

Εκτίμηση μη-γραμμικών χαρακτηριστικών

Μη-γραμμικά χαρακτηριστικά ή **αναλλοίωτα μέτρα**

- **Διάσταση**

1. Ευκλείδεια
2. Τοπολογική
3. Μορφοκλασματική

(συσχέτισης, πληροφορίας, μέτρησης κουτιών, ...)

- **Εκθέτες Lyapunov**

(μεγαλύτερος, όλο το φάσμα)

- **Εντροπία**

Διάσταση συσχέτισης ν

Η διάσταση συσχέτισης ν χαρακτηρίζει τη μορφοκλασματική δομή του ελκυστή (αυτο-ομοιότητα σε διάφορες κλίμακες) χρησιμοποιώντας την πυκνότητα των σημείων του ελκυστή στο χώρο των καταστάσεων

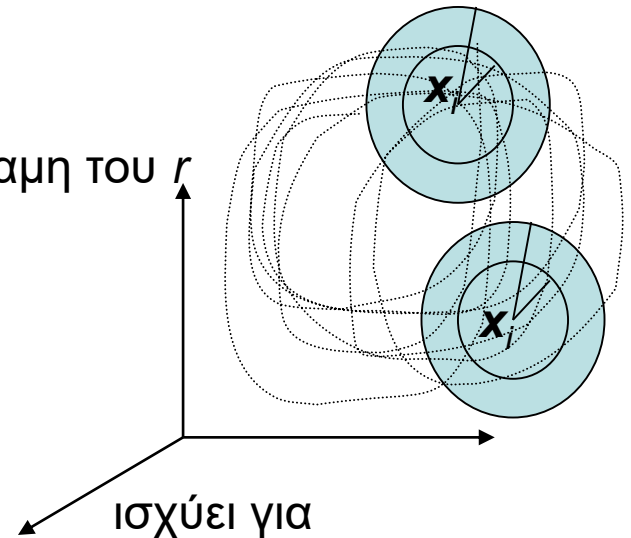
Η βασική ιδέα είναι πως η πιθανότητα δύο σημείων να βρίσκονται σε απόσταση μικρότερη του r

$$P(\|\mathbf{x}_i - \mathbf{x}_j\| < r)$$

μεταβάλλεται ως προς r ανάλογα με κάποια δύναμη του r

μ_i : αριθμός των σημείων που βρίσκονται μέσα σε σφαίρα με ακτίνα r και κέντρο \mathbf{x}_i

$$\langle \mu_i \rangle_x \approx P(\|\mathbf{x}_i - \mathbf{x}_j\| < r)$$



νόμο κλιμάκωσης (scaling law) $\langle \mu_i \rangle_x \sim r^\nu$

ισχύει για

$$r \rightarrow 0$$

$$N \rightarrow \infty$$

ν ακέραιος \Rightarrow ελκυστής είναι συνήθης γεωμετρικό αντικείμενο

ν μη-ακέραιος \Rightarrow ελκυστής είναι μορφοκλασματικό αντικείμενο

Εκτίμηση της διάστασης συσχέτισης ν

χρονοσειρά $\{x_i\}$, $i = 1, \dots, N + (m-1)\tau$

ανακατασκευή $\{x_i\}$, $i = 1, \dots, N$

Εκτίμηση του $\langle \mu_i \rangle_x$

Άθροισμα συσχέτισης

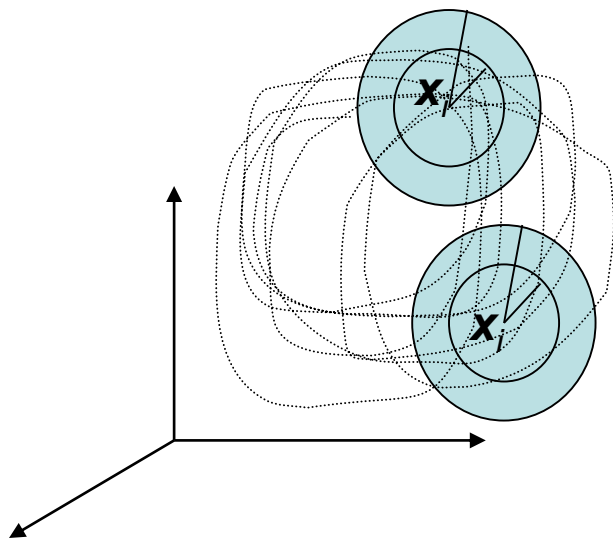
$$C(r) = \frac{2}{N(N-1)} \sum_{i=1}^N \sum_{j=i+1}^N \Theta(r - \|x_i - x_j\|)$$

Heaviside συνάρτηση $\Theta(x) = \begin{cases} 0 & \text{όταν } x \leq 0 \\ 1 & \text{όταν } x > 0 \end{cases}$

Νόμος κλιμάκωσης $C(r) \propto r^\nu$ για r μικρό

Εκτίμηση $\nu = \frac{d \log C(r)}{d \log r}$ για εύρος τιμών του r

Σύγκλιση του $\nu(m)$ για m ικανοποιητικά μεγάλο

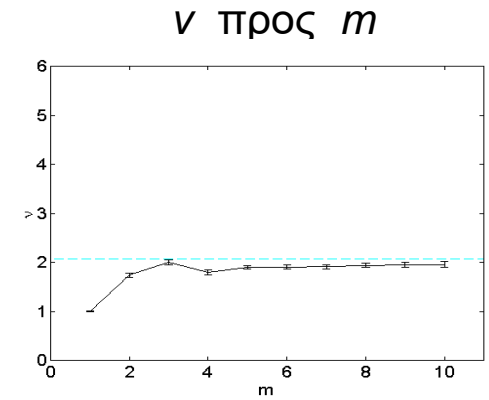
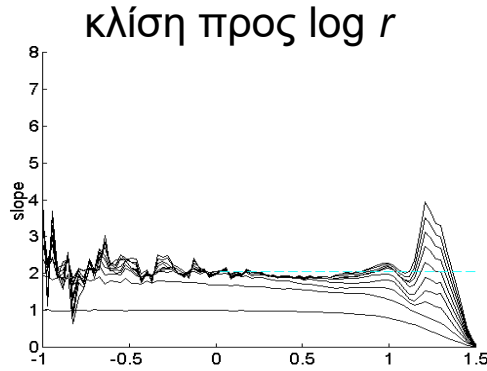
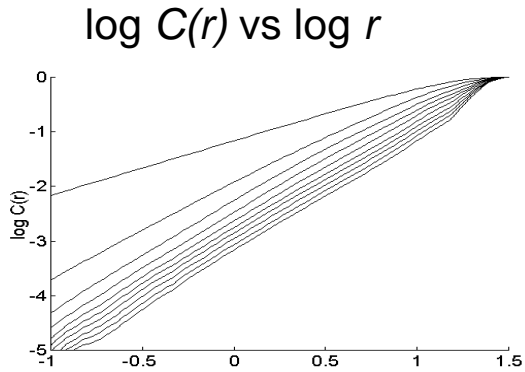


