Χρονοσειρές - Μάθημα 5
Εκτίμηση μοντέλου MA(q)

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 **?**

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

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

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

?

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

Προσαρμογή διαδικασίας (μοντέλο) MA(q) ➡ εκτίμηση παραμέτρων $\theta_1, \theta_2, \ldots, \theta_q, \sigma_Z^2$

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

Μέθοδος ροπών

Διασπορά $\sigma_X^2 = (1+\theta_1^2 + \cdots \theta_q^2)\sigma_Z^2$

Αυτοσυσχέτιση

$$\rho_\tau = \begin{cases} \dfrac{-\theta_\tau + \theta_1\theta_{\tau+1} + \cdots + \theta_{q-\tau}\theta_q}{1+\theta_1^2+\theta_2^2+\cdots+\theta_q^2} & \tau = 1,2,\ldots,q \\ 0 & \tau > q \end{cases}$$

Μη-γραμμικό σύστημα εξισώσεων ως προς τις παραμέτρους $\theta_1, \theta_2, \ldots, \theta_q$

Εκτίμηση των $\rho_1, \rho_2, \ldots, \rho_q, \sigma_X^2 \Rightarrow \underbrace{r_1, r_2, \ldots, r_q}, s_X^2$

Αλγόριθμος innovation

Μέθοδος ελαχίστων τετραγώνων

Προσαρμογή μοντέλου MA(q) στα δεδομένα

$\Downarrow$

Ελαχιστοποίηση αθροίσματος τετραγώνων των σφαλμάτων προσαρμογής

$$\min S(\mu, \theta_1, \ldots \theta_q) = \min \sum_{t=q+1}^{n} (x_t - \mu + \theta_1 z_{t-1} + \ldots + \theta_q z_{t-q})^2$$ ως προς $\mu, \theta_1, \theta_2, \ldots, \theta_q$

Αριθμητική μέθοδος βελτιστοποίησης

$\Downarrow$

$\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_q$

MA(1) $X_t - \mu = Z_t - \theta Z_{t-1}$

Μέθοδος ροπών

$$\rho_\tau = \begin{cases} \dfrac{-\theta}{1+\theta^2} & \tau = 1 \\ 0 & \tau \geq 2 \end{cases}$$

$$|r_1| > 0.5 \;\Rightarrow\; \hat{\theta} = \frac{r_1}{|r_1|}$$

$$|r_1| \leq 0.5 \;\Rightarrow\; r_1\hat{\theta}^2 + \hat{\theta} + r_1 = 0 \;\Rightarrow\; \hat{\theta}_{1,2} = \frac{-1 \pm \sqrt{1-4r_1^2}}{2r_1}$$

Επιλέγουμε τη λύση $|\hat{\theta}| < 1$ που δίνει αντιστρεψιμότητα

$$\sigma_X^2 = (1+\theta^2 + \cdots \theta^q)\sigma_Z^2 \qquad s_Z^2 = \frac{s_X^2}{1+\hat{\theta}^2}$$

Μέθοδος ελαχίστων τετραγώνων

Υποθέτω $z_0 = 0$ (και $\mu = 0$) $\qquad z_t = x_t + \theta z_{t-1}$

$$z_1 = x_1$$

$$z_2 = x_2 + \theta z_1 = x_2 + \theta x_1$$

$$z_3 = x_3 + \theta z_2 = x_3 + \theta(x_2 + \theta x_1) = x_3 + \theta x_2 + \theta^2 x_1$$

$$\vdots \qquad \vdots \qquad \vdots \qquad \vdots$$

$$z_n = x_n + \theta z_{n-1} = x_n + \theta x_{n-1} + \theta^2 x_{n-2} + \cdots + \theta^{n-2} x_2 + \theta^{n-1} x_1$$

$$\min \sum_{t=1}^{n} z_t^2 = \min\left\{x_1^2 + (x_2 + x_1\theta)^2 + (x_3 + x_2\theta + x_1\theta^2)^2 \ldots + (x_n + x_{n-1}\theta + \cdots + x_1\theta^{n-1})^2\right\}$$

