διαδικασία κινούμενου μέσου τάξης q, MA(q)

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + Z_t \qquad Z_t \sim \mathrm{WN}(0,\sigma_Z^2) \qquad X_t = Z_t - \theta_1 Z_{t-1} - \theta_2 Z_{t-2} - \cdots - \theta_q Z_{t-q}$$

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + Z_t - \theta_1 Z_{t-1} - \theta_2 Z_{t-2} - \cdots - \theta_q Z_{t-q}$$

$$X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} - \cdots - \phi_p X_{t-p} = Z_t - \theta_1 Z_{t-1} - \theta_2 Z_{t-2} - \cdots - \theta_q Z_{t-q}$$

$$\phi(B) X_t = \theta(B) Z_t$$

$$X_t = \frac{\theta(B)}{\phi(B)} Z_t \qquad \frac{\phi(B)}{\theta(B)} X_t = Z_t$$

Στασιμότητα: καθορίζεται από το AR μέρος

Αντιστρεψιμότητα: καθορίζεται από το MA μέρος

Αυτοσυσχέτιση:

$$X_{t-\tau} X_t = X_{t-\tau}(\phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + Z_t - \theta_1 Z_{t-1} - \theta_2 Z_{t-2} - \cdots - \theta_q Z_{t-q})$$

$$\gamma_X(\tau) = \phi_1 \gamma_X(\tau-1) + \cdots + \phi_p \gamma_X(\tau-p) + \mathrm{E}[X_{t-\tau} Z_t] - \theta_1 \mathrm{E}[X_{t-\tau} Z_{t-1}] - \cdots - \theta_q \mathrm{E}[X_{t-\tau} Z_{t-q}]$$

Για $\tau > q$ $\quad \gamma_\tau = \phi_1 \gamma_{\tau-1} + \cdots + \phi_p \gamma_{\tau-p} \qquad \rho_\tau = \phi_1 \rho_{\tau-1} + \cdots + \phi_p \rho_{\tau-p}$ $\quad$ όπως για AR(p)

Για $\tau \le q$ $\quad$ μίξη της αυτοσυσχέτισης για AR(p) και MA(q)

Μικτή διαδικασία ARMA(1,1)

$$X_t = \phi X_{t-1} + Z_t - \theta Z_{t-1} \qquad (1-\phi B)X_t = (1-\theta B)Z_t \qquad X_t = \frac{(1-\theta B)}{(1-\phi B)} Z_t$$

Συνθήκη στασιμότητας: $|\phi| < 1$

Συνθήκη αντιστρεψιμότητας: $|\theta| < 1$

Αυτοσυσχέτιση:

$$X_{t-\tau} X_t = X_{t-\tau}(\phi X_{t-1} + Z_t - \theta Z_{t-1})$$

$$\gamma_\tau = \phi \gamma_{\tau-1} + \mathrm{E}[X_{t-\tau} Z_t] - \theta \mathrm{E}[X_{t-\tau} Z_{t-1}]$$

$\tau = 0 \qquad \sigma_X^2 = \gamma_0 = \phi\gamma_1 + \sigma_Z^2 - \theta(\phi - \theta)\sigma_Z^2$

$\tau = 1 \qquad \gamma_1 = \phi\gamma_0 - \theta\sigma_Z^2 \qquad \boxed{?}$

$$\gamma_0 = \frac{1+\theta^2 - 2\phi\theta}{1-\phi^2}\sigma_Z^2$$

$$\gamma_1 = \frac{(\phi-\theta)(1-\phi\theta)}{1-\phi^2}\sigma_Z^2$$

$\tau > 1 \qquad \gamma_\tau = \phi\gamma_{\tau-1}$ όπως για AR(1)

$$\rho_\tau = \begin{cases} \dfrac{(\phi-\theta)(1-\phi\theta)}{1+\theta^2 - 2\phi\theta} & \tau = 1 \\ \phi\rho_{\tau-1} & \tau \geq 2 \end{cases}$$

11 **Συνθήκες αντιστρεψιμότητας και στασιμότητας για τους συντελεστές φ και θ, καθώς και για τις αυτοσυσχετίσεις και μερικές αυτοσυσχετίσεις διαδικασίας ARMA(1,1)**

Μερική αυτοσυσχέτιση $\phi_{\tau\tau}$: Φθίνει με την υστέρηση όπως για MA(1)