Αν ν μικρό και μη-ακέραιο και σύστημα αιτιοκρατικό

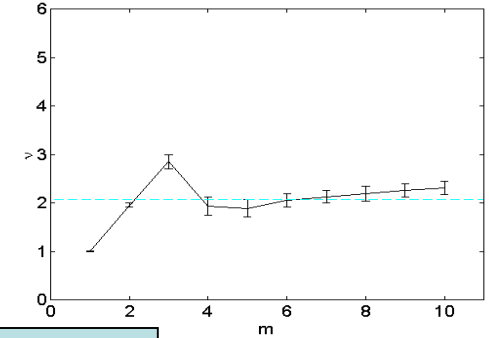
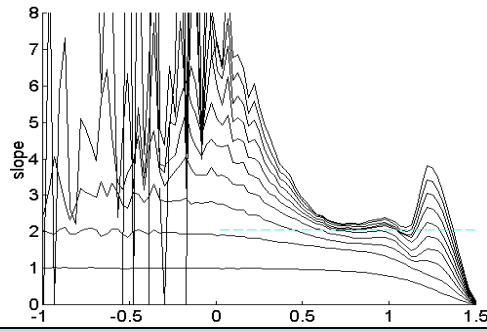
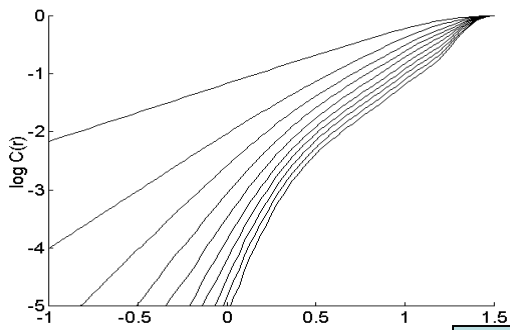


χαμηλή διάσταση και μορφοκλασματική (χασοτική) δομή

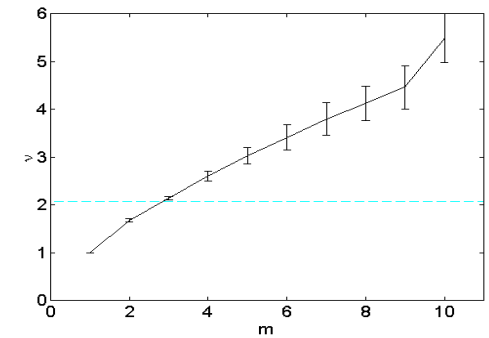
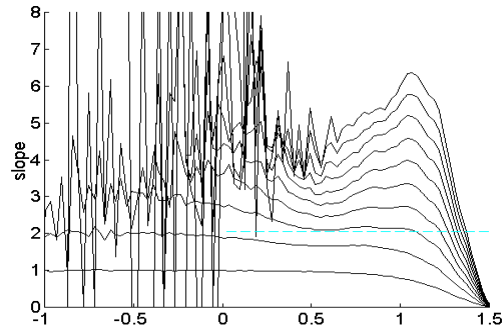
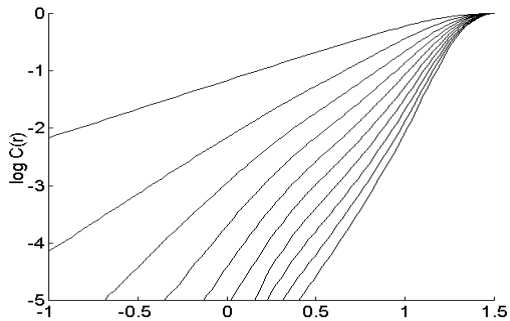
x-Lorenz χωρίς θόρυβο, $\tau=2$



x-Lorenz + 10% θόρυβο παρατήρησης, $\tau=2$



x-Lorenz + 10% θόρυβο παρατήρησης, $\tau=10$

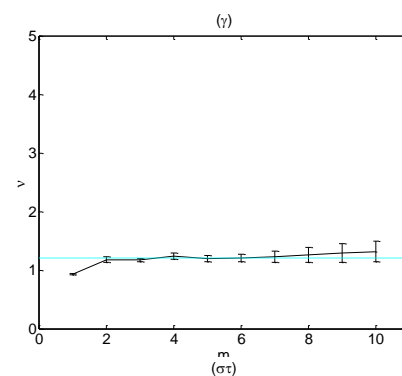
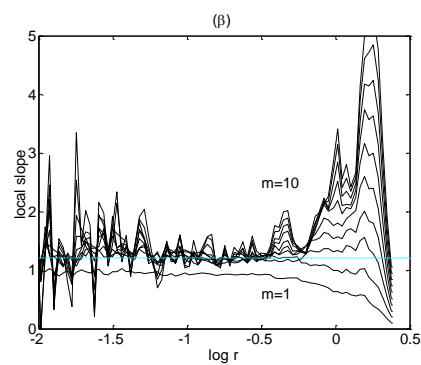
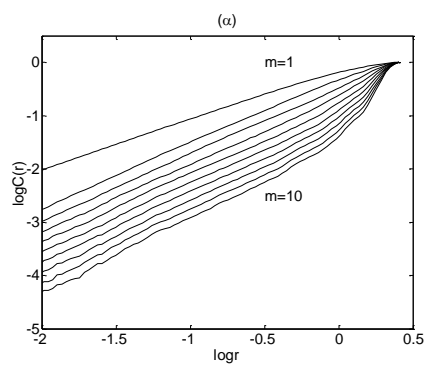


Η εκτίμηση του v επηρεάζεται από τους παρακάτω παράγοντες:

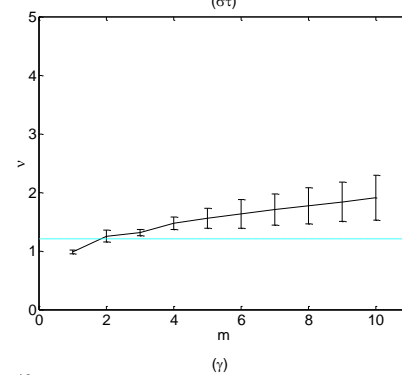
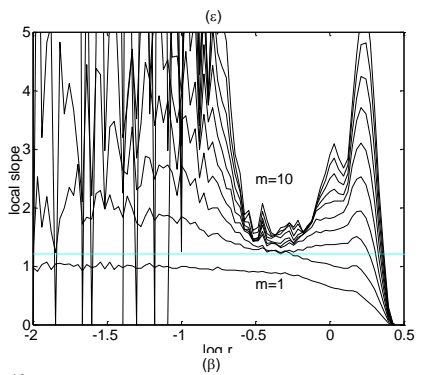
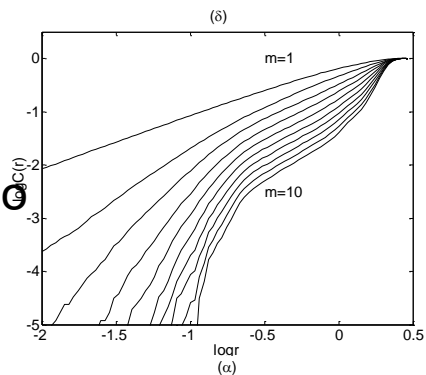
- χρόνος συσχέτισης $\rightarrow |i - j| < w$
- επιλογή του τ και m
- θόρυβος
- μήκος χρονοσειράς

