8/11/2004

Τ.Ε.Ι. ΠΑΤΡΑΣ
ΣΧΟΛΗ ΔΙΟΙΚΗΣΗΣ ΚΑΙ ΟΙΚΟΝΟΜΙΑΣ
ΤΜΗΜΑ ΕΠΙΧΕΙΡΗΜΑΤΙΚΟΥ ΣΧΕΔΙΑΣΜΟΥ & ΠΛΗΡΟΦΟΡΙΑΚΩΝ ΣΥΣΤΗΜΑΤΩΝ
Η' ΕΞΑΜΗΝΟ

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

Θέμα : «Μελέτη, ανάλυση, σχεδίαση και υλοποίηση (Τεχνικών) αλγορίθμων εκτίμησης και πρόβλεψης για Οικονομικό Χαρτοφυλάκιο»

Επιβλέπων :
Δρ. Ντεμίρης Ηλίας, Μπεληγιάννης Γρηγόριος

Σπουδαστές :
Λότσιου Γεωργία
Παππάς Πέτρος

Πίνακας Περιεχομένων

1. Εισαγωγή

1. Εισαγωγή στις χρονικές σειρές

Χρονική σειρά είναι μια οικογένεια τυχαίων μεταβλητών x_t , t E T , όπου T είναι μια χρονική περίοδος ή υποσύνολο του χώρου. Αν T συνεχές , η χρονική σειρά λέγεται συνεχής, ενώ αν T διακριτό , η σειρά λέγεται διακριτή.

Η εργασία θα ασχοληθεί με διακριτές σειρές στις οποίες μάλιστα οι παρατηρήσεις λαμβάνονται σε χρονικές στιγμές που ισαπέχουν μεταξύ τους. Αν λοιπόν έχουμε N διαδοχικές τιμές, τότε γράφουμε $X_1, X_2, \ldots X_N$ για να δηλώσουμε παρατηρήσεις που έγιναν στους χρόνους $\tau_o + h$, $\tau_o + 2h$, $\tau_o + Nh$. Το τ_o και το h μπορεί να είναι σημαντικά αλλά και άνευ σημασίας. Γενικά x_t είναι η παρατήρηση στο χρόνο t.

Ακόμη ξέρουμε ότι ένα στατιστικό φαινόμενο που εκτυλίσσεται μέσα στο χρόνο, σύμφωνα με τους νόμους των πιθανοτήτων, καλείται στοχαστική διαδικασία. Η χρονική σειρά μπορεί να θεωρηθεί σαν στοχαστική διαδικασία με πεπερασμένο πλήθος παρατηρήσεων.

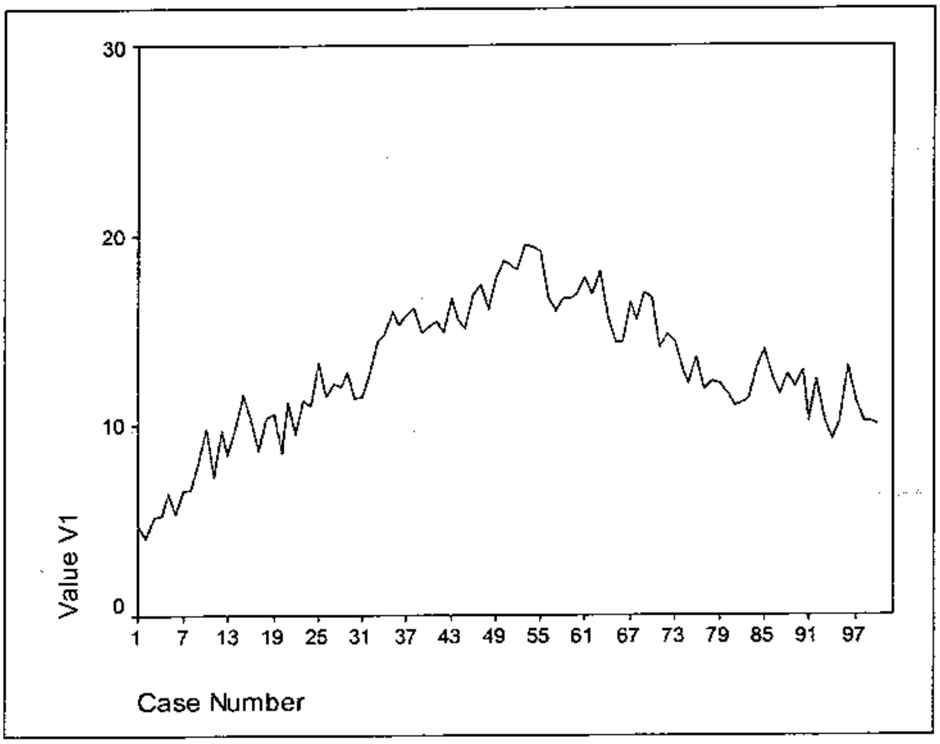
Τέλος, θα μπορούσαμε να πούμε απλά ότι η χρονική σειρά είναι ένα σύνολο παρατηρήσεων που λαμβάνονται παίρνοντας μετρήσεις μιας τυχαίας μεταβλητής, σε κανονικά χρονικά διαστήματα.

Τα υπολογιστικά προβλήματα που προκύπτουν από τη μελέτη των χρονικώ νσειρών είναι τεράστια και η εμφάνιση ηλεκτρονικών υπολογιστών έδωσε μεγάλη ώθηση στην ανάλυση και επεξεργασία τέτοιων σειρών . Η μελέτη χρονικών παρουσιάζει πρακτικό και θεωρητικό ενδιαφέρον.

Επειδή υπάρχουν αρκετές θεωρητικές δυσκολίες και πολύπλοκα προγράμματα, η μελέτη χρονικών σειρών, ήταν προνόμιο λίγων .Όμως με την εμφάνιση του βιβλίου Box – Jenkins (1976) η θεωρία συνδέθηκε με την πράξη και το θέμα έγινε ευρύτατα γνωστό.

Στην οικονομία, κλιματολογία, Βιολογία, Φυσική, Οικολογία, Ωκεανογραφία κλπ έχουμε παρατηρήσει που μπορούν αν χαρακτηριστούν ως χρονικές σειρές.

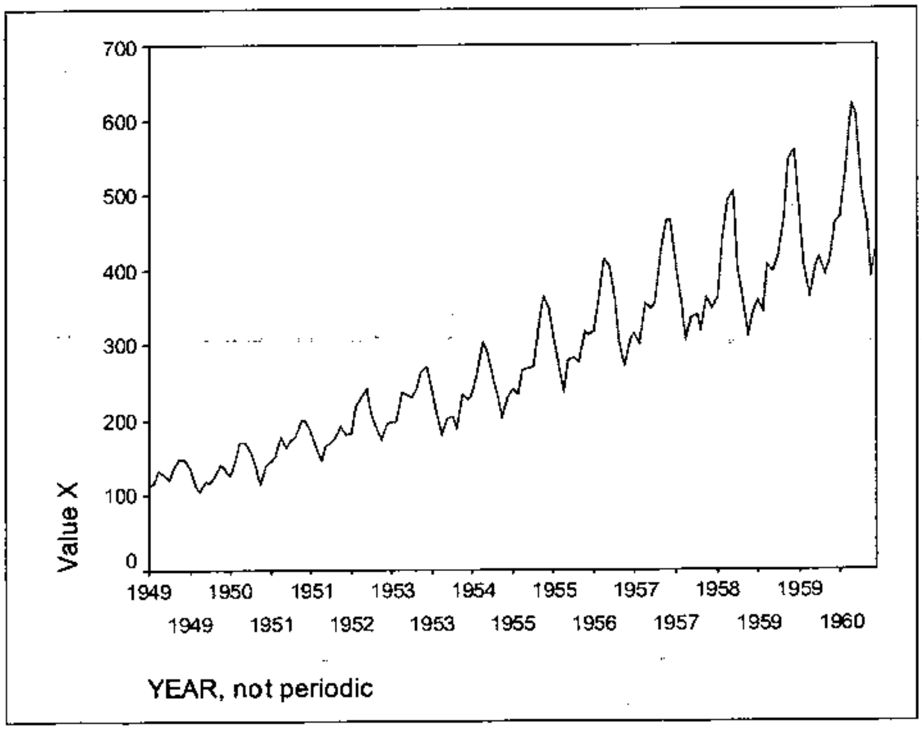
Τα παρακάτω είναι μερικά παραδείγματα χρονικών σειρών.



1° Παράδειγμα

Ας υποθέσουμε ότι το 1° παράδειγμα είναι ένα μέρος μιας κλασικής χρονοσειράς που αναφέρεται στις πωλήσεις ενός βιομηχανικού προϊόντος στους τελευταίους 100 μήνες .

Στο δεύτερο παράδειγμα έχουμε Επιβάτες αερογραμμών (σε χιλιάδες) από τον Ιανουάριο του 1949 έως και τον Δεκέμβριο του 1959. Παρατηρούμε ότι αυτή η σειρά εμφανίζει περιοδικότητα και μια συνεχή άνοδο, δηλαδή μια τάση (trend)



2° Παράδειγμα

2. Οι κυριότερες συνιστώσες μιας χρονολογικής σειράς

Οι χρονολογικές σειρές παρουσιάζουν μεταβολές με μορφή και ένταση που κάθε φορά διαφέρει.