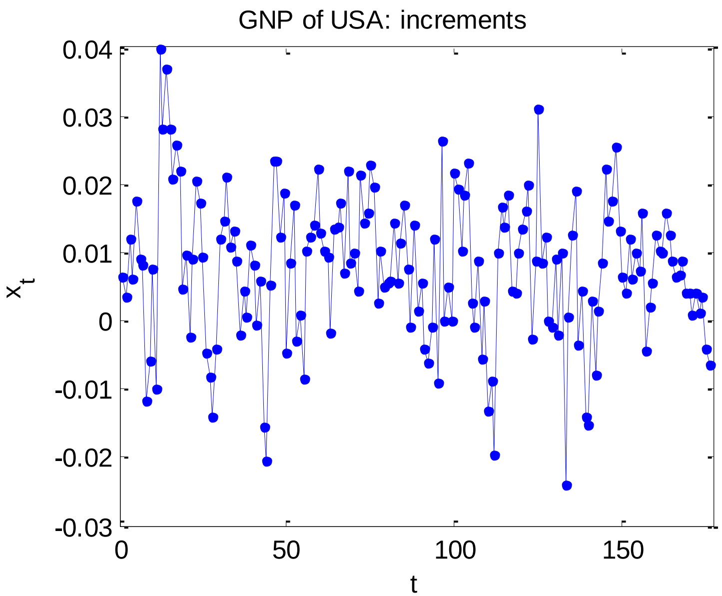
$$\min\left\{a_0 + a_1\theta + \cdots + a_{2n-2}\theta^{2n-2}\right\}$$

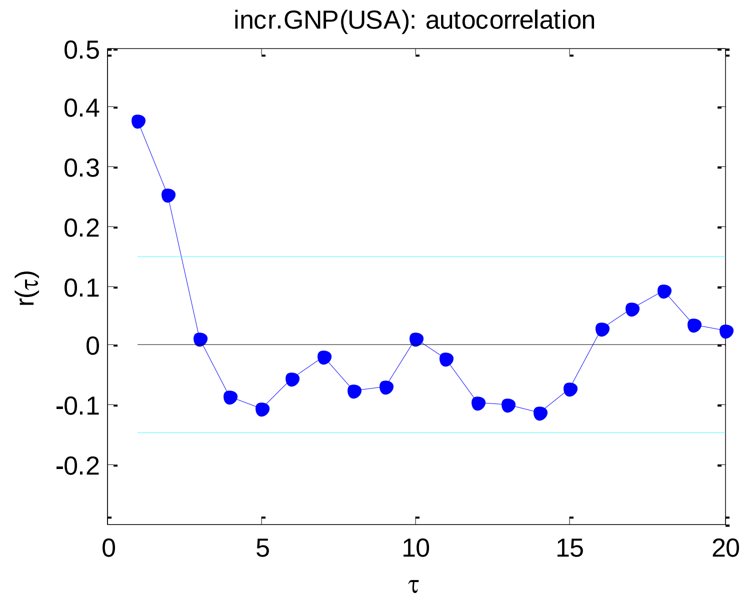
$2n$-3 λύσεις, θα πρέπει να επιλέξω λύση $|\hat{\theta}| < 1$

⇓

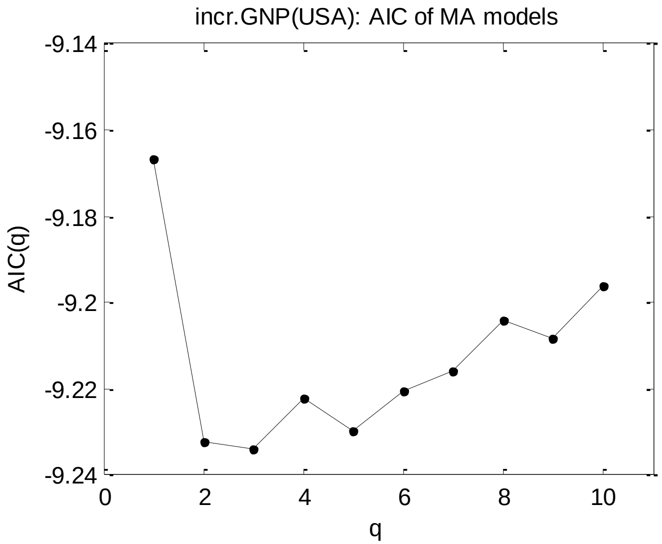
προγραμματισμός λύσης ελαχίστων τετραγώνων με περιορισμούς για αντιστρεψιμότητα

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





τάξη
MA μοντέλου ?



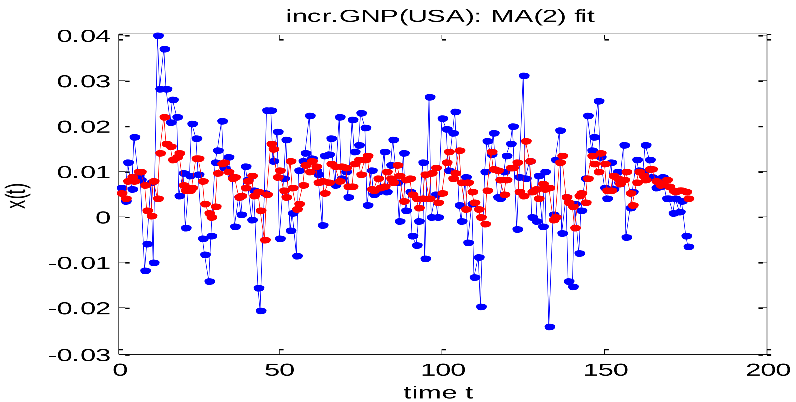
MA(2) ?