$n=924$

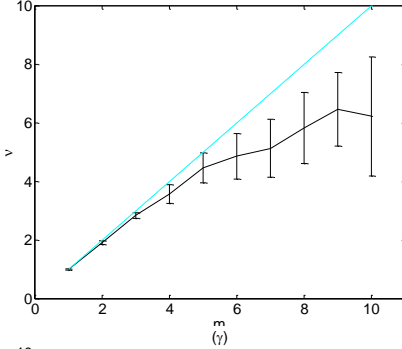
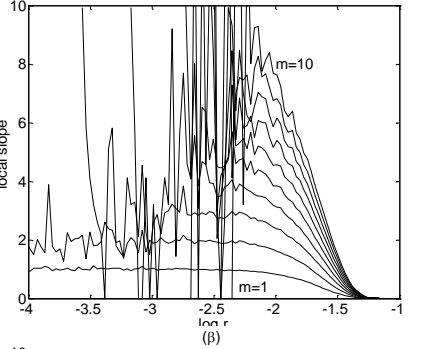
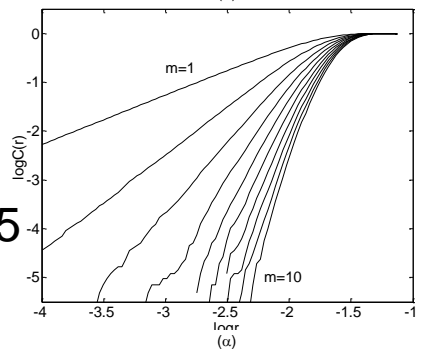
Henon



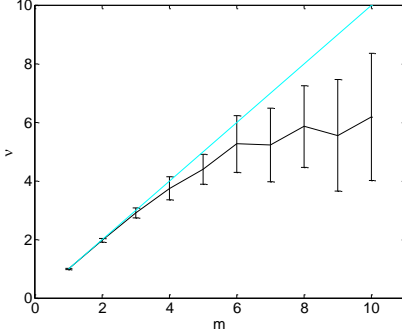
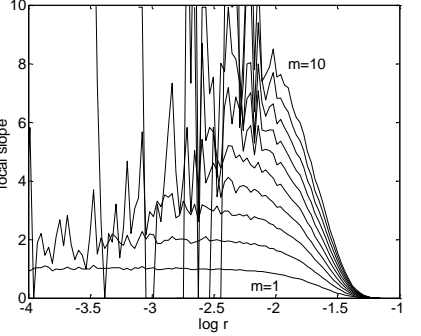
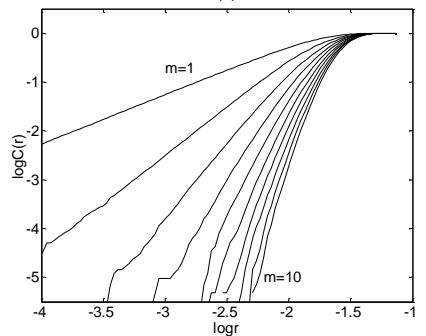
Henon
+ 10% λευκό θόρυβο



Αποδόσεις
Γενικού δείκτη
1.1.2005 – 20.9.2005

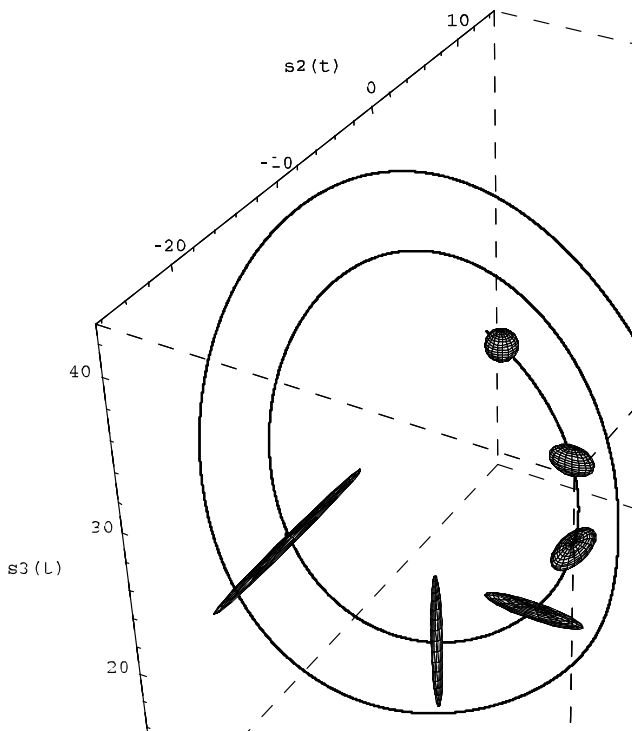


λευκός θόρυβος



Εκθέτες Lyapunov

Οι εκθέτες Lyapunov μετράνε το μέσο βαθμό απόκλισης και σύγκλισης των τροχιών στον ελκυστή, στις κατευθύνσεις του τοπικά αναλυμένου χώρου καταστάσεων



Lyapunov spectrum: $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m$

$\lambda_i > 0 \rightarrow$ απόκλιση

$\lambda_i < 0 \rightarrow$ σύγκλιση

$\lambda_i = 0 \rightarrow$ κατεύθυνση ροής

Σύστημα απώλειας ενέργειας: $\sum_{i=1}^m \lambda_i < 0$

Αν $\lambda_1 > 0$ και το σύστημα είναι αιτιοκρατικό



χάος

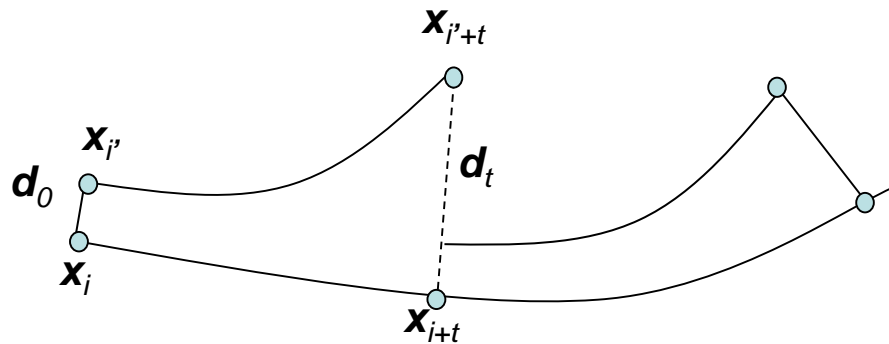
Μέγιστος εκθέτης Lyapunov λ_1

Αρχική απόσταση $\mathbf{d}_0 = \mathbf{x}_i - \mathbf{x}_{i'}$ δύο κοντινών τροχιών θα πρέπει να μεγαλώνει εκθετικά με το χρόνο.

Μετά από χρόνο t : $\mathbf{d}_t = \mathbf{x}_{i+t} - \mathbf{x}_{i'+t}$

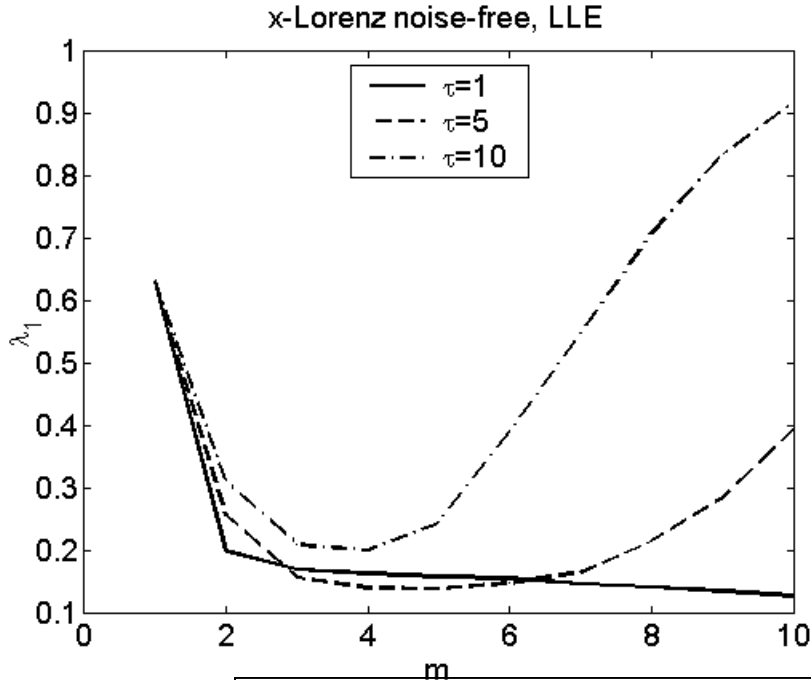
Αν $\delta_t \approx \delta_0 e^{\lambda_1 t} \implies \lambda_1$ είναι ο μέγιστος εκθέτης Lyapunov