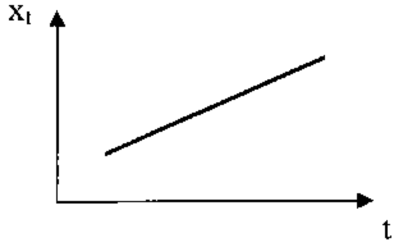
Εξετάζοντας τις μεταβολές αυτές, που ονομάζονται κινήσεις της μεταβλητής x_t σε συνάρτηση με το χρόνο t μιας χρονολογικής σειράς, διακρίνουμε κυρίως τα ακόλουθα είδη κίνησης :

I. Μακροχρόνια τάση ή γενική τάση (trend)
II. Τις περιοδικές μεταβολές
III. Τις κυκλικές μεταβολές
IV. Τις άρρυθμες ή ακανόνιστες ή απρόοπτες μεταβολές

Α. Μακροχρόνια τάση

Αν για μεγάλη περίοδο οι τιμές μιας χρονολογικής σειράς τείνουν να αυξηθούν ή να μειωθούν, τότε λέμε ότι η σειρά των παρατηρήσεων παρουσιάζει μακροχρόνια τάση. Δηλαδή, *τάση* είναι η μακροχρόνια αύξηση ή μείωση που παρατηρειται στα δεδομένα.

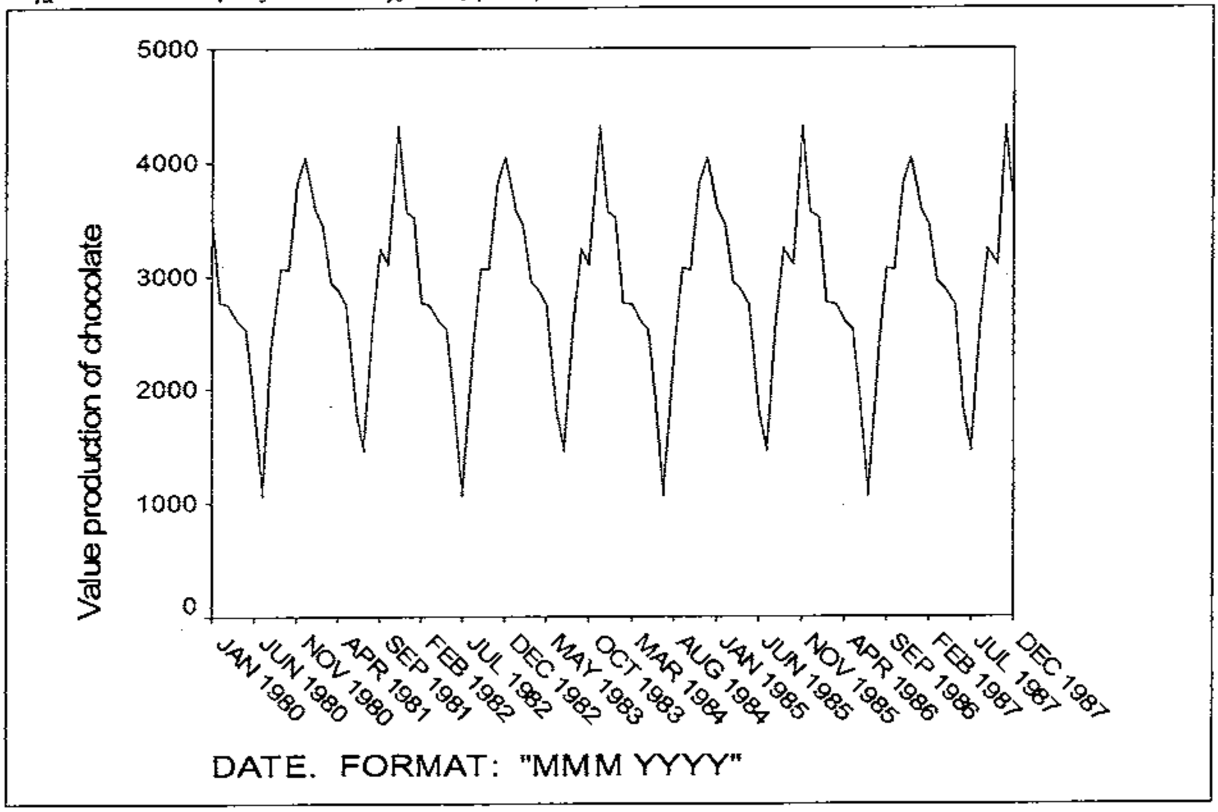
Το χαρακτηριστικό της μακροχρόνιας τάσης είναι ότι έχουμε μακροχρόνια και σταθερή κίνηση των οικονομικών μεγεθών που επηρεάζεται από γενικότερους παράγοντες



Β. Τις περιοδικές μεταβολές

Οι περιοδικές μεταβολές είναι εκείνες που επαναλαμβάνονται κατά ορισμένα χρονικά διαστήματα μέσα σε ορισμένη χρονική περίοδο.

Οι πιο συχνές περιοδικές μεταβολές είναι εκείνες που συμβαίνουν μέσα σε ένα χρόνο και ονομάζονται *εποχιακές μεταβολές.*



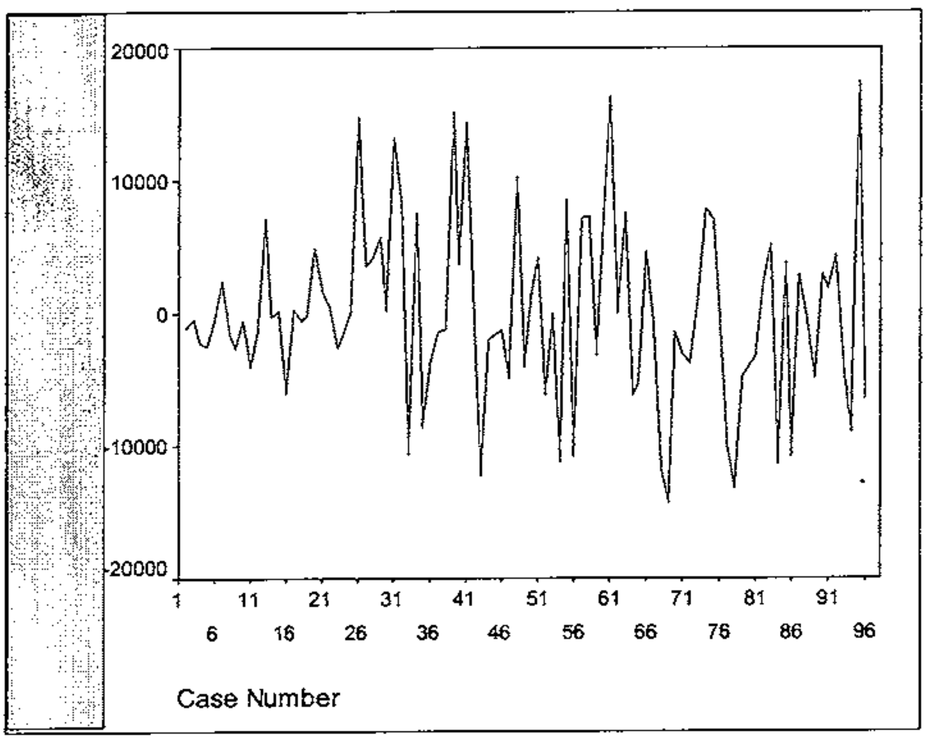
Γ. Κυκλικές μεταβολές

Κυκλικές μεταβολές ή κυκλικές διακυμάνσεις είναι οι ταλαντώσεις γύρω από μια γραμμή ή καμπύλη τάσης σε μια μακροχρόνια περίοδο

Οι κυκλικές μεταβολές διαφέρουν από τις περιοδικές γιατί είναι διάρκειας μεγαλύτερης από ένα έτος και δεν παρουσιάζουν, γενικά , κανονική περιοδικότητα.

Δ. Ακανόνιστες μεταβολές

Οι ακανόνιστες μεταβολές διακρίνονται σε δύο κατηγορίες : στις *συμπτωματικές* και στις *τυχαίες*. Οι συμπτωματικές μεταβολές προέρχονται από εξαιρετικά και απρόβλεπτα γεγονότα, όπως είναι οι σεισμοί , οι θύελλες , οι απεργίες, οι επιδημίες, οι πόλεμοι κλπ. ενώ οι τυχαίες μεταβολές οφείλονται σε πολυάριθμους και άγνωστους παράγοντες ή όπως συνήθως λέγεται στην τύχη.



3. Γιατί συγκεντρώνουμε δεδομένα χρονολογικών σειρών

Άραγε, η συγκέντρωση τέτοιων δεδομένων σε τι είδους ερωτήσεις θα μπορούσε να δώσει απάντηση ;

Μια αιτία γι αυτή τη συγκέντρωση θα μπορούσε να είναι η αναζήτηση ενός απλού μαθηματικού μοντέλου που να εξηγεί τη συμπεριφορά της σειράς . Μια άλλη αιτία θα μπορούσε να είναι η χρησιμοποίηση αυτού μοντέλου για πρόβλεψη, πράγμα πολύ σημαντικό για τις επιχειρήσεις και την επιστήμη.

Τα βήματα που ακολουθούμε για την ανάλυση της χρονικής σειράς είναι

1. Εξασφαλίζουμε ότι η σειρά έχει σταθερή συμπεριφορά (στατικότητα ή αλλιώς στασιμότητα της σειράς)
2. εξασφαλίζουμε ότι δεν υπάρχουν περιοδικότητες
3. εξασφαλίζουμε κατάλληλο μοντέλο που να προσαρμόζεται στα δεδομένα μας
4. εκτιμούμε άγνωστες παραμέτρους του μοντέλου μας και τέλος
5. χρησιμοποιούμε το μοντέλο για προβλέψεις

Η ανάλυση χρονικών σειρών μπορεί να γίνει με δύο διαφορετικούς τρόπους. Ο πρώτος υποθέτει ότι κάθε παρατήρηση στο χρόνο t είναι συνάρτηση της συμπεριφοράς της χρονοσειράς στους προηγούμενους χρόνους, ενώ ο δεύτερος, υποθέτει ότι κάθε παρατήρηση, είναι συνάρτηση ημίτονων και συνημίτονων με διαφορετικές συχνότητες.

