

Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Η/Υ, Θέμα εργασίας στο μάθημα «Χρονοσειρές»

Οδηγίες:

Μπορούν να γίνουν ομάδες ως και δύο ατόμων. Ο κάθε φοιτητής ή ομάδα φοιτητών θα πρέπει να στείλει e-mail στο διδάσκοντα, δηλώνοντας ονοματεπώνυμο και έγκυρο e-mail για κάθε μέλος της ομάδας, για να πάρει έναν αριθμό από 1 ως 15, που θα αντιστοιχεί σε δεδομένα της εργασίας. Σχετικά με την παράδοση της εργασίας θα πρέπει:

- Το κείμενο της αναφοράς της ανάλυσης που ζητείται να είναι γραμμένο σε κάποιο πρόγραμμα επεξεργασίας κειμένου (Word, LaTeX).
- Τα προγράμματα που ενδεχομένως δημιούργησε ο φοιτητής ή η ομάδα φοιτητών, τα γραφήματα και οι πίνακες θα πρέπει να παρουσιάζονται στο σημείο του κειμένου που αναφέρονται.
- Η εργασία θα πρέπει να σταλεί ως συνημμένο αρχείο με e-mail στη διεύθυνση dkugiu@gen.auth.gr ως την **Πέμπτη 3/3/2016**. Θα ακολουθήσει ενδεχόμενη **προφορική παρουσίαση** της εργασίας σε μέρα/ώρα που θα καθορισθεί.

Γενικά - δεδομένα

Η υπολογιστική εργασία είναι στην ανάλυση σεισμολογικών δεδομένων (χρονοσειρών). Στην ιστοσελίδα του μαθήματος δίνονται 15 αρχεία χρονοσειρών, με όνομα $x<i>i>.dat$, όπου i είναι δείκτης από 1 ως 15. Το κάθε αρχείο έχει παρατηρήσεις από την κατακόρυφη συνιστώσα της επιφανειακής διαταραχής του εδάφους (mechanic vibration), γνωστής και ως σεισμικό κύμα (seismic wave), που καταγράφηκαν από ένα σεισμολογικό σταθμό. Οι παρατηρήσεις έχουν συλλεχθεί με σταθερό χρόνο δειγματοληψίας 1/125 sec για 16 sec, δηλαδή το κάθε αρχείο έχει 2000 παρατηρήσεις.

Ο κάθε φοιτητής η ομάδα φοιτητών θα μελετήσει τη χρονοσειρά που του / τους αντιστοιχεί. Θεωρούμε πως όλες οι χρονοσειρές είναι στάσιμες.

Πρώτο στάδιο - Γραμμική ανάλυση για το έτος 1989, 2000 και 2011

Θα πρέπει να αντιμετωπίσετε τα παρακάτω ερωτήματα:

1. Είναι η χρονοσειρά λευκός θόρυβος (white noise) ή υπάρχουν σημαντικές αυτοσυσχετίσεις; Για αυτό θα πρέπει να συμπεριλάβετε και κατάλληλο έλεγχο υπόθεσης με βάση την αυτοσυσχέτιση.
2. Σε συνέχεια του ερωτήματος 1 και αν το επιτρέπει το συμπέρασμα σας στο ερώτημα 1, να γίνει διερεύνηση για κατάλληλη τάξη αυτοπαλινδρομούμενου μοντέλου (AR) για τη χρονοσειρά των 1900 πρώτων παρατηρήσεων. Έστω p η επιλεγμένη τάξη. (matlab: ar και aic)
3. Να γίνει πρόβλεψη με το μοντέλο AR(p) για ένα χρονικό βήμα μπροστά ($T=1$) για τις τελευταίες 100 παρατηρήσεις (δηλαδή να γίνουν οι προβλέψεις \hat{x}_{t+1}) και να υπολογισθεί το $nmse$. (matlab: predict)
4. Οι προβλέψεις στο 3 να γίνουν επίσης με μοντέλο μηδενικής τάξης για υστέρηση και διάσταση εμβύθισης $m=p$ (όση και η τάξη του AR). (matlab: χρήση ειδικής συνάρτησης,δες εργαστήριο μαθήματος)
5. Να γίνει σχεδιάγραμμα των προβλέψεων ενός βήματος για τα δύο μοντέλα (AR και τοπικό μοντέλο) καθώς και των αντίστοιχων πραγματικών τιμών (σε ένα σχήμα).
6. Να συγκριθούν τα δύο μοντέλα και να σχολιαστεί η χρησιμότητα τους για πρόβλεψη με βάση τα αποτελέσματα από τα παραπάνω 3 ερωτήματα.

