**Ανάλυση Χρονοσειρών**

**Υπολογιστική Εργασία 2018-2019**

Σαμπαζιώτης Σάββας, 7974

Κεχαγιάς Ανδρέας, ΧΧΧΧ

**Εισαγωγή**

Το κείμενο αυτό αποτελεί την αναφορά στην υπολογιστική εργασία του μαθήματος Ανάλυσης Χρονοσειρών, του ακαδημαϊκού έτους 2018-2019. Παρακάτω παρουσιάζονται οι δύο διαφορετικές μελέτες που προδιαγράφονται στην εκφώνηση της εργασίας, πάνω στη δεδομένη χρονοσειρά ***dat19.mat.***



Η παραπάνω χρονοσειρά διαχωρίστηκε συνολικά σε 11 μη-επικαλυπτόμενα παράθυρα των 250 δειγμάτων τα οποία παρουσιάζονται παρακάτω.



**Μελέτη Α**

1. **Απαλοιφή Τάσης**

Πρωταρχική ανάγκη της ανάλυσης είναι η μετατροπής της κάθε χρονοσειράς σε στάσιμο σήμα. Για κάθε χρονοσειρά, δοκιμάστηκαν διάφορες μέθοδοι μετασχηματισμού. Η τελικές στάσιμες χρονοσειρές προκύπτουν από την μέθοδο του Κινούμενου Μέσου τάξης 21.



Ο ακριβής μετασχηματισμός που εφαρμόστηκε περιγράφεται στον παρακάτω ψευδοκώδικα. Η υλοποίηση και παρουσίαση γίνεται στο m-script *“convert2Stationary.m”*

|  |  |
| --- | --- |
| Αλγόριθμος | Σχόλια |
|  | *Φόρτωση χρονοσειράς dat19* |
|  | *Αφαίρεση εξομαλυμένου σήματος (εκτιμώμενη τάση) από το πρωταρχικό.* |
| 1. *Για κάθε*     1. *b=k\*250* | *Διαχωρισμός Χρονοσειράς σε 11 μη-επικαλυπτόμενα παράθυρα χρονοσειρών , 250 δειγμάτων η κάθε μία.* |
|  | *Αφαίρεση εξοκύμενων τιμών βάσει του κανόνα της διασποράς: Τα δείγματα εκτός των ορίων , αντικαθίστανται με τιμές από γραμμική παρεμβολή των γειτονικών του.* |
|  | *Απαλοιφή μέσου όρου* |
|  | *Κανονικοποίηση κατά διασπορά* |
| 1. *Τέλος* |  |

Όπως παρατηρεί κανείς, επιλέχθηκε να γίνει η απαλοιφή της συνολικής τάσης και έπειτα να γίνει ο διαχωρισμός στα παράθυρα των 250 χρονοσειρών. Με αυτόν τον τρόπο, ελαχιστοποιήθηκε το πλήθος των δειγμάτων που θα «παραμελούνταν» από το φιλτράρισμα, καθώς η τεχνική Κινούμενου Μέσου τάξης , παράγει τοπικές τιμές στο διάστημα και μόνο, αφήνοντας δείγματα αφιλτράριστα.

**Εναλλακτικές τεχνικές που εξετάσθηκαν**

Όπως αναφέρθηκε αρχικά, πραγματοποιήθηκαν και επιπλέον δοκιμές απαλοιφής της τάσης, όπως εκτίμηση της τάσης με MA φιλτράρισμα μικρότερης τάξης, η εφαρμογή πρώτων και δεύτερων διαφορών, μέχρι και η εξέταση στασιμότητας με το Augmented Dickey Fuller Test***.*** Ωστόσο, το ADF test αποτελούσε ιδιαίτερη διερεύνηση για τη επιλογή κατάλληλου πλήθους lags, με αποτέλεσμα να εγκαταλειφτεί ως μετρική στασιμότητας.

Το βασικό δίλλημα ήταν μεταξύ των απλών πρώτων διαφορών και αυτή του κινούμενου μέσου. Εν τέλει, επιλέχθηκε η ΜΑ βάσει της παρακάτω παρατήρησης στις δειγματικές μερικές αυτοσυσχετίσεις που προκύπτουν:



* Παρατηρούμε, πως στη περίπτωση των πρώτων διαφορών, η χρονοσειρά είναι πλέον στάσιμη μεν, αλλά με ενισχυμένο θόρυβο δε. Το γεγονός αυτό αντικατοπτρίζεται και στην δειγματική αυτοσυσχετίσεις και μερική αυτοσυσχέτιση, η οποία προσεγγίζει αυτές των idd διαδικασιών.
* Αντίθετα, η MA(21) τεχνική, περιορίζει τον θόρυβο και αναδεικνύει με όμορφο τρόπο την χρήσιμη πληροφορία την οποία αποζητάμε να προβλέψουμε και να περιγράψουμε. Επίσης, η μερική αυτοσυσχέτιση φανερώνει την τάξη AR μοντέλου που θα μελετηθεί στην συνέχεια.

Βάσει των παραπάνω συμπερασμάτων, επιλέχθηκε η MA μέθοδος, καθώς αναμένεται να διευκολύνει σημαντικά την επιλογή κατάλληλου γραμμικού μοντέλου μετέπειτα.

Θεωρητικά, αναμένεται η χρήση ενός ARMA(p,q) μοντέλου με υψηλό p και με πρώτες διαφορές για απαλοιφή τάσης, να έχει παρόμοια απόδοση με AR μοντέλου με MA απαλοιφή τάσης. Ωστόσο, για τον προαναφερθέντα λόγο, επιλέχθηκε η τακτική που θα διευκολύνει την μοντελοποίηση μετέπειτα.