Όταν χρησιμοποιούμε τον πρώτο τρόπο λέμε ότι κάνουμε ανάλυση χρονικών σιρών σε πεδίο χρόνων (time domain) ενώ στην άλλη περίπτωση, κάνουμε ανάλυση χρονικών σειρών «πεδίο συχνοτήτων» (frequency domain) .

Εδώ θα ασχοληθούμε με την ανάλυση χρονικών σειρών μόνο σε πεδίο χρόνων

Ιδιαίτερα θα ασχοληθούμε με τα εξής τρία μοντέλα :

i. Αυτοπαλινδρομούμενα (autoregressive) τάξης p , που συμβολίζονται AR(p)
ii. Κινούμενου μέσου (moving average) τάξης q , που συμβολίζονται MA (q) και
iii. Μεικτά τάξης (p, q) που τα συμβολίζουμε ARMA (p, q)

Αν τα δεδομένα μας χρειάζονται κάποια εξομάλυνση τότε η ανάλυση χρονικών σειρών απαιτεί περισσότερο σύνθετα μοντέλα .

Σήμερα υπάρχουν έτοιμα προγράμματα με τη βοήθεια των οποίων μπορούμε να κάνουμε την ανίχνευση του μοντέλου , την εκτίμηση των παραμέτρων του καθώς και την πρόβλεψη.

Το στατιστικό πακέτο SPSS διαθέτει προγράμματα χρονικών σειρών. Με ένα τέτοιο πρόγραμμα ασχολήθηκε η εργασία . Χρησιμοποιώντας τη Matlab, υλοποιήθηκε ένα μενού επιλογών που έδινε τη δυνατότητα στο χρήστη να αναλύει χρονοσειρές με τη βοήθεια του αλγορίθμου των Box – Jenkins και των φίλτρων Kalman.

2. Προβλέψεις

1. Η φύση και η χρήση των προβλέψεων

Η πρόβλεψη είναι ένα στοιχείο – κλειδί στη λήψη αποφάσεων. Το γεγονός αυτό δεν αποτελεί έκπληξη αφού η απόλυτη αποτελεσματικότητα οποιασδήποτε απόφασης στηρίζεται πάνω στα γεγονότα που ακολουθούν την απόφαση. Η ικανότητα να προβλέψουμε τις ανεξέλεγκτες πλευρές αυτών των γεγονότων στηρίζεται κατά πολύ στην δυνατότητά μας να διατυπώνουμε μαθηματικά υποδείγματα.

Για αυτόν το λόγο, τα συστήματα διοίκησης για το σχεδιασμό και τον έλεγχο τυπικά περιλαμβάνουν μια συνάρτηση πρόβλεψης. Τα ακόλουθα είναι παραδείγματα καταστάσεων όπου οι προβλέψεις είναι χρήσιμες.

Διαχείριση αποθήκης. Στον έλεγχο των αποθηκών αγορασμένων ειδών σε αεροπορικές εγκαταστάσεις είναι απαραίτητη μια εκτίμηση του βαθμού χρήσης για κάθε είδος προκειμένου να αποφασιστούν οι ποσότητες προμήθειας . Επιπροσθέτως, απαιτείται μια εκτίμηση του λάθος πρόβλεψης στις προμήθειες για να καθοριστούν τα επίπεδα επαναπαραγγελίας.

Σχεδιασμός παραγωγής. Για να σχεδιαστεί η παραγωγή μιας γραμμής προϊόντος, πιθανώς να είναι απαραίτητη μια πρόβλεψη των πωλήσεων ανά μονάδα προϊόντος για κάποιο αριθμό μηνών στο μέλλον. Αυτές οι προβλέψεις, δηλαδή για προϊόντα που έχουν διατεθεί μπορούν αργότερα να μετατραπούν σε απαιτήσεις για ημιτελή προϊόντα, υλικά κτλ. έτσι ώστε ολόκληρο το σύστημα παραγωγής να προγραμματιστεί.

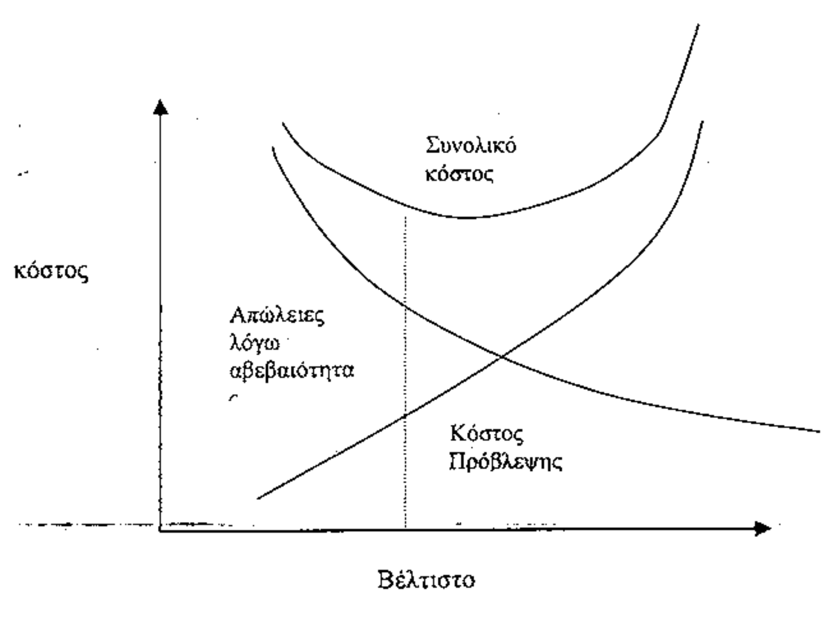
Χρηματοοικονομικός σχεδιασμός. Ένας οικονομικός μάνατζερ ενδιαφέρεται για τη ροή χρήματος που θα έχει μια επιχείρηση στο χρόνο. Είναι λογικό τότε, ο μάνατζερ να επιθυμεί μια πρόβλεψη των χρηματικών εισροών – εκροών για μελλοντικές χρονικές περιόδους σαν μια βοήθεια στη λήψη αποφάσεων

Προγραμματισμός προσωπικού. Ο μάνατζερ μιας ταχυδρομικής εταιρείας χρειάζεται μι πρόβλεψη της ποσότητας των επιστολών που διαχειρίζεται η εταιρεία για να προγραμματίσει τον αριθμό του προσωπικού και τον εξοπλισμό που θα χρησιμοποιήσει.

Σχεδιασμός εγκαταστάσεων : Γενικά, οι αποφάσεις για τη δημιουργία νέων εγκαταστάσεων απαιτεί μια μακροχρόνια πρόβλεψη των δραστηριοτήτων που χρησιμοποιούν τις εγκαταστάσεις. Αυτό είναι πολύ σημαντικό στο σχεδιασμό μιας εγκατάστασης όπως επίσης και για να δικαιολογήσει το κόστος που θα απαιτηθεί για αυτήν την επένδυση.

Έλεγχος επεξεργασίας. Η πρόβλεψη μπορεί επίσης να είναι ένα σημαντικό μέρος ενός συστήματος ελέγχου επεξεργασίας. Παρακολουθώντας διάφορες μεταβλητές και χρησιμοποιώντας αυτές για να προβλεφθεί η μελλοντική συμπεριφορά μιας διαδικασίας, μπορεί να καθοριστεί ο βέλτιστος χρόνος και ο βαθμός μιας πράξης ελέγχου. Για παράδειγμα μια μονάδα χημικής επεξεργασίας μπορεί να γίνει λιγότερο αποτελεσματική καθώς περνάνε οι ώρες συνεχούς δραστηριότητας. Προβλέποντας την απόδοση της μονάδας χρησιμεύει στο σχεδιασμό του χρόνου τερματισμού και επανεξέτασης.

Για αυτά τα παραδείγματα καταλαβαίνουμε πως σκοπός της πρόβλεψης είναι να μειωθεί το ρίσκο στη λήψη αποφάσεων. Οι προβλέψεις είναι συνήθως λάθος αλλά το μέγεθος των λαθών πρόβλεψης θα εξαρτηθεί από το σύστημα πρόβλεψης που θα χρησιμοποιηθεί . Δεσμεύοντας όλο και περισσότερους πόρους στην πρόβλεψη θα πρέπει να είμαστε σε θέση να βελτιώνουμε την ακρίβεια πρόβλεψης και συνεπώς να εξαλείφουμε κάποιες από τις απώλειες που θα είναι απόρροια της αβεβαιότητας στη διαδικασία λήψης αποφάσεων. Αυτή η ιδέα απεικονίζεται στο παρακάτω διάγραμμα όπου καθώς το κόστος της πρόβλεψης αυξάνεται οι απώλειες που συνδέονται με με το ρίσκο μειώνονται. Σε κάποιο σημείο της προσπάθειας πρόβλεψης το σύνολο σε αυτά τα κόστη είναι ελάχιστο.