Μια ARMA(p,q) διαδικασία με μικρά p,q, παρουσιάζει συσχετίσεις (ρ_τ και $\phi_{\tau\tau}$) που επιτυγχάνονται από AR(p) μόνο για μεγάλη τάξη p, ή από MA(q) μόνο για μεγάλη τάξη q

Εκτίμηση μοντέλων τύπου AR, MA, ARMA

(στάσιμη) χρονοσειρά
(στοχαστική διαδικασία) $\{X_t\}_{t=-\infty}^{\infty}$

(στάσιμη) χρονοσειρά
n παρατηρήσεων $\{x_1, x_2, \ldots, x_n\}$

μέση τιμή μ

Αυτοδιασπορά

$$\gamma_\tau = \gamma(\tau) = \mathrm{E}\left[(X_t - \mu)(X_{t-\tau} - \mu)\right]$$

$$= \mathrm{E}\left[X_t X_{t-\tau}\right] - \left(\mathrm{E}\left[X_t\right]\right)^2$$

Αυτοσυσχέτιση

$$\rho_\tau = \rho(\tau) = \frac{\gamma(\tau)}{\gamma(0)}$$

Δειγματική μέση τιμή $\bar{x} = \dfrac{1}{n}\sum_{t=1}^{n} x_t$

Δειγματική αυτοδιασπορά

$$c_\tau = c(\tau) = \frac{1}{n}\sum_{t=\tau+1}^{n}(x_t x_{t-\tau} - \bar{x}^2)$$

$$\tau = 0, 1, \ldots, n-1$$

Δειγματική αυτοσυσχέτιση

$$r_\tau = r(\tau) = \frac{c(\tau)}{c(0)}$$

στοχαστική διαδικασία AR(p)

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + Z_t$$

$$Z_t \sim \mathrm{WN}(0, \sigma_Z^2)$$

στοχαστική διαδικασία MA(q)

$$X_t = Z_t - \theta_1 Z_{t-1} - \theta_2 Z_{t-2} - \cdots - \theta_q Z_{t-q}$$

στοχαστική διαδικασία ARMA(p,q)

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + Z_t$$

$$- \theta_1 Z_{t-1} - \theta_2 Z_{t-2} - \cdots - \theta_q Z_{t-q}$$

Εκτίμηση διαδικασίας (μοντέλο)

- AR, MA ή ARMA **?** **άλλο μοντέλο ?**
- τάξη p ή/και q **?**
- εκτίμηση παραμέτρων μοντέλου **?**

$\mathrm{AR}(p): \phi_1, \phi_2, \ldots, \phi_p, \sigma_Z^2$

$\mathrm{MA}(q): \theta_1, \theta_2, \ldots, \theta_q, \sigma_Z^2$

$\mathrm{ARMA}(p,q): \phi_1, \phi_2, \ldots, \phi_p, \theta_1, \theta_2, \ldots, \theta_q, \sigma_Z^2$

$\boxed{?}$

Εκτίμηση μοντέλου AR(p)

Υποθέτω στοχαστική διαδικασία AR(p) για τη χρονοσειρά $\{x_1, x_2, \ldots, x_n\}$

Προσαρμογή διαδικασίας (μοντέλο) AR(p) $\Rightarrow$ εκτίμηση παραμέτρων $\phi_1, \phi_2, \ldots, \phi_p, \sigma_Z^2$

Μέθοδος ροπών (μέθοδος Yule-Walker, YW)

Εκτίμηση των παραμέτρων από τις δειγματικές αυτοσυσχετίσεις

$$r_1, r_2, \ldots, r_p, s_X^2 \quad \Rightarrow \quad \hat{\phi}_1, \hat{\phi}_2, \ldots, \hat{\phi}_p, s_Z^2$$

Εξισώσεις Yule-Walker

$$\begin{bmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{p-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{p-2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho_{p-1} & \rho_{p-2} & \rho_{p-3} & \cdots & 1 \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_p \end{bmatrix} = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_p \end{bmatrix}$$

$$\boldsymbol{P}_p \boldsymbol{\phi} = \boldsymbol{\rho}_p$$

$$\sigma_X^2 = \frac{\sigma_Z^2}{1 - \phi_1\rho_1 - \phi_2\rho_2 - \cdots - \phi_p\rho_p}$$

