Χρονοσειρές

Δημήτρης Κουγιουμτζής

10 Μαΐου 2011

Χρονοσειρές - Εισαγωγή

Μοντέλο γραμμικής πολλαπλής παλινδρόμησης:
$x_1, x_2, \ldots, x_k$ ανεξάρτητες μεταβλητές,
y: εξαρτημένη μεταβλητή

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \epsilon,$$
$$\mathrm{E}[\epsilon] = 0, \quad \sigma_e^2 = \mathrm{Var}[\epsilon]$$

ή γενικά

$$y = f(x_1, x_2, \ldots, x_k) + \epsilon, \quad \text{ή} \quad y = f(x_1, x_2, \ldots, x_k, \epsilon)$$

Μοντέλο Χρονοσειρών: Χρονοσειρά $\{x_t\}_{t=1}^n = \{x_1, \ldots, x_n\}$
$y \longrightarrow x_t$
$x_1 \longrightarrow x_{t-1}, \; x_2 \longrightarrow x_{t-2}, \; \ldots, \; x_k \longrightarrow x_{t-k}$

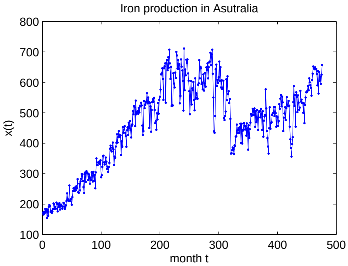
$$x_t = f(x_{t-1}, x_{t-2}, \ldots, x_{t-k}) + \epsilon, \quad \text{ή} \quad x_t = f(x_{t-1}, x_{t-2}, \ldots, x_{t-k}, \epsilon)$$

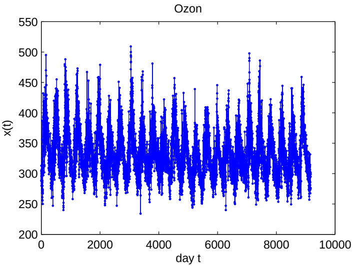
Χαρακτηριστικά χρονοσειρών

Χρονοσειρά $\{x_t\}_{t=1}^{n} = \{x_1, \ldots, x_n\}$
χρόνος δειγματοληψίας σταθερός

1. **Στασιμότητα**, πηγές μη-στασιμότητας:
   τάσεις, **περιοδικότητα** (εποχικότητα)
2. **Αιτιοκρατία** και **στοχαστικότητα**:
   Διερεύνηση και ταύτιση του αιτιοκρατικού μέρους του συστήματος που παράγει τη χρονοσειρά;
3. **Γραμμικότητα** και **μη-γραμμικότητα**
   - Γραμμικά αιτιοκρατικά δυναμικά συστήματα έχουν απλές λύσεις (σταθερό σημείο, περιοδικά σημεία ή τροχιές)
   - Εντοπισμός μη-γραμμικότητας είναι δύσκολος σε μια παρατηρούμενη στοχαστική σύστημα.
   Άρα τα πιθανά συστήματα που ενδιαφερόμαστε να διερευνήσουμε:
   γραμμική στοχαστική διαδικασία
   μη-γραμμικό δυναμικό (πιθανώς χαοτικό) σύστημα

Στασιμότητα, τάση και περιοδικότητα





Στασιμότητα ;

αυστηρή στασιμότητα

σππ $f(x_t, x_{t+1}, \ldots, x_{t+\tau})$ σταθερή για κάθε t και τ

ασθενής στασιμότητα

$\mathrm{E}[x_t]$, $\mathrm{Var}[x_t]$ και $\mathrm{Cov}[x_t, x_{t+\tau}]$ σταθερά για κάθε t και τ