Ταυτόχρονα, θα πρέπει ν σημειωθεί αυτό το υπόδειγμα βασίζεται στην υπόθεση του οριακού κόστους. Αυτό σημαίνει πως κάθε επιπλέον Ευρώ (€) που ξοδεύεται στην πρόβλεψη έχει ως αποτέλεσμα σε μια μικρότερη μείωση στην απώλεια ρίσκου απ' ότι το προηγούμενο ευρώ. Πιο απλά, όσο περισσότερα Ευρώ δαπανά στην πρόβλεψη ένας οργανισμός / εταιρεία τόσο λιγότερες απώλειες θα έχει. Πέρα από κάποιο σημείο όμως επιπρόσθετοι πόροι στην πρόβλεψη δε θα επιφέρουν καμία βελτίωση. Αυτό συμβαίνει διότι θα είναι αδύνατο να μειωθεί το μέσο σφάλμα πρόβλεψης κάτω από ένα δεδομένο επίπεδο ανεξαρτήτως της πολυπλοκότητα της μεθόδου που χρησιμοποιείται.

Επειδή όμως η πρόβλεψη δεν μπορεί ποτέ να εξαλείψει τελείως το ρίσκο, πρέπει η ομάδα λήψης απόφασης να σκεφτεί πολύ καλά την αβεβαιότητα που υπόκειται στην πρόβλεψη. Συχνά η απόφαση συνδέεται με την πρόβλεψη με την ακόλουθη σχέση

ΠΡΑΓΜΑΤΙΚΗ ΑΠΟΦΑΣΗ = ΑΠΟΦΑΣΗ ΠΟΥ ΛΑΜΒΑΝΕΤΑΙ ΜΕ ΤΗΝ ΠΡΟΫΠΟΘΕΣΗ ΟΤΙ Η ΠΡΟΒΛΕΨΗ ΕΙΝΑΙ ΣΩΣΤΗ

+ ΑΝΕΚΤΙΚΟΤΗΤΑ ΓΙΑ ΤΟ ΣΦΑΛΜΑ ΠΡΟΒΛΕΨΗΣ

Αυτό σημαίνει ότι το σύστημα πρόβλεψης πρέπει να παρέχει μια περιγραφή του σφάλματος πρόβλεψης όπως επίσης και της ίδιας της πρόβλεψης. Θα ήταν ιδανικό το σύστημα πρόβλεψης να οδηγούσε σε μια εκτίμηση της κατανομής της πιθανότητας της μεταβλητής που διερευνάται. Κατά συνέπεια το ρίσκο περιλαμβάνεται στη διαδικασία λήψης αποφάσεων.

Η πρόβλεψη δεν ο αυτοσκοπός. Είναι μάλλον ένα μέσο που μας βοηθά να φτάσουμε στο τέλος. Το σύστημα πρόβλεψης είναι ένα κομμάτι ενός ευρύτερου συστήματος, του συστήματος μάνατζμεντ και σαν υποσύνολο αλληλεπιδρά με άλλα συστήματα για να επιφέρουν το επιθυμητό αποτέλεσμα.

2. Ορίζοντας το πρόβλημα πρόβλεψης

Έστω ότι χρησιμοποιούμε την πρόβλεψη για το σχεδιασμό και προγραμματισμό της παραγωγής ή τον έλεγχο της αποθήκης, οπότε το ενδιαφέρον μας περιστρέφεται γύρω από τις μελλοντικές πωλήσεις των προϊόντων ή τη χρήση του α' υλών. Τυπικά, θα αναφερόμαστε στη μεταβλητή «ζήτηση».

Για να ορίσουμε το πρόβλημα της πρόβλεψης θα πρέπει πρώτα να προσδιορίσουμε το πρόβλημα λήψης μιας απόφασης. Η πληροφορία που προέρχεται από τη διαδικασία πρόβλεψης ουσιαστικά προορίζεται για να βελτιώσει τις συνθήκες κάτω από τις οποίες λαμβάνεται μια επιχειρηματική απόφαση.. Για αυτό το λόγο η φύση των αποφάσεων που θα παρθούν θα καθορίσει τα χαρακτηριστικά του συστήματος πρόβλεψης. Μια μελέτη του προβλήματος λήψης αποφάσεων θα μας βοηθήσει να απαντήσουμε ερωτήματα όπως η μεταβλητή που είναι αν προβλεφθεί, ο χρόνος που θα περιληφθεί και η ακρίβεια που θα απαιτηθεί.

Σημαντική παράμετρος του προβλήματος είναι το επίπεδο της λεπτομέρειας. Ένα σύστημα σχεδιασμού παραγωγής μπορεί να απαιτήσει μια πρόβλεψη της ζήτησης σε μονάδες για κάθε έτοιμο προϊόν προκειμένου να καθοριστεί η πολιτική αποθεμάτων. Από την άλλη πλευρά, για τη σύνταξη του ετήσιου προϋπολογισμού ο διευθυντής πωλήσεων ίσως χρειαστεί μια πρόβλεψη των συνολικών πωλήσεων.

Επιπλέον, οι αποφάσεις περιλαμβάνουν τρία σημαντικά στοιχεία για το χρόνο: την περίοδο πρόβλεψης, τον ορίζοντα πρόβλεψης και το διάστημα πρόβλεψης Η *περίοδος πρόβλεψης* είναι η βασική μονάδα χρόνου για τις οποίες διενεργούνται οι προβλέψεις. Για παράδειγμα μπορεί να θέλουμε να προβλέψουμε τη ζήτηση για ένα μια εβδομάδα οπότε σε αυτήν την περίπτωση η περίοδος είναι η εβδομάδα. Ο *ορίζοντας πρόβλεψης* είναι το πλήθος των περιόδων που καλύπτονται από την πρόβλεψη. Αν για παράδειγμα ο ορίζοντας θέλουμε πρόβλεψη για 10 εβδομάδες, τότε η περίοδος είναι η εβδομάδα και ο ορίζοντας το 10. Τελευταίο, το διάστημα πρόβλεψης είναι η συχνότητα με την ετοιμάζονται νέες προβλέψεις. Συχνά το διάστημα πρόβλεψης ταυτίζεται με την περίοδο πρόβλεψης, συνεπώς οι προβλέψεις επαναλαμβάνονται σε κάθε περίοδο χρησιμοποιώντας όλο και πιο πρόσφατα δεδομένα. Αν ο ορίζοντας έχει πάντα το ίδιο μήκος , έστω περίοδοι Τ, και η πρόβλεψη επαναλαμβάνεται κάθε περίοδο, τότε λέμε ότι λειτουργούμε σε επίπεδο «κινητού ορίζοντα».

Η περίοδος και ο ορίζοντας πρόβλεψης εξαρτώνται από τη διαδικασία λήψης απόφασης. Για να αξίζει μια πρόβλεψη θα πρέπει ο ορίζοντας πρόβλεψης να μην είναι μικρότερος από το χρόνο που απαιτείται για να παρθεί μια απόφαση. Το πόσο μακροχρόνια θα είναι η πρόβλεψη θα εξαρτηθεί σε σημαντικό βαθμό από τη φύση του προβλήματος.

Το διάστημα πρόβλεψης συχνά καθορίζεται από το πληροφοριακό σύστημα που επεξεργάζεται τη μεταβλητή που είναι να προβλεφθεί. Αν οι πωλήσεις καταγράφονται μηνιαία τότε και το διάστημα πρόβλεψης θα πρέπει να είναι σε μηναία βάση και όχι για παράδειγμα σε εβδομαδιαία.

Μια σημαντική πτυχή του προβλήματος πρόβλεψης σχετίζεται με την απαιτούμενη μορφή της πρόβλεψης. Μας εξυπηρετεί να θεωρήσουμε την μεταβλητή πρόβλεψης σαν μια τυχαία μεταβλητή που έχει μια άγνωστη κατανομή πιθανότητας και της οποίας τις παραμέτρους (μέσος, διάμεσος, τυπικό απόκλιση, διάστημα εμπιστοσύνης) επιθυμούμε να εκτιμήσουμε . Λιγότερο συχνό θα είναι το φαινόμενο να εκτιμήσουμε την ίδια την κατανομή (π.χ. κανονική, Poisson) . Συνήθως η πρόβλεψη θα περιλαμβάνει τα ακόλουθα : 1) μια εκτίμηση της αναμενόμενης τιμής της μεταβλητής συν την εκτίμηση του τυπικής απόκλισης του σφάλματος πρόβλεψης και 2) ένα διάστημα που θα έχει με δεδομένη πιθανότητα (έστω 95 %) την πραγματική μελλοντική τιμή. Δηλαδή, κατά 95 % οι πωλήσεις του μήνα Μάρτιος 2005 θα διακυμανθούν από 4.250 € έως 4.480 € . Αυτό το διάστημα καλείται *διάστημα πρόβλεψης*.

Επιπροσθέτως, σε ορισμένες περιπτώσεις μπορεί να μην μας ενδιαφέρει τόσο η πρόβλεψη της τιμής της μεταβλητής όσο οι αλλαγές που θα επέλθουν στη διαδικασία (αύξηση – μείωση). Ειδικά αυτή η περίπτωση συναντιέται στη διαδικασία ελέγχου όταν θέλουμε να προβλέψουμε τη χρονική στιγμή στην οποία μια διαδικασία θα τείνει να βρίσκεται εκτός ελέγχου.

Η ακρίβεια της πρόβλεψης θα επηρεάσει και το σύστημα πρόβλεψης που θα υιοθετηθεί. Ένα πολύ σημαντικό χαρακτηριστικό του συστήματος διοίκησης είναι η ικανότητά του να επιτυγχάνει τη βέλτιστη απόδοση μέσα σε ένα περιβάλλον αβεβαιότητας. Η βελτίωση της ακρίβειας θα μειώσει την αβεβαιότητα αλλά οι απαιτούμενοι πόροι μπορεί να μην συμφέρουν την επιχείρηση από οικονομικής πλευράς.