Εκτίμηση των $\rho_1, \rho_2, \ldots, \rho_p, \sigma_X^2 \Rightarrow r_1, r_2, \ldots, r_p, s_X^2$ και αντικατάσταση …

$$\begin{bmatrix} 1 & r_1 & r_2 & \cdots & r_{p-1} \\ r_1 & 1 & r_1 & \cdots & r_{p-2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ r_{p-1} & r_{p-2} & r_{p-3} & \cdots & 1 \end{bmatrix} \begin{bmatrix} \hat{\phi}_1 \\ \hat{\phi}_2 \\ \vdots \\ \hat{\phi}_p \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_p \end{bmatrix}$$

$$\boldsymbol{R}_p \hat{\boldsymbol{\phi}} = \boldsymbol{r}_p$$

$$s_X^2 = \frac{s_Z^2}{1 - \hat{\phi}_1 r_1 - \hat{\phi}_2 r_2 - \cdots - \hat{\phi}_p r_p}$$

$$\Rightarrow \quad \hat{\boldsymbol{\phi}} = \boldsymbol{R}_p^{-1} \boldsymbol{r}_p \qquad s_Z^2 = s_X^2 (1 - \hat{\phi}_1 r_1 - \hat{\phi}_2 r_2 - \cdots - \hat{\phi}_p r_p)$$

$\{x_1, x_2, \ldots, x_n\}$ με μέση τιμή μ

Γενική μορφή μοντέλου AR(p) $X_t - \mu = \phi_1(X_{t-1}-\mu) + \phi_2(X_{t-2}-\mu) + \cdots + \phi_p(X_{t-p}-\mu) + Z_t$

Μέθοδος ελαχίστων τετραγώνων (ordinary least squares, OLS)

Προσαρμογή μοντέλου AR(p) στα δεδομένα

Ελαχιστοποίηση αθροίσματος τετραγώνων των σφαλμάτων προσαρμογής

$$\min S(\mu,\phi_1,\ldots\phi_p) = \min \sum_{t=p+1}^{n} \left(x_t - \mu - \phi_1(x_{t-1}-\mu) - \ldots - \phi_p(x_{t-p}-\mu)\right)^2$$ ως προς $\mu,\phi_1,\phi_2,\ldots,\phi_p$

$$\hat{z}_t = (x_t - \hat{\mu}) - \hat{\varphi}_1(x_{t-1}-\hat{\mu}) - \ldots - \hat{\varphi}_p(x_{t-p}-\hat{\mu}), \quad t = p+1,\ldots,n$$

$$\hat{\mu},\hat{\phi}_1,\hat{\phi}_2,\ldots,\hat{\phi}_p \quad s_Z^2 = \frac{1}{n-p}\sum_{t=p+1}^{n}\hat{z}_t^2$$

$$\hat{\mu} \simeq \bar{x} = \frac{1}{n}\sum_{t=1}^{n} x_t$$

AR(1) $X_t - \mu = \phi_1(X_{t-1}-\mu) + Z_t$

$$S(\mu,\phi_1) = \sum_{t=2}^{n}\left(x_t - \mu - \phi_1(x_{t-1}-\mu)\right)^2$$

$$\hat{\mu} = \frac{\bar{x}_{(2)} - \hat{\varphi}\bar{x}_{(1)}}{1-\hat{\varphi}} \qquad \bar{x}_{(1)} = \frac{1}{n-1}\sum_{t=2}^{n} x_{t-1} \qquad \bar{x}_{(2)} = \frac{1}{n-1}\sum_{t=2}^{n} x_t \qquad \hat{\mu} \simeq \bar{x}$$

$$c_1 = \frac{1}{n}\sum_{t=2}^{n}(x_t-\bar{x})(x_{t-1}-\bar{x})$$

$$\hat{\phi} = \frac{\sum_{t=2}^{n}(x_t-\hat{\mu})(x_{t-1}-\hat{\mu})}{\sum_{t=2}^{n}(x_t-\hat{\mu})^2} \simeq \frac{\sum_{t=2}^{n}(x_t-\bar{x})(x_{t-1}-\bar{x})}{\sum_{t=2}^{n}(x_t-\bar{x})^2}$$

$$c_0 = \frac{1}{n}\sum_{t=1}^{n}(x_t-\bar{x})^2$$

$$\hat{\varphi} \simeq \frac{c_1}{c_0} = r_1$$

Άλλες μέθοδοι εκτίμησης παραμέτρων AR(p)

- Προς τα εμπρός και πίσω (backward – forward approach, FB)

- Αλγόριθμος του Burg (Burg)