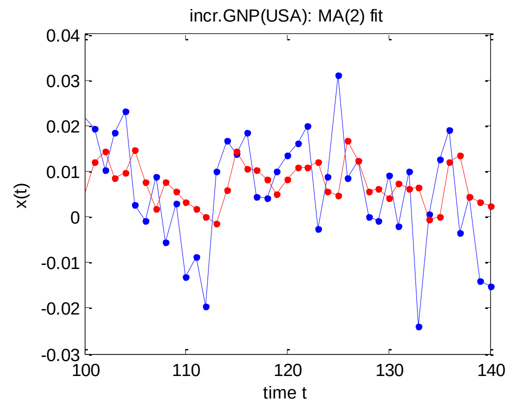
εκτίμηση παραμέτρων $\bar{x} = 0.0077$

OLS → $\hat{\theta}_1 = -0.312 \quad \hat{\theta}_2 = -0.272$

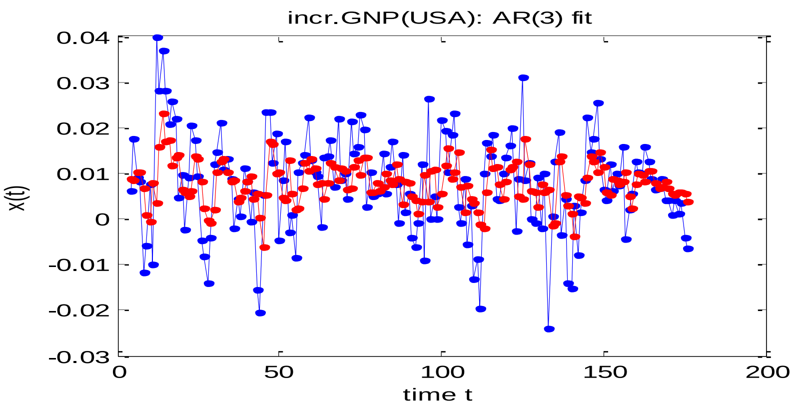
διασπορά σφαλμάτων (υπολοίπων) $s_z^2 = 0.000097 \quad s_z = 0.00983$

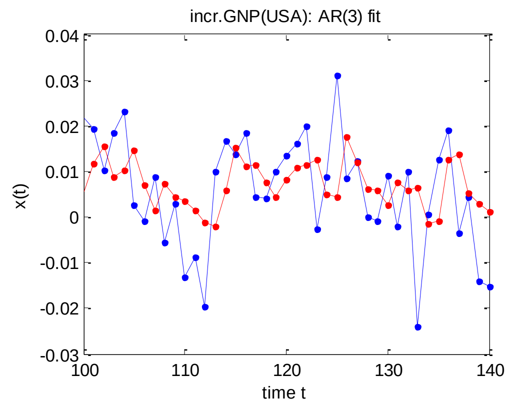
προσαρμοσμένο MA(2) $x_t = 0.0077 + z_t + 0.312 z_{t-1} + 0.272 z_{t-2} \qquad t = 1, \ldots, 176$





προσαρμογή
με MA(2)





προσαρμογή
με AR(3)

Διάγνωση καταλληλότητας μοντέλου

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

Εκτίμηση μοντέλου ARMA(p,q)

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 **?**

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

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

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

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

Προσαρμογή διαδικασίας (μοντέλο) ARMA(p,q)

⇓

εκτίμηση παραμέτρων $\phi_1, \phi_2, \ldots, \phi_p, \theta_1, \theta_2, \ldots, \theta_q, \sigma_Z^2$

Μέθοδος ροπών και μέθοδος ελαχίστων τετραγώνων όπως για MA(q)

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

Μέθοδος ροπών

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

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

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

Επίλυση συστήματος εξισώσεων ως προς ϕ, θ **?**

$$s_Z^2 = \frac{1-\hat{\phi}^2}{1+\hat{\theta}^2-2\hat{\phi}\hat{\theta}} s_X^2$$