Παράλληλα, ένας σημαντικός παράγοντας που επιδρά στο πρόβλημα πρόβλεψης είναι η διαδικασία που παράγει τη μεταβλητή. Αν η διαδικασία είναι σταθερή, είτε με την έννοια ότι εξελίσσεται κάτω από σταθερές συνθήκες είτε με την έννοια ότι αλλάζει αργά στο χρόνο, τότε το σύστημα πρόβλεψης θα είναι πολύ διαφορετικό από ένα σύστημα που θα μελετά μια εξαιρετικά ασταθή με σημαντικές αλλαγές διαδικασία. Στην πρώτη περίπτωση το σύστημα θα έκανε εκτεταμένη χρήση ιστορικών δεδομένων για να προβλέψει το μέλλον ενώ στη δεύτερη το σύστημα θα επιχειρούσε να ανιχνεύσει τις αλλαγές.

Ταυτόχρονα, στο προσδιορισμό της πρόβλεψης είναι πολύτιμη η διαθεσιμότητα των δεδομένων. Η ποσότητα, η ακρίβεια και η καταλληλότητα αυτών των πληροφοριών είναι ζωτικής σημασίας για την πρόβλεψη. Ένα κλασικό παράδειγμα είναι το πρόβλημα της πρόβλεψης της ζήτησης για ένα βιομηχανικό προϊόν όταν η επιχείρηση κρατάει αρχεία ανά παραγγελίες και παραλαβές. Κανένα δεν απεικονίζει τη ζήτηση αφού οι παραγγελίες κλείνονται πριν από την επιθυμητή περίοδο διανομής και οι παραλαβές μπορεί να γίνουν σε περίοδο διαφορετική από αυτή που θέλει ο πελάτης. Η εταιρεία τότε αν θέλει πληροφορίες για το πότε εκτελούνται οι παραγγελίες, θα πρέπει να δημιουργήσει ένα ειδικό αρχείο που θα τις καταχωρεί.

Μια σοβαρή σύγχυση που συναντιέται συχνά στις προβλέψεις είναι η διάκριση του «τι μπορεί να πουληθεί» και του «τι θα πουληθεί» Η πρώτη έννοια έχει να κάνει με τις ευκαιρίες που είναι διαθέσιμες στην εταιρεία και η δεύτερη με τους περιορισμούς, αποφάσεις διοίκησης απεικονίζοντας ένα στόχο. Θα ήταν λοιπόν πιο εύστοχο να αποκαλούμε τότε την πρόβλεψη ως ένα προϋπολογισμό.

Οι υπολογιστικοί περιορισμοί που τίθενται στο σύστημα πρόβλεψης έχουν να κάνουν με το πλήθος των μεταβλητών που διερευνούμε . Όσο μεγαλύτερο είναι το πλήθος των μεταβλητών τόσο αυξάνει η πολυπλοκότητα του προβλήματος.

Δύο πολύ σημαντικοί παράγοντες που διαδραματίζουν ρόλο στον προσδιορισμό του προβλήματος πρόβλεψης είναι οι ικανότητες και τα ενδιαφέροντα των ανθρώπων που θα διενεργήσουν την πρόβλεψη. Η διαδικασία πρόβλεψης προϋποθέτει τη συμμετοχή των μάνατζερ. Αν όμως οι μάνατζερ είναι πεπεισμένοι ότι η πρόβλεψη δε θα τους βοηθήσει τότε και εκείνοι δεν θα τη χρησιμοποιήσουν.

3. Γιατί αμφισβητούμε την πρόβλεψη

Παρά το γεγονός ότι η πρόβλεψη επηρεάζει σημαντικά τη διαδικασία λήψης αποφάσεων υπάρχουν μερικοί που αμφισβητούν την αξιοπιστία ενός συστήματος πρόβλεψης.. Οι λόγοι που δικαιολογούν αυτό την αμφισβήτηση έχουν να κάνουν με το γεγονός ότι μερικές εταιρείες, σύμφωνα με τον Jeffrey Jarret (Μέθοδοι Προβλέψεων για οικονομικές - επιχειρηματικές αποφάσεις – 1987), εξακολουθούν να πιστεύουν πως το μέλλον δεν επιφυλάσσει σημαντικές αλλαγές και δε θα υπάρξει αρκετός χρόνος , στο μέλλον, που θα επιτρέπει σε μια εταιρεία να αντιδράσει στις αλλαγές των γεγονότων.

Παράλληλα, σύμφωνα πάντα με τον ίδιο συγγραφέα κάποιες εταιρείες δε χρησιμοποιούν την γιατί απλά αγνοούν την ανάγκη της ή επειδή η εταιρεία φάνηκε τυχερή με το να μην παρουσιάσει υπερβολικά αποθέματα ή μια καταστροφική πτώση στις πωλήσεις της.

Εντούτοις , όλοι οι οργανισμοί λειτουργούν μέσα σε μια ατμόσφαιρα αβεβαιότητας σε ότι αφορά τα μελλοντικά γεγονότα και είναι υποχρεωμένοι να παίρνουν αποφάσεις με δεδομένη αυτή την αβεβαιότητα.

4. Ταξινόμηση των προβλέψεων

Το κριτήριο για την ταξινόμηση των προβλέψεων είναι, υπογραμμίζει ο Jeffrey Jarret στο βιβλίο του Μέθοδοι Προβλέψεων για οικονομικές - επιχειρηματικές αποφάσεις (1987), αυτό που ονομάζεται χρονικός ορίζοντας. Με τον όρο χρονικός ορίζοντας (lead time) εννοούμε την απόσταση σε χρόνο ανάμεσα στο σημείο στο οποίο γίνεται η πρόβλεψη και το χρονικό σημείο στο οποίο αυτή αναφέρεται. Αν για παράδειγμα μια εταιρεία παραγωγής ενδυμάτων προετοιμάζει μια πρόβλεψη τον Ιούνιο του 2004 για τις πρώτες ύλες που θα χρειαστεί τον Σεπτέμβριο του 2004 τότε ο χρονικός ορίζοντας της πρόβλεψης λέγεται ότι είναι 2 μήνες .

Οπότε με βάση αυτό το κριτήριο οι προβλέψεις ταξινομούνται σε βραχυπρόθεσμες , μεσοπρόθεσμες και μακροπρόθεσμες . Για την ταξινόμηση πρέπει να λαμβάνονται υπόψη ο ρυθμός μεταβολής των χρονοσειρών και το είδος της εταιρείας. Για παράδειγμα, διαφορετικό χρονικό ορίζοντα έχει η ζήτηση για νωπά λαχανικά και διαφορετικό μια πρόβλεψη ζήτησης για μια μονάδα παραγωγής ηλεκτρικής ενέργειας.

Ταυτόχρονα όσο αφορά το ρυθμό μεταβολής της χρονικής σειράς μια εταιρεία παραγωγής παγωτών θα έχει μεγαλύτερη ζήτηση στα προϊόντα της στους θερινούς μήνες.

3. Γραμμικά – μη γραμμικά μοντέλα

1. Ορισμοί

Η μορφή του μοντέλου πρόβλεψης μπορεί να είναι οποιοδήποτε , εδώ όμως θα ασχοληθούμε μόνο με το γραμμικό μοντέλο . Το γενικό γραμμικό μοντέλο είναι το

$$Y = \beta_o + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k + \varepsilon \qquad (1)$$

Όπου : Y είναι η εξαρτημένη (dependent) μεταβλητή ή απόκριση (response) .

$X_1, X_2, \ldots, X_k$ είναι οι k «ανεξάρτητες» (independent) ή «προβλέπουσες» (predictor) μεταβλητές

$\beta_o, \beta_1, \ldots, \beta_k$ είναι (k + 1) άγνωστες παράμετροι (συντελεστές παλινδρόμησης) που ζητείται να εκτιμηθούν.

Για τις μεταβλητές $X_1, X_2, \ldots, X_k$ η ονομασία «ανεξάρτητες» δεν σημαίνει ότι είναι πράγματι ανεξάρτητες. Μπορεί για παράδειγμα να ισχύει $X_2 = X_1^2$ ή $X_3 = X_1 + X_2$. Ο λόγος στον οποίο οφείλεται αυτή η ονομασία είναι, ότι το ζητούμενο συνήθως είναι , πως οι τιμές αυτών των μεταβλητών επηρεάζουν τις τιμές της εξαρτημένης μεταβλητής και μπορούν να ελέγχονται από τον ερευνητή.

Το σφάλμα ε , περιέχει κάθε απόκλιση της πραγματικής κατάστασης από το μοντέλο. Έτσι εκτός από τα πιθανά σφάλματα μετρήσεων, περιέχει επίσης και σφάλματα προσαρμογής, που οφείλονται είτε σε παραλείψεις μεταβλητών είτε σε χρήση μεταβλητών που δε σχετίζονται με την Y.

Η δυνατότητα των προβλεπουσών μεταβλητών να συσχετίζονται μεταξύ τους διευρύνει τις περιπτώσεις εφαρμογής του μοντέλου (1) . Πράγματι με κατάλληλους μετασχηματισμούς «μη – γραμμικά» ή και «εκθετικά» μοντέλα ανάγονται στο γενικό γραμμικό μοντέλο όπως φαίνεται παρακάτω. Πράγματι :

Θέτοντας $X_j = x^j$, $j = 1, \ldots k$ το πολυωνυμικό μοντέλο

$$Y = \beta_o + \beta_1 x + \beta_2 x^2 + \ldots + \beta_k x^k + \varepsilon$$

ανάγεται αμέσως στο γραμμικό μοντέλο (1)

Όμοια θέτοντας $X_1 = x$, $X_2 = z$, $X_3 = x^2$, $X_4 = xz$, το μοντέλο

$Y = \beta_o + \beta_1 x + \beta_2 z + \beta_3 x^2 + \beta_4 x z + \varepsilon$

ανάγεται πάλι στο (1).