Υπολογισμός:
$$\lambda_1 = \frac{1}{Nt} \sum_{j=1}^N \ln \frac{\delta_{t,j}}{\delta_{0,j}}$$

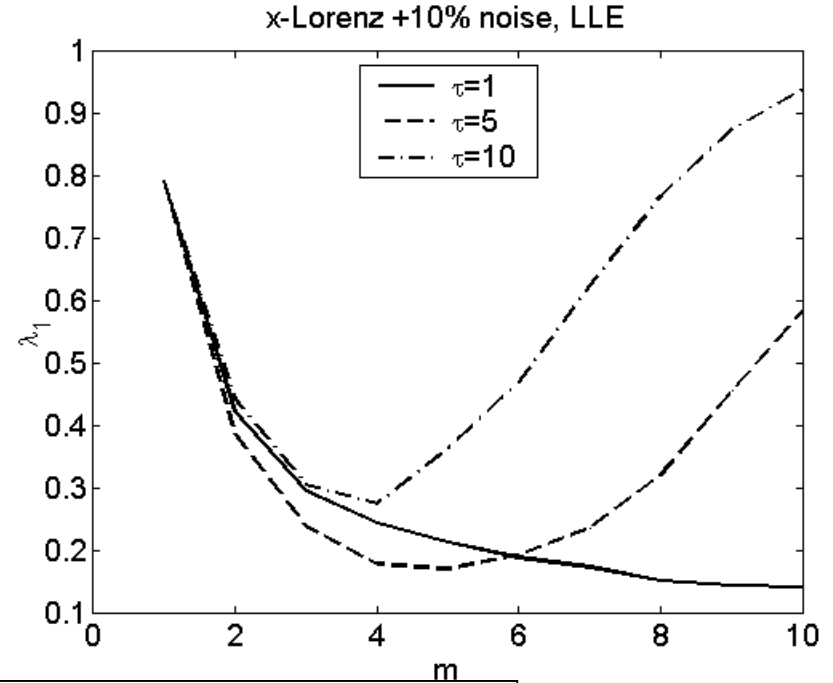


Παράδειγμα: x-Lorenz

χωρίς θόρυβο



με 10%-θόρυβο



Η εκτίμηση του λ_1 εξαρτάται από: τ , m , θόρυβο

Άλλα χαρακτηριστικά / μέτρα:

27

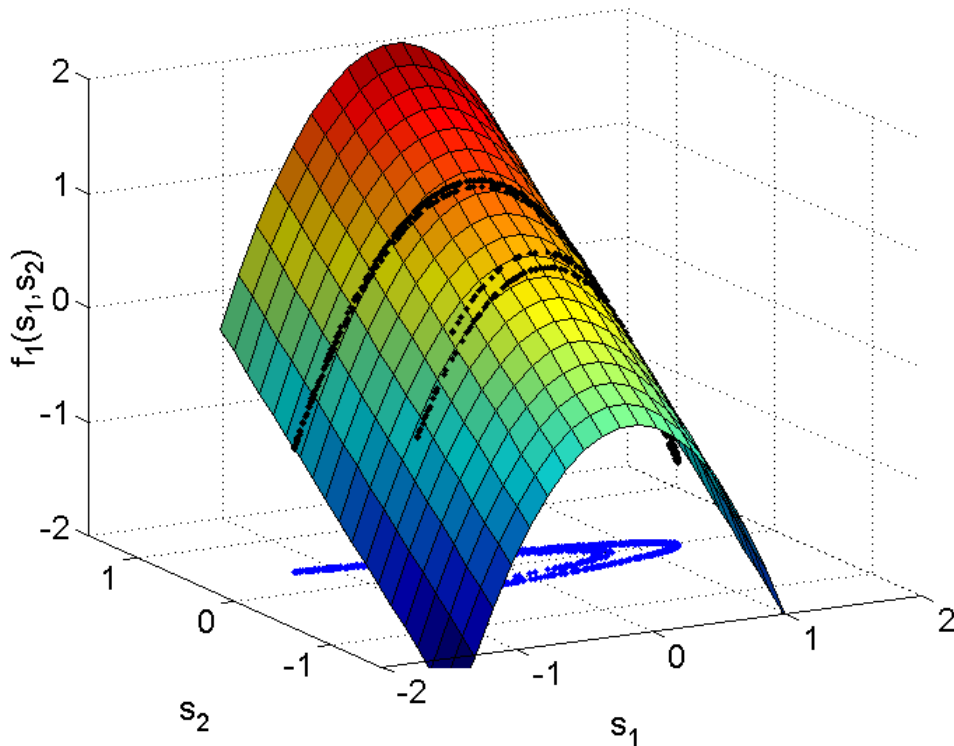
Προσεγγιστική εντροπία (approximate entropy): παρουσίαση, παραδείγματα, εφαρμογές

28

Κάποιο άλλο μη-γραμμικό χαρακτηριστικό / μέτρο

Μοντέλα πρόβλεψης

Το σύστημα που παράγει τη χρονοσειρά: $s_{i+1} = f(s_i)$



απεικόνιση Henon

$$s_{1,i+1} = 1 - 1.4s_{1,i}^2 + s_{2,i}$$

$$s_{2,i+1} = 0.3s_{1,i}$$

$$s_i \xrightarrow{f} s_{i+1}$$

$$(s_{1,i}, s_{2,i}) \xrightarrow{f_1} s_{1,i+1}$$

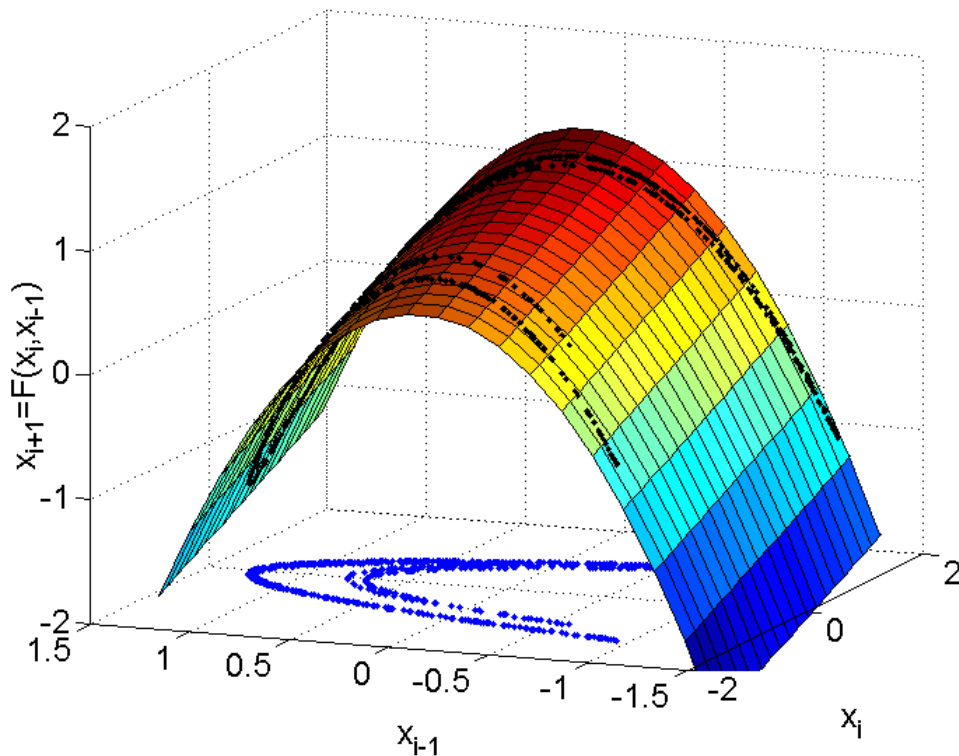
$$(s_{1,i}, s_{2,i}) \xrightarrow{f_2} s_{2,i+1}$$