15 **Η συνάρτηση αντίστροφης αυτοσυσχέτισης (inverse autocorrelation)**

Μέθοδος ελαχίστων τετραγώνων

16 **Μέθοδοι διερεύνησης της επάρκειας μοντέλου**

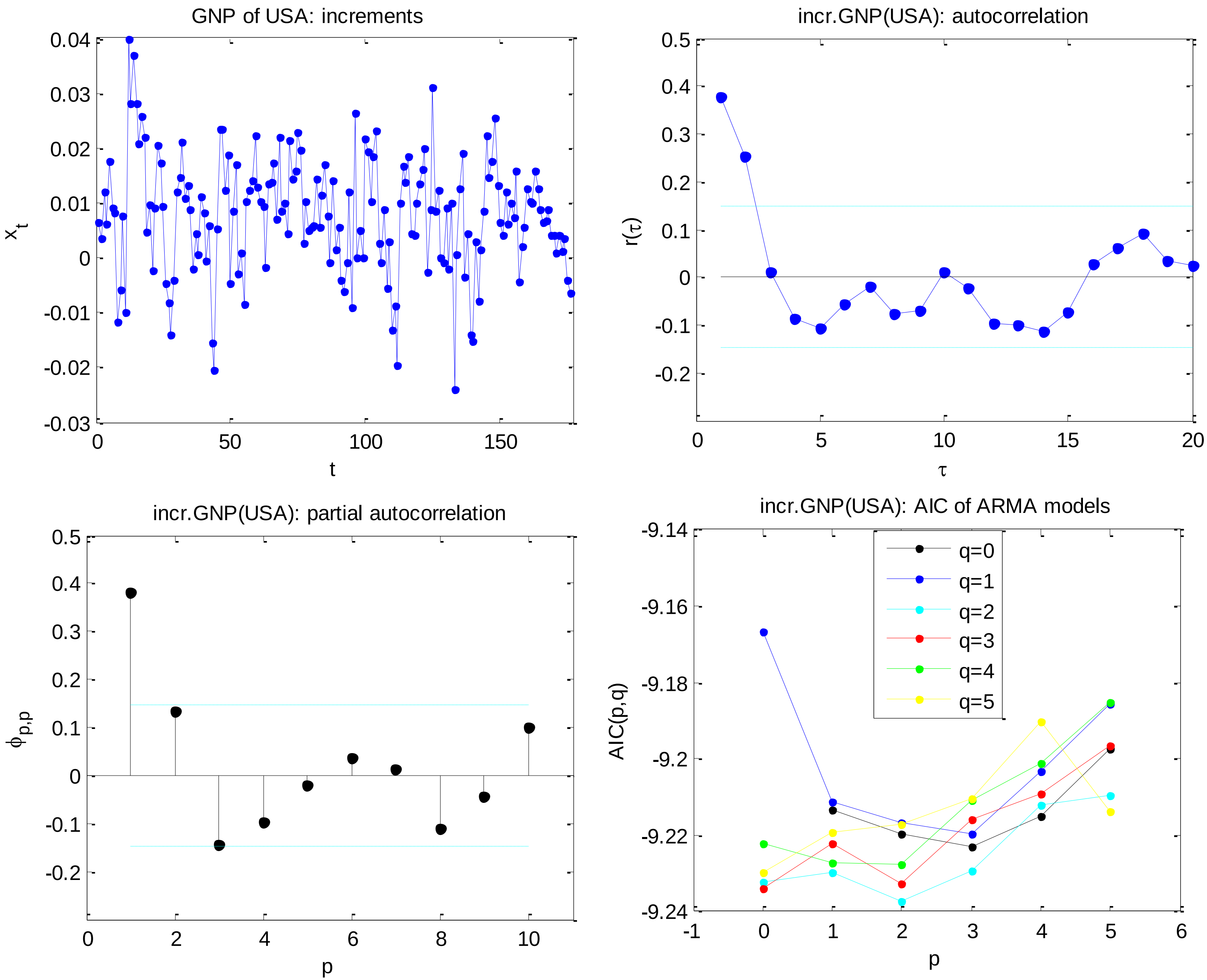
Υποθέτω $z_0 = 0$ (και $x_0 = \mu = 0$)

$$z_1 = \boxed{x_1}$$
$$z_2 = x_2 - \phi x_1 + \theta z_1 = \boxed{x_2 + (\theta-\phi)x_1}$$
$$z_3 = x_3 - \phi x_2 + \theta z_2 = \boxed{x_3 + (\theta-\phi)x_2 + \theta(\theta-\phi)x_1}$$
$$\vdots \qquad \vdots \qquad \vdots \qquad \vdots$$
$$z_n = x_n - \phi x_{n-1} + \theta z_{n-1} = \boxed{x_n - (\theta-\phi)x_{n-1} + \theta(\theta-\phi)x_{n-2} + \cdots + \theta^{n-2}(\theta-\phi)x_1}$$

$$\min \sum_{t=1}^{n} z_t^2$$

προγραμματισμός λύσης ελαχίστων τετραγώνων με περιορισμούς για αντιστρεψιμότητα και στασιμότητα

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).