Ακόμη και το εκθετικό μοντέλο $y = 2^{\alpha + \beta x + \gamma x\text{^}2 + \varepsilon}$ ανάγεται στο (1), αρκεί να θέσουμε $y = \log_2 y$, $X_1 = x$, $X_2 = x^2$, $\beta_o = \alpha$, $\beta_1 = \beta$, $\beta_2 = \gamma$.

Το ίδιο και για το μοντέλο $y = \alpha\ \varepsilon\ x_1^{\ \beta}\ x_2^{\ \gamma}$ όπου αρκεί να θέσουμε $Y = \ln y$, $X_1 = \ln x_1$, $X_2 = \ln x_2$, $\beta_o = \ln\alpha$, $\beta_1 = \beta$, $\beta_2 = \gamma$, $e = \ln \varepsilon$

Υπάρχουν βέβαια και μοντέλα όπως π.χ. τα $y = \beta_o x + \dfrac{1}{\beta_1 + \beta_2 x}$ ή $y = A e^{Bx} + \Gamma e^{\Delta x}$ που δεν είναι γραμμικά και ούτε μπορούν να αναχθούν σε γραμμικά με κάποιους μετασχηματισμούς.

Για να μπορέσουμε να χρησιμοποιήσουμε μεθόδους στατιστικής ανάλυσης για τον υπολογισμό των παραμέτρων β_i $i = 0, 1, \ldots, k$ υποθέτουμε ότι τα σφάλματα ε_i είναι τυχαίες μεταβλητές, ασυσχέτιστες με κοινή διασπορά σ^2. Όταν θέλουμε να κάνουμε ελέγχους υποθέσεων για τα β_i, τότε υποθέτουμε ότι τα σφάλματα είναι ανεξάρτητες τυχαίες μεταβλητές με κοινή κανονική κατανομή $N(0, \sigma^2)$.

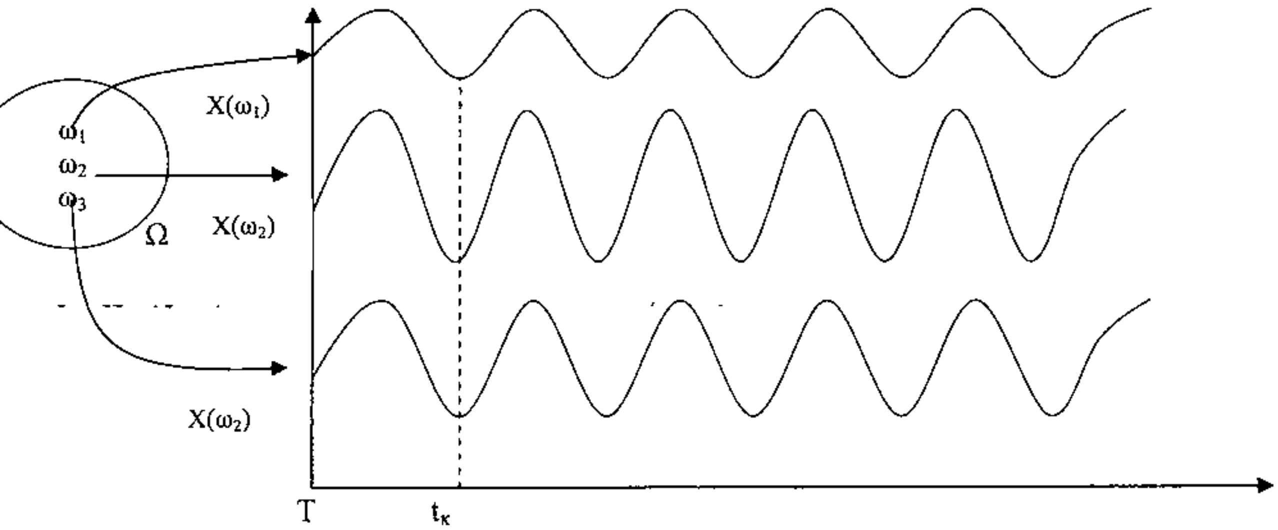
4. Στοχαστική διαδικασία

1. Εισαγωγή

Έστω ότι έχουμε μια τυχαία μεταβλητή. Μια τυχαία μεταβλητή είναι μια συνάρτηση που παίρνει γεγονότα και τα απεικονίζει στην ευθεία των πραγματικών αριθμών όπου γεγονός είναι ένα υποσύνολο του δειγματικού χώρου και συμβολίζεται με ω. Δειγματικός χώρος είναι το σύνολο όλων των δυνατών αποτελεσμάτων ενός πειράματος τύχης και συμβολίζεται με Ω..

Για παράδειγμα έχουμε ένα σάκο με μπάλες τριών χρωμάτων, κόκκινου, κίτρινου και πράσινου. Αν τραβήξουμε μια μπάλα κόκκινου χρώματος κερδίζουμε 10 €, κίτρινου 15 €, και πράσινου 25 €. Σε αυτήν την περίπτωση γεγονός είναι το χρώμα . Δηλαδή ω_1 = κόκκινο , ω_2 = κίτρινο , ω_3 = πράσινο , και τιμή της τυχαίας μεταβλητής είναι αντιστοίχως $X(\omega_1)$ = 10 €, $X(\omega_2)$ = 15 €, $X(\omega_3)$ = 25 €. Άρα για ένα συγκεκριμένο $\omega \in \Omega$ η ποσότητα $X(\omega)$ είναι ένας αριθμός.

Ο αριθμός αυτός θα μπορούσε να εκφράζει την κατάσταση ενός «συστήματος» , π.χ. τη θέση ενός σωματιδίου κατά τη χρονική στιγμή t , $t \in T \in R_+$. Για αυτό το λόγο είναι προτιμότερο να γράφουμε $X(t, \omega)$. Το ω εδώ εκφράζει ένα συγκεκριμένο σωματίδιο. Καθώς αυξάνει ο χρόνος t , το σωματίδιο κινείται και η τροχιά που διαγράφει εκφράζεται από τη συνάρτηση $X(t, \omega)$, $t \in T$. Όταν η επιλογή του σωματιδίου είναι τυχαία ή /και η συμπεριφορά του τυχαία, η συνάρτηση $X(t, \omega)$, $t \in T$ είναι μια τυχαία συνάρτηση η οποία ονομάζεται στοχαστική ανέλιξη ή στοχαστική διαδικασία.



Από τα παραπάνω προκύπτει ότι η X(t, ω) μπορεί να έχει τις ακόλουθες ερμηνείες :

1. Είναι μια μια οικογένεια συναρτήσεων X (t, ω) με $t \in T$ και $\omega \in \Omega$. Εδώ t και ω είναι τυχαίες μεταβλητές.
2. Για σταθερό ω, η X(t, ω) είναι μια συνάρτηση του χρόνου t που εκφράζει μια συγκεκριμένη τροχιά ή διαφορετικά μια «πραγματοποίηση» (realization) της στοχαστικής διαδικασίας.
3. Για σταθερό t, η η X(t, ω) είναι μια τυχαία μεταβλητή που εκφράζει την κατάσταση της στοχαστικής διαδικασίας τη χρονική στιγμή t .
4. Για σταθερό t και ω η X (t, ω) είναι ένας αριθμός.

Όταν το σύνολο Τα συμπίπτει με το R_+ ή κάποιο συνεχές υποσύνολό του , τότε η διαδικασία X(t) ονομάζεται διαδικασία «συνεχούς χρόνου» και όταν το Τ είναι το σύνολο των ακεραίων ή κάποιο υποσύνολο του R_+ με αριθμήσιμο πλήθος στοιχείων, τότε η διαδικασία ονομάζεται διαδικασία «διακριτού χρόνου». Παρόμοια αν το σύνολο των τιμών της X(t) συμπίπτει με το R ή με κάποιο συνεχές υποσύνολό του, η X(t) ονομάζεται διαδικασία «συνεχών καταστάσεων» διαφορετικά ονομάζεται διαδικασία «διακριτών καταστάσεων».

Οι στοχαστικές διαδικασίες διακρίνονται σε «στάσιμες» και «μη στάσιμες». Για την πρώτη κατηγορία έχουμε «στάσιμες με την ευρεία έννοια» και τις «αυστηρά στάσιμες» στοχαστικές διαδικασίες.