Μοντέλα πρόβλεψης

Το σύστημα που παράγει τη χρονοσειρά: $s_{i+1} = f(s_i)$ άγνωστο

Το ανακατασκευασμένο σύστημα από τη χρονοσειρά: $x_{i+1} = F(x_i)$ εκτίμηση ?

Το πρόβλημα της μοντελοποίησης / πρόβλεψης χρονοσειρών:
δίνονται x_1, x_2, \dots, x_i , εκτίμηση / πρόβλεψη του x_{i+1}



Ανακατασκευή του χώρου
καταστάσεων

με τη μέθοδο των υστερήσεων:

$$\mathbf{x}_i = [x_i, x_{i-1}, \dots, x_{i-(m-1)}]^T$$

Η συνάρτηση που μας ενδιαφέρει να
εκτιμήσουμε για χρονοσειρές είναι:

$$x_{i+1} = F(x_i) \quad F: \mathbb{P}^m \mapsto \mathbb{P}^m$$



$$x_{i+1} = F(x_i) \quad F: \mathbb{P}^m \mapsto \mathbb{P}$$

$$m = 2, \tau = 1 \quad x_{i+1} = F(x_i, x_{i-1})$$

Μη-γραμμικά μοντέλα πρόβλεψης

- *Σφαιρικά μοντέλα*, π.χ. πολυώνυμα
η F έχει αναλυτική μορφή και ίδια για όλο το πεδίο ορισμού

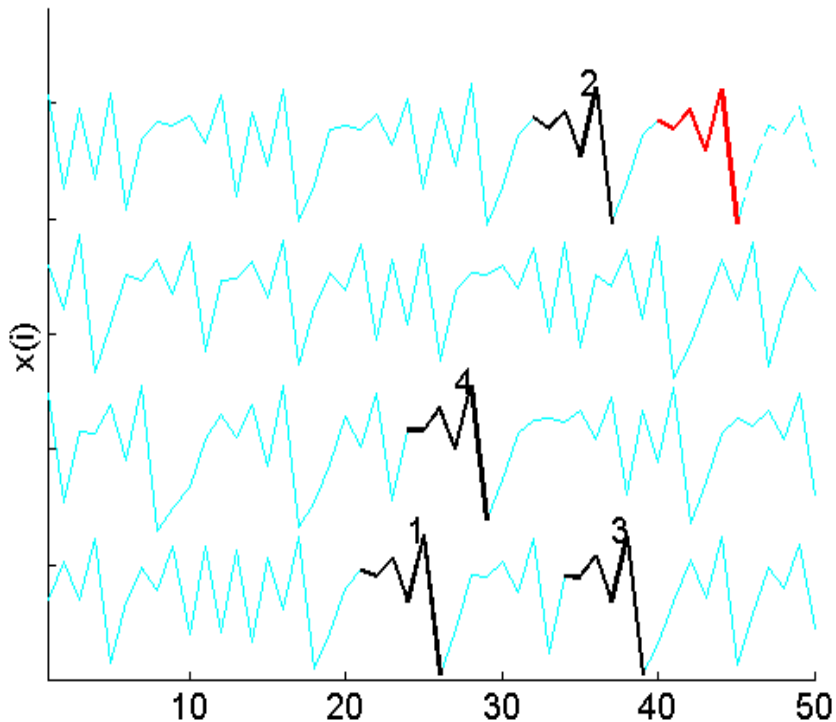
- *Τοπικά μοντέλα*, π.χ. τοπικό γραμμικό μοντέλο
η F ορίζεται διαφορετικά για κάθε σημείο του χώρου

- *Ημι-τοπικά μοντέλα*, π.χ. νευρωνικά δίκτυα
η F έχει μια μορφή που ορίζεται ως σταθμισμένο άθροισμα τοπικών στοιχειωδών συναρτήσεων

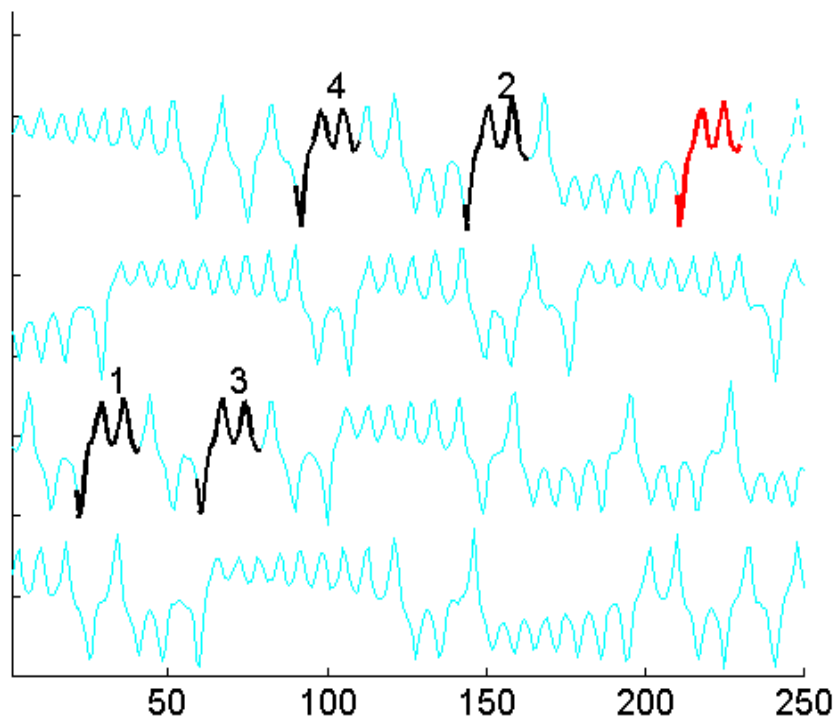
Πρόβλεψη με την εύρεση όμοιων τμημάτων της χρονοσειράς

Πρόβλεψη για χρόνο $i+T$ από τις εικόνες για T χρονικά βήματα «όμοιων» τμημάτων της χρονοσειράς από το παρελθόν

Henon + 5% noise: Analogue method



Lorenz + 5% noise: Analogue method



Τοπικά Μοντέλα Πρόβλεψης

Εφαρμογή της ιδέας «όμοιων» τμημάτων:
τμήματα χρονοσειράς → ανακατασκευασμένα σημεία