τάξη ARMA μοντέλου ?

ARMA(2,2) ?

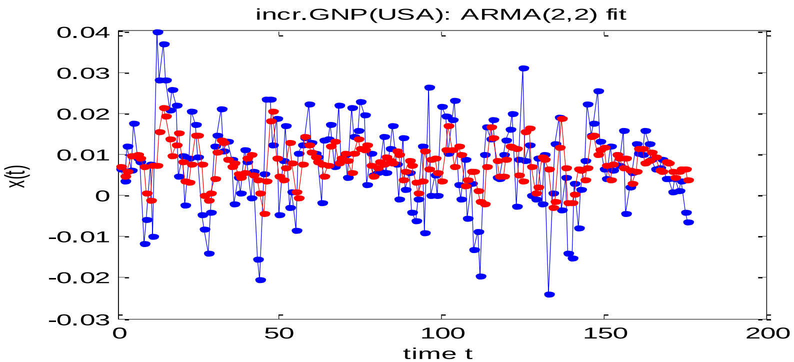
εκτίμηση παραμέτρων $\bar{x} = 0.0077$

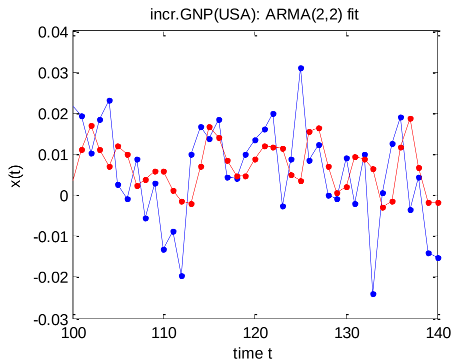
OLS → $\hat{\phi}_1 = 0.614$ $\hat{\phi}_2 = -0.455$ $\hat{\theta}_1 = 0.301$ $\hat{\theta}_2 = -0.600$

διασπορά σφαλμάτων (υπολοίπων) $s_z^2 = 0.000097$ $s_z = 0.00983$

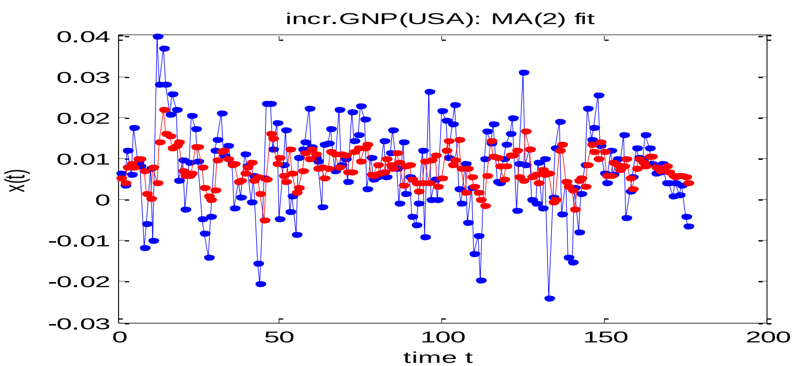
$t = 1, \ldots, 176$

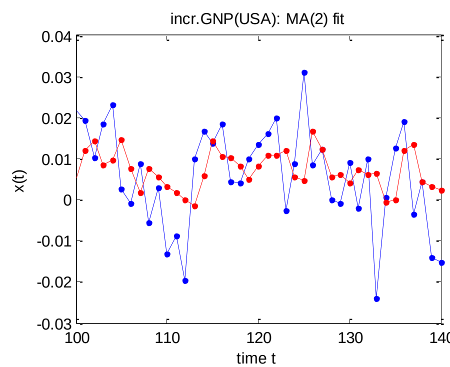
προσαρμοσμένο ARMA(2,2) $\hat{x}_t = 0.0065 + 0.614 x_{t-1} - 0.455 x_{t-2} + z_t - 0.301 z_{t-1} + 0.600 z_{t-2}$