Στασιμότητα υπό την ευρεία έννοια

Μια στοχαστική διαδικασία ονομάζεται στάσιμη με την ευρεία έννοια (weak sense stationarity) ή αλλιώς στάσιμη β' τάξης όταν

A. η μέση τιμή της είναι σταθερή (στάσιμη α' τάξης) και

B. η συνάρτηση αυτοσυσχέτισης εξαρτάται μόνο από τη διαφορά $\tau = t_1 - t_2$

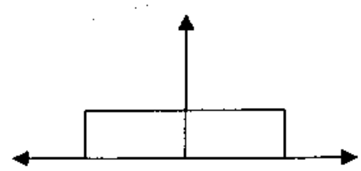
δηλαδή $\mu (t) = E (x(t)) = \mu$

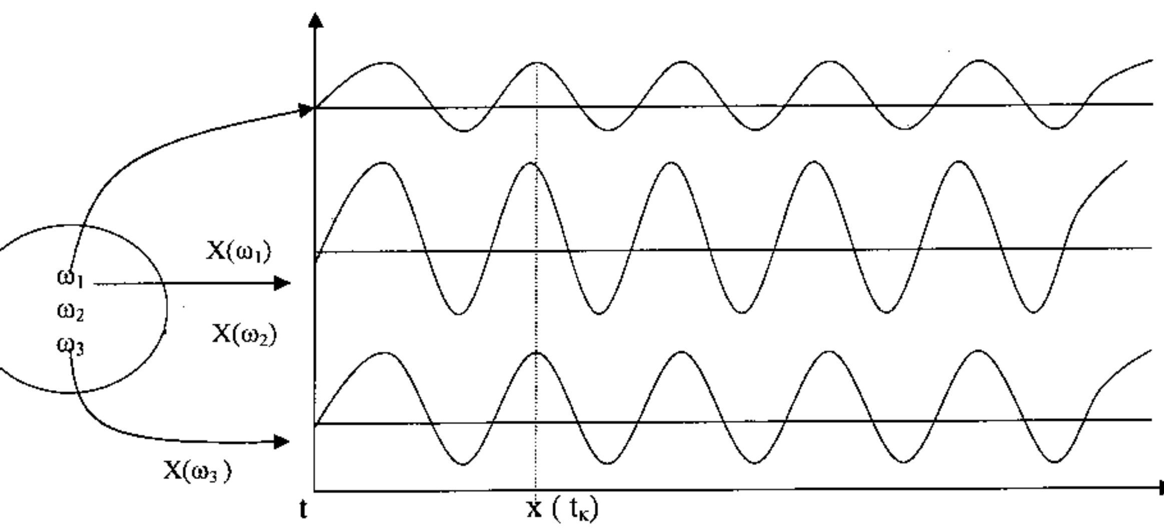
$R(t_1 , t_2) = r (\tau)$ με $\tau = t_1 - t_2$

2. Παραδείγματα

Α. Παράδειγμα στασιμότητας 1ης τάξης

Έστω ότι έχουμε τη διαδικασία $x(t) = A\sin(\Omega_0 t)$ όπου $A \sim U[-1, 1]$





Τα αποτελέσματα του πειράματος τύχης είναι τα ω_i . Ανάλογα με τη τιμή της τ.μ. (τυχαίας μεταβλητής) αυξομειώνεται το πλάτος Α. Δηλαδή το πλάτος επηρεάζεται με τυχαίο τρόπο. Πριν εκτελέσουμε το πείραμα δε γνωρίζουμε το πλάτος.

Αν παγώσουμε το χρόνο δημιουργείται μια νέα τυχαία μεταβλητή η $x(t_\kappa)$. Για κάθε χρονική στιγμή η στοχαστική διαδικασία επιστρέφει μια τυχαία μεταβλητή. Για $t = 0$ η στοχαστική διαδικασία παίρνει με πιθανότητα 100% την τιμή 0 καθώς για οποιοδήποτε αποτέλεσμα του πειράματος τύχης το $X(\omega_i) = 0$.

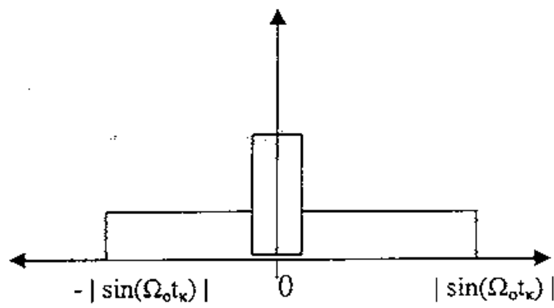
Το ερώτημα που δημιουργείται τώρα είναι ποια είναι η συνάρτηση πυκνότητας πιθανότητας για $t = 0$ και ποια για $t = \kappa$. Να σημειώσουμε πως συνάρτηση πυκνότητα πιθανότητας $f(x)$ (σ.π.π.) είναι η παράγωγος της συνάρτησης κατανομής $F(x)$ όπου $F(x)$ είναι η πιθανότητα η τ.μ. να πάρει μια τιμή μικρότερη ή ίση από έναν αριθμό x.

Δηλαδή $F(x) = P\{X \leq x\} = \int_{-\infty}^{x} f(x)dx$

Δύο είναι οι ιδιότητες της f (x)

1. $f(x) \geq 0$
2. $\int_{-\infty}^{\infty} f(x)dx = 1$

Ποια είναι η πυκνότητα πιθανότητας για t = κ ;



Όσο μειώνεται ο χρόνος τόσο αυξάνεται το ύψος του παραλληλογράμμου

Αξιοποιούμε τη 2η ιδιότητα της σ.π.π.

$$\int_{-\infty}^{\infty} f(x)dx = 1$$

$$\int_{-|\sin(\Omega_0 t_\kappa|}^{|\sin(\Omega_0 t_\kappa|} \lambda dx = 1$$

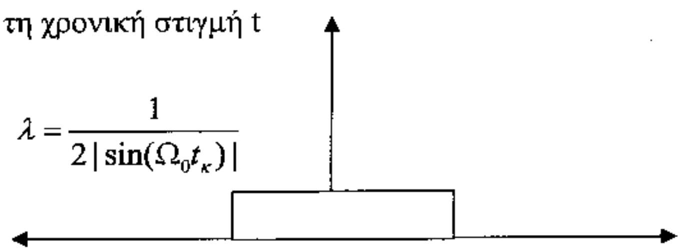
$$\lambda \int_{-|\sin(\Omega_0 t_\kappa|}^{|\sin(\Omega_0 t_\kappa|} dx = 1$$

$$\lambda \left[x \right]_{|\sin(\Omega_0 t_\kappa|}^{-|\sin(\Omega_0 t_\kappa|} = 1$$

$$2\lambda|\sin(\Omega_0 t_\kappa \lambda)|=1$$

$$\lambda=\frac{1}{2|\sin(\Omega_0 t_\kappa \lambda)|}$$

Η τιμή που παίρνει η σ.π.π τη χρονική στιγμή t



Κάθε μία από τις σ.π.π. αντιστοιχεί σε μια τ.μ.

Στοχαστική διαδικασία και επομένως χρονοσειρά σημαίνει πώς όλες οι ροπές είναι συναρτήσεις του χρόνου.

Πάμε να υπολογίσουμε τη μέση τιμή της διαδικασίας x (t) .

$$\mathrm{E}\{x_\kappa\} = \int_{-\infty}^{\infty} x_\kappa f(x)dx$$

$$f(x) = \begin{cases} \lambda = \dfrac{1}{2|\sin(\Omega_0 t_\kappa)|}, & \text{Για κάθε } x \leq |\sin(\Omega_0 t_\kappa)| \\ 0, & \text{αλλού} \end{cases}$$

$$\mathrm{E}\{x_\kappa\} = \int_{-\infty}^{\infty} x_\kappa \frac{1}{2|\sin(\Omega_0 t_\kappa)|}dx$$

$$= \int_{-|\sin(\Omega_0 t_\kappa|}^{|\sin(\Omega_0 t_\kappa|} x_\kappa \frac{1}{2|\sin(\Omega_0 t_\kappa)|} dx$$

$$= \lambda \int_{-|\sin(\Omega_0 t_\kappa|}^{|\sin(\Omega_0 t_\kappa|} x_\kappa dx$$

$$= \lambda \left[\frac{x^2}{2} \right]_{|\sin(\Omega_0 t_\kappa|}^{-|\sin(\Omega_0 t_\kappa|}$$

$$= \frac{\lambda}{2} (\sin^2(\Omega_0 t_\kappa) - \sin^2(\Omega_0 t_\kappa)) = 0$$

Άρα $E\{x_\kappa\} = 0$ δηλαδή για κάθε χρονική στιγμή η μέση τιμή είναι 0.

Ποια είναι όμως η $E\{x_\kappa\}$ για $\kappa = 0$;

$$E\{x_\kappa\} = \int_{-\infty}^{\infty} x_\kappa f(x) dx$$

$$E\{x_\kappa\} = \int_{-\infty}^{\infty} x_\kappa \delta(x_\kappa) dx = 0$$

Η μέση τιμή είναι ανεξάρτητη του χρόνου, άρα η στοχαστική διαδικασία είναι στάσιμη $1^{ης}$ τάξης.

$\mu(t) = C$ (Η μέση τιμή ισούται με μια σταθερά)

Β. Παράδειγμα στασιμότητας 2^{ης} τάξης

Έστω ότι έχουμε τη διαδικασία x(t) = Asin(Ω_ot + φ) όπου A ~ U [-1, 1] και φ ~ U [0, 2π] με A, φ ανεξάρτητες τυχαίες μεταβλητές.

Παρατηρούμε ότι σε σχέση με το προηγούμενο παράδειγμα μετατοπίζεται το ημίτονο κατά φ.

Υπολογίζουμε τη μέση τιμή της x(t)

$$\mathrm{E}\{xt_\kappa\} = \int_{-\infty}^{\infty} x_{t\kappa} f(a)da \int_{-\infty}^{\infty} f(\mathrm{B})d\mathrm{B}$$

Ολοκλήρωση ως προς α

$$= \int_0^{2\pi} [[\int_{-1}^{1} \alpha \sin(\Omega_o t + B)d\alpha]dB$$

Ολοκλήρωση ως προς α ανεξάρτητη του φ

$$= \int_0^{2\pi} \sin(\Omega_o t + B)(\int_{-1}^{1} ada)]d \quad {}^{1}_{-1}$$

$$= \int_0^{2\pi} \sin(\Omega_o t + B)dB \left[\frac{a^2}{2} \right]$$

= 0

Η μέση τιμή είναι ανεξάρτητη του χρόνου, άρα η στοχαστική διαδικασία είναι στάσιμη 1^{ης} τάξης.