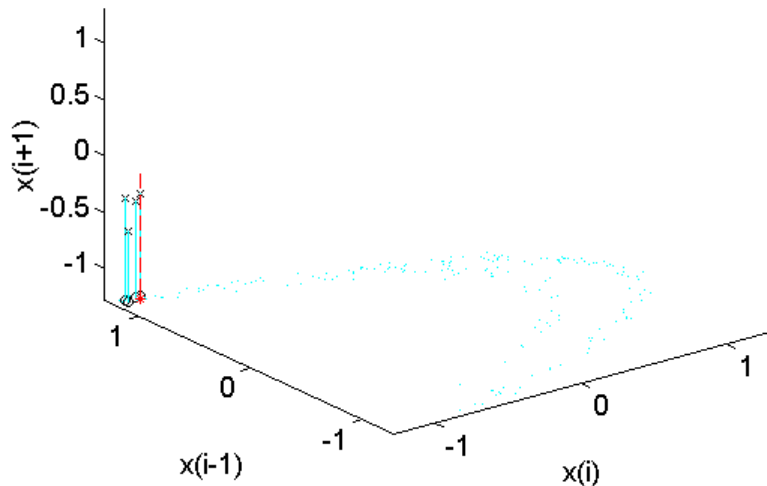
Κοντινότερα γειτονικά σημεία του $\mathbf{x}_i: \{x_{i(1)}, x_{i(2)}, \dots, x_{i(K)}\}$

Πρόβλεψη του x_{i+T} από τις εικόνες των γειτόνων: $\{x_{i(1)+T}, x_{i(2)+T}, \dots, x_{i(K)+T}\}$

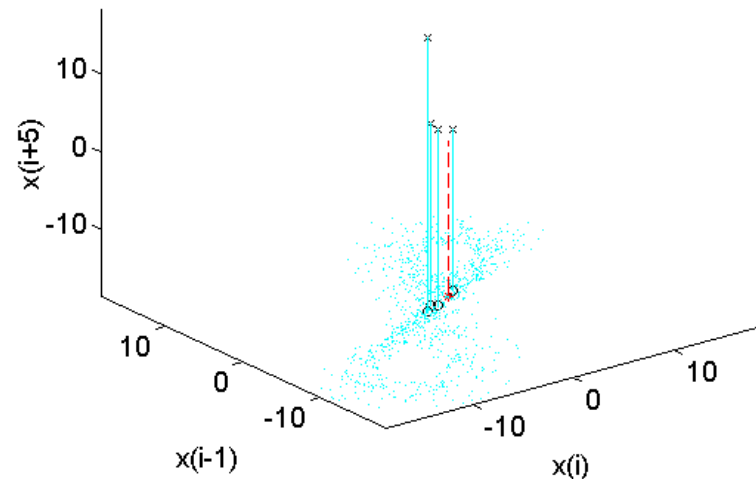
Πρόβλεψη μηδενικής τάξης: $\hat{x}_{i+T} \equiv x_i(T) = x_{i(1)+T}$

Πρόβλεψη μέσου όρου: $x_i(T) = \frac{1}{K} \sum_{j=1}^K x_{i(j)+T}$

Henon + 5% noise: State space prediction



Lorenz + 5% noise: State space prediction



Τοπικά γραμμική πρόβλεψη

Υποθέτουμε ότι για τη «γειτονιά» του \mathbf{x}_i ισχύει το τοπικά γραμμικό μοντέλο :

$$\begin{aligned}x_{i+1} &= F(\mathbf{x}_i) = F(x_i, x_{i-\tau}, \dots, x_{i-(m-1)\tau}) \\ &= a_0 + a_1 x_i + a_2 x_{i-\tau} + \dots + a_m x_{i-(m-1)\tau} \\ &= a_0 + \mathbf{a}' \mathbf{x}_i\end{aligned}$$

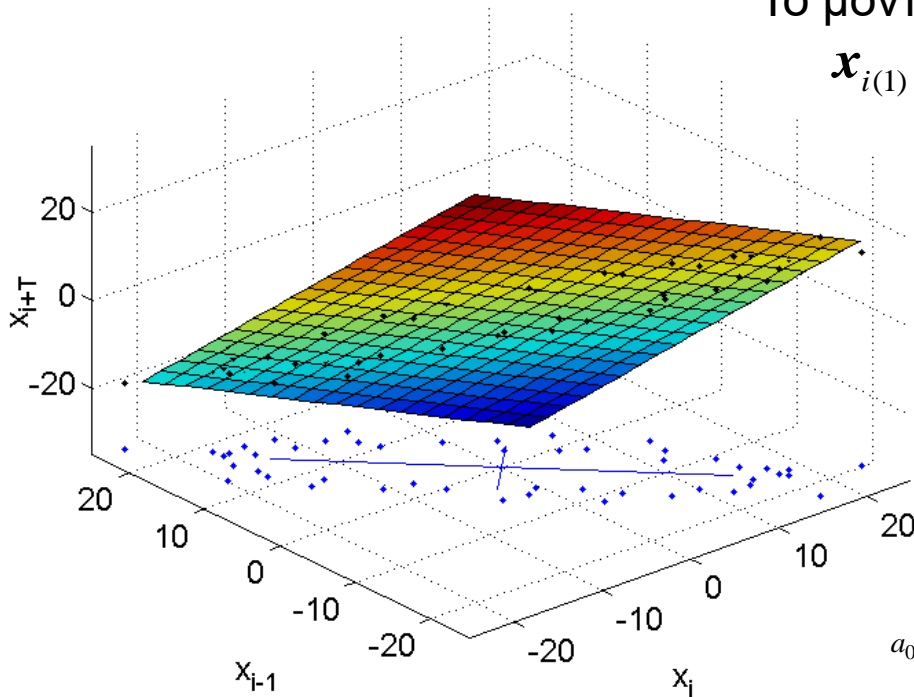
Το μοντέλο ισχύει για τα
 $\mathbf{x}_{i(1)}, \mathbf{x}_{i(2)}, \dots, \mathbf{x}_{i(K)}$