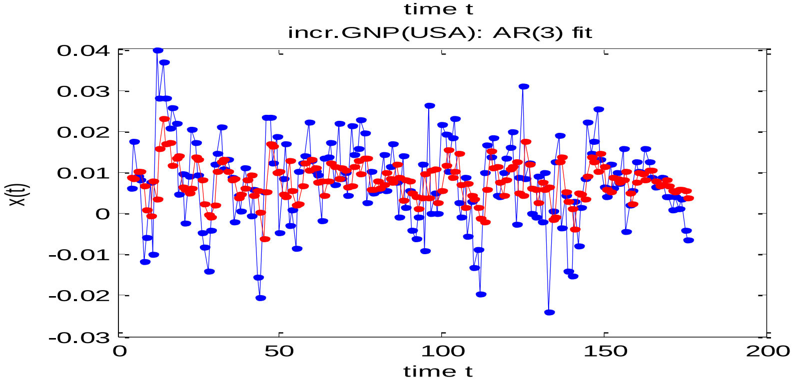


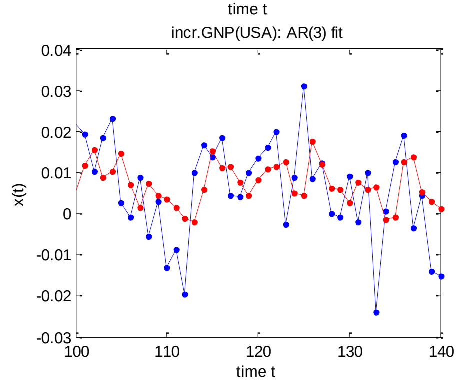
προσαρμογή με ARMA(2,2)





προσαρμογή με MA(2)





προσαρμογή με AR(3)

Μοντέλο χρονοσειράς με τάση (ARIMA)

$\{Y_t\}_{t=1}^{\infty}$ **τυχαίος περίπατος** (μη-στάσιμη διαδικασία)

$Y_t = Y_{t-1} + X_t = X_1 + X_2 + \cdots X_t$ $\{X_t\}_{t=-\infty}^{\infty}$ iid $\mathrm{E}[X_t]=0$ $\mathrm{E}\left[X_t^2\right]=\sigma^2$

διαδικασία AR(1) για $\phi = 1$

Πρώτες διαφορές: $X_t = (1-B)Y_t = Y_t - Y_{t-1}$ διαδικασία iid

$\{Y_t\}_{t=1}^{\infty}$ μη-στάσιμη διαδικασία **που παρουσιάζει τάση**

πρώτες διαφορές: $X_t = Y_t - Y_{t-1}$ στάσιμη διαδικασία **?** **ΝΑΙ** **ΟΧΙ**

διαφορές δεύτερης τάξης: $X'_t = X_t - X_{t-1} = Y_t - 2Y_{t-1} + Y_{t-2}$ στάσιμη διαδικασία **?** **ΟΧΙ** **ΝΑΙ**

$X_t = \nabla^d Y_t$ $\{X_t\}_{t=1}^{\infty}$ στάσιμη μετά από διαφορές d τάξης:
$= (1-B)^d Y_t$

$\boxed{\text{Συνήθως } d = 1}$

AR(p), MA(q), ARMA(p,q) **?**

$\{Y_t\}_{t=1}^{\infty}$ μη-στάσιμη διαδικασία **ARIMA(p,d,q)**

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

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

$$\phi(B)\nabla^d Y_t = \theta(B) Z_t$$

$$\phi(B)(1-B)^d Y_t = \theta(B) Z_t$$

Το πολυώνυμο $\phi(B)(1-B)^d$ έχει μια ρίζα =1 και όλες τις άλλες εκτός του μοναδιαίου κύκλου

Προσαρμογή μοντέλου ARIMA (διαδικασία Box-Jenkins)

χρονοσειρά παρατηρήσεων $\{y_1, y_2, \ldots, y_n\}$

? διάγραμμα ιστορίας (γράφημα χρονοσειράς)
αυτοσυσχέτιση (ισχυρή και φθίνει πολύ αργά)
άλλο?

ένδειξη πως έχει τάση

διαφορές τάξης d $\quad x_t = (1-B)^d y_t$
άλλο?

στάσιμη χρονοσειρά $\{x_1, x_2, \ldots, x_n\}$

τάξη μοντέλου
εκτίμηση παραμέτρων μοντέλου

προσαρμογή μοντέλου AR(p), MA(q), ARMA(p,q)

διαγνωστικός έλεγχος ?

επάρκεια (adequacy) του μοντέλου

κατάλληλο μοντέλο ARMA(p,q) για $\{x_1, x_2, \ldots, x_n\}$

με τον αντίστροφο μετασχηματισμό του $x_t = (1-B)^d y_t$

έχουμε το μοντέλο ARIMA(p, d,q) για $\{y_1, y_2, \ldots, y_n\}$

αν η αυτοσυσχέτιση φθίνει στο 0

η χρονοσειρά είναι στάσιμη

αν η αυτοσυσχέτιση είναι στατιστικά ασήμαντη

είναι iid ? έλεγχος ανεξαρτησίας ?