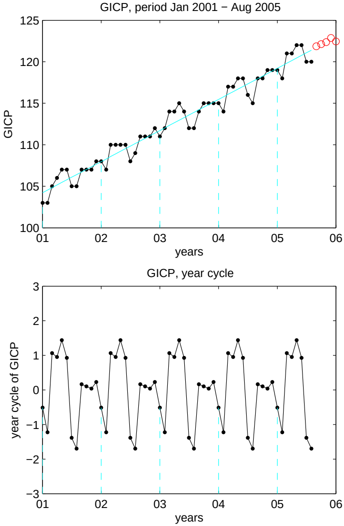
Απαλοιφή τάσης και περιοδικότητας

Διαχωρισμός της χρονοσειράς

$$x_t = \mu_t + s_t + y_t,$$

- μ_t: συνιστώσα τάσης
- s_t: συνιστώσα περιοδικότητας (ή εποχικότητας) για περίοδο d $(s_{t-d} = s_t)$
- y_t: χρονοσειρά των υπολοίπων

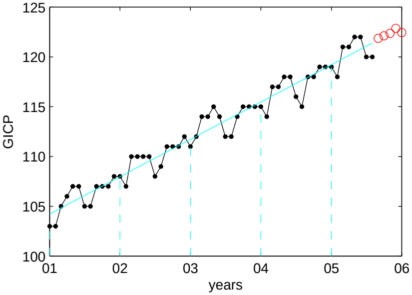
Παράδειγμα: GICP



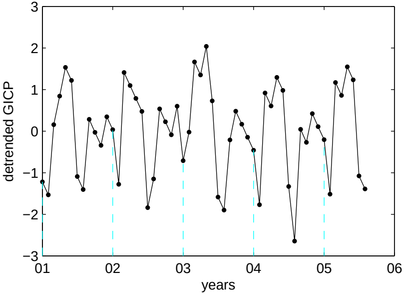
Παράδειγμα: GICP (συνέχεια)

$\{x_t\}_{t=1}^{56}$

Εκτίμηση τάσης: $\mu_t = 103.9 + 0.31t$

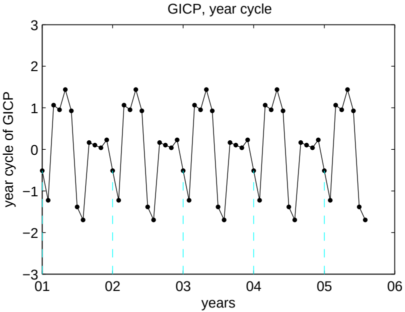


$x'_t = x_t - \mu_t$

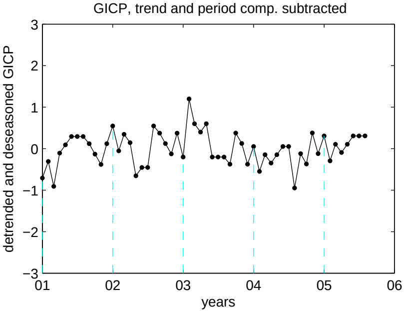


Παράδειγμα: GICP (συνέχεια)

$\{s_t\}_{t=1}^{56}$ για $d = 12$



$y_t = x'_t - s_t = x_t - \mu_t - s_t$



Παράδειγμα: GICP (συνέχεια)

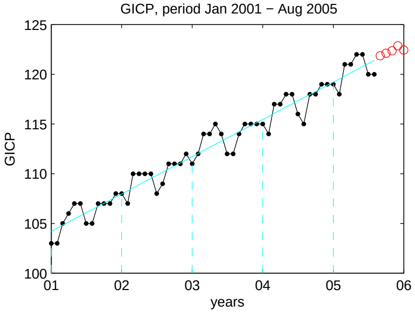
Πρόβλεψη για x_{n+1} ; $n+1 = 57$, Σεπτέμβριος 2005

Προέκταση τάσης για $n+1$:

$\mu_{n+1} = 103.9 + 0.31(n+1) \quad \Rightarrow \quad \mu_{57} = 103.9 + 0.31 \cdot 57 = 121.7$

Εκτίμηση περιοδικής συνάρτησης: $s_{n+1} = s_{57} = s_9 = 0.16$

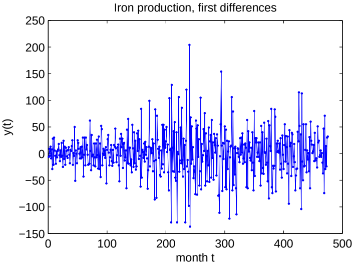
Πρόβλεψη: $\hat{x}_{57} = x_{56}(1) = \mu_{57} + s_{57} = 121.7 + 0.16 = 121.86$