- Μέθοδος μέγιστης πιθανοφάνειας (maximum likelihood, ML)
  - υπό συνθήκη (conditioned)
  - χωρίς συνθήκη (unconditioned)

ML εκτίμηση είναι η βέλτιστη, οι άλλες μέθοδοι την προσεγγίζουν

Η OLS και ML συμπίπτουν όταν η χρονοσειρά είναι από Γκαουσιανή διαδικασία

Ασυμπτωτικά (για μεγάλο n) συγκλίνουν οι εκτιμήσεις με όλες τις μεθόδους

Η YW έχει την πιο αργή σύγκλιση στη ML

12 **Ο αλγόριθμος Burg για την εκτίμηση παραμέτρων μοντέλου AR(p)**

13 **Ο αλγόριθμος μέγιστης πιθανοφάνειας για την εκτίμηση παραμέτρων μοντέλου AR(p)**

Προσδιορισμός τάξης p μοντέλου AR

κριτήριο μερικής αυτοσυσχέτισης

μερική αυτοσυσχέτιση για υστέρηση τ: συσχέτιση των x_t, $x_{t-\tau}$

χωρίς τη συσχέτιση με $x_t, \underbrace{x_{t-1}, \ldots, x_{t-\tau+1}}, x_{t-\tau}$

$$x_t = \phi_{1,1} x_{t-1} + z_t$$

$$x_t = \phi_{1,2} x_{t-1} + \phi_{2,2} x_{t-2} + z_t$$

$$x_t = \phi_{1,3} x_{t-1} + \phi_{2,3} x_{t-2} + \phi_{3,3} x_{t-3} + z_t$$

εκτίμηση του $\phi_{\tau\tau}$ για μοντέλο AR(τ)

Η τάξη είναι p αν $\hat{\phi}_{p,p} \neq 0$ και $\hat{\phi}_{\tau,\tau} = 0$ για $\tau > p$

(πτώση από μη-μηδενική σε μηδενική μερική αυτοσυσχέτιση)

κριτήρια με βάση το σφάλμα προσαρμογής

- κριτήριο πληροφορίας του Akaike
  Akaike information criterion (AIC)

$$\mathrm{AIC}(p) = \ln(s_z^2) + \frac{2p}{n}$$

- κριτήριο Μπεϋζιανής πληροφορίας (Scwartz)
  Bayesian information criterion (BIC)

$$\mathrm{BIC}(p) = \ln(s_z^2) + \frac{p\ln(n)}{n}$$

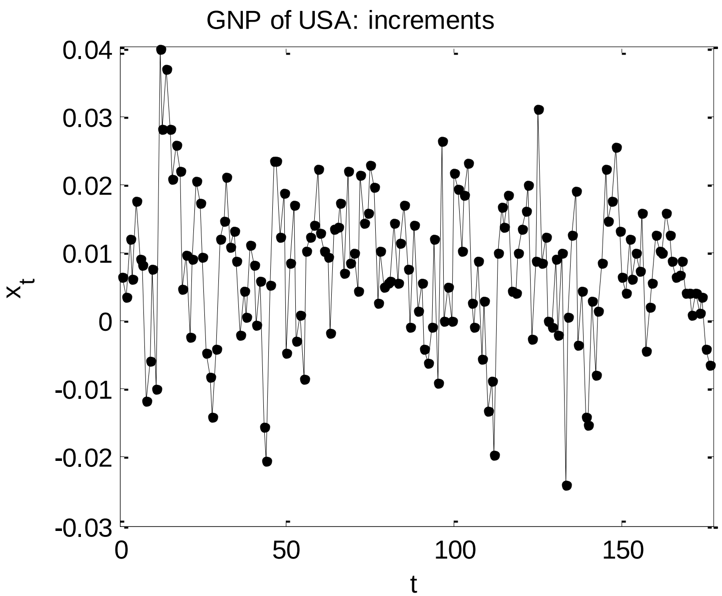
- κριτήριο τελικού σφάλματος πρόβλεψης
  Final prediction error (FPE)