$$x_{i(1)+T} = a_0 + \mathbf{a}' \mathbf{x}_{i(1)}$$

$$x_{i(2)+T} = a_0 + \mathbf{a}' \mathbf{x}_{i(2)}$$

⋮

$$x_{i(K)+T} = a_0 + \mathbf{a}' \mathbf{x}_{i(K)}$$

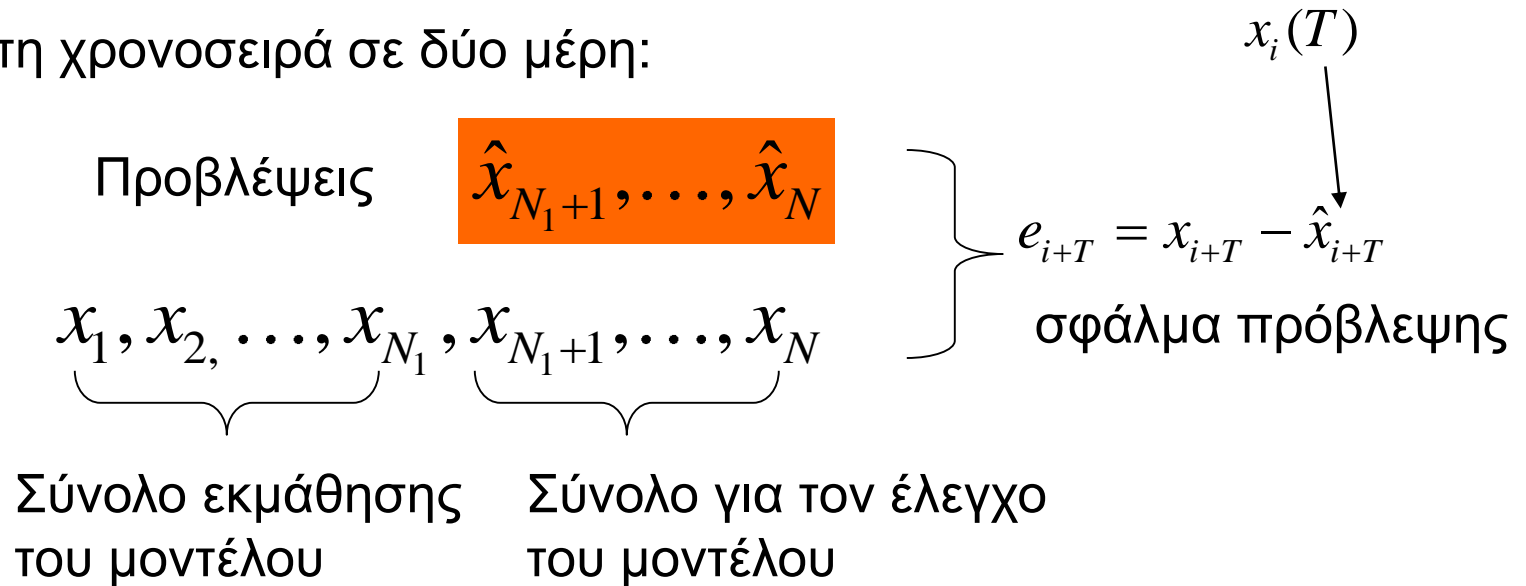


Εκτίμηση παραμέτρων a_0, a_1, \dots, a_m
(μέθοδος ελαχίστων τετραγώνων)

$$\min_{a_0, a_1, \dots, a_m} \sum_{j=1}^K (x_{i(j)+T} - a_0 - a_1 x_{i(j)} - \dots - a_m x_{i(j)-(m-1)\tau})^2$$

Εκτίμηση λάθους πρόβλεψης

Χωρίζουμε τη χρονοσειρά σε δύο μέρη:



Στατιστική για το σφάλμα πρόβλεψης

$$\text{NRMSE}(T) = \sqrt{\frac{\frac{1}{N-T-N_1} \sum_{t=N_1+1}^{N-T} (x_{t+T} - \hat{x}_{t+T})^2}{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2}}$$

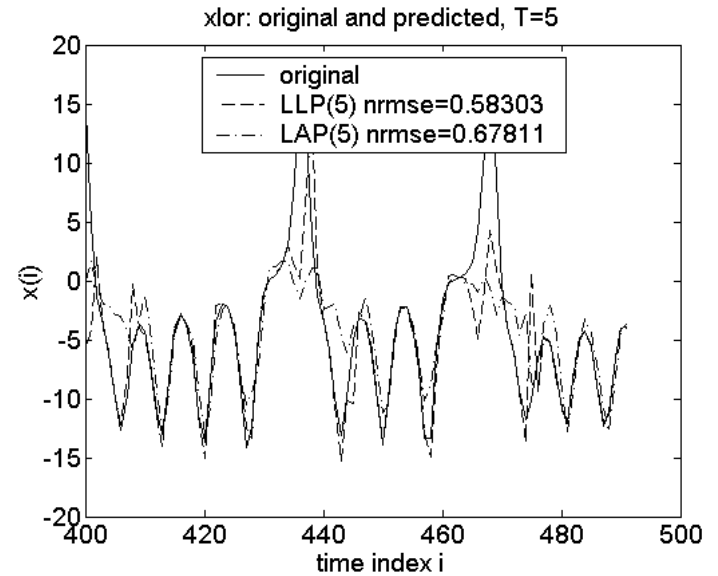
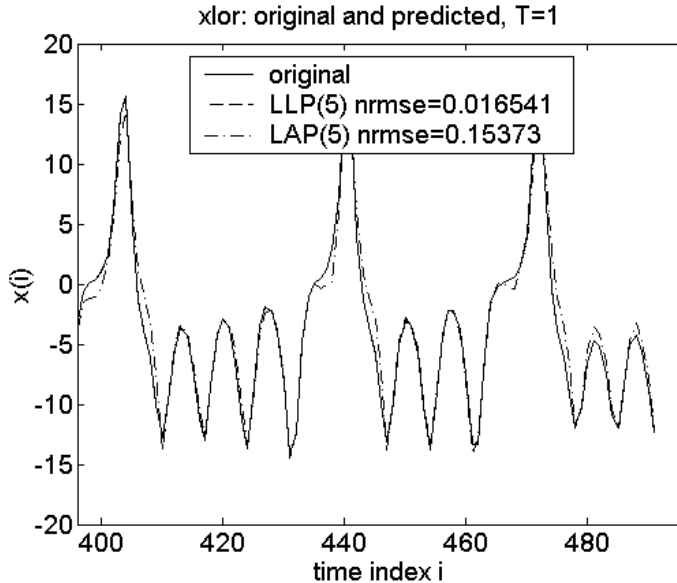
Παράδειγμα: x-Lorenz

Πρόβλεψη με:

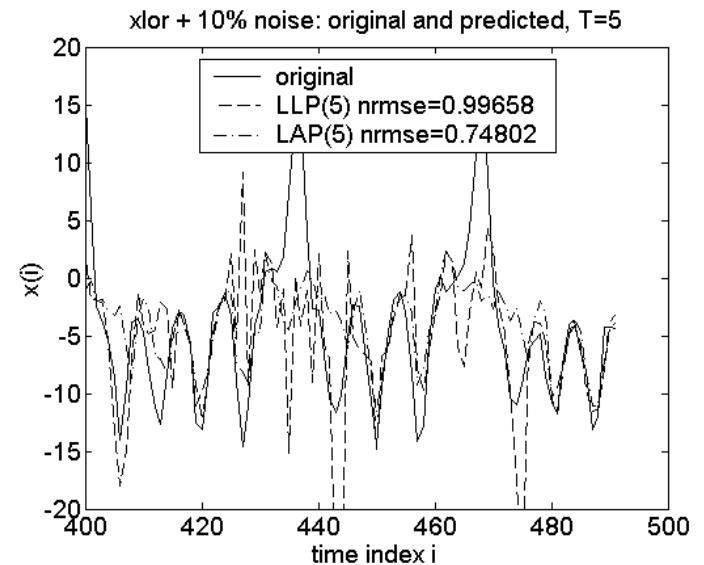
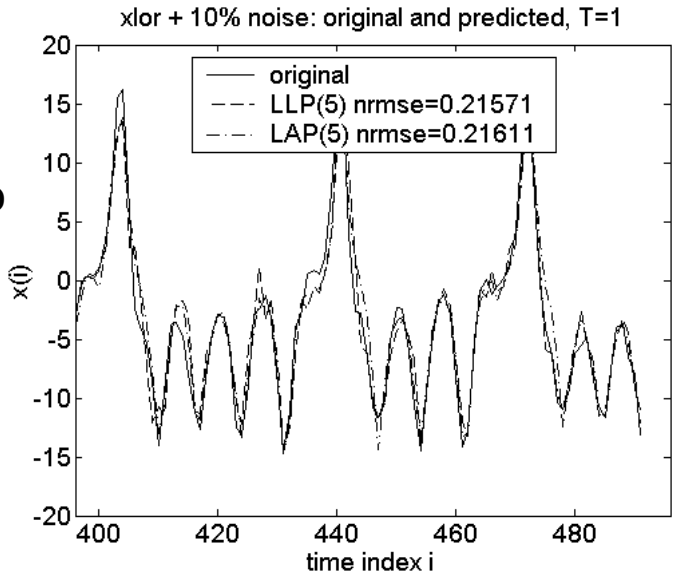
- τοπικό γραμμικό μοντέλο πρόβλεψης (LLP)
- πρόβλεψη τοπικού μέσου όρου (LAP)