ΝΑΙ

ΣΤΟΠ

ΟΧΙ

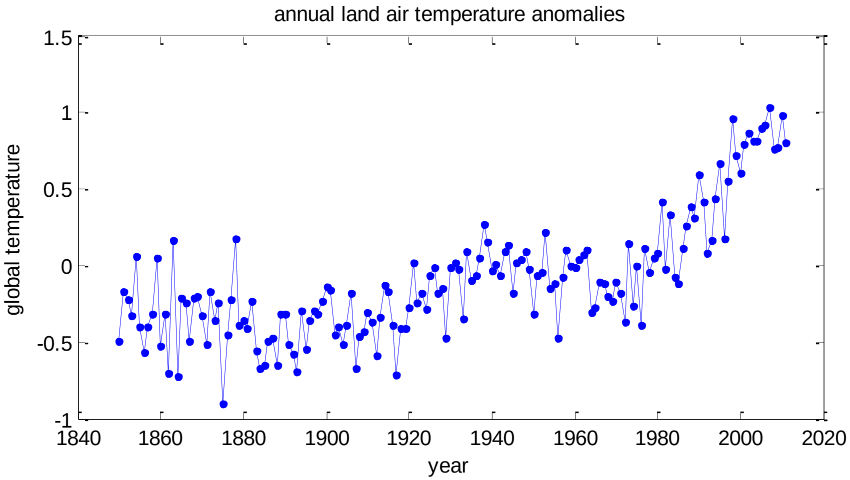
μη-γραμμικό μοντέλο?

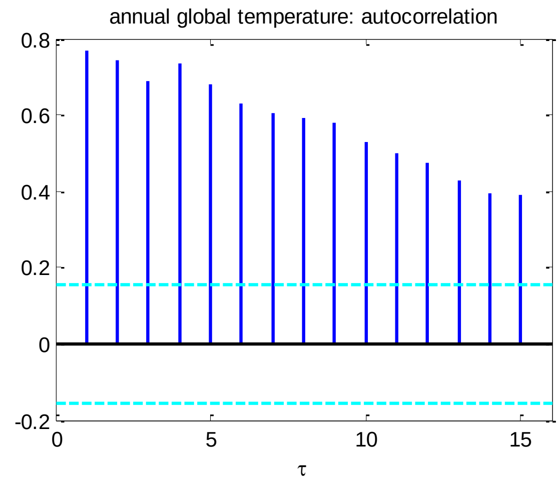
πρόβλεψη?

Παράδειγμα Ετήσιος δείκτης για τη θερμοκρασία της γης (ανωμαλία στη θερμοκρασία εδάφους στο βόρειο ημισφαίριο σε πλέγμα 5^o x 5^o), περίοδος 1850-2011