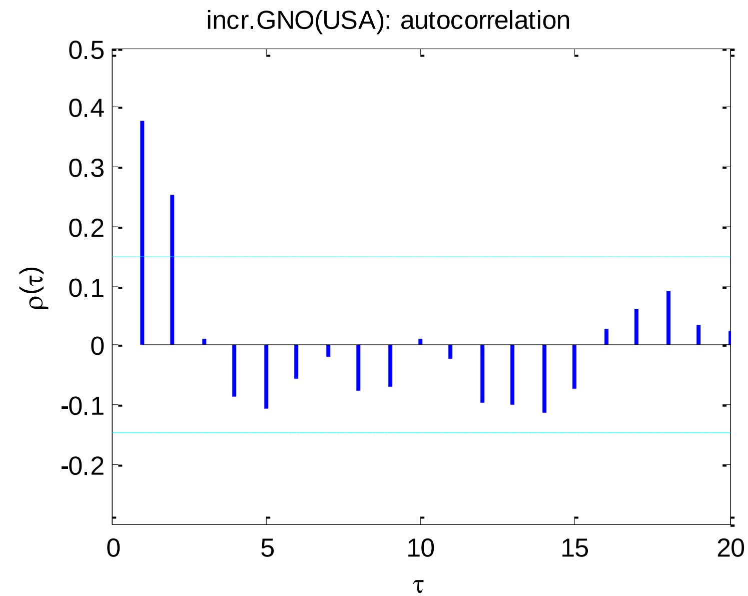
$$\mathrm{FPE}(p) = s_z^2 \frac{n+p}{n-p}$$

14 Ιδιότητες και σύγκριση των κριτηρίων πληροφορίας AIC, BIC και FPE

Παράδειγμα

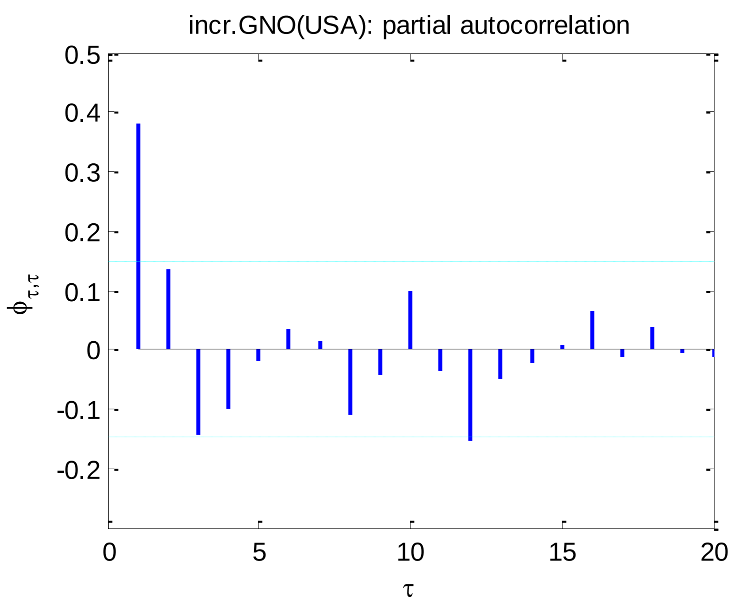
Ρυθμός μεταβολής του ακαθάριστου εθνικού προϊόντος (ΑΕΠ) των ΗΠΑ (τετραμηνιαίες τιμές, 2^o τετράμηνο 1947 – 1^o τετράμηνο 1991). Η εποχικότητα έχει διορθωθεί (αφαιρώντας τον εποχικό κύκλο).

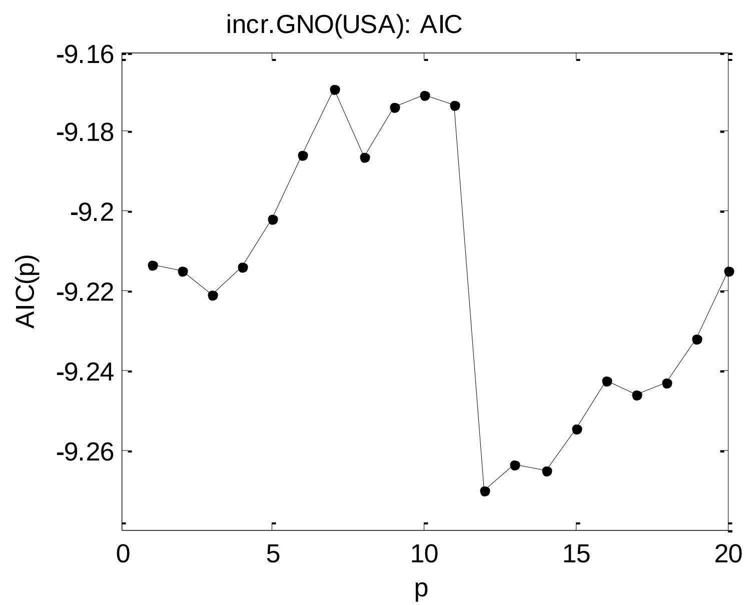




στάσιμη ?