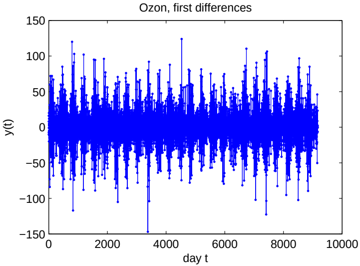


Απαλοιφή τάσης με πρώτες διαφορές

στοχαστική τάση: δε μπορεί να περιγραφεί ως κάποια γνωστή συνάρτηση του χρόνου

πρώτες διαφορές: $y_t = \nabla x_t = x_t - x_{t-1}$





Συσχέτιση σε χρονοσειρά, Λευκός θόρυβος

$x_t, x_{t+1}, \ldots, x_{t+\tau}$ **ανεξάρτητες τυχαίες μεταβλητές με ίδια κατανομή (iid)**:

- $f_{x_t}(x) = f_{x_{t+1}}(x) = \ldots = f_{x_{t+\tau}}(x) = f(x)$
- $f_{x_t, x_{t+j}}(x, y) = f_{x_t}(x) f_{x_{t+j}}(y)$ για $j = 1, 2, \ldots$

Μια iid χρονοσειρά είναι εντελώς τυχαία (δεν περιέχει αυτοσυσχετίσεις γραμμικές ή μη-γραμμικές)

Μια iid χρονοσειρά λέγεται και **λευκός θόρυβος** $\mathrm{WN}(0, \sigma_\epsilon^2)$

Αν τα iid στοιχεία της χρονοσειράς ακολουθούν Γκαουσιανή κατανομή, τότε η χρονοσειρά λέγεται **Γκαουσιανός λευκός θόρυβος**

Τυχαίος περίπατος

τυχαίος περίπατος:

$$x_t = x_{t-1} + \epsilon_t, \quad \epsilon_t \sim \mathrm{WN}(0, \sigma_\epsilon^2)$$

Αρχίζοντας από κάποια τιμή μ (για $t = 0$):

$$x_t = \mu + \sum_{j=1}^{t} \epsilon_j,$$

Ο τυχαίος περίπατος είναι μη-στάσιμη χρονοσειρά με στοχαστικές τάσεις.

Παίρνοντας τις πρώτες διαφορές προκύπτει η στάσιμη χρονοσειρά του λευκού θορύβου.

Γραμμική συσχέτιση

Δύο τ.μ. X, Y:

- συνδιασπορά $\sigma_{XY} \equiv \mathrm{Cov}[X, Y]$
- συντελεστής συσχέτισης $\rho_{XY} \equiv \mathrm{Corr}[X, Y]$

Χρονοσειρά $\{x_t\}$:

- αυτοδιασπορά $\mathrm{Cov}[x_t, x_{t-\tau}]$ για κάποια υστέρηση τ
- **(γραμμική) αυτοσυσχέτιση** για κάποια υστέρηση τ: $\rho_\tau = \mathrm{Corr}[x_t, x_{t-\tau}]$

Εκτίμηση της αυτοσυσχέτισης

$$\hat{\rho}_\tau = r_\tau = \mathrm{Corr}(x_t, x_{t-\tau}) = \frac{\sum_{t=\tau+1}^{n}(x_t - \bar{x})(x_{t-\tau} - \bar{x})}{\sum_{t=\tau+1}^{n}(x_t - \bar{x})^2}$$

Παραδείγματα αυτοσυσχέτισης

$$x_t \sim \mathrm{WN}(0, \sigma_x^2) \implies r_\tau \sim \mathrm{N}(0, 1/n)$$

'στατιστικά μηδενική' αυτοσυσχέτιση αν $r_\tau \in [-2/\sqrt{n}, 2/\sqrt{n}]$