Πηγή: http://www.cru.uea.ac.uk/cru/data/temperature

$\{y_1, y_2, \ldots, y_n\}$ πραγματικές μετρήσεις

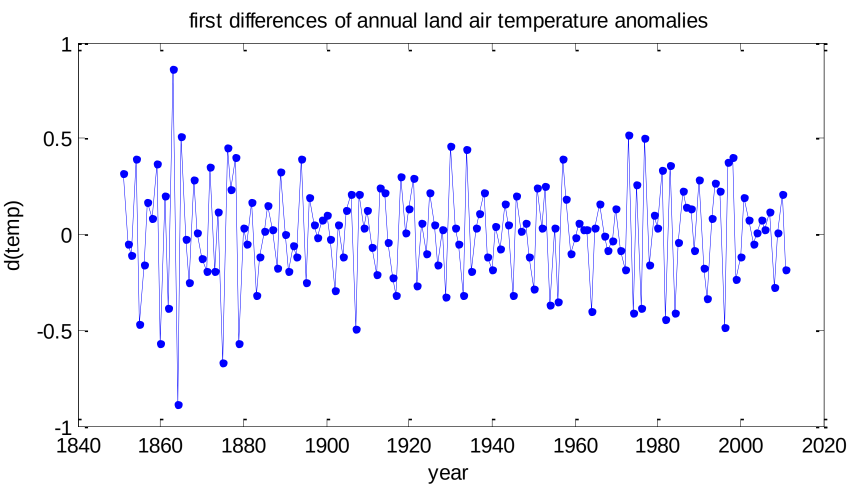


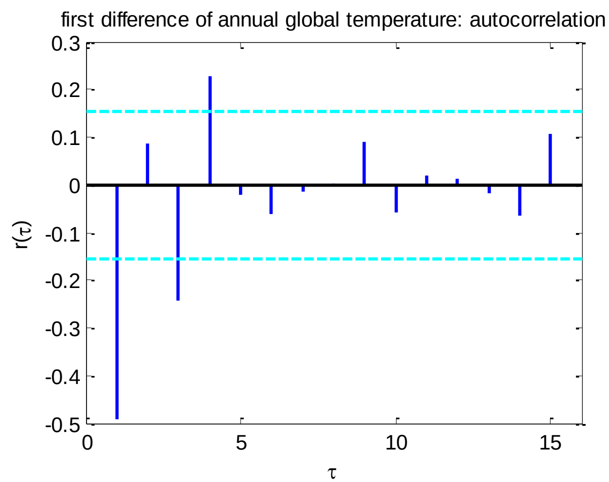


στάσιμη χρονοσειρά?

ΟΧΙ

$\{x_1, x_2, \ldots, x_n\}$ πρώτες διαφορές





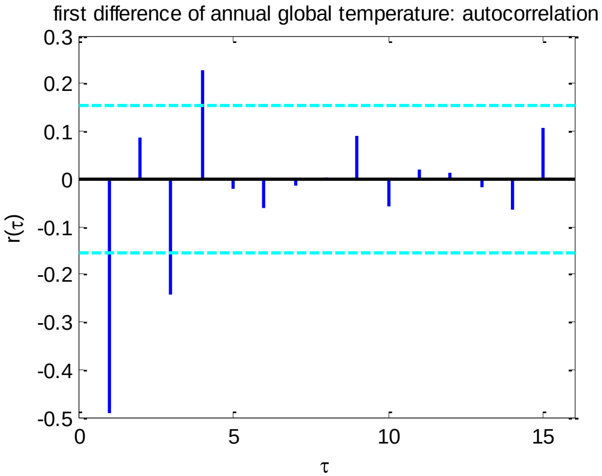
στάσιμη χρονοσειρά?

ΝΑΙ

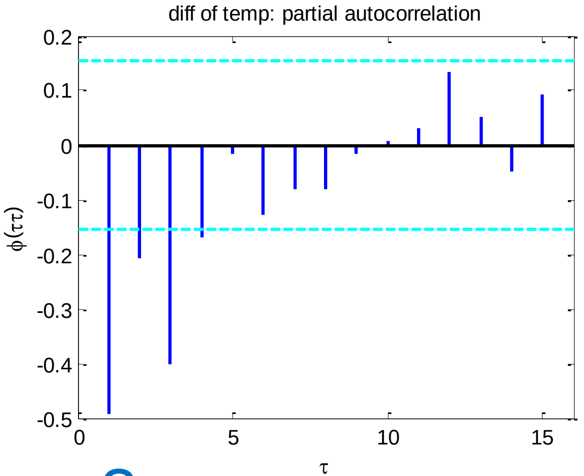
Μοντέλο για τη χρονοσειρά $\{x_1, x_2, \ldots, x_n\}$?

17 **Μοντέλα ARFIMA (ή FARIMA)**

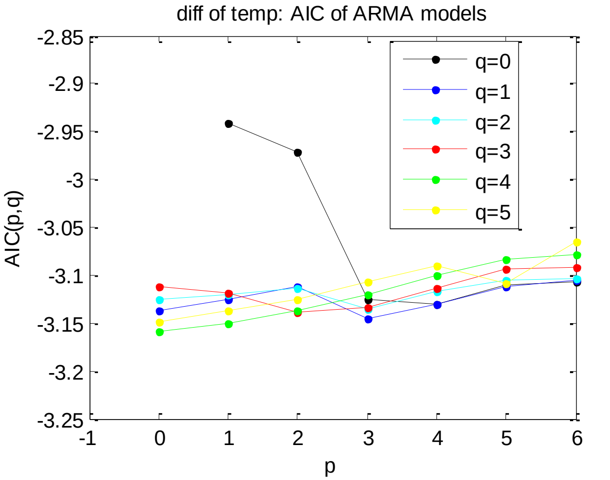
αυτοσυσχέτιση



μερική αυτοσυσχέτιση



κριτήριο AIC



Πιο κατάλληλη μορφή μοντέλου ?

προσαρμογή MA(4) ($\bar{x} = 0.008$)

Μοντέλο για τη χρονοσειρά $\{y_1, y_2, \ldots, y_n\}$

ARIMA(0,1,4) $(1-B)Y_t = \theta_4(B)Z_t$

$$x_t = 0.008 + z_t - 0.758z_{t-1} - 0.022z_{t-2} - 0.219z_{t-3} + 0.275z_{t-4} \qquad s_z^2 = 0.0414 \qquad s_z = 0.2035$$