συσχετίσεις?





τάξη
AR μοντέλου ?

AR(3) ?

εκτίμηση παραμέτρων

$$x_t - \hat{\mu} \to x_t \qquad \hat{\mu} = 0.0077$$

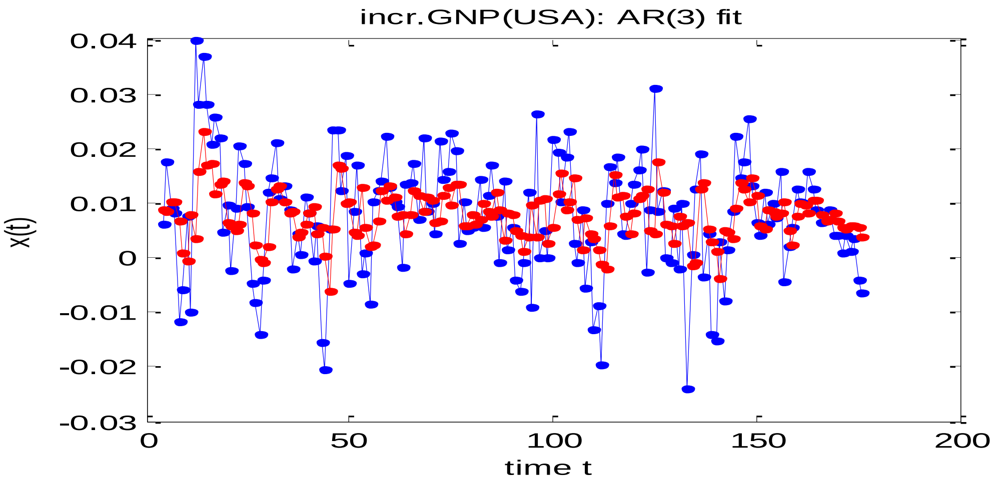
OLS → $\hat{\phi}_1 = 0.35 \quad \hat{\phi}_2 = 0.18 \quad \hat{\phi}_3 = -0.14$

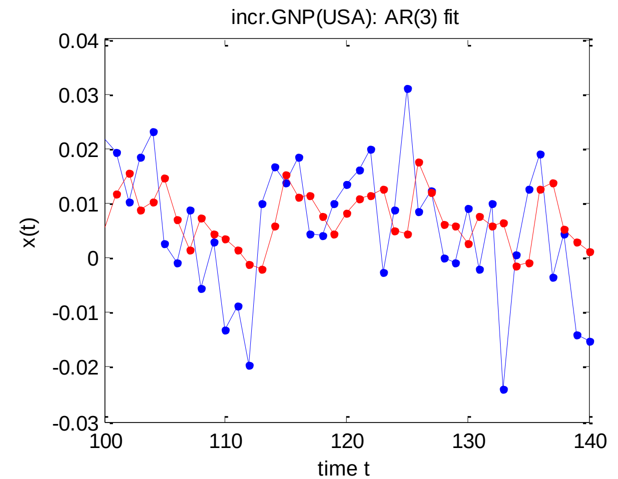
$$\hat{\phi}_0 = \hat{\mu}\left(1 - \hat{\phi}_1 - \hat{\phi}_2 - \hat{\phi}_3\right) = 0.0047$$

εκτίμηση $\hat{x}_t = 0.0047 + 0.35x_{t-1} + 0.18x_{t-2} - 0.14x_{t-3} \qquad t = 4,\ldots,176$

σφάλματα ή υπόλοιπα (residual) εκτίμησης $\hat{z}_t = x_t - \hat{x}_t$ $\quad s_z^2 = \hat{\sigma}_z^2 = 0.0000989$ $\quad s_z = \hat{\sigma}_z = 0.0098$

προσαρμοσμένο AR(3) $x_t = 0.0047 + 0.35x_{t-1} + 0.18x_{t-2} - 0.14x_{t-3} + z_t$





Διάγνωση καταλληλότητας μοντέλου

είναι τα υπόλοιπα ανεξάρτητα → έλεγχο ανεξαρτησίας στα $\left\{\hat{z}_t\right\}_{t=p+1}^{n}$