Δεύτερο στάδιο – Γραμμική και Μη-γραμμική ανάλυση για το έτος που σας αντιστοιχεί

Στο δεύτερο στάδιο ανάλυσης, θα διερευνήσετε αν η χρονοσειρά έχει γραμμικές ή/και μη-γραμμικές (αυτο)συσχετίσεις και πως αυτό ανιχνεύεται αν θεωρήσουμε μια μικρή σειρά (500 παρατηρήσεων, τη δεύτερη χρονοσειρά) ή όλη τη διαθέσιμη σειρά (την πρώτη χρονοσειρά). Κάποια από τα βήματα της ανάλυσης που μπορείτε να κάνετε για κάθε μια από τις δύο χρονοσειρές είναι:

1. Σχεδιάγραμμα της χρονοσειράς (matlab: συνάρτηση plot, pser, pserlarge).
2. Στατιστικός έλεγχος ανεξαρτησίας με βάση την αυτοσυσχέτιση (Portmanteau test). (matlab: χρήση συνάρτησης portmanteauLB).
3. Εκτίμηση της υστέρησης τ που δίνει το κριτήριο της αμοιβαίας πληροφορίας (matlab: χρήση συνάρτησης mutualinformation).
4. Εκτίμηση της διάστασης εμβύθισης m που σας δίνει το κριτήριο των ψευδών κοντινότερων γειτόνων (false nearest neighbors) χρησιμοποιώντας ως υστέρηση αυτή που βρήκατε παραπάνω (matlab: χρήση συνάρτησης falsenearest).
5. Προσαρμογή με τοπικό μοντέλο μέσου όρου και τοπικό γραμμικό μοντέλο για υστέρηση τ και διάσταση εμβύθισης m που εκτιμήσατε παραπάνω (και τον ίδιο αριθμό γειτονικών σημείων) (matlab: χρήση συνάρτησης localfitnmse και localpredictnmse). Σύγκριση με την προσαρμογή με γραμμικό μοντέλο.

6. Εκτίμηση της διάστασης συσχέτισης για διαστάσεις $m=1, \dots, 10$, χρησιμοποιώντας την υστέρηση από το κριτήριο της αμοιβαίας πληροφορίας με κατάλληλα σχήματα (matlab: χρήση συνάρτησης `correlationdimension`).

Με βάση τα αποτελέσματα από τα μέτρα θα πρέπει να σχολιάσετε για τη μορφή του συστήματος της χρονοσειράς, δηλαδή αν είναι πλήρως στοχαστικό ή όχι και αν είναι γραμμικό ή μη-γραμμικό.

Στις απαντήσεις θα πρέπει να συμπεριλάβετε πίνακες αποτελεσμάτων και σχήματα, όπου είναι απαραίτητο.

Μεταφορά σχημάτων από το matlab σε πρόγραμμα επεξεργασίας κειμένου:

Για τη μεταφορά σχήματος από το matlab στο πρόγραμμα επεξεργασίας κειμένου θα πρέπει να αποθηκεύσετε το σχήμα σε αρχείο, π.χ. χρησιμοποιώντας το μενού στο παράθυρο του σχήματος (File → Export). Για μεταφορά στο Word υπάρχουν διάφορες κατάλληλες μορφές αρχείου εικόνας, η πιο απλή είναι “Enhanced Metafile” (*.emf). Για μεταφορά στο LaTeX η πιο κατάλληλη μορφή είναι “Encapsulated Postscript” (*.eps). Η φόρτωση του αρχείου δεδομένου μπορεί εύκολα να γίνει με την εντολή `load` και η αποθήκευση δεδομένων (μεταβλητών) με την εντολή `save`. Δίνοντας “help <function>” μπορεί κάποιος να δει τη βοήθεια για κάθε συνάρτηση / εντολή.