$$\tau = 1, m = 5, K = 11$$

χωρίς θόρυβο



με 10%-θόρυβο

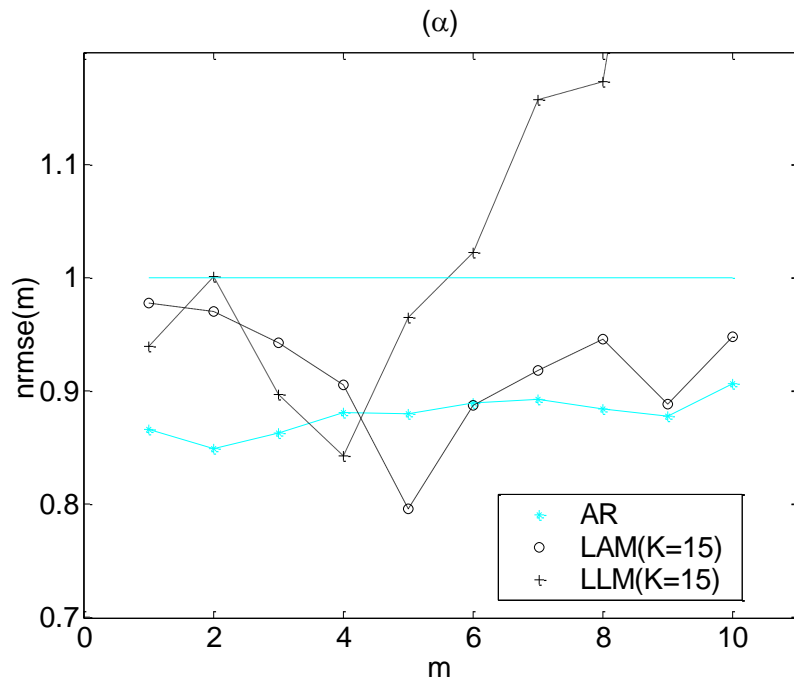


τετραμηνιαίος ρυθμός μεταβολής ΑΕΠ των ΗΠΑ την περίοδο 1947 – 1991

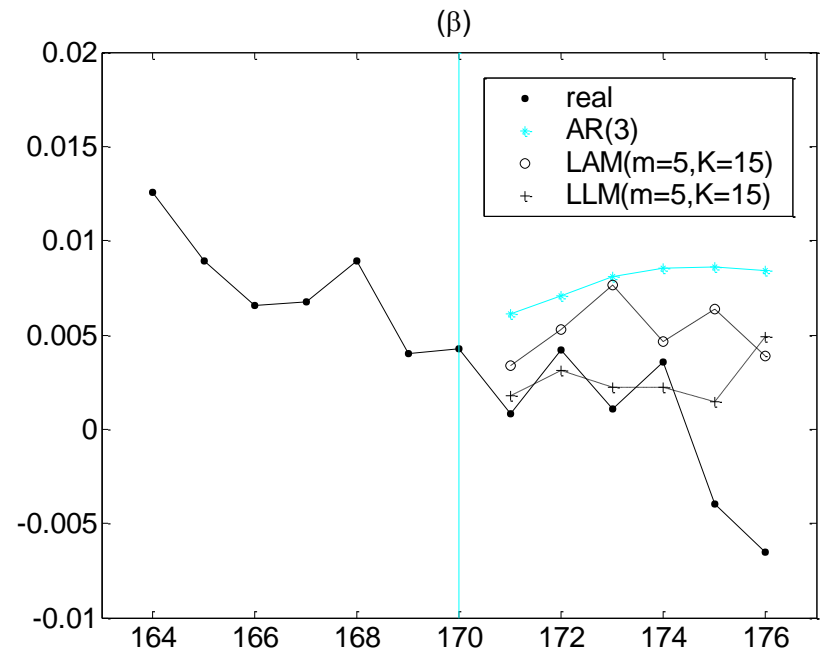
Πρόβλεψη με

- γραμμικό μοντέλο AR
- τοπικό μοντέλο μέσου όρου, LAM
- τοπικό γραμμικό μοντέλο, LLM

Σφάλμα πρόβλεψης (nrmse) για τα 30 τελευταία τετράμηνα



Προβλέψεις με αφετηρία το πρώτο τετράμηνο του 1989 και ορίζοντα τα 6 τελευταία τετράμηνα



Γενικός δείκτης ΧΑΑ την περίοδο 1.1.2002 – 20.9.2005

Πρόβλεψη με

- γραμμικό μοντέλο AR
- τοπικό μοντέλο μέσου όρου, LAM

30

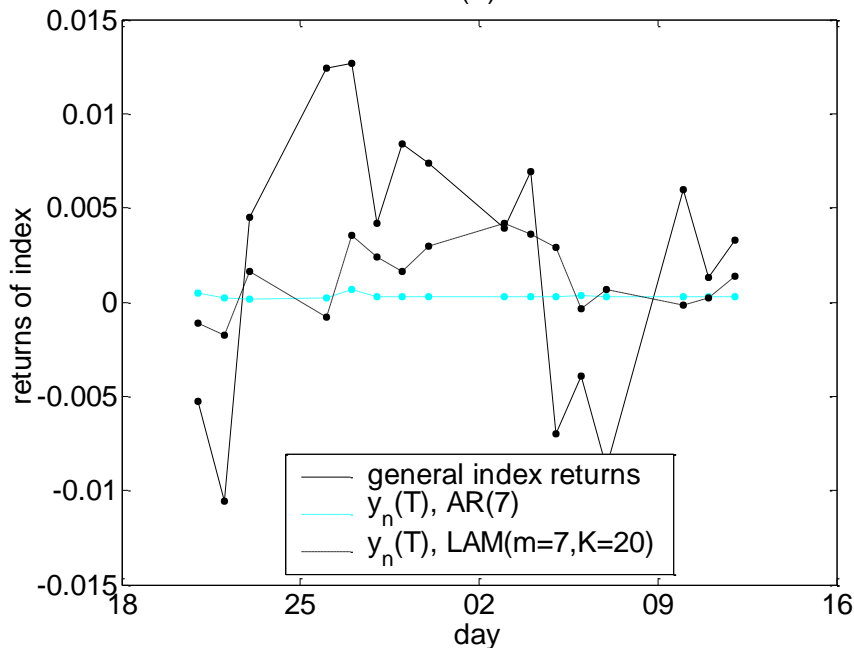
Ανάλυση χρονοσειρών με το διαχωρισμό σε εμπειρικούς τρόπους ταλάντωσης (empirical mode decomposition)

Πρόβλεψη με αφετηρία την 20.9.2005 και για 1 μέχρι 16 μέρες μπροστά

$$\text{Αποδόσεις } y_t = \frac{x_t - x_{t-1}}{x_{t-1}}$$

γενικού δείκτη

(α)



Γενικός δείκτης

(β)