μ (x(t)) = 0

Για να έχουμε στασιμότητα β ' τάξης θα πρέπει η διαδικασία να είναι στάσιμη α ' τάξης, κάτι που αποδείξαμε προηγουμένως , και επιπλέον να ισχύει

$\mathbf{E}\{ x_t \; x_{t+\kappa} \} = \gamma_\kappa$

E { Asin(Ω_ot + φ) Asin(Ωo (t + κ) + φ) }=

E { A^2sin(Ω_ot + φ) sin(Ωo (t + κ) + φ) }=

Έστω α = $\Omega_o t$ + φ και β = Ωo (t + κ) + φ

Αξιοποιούμε την ιδιότητα των συνημίτονων

2sin(α)sin(b)=cos(α – β) – cos(α + β)

$$\frac{1}{2}\ E\ \{\ A^2\ [\ \cos(\ \Omega_o t+\varphi\ -\ \Omega_o\ (t+\kappa\ )-\varphi\ )\ -\cos(\Omega o\ (t+\varphi\ )+\Omega_o\ (t+\kappa\ )+\varphi)\ ]\ \}=$$

$$\frac{1}{2}\ E\ \{\ A^2\ [\ \cos(\ -\Omega_o\kappa)-\cos(\ 2\Omega ot+2\varphi+\Omega_o\kappa\ )]\ \}=$$

$$\frac{1}{2}\ E\ \{\ A^2\ [\ \cos(\ \Omega_o\kappa)-\cos(\ \Omega o(2t+\kappa\ )+2\varphi\ )]\ \}=$$

Αξιοποιούμε την ιδιότητα των ημιτόνων

$\cos(\alpha+\beta)=\cos(\alpha)\cos(\beta)\ -\ \sin(\alpha)\sin(b)$

με $\alpha = 2t+\kappa$ και $\beta = 2\varphi$

$$\frac{1}{2}\ E\ \{\ A^2\ [\ \cos(\ \Omega_o\kappa)-[\ \cos(\ \Omega_o(2t+\kappa\ )\cos\ (2\varphi)-\sin(\Omega_o\ (\ 2t+\kappa\ )\ \sin(2\varphi)\ )]\ ]\ \}=$$

$$\frac{1}{2}\ E\ \{\ A^2\ [\ \cos(\ \Omega_o\kappa)\}\ -\ \frac{1}{2}\ E\{\ A^2\ [\ \cos(\ \Omega_o(2t+\kappa\ )\cos\ (2\varphi)\ -\ \sin(\Omega_o\ (\ 2t+\kappa\ )$$
$$\sin(2\varphi)\ )]\ ]\ \}=$$

$$\frac{1}{2}\ \cos(\ \Omega_o\kappa)\}\ E\{\ A^2\ \}-\frac{1}{2}\ E\{\ A^2\ [\ \cos(\ \Omega_o(2t+\kappa\ )\cos\ (2\varphi)\ \}\ +$$

$$\frac{1}{2}\ E\{\ A^2\ \sin(\Omega_o\ (t+\kappa)\ \sin(2\varphi)\ )\ \}=$$

$\cos(\ \Omega_o(2t+\kappa\ )$ & $\sin(\Omega_o\ (t+\kappa)$ σταθερές ποσότητες

$$\frac{1}{2}\ \cos(\ \Omega_o\kappa)\}\ \int_{-1}^{1} Af(A)dA\ -\frac{1}{2}\ [\ \cos(\ \Omega_o(2t+\kappa\ )\ E\{\ A^2\cos\ (2\varphi)\ \}\ +$$

$$\frac{1}{2}\ \sin(\Omega_o\ (t+\kappa)\ E\{\ A^2\ \sin(2\varphi)\}\ =$$

$$\frac{1}{2}\ \cos(\ \Omega_o\kappa)\}\left[\frac{A^3}{3}\right]_{-1}^{1}\ -\frac{1}{2}\ [\ \cos(\ \Omega_o(2t+\kappa\ )\ E\{\ A^2\cos\ (2\varphi)\ \}\ +$$

$$\frac{1}{2}\ \sin(\Omega_o\ (t+\kappa)\ E\{\ A^2\ \sin(2\varphi)\}\ =$$

$$\frac{1}{3}\ \cos(\ \Omega_o\kappa)$$

άρα η συνδιασπορά εξαρτάται μόνο από το κ και όχι από το χρόνο.

3. Συντελεστής συσχέτισης

$$\rho = \frac{\mathrm{E}(\mathrm{x}-\mu_X)\,(\mathrm{y}-\mu_y)}{\sqrt{\mathrm{E}(\mathrm{x}-\mu_X)^2}\,\sqrt{\mathrm{E}(\mathrm{y}-\mu_y)^2}}$$

Ο αριθμητής δηλώνει τη συνδιασπορά (covariance) και ο παρανομαστής το γινόμενο της τυπικών αποκλίσεων των τ.μ. x και y

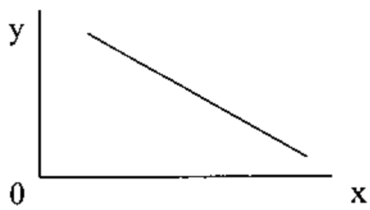
Όταν η συνδιακύμανση είναι θετικός αριθμός, οι μεταβλητές X και Y είναι θετικά συσχετισμένες και λέμε ότι *μεταβάλλονται ομόρροπα.*

Όταν όμως η συνδιακύμανση είναι αρνητικός αριθμός, οι μεταβλητές X και Y είναι αρνητικά συσχετισμένες και λέμε ότι *μεταβάλλονται αντίρροπα.*

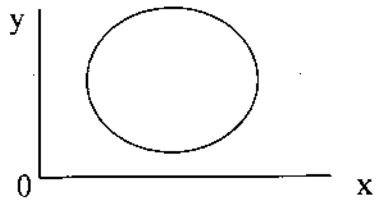
Ο συντελεστής συσχέτισης ρ μας δείχνει το βαθμό εξάρτησης ανάμεσα σε δύο τυχαίες μεταβλητές y και x.

i) Αν ρ = 1 τότε έχουμε πλήρη γραμμική θετική συσχέτιση

ii) Αν ρ = -1 τότε έχουμε πλήρη γραμμική αρνητική συσχέτιση



iii) Αν ρ = 0 τότε οι x και y καλούνται ασυσχέτιστες



Θα αποδείξουμε πως ο συντελεστής συσχέτισης 2 τυχαίων μεταβλητών x και y παίρνει τιμές από – 1 έως και 1 .

$$-1 \le \rho_{x,y} \le 1$$

Απόδειξη :

Θέτουμε μια τυχαία μεταβλητή $z = \alpha(x - \mu_x) + (y - \mu_y)$

Η μέση τιμή της z είναι

$E(z) = E\{\alpha(x - \mu_x) + (y - \mu_y)\}$ όπου E το expectation (αναμενόμενη μέση τιμή)

$= \alpha E\{(x - \mu_x)\} + E\{(y - \mu_y)\}$

$= \alpha E\{x\} - \alpha E\{\mu_x\} + E\{y\} - E\{\mu_y\}$

$= \alpha \mu_x - \alpha \mu_x + \mu_y - \mu_y$

$= 0$

αφού η μέση αναμενόμενη τιμή του μ_x και μ_y είναι η μέση τιμή ενός σταθερού όρου άρα μ_x και μ_y

Η διασπορά της z είναι

$\sigma^2{}_z = E\{(z - \mu)^2\} = E\{(z - 0)^2\} = E\{(z)^2\}$

$= E(\alpha(x - \mu_x) + (y - \mu_y))2$

$= \alpha^2 E((x - \mu_x)^2 + 2\alpha E((x - \mu_x)(y - \mu_y)) + E((y - \mu_y))^2$

$= \alpha^2 \sigma_x{}^2 + 2\alpha E((x - \mu_x)(y - \mu_y)) + \sigma_y{}^2$

$\{ y = \alpha x^2 + \beta x + \gamma$ παραβολή $\}$

Αφού η διασπορά είναι πάντα ένας θετικός αριθμός άρα $\Delta < 0$

Όποτε $\beta^2 - 4\alpha\gamma < 0$

Δηλαδή

$4E((x - \mu_x)(y - \mu_y))^2 - 4\sigma_x{}^2\sigma_y{}^2 < 0$

$4E((x - \mu_x)(y - \mu_y))^2 < 4\sigma_x{}^2\sigma_y{}^2$

$$\frac{E((x-\mu_X)(y-\mu_y))^2}{\sigma_X^2\sigma_y^2} < 1$$

$$\left(\frac{E((x-\mu_X)(y-\mu_y))}{\sigma_X\sigma_y}\right)^2 < 1$$

$\rho^2{}_{x,y} < 1$

$|\rho_{x,y}| < 1$

$-1 \le \rho_{x,y} \le 1$

Μήτρα των συντελεστών συσχέτισης

Έχουμε τρεις τυχαίες μεταβλητές :

X_1

X_2

X_3

Η μέση τιμή τους είναι $\mu_1 = E\{X_1\}$, $\mu_2 = E\{X_2\}$ και $\mu_3 = E\{X_3\}$ αντιστοίχως

$X_1 - \mu_1 = E\{X_1\}$,

Εστω Α ο παρακάτω πίνακας $X - \mu$

$$X - \mu = \begin{bmatrix} X_1 - \mu_1 \\ X_2 - \mu_2 \\ X_3 - \mu_3 \end{bmatrix}$$

Πολλαπλασιάζουμε με τον ανάστροφό του Α' και έχουμε

$$(X - \mu) * (X - \mu)^T = \begin{bmatrix} X_1 - \mu_1 \\ X_2 - \mu_2 \\ X_3 - \mu_3 \end{bmatrix} * [X_1 - \mu_1 \quad X_2 - \mu_2 \quad X_3 - \mu_3]$$

Πολλαπλασιάζουμε ένα πίνακα 3 * 1 με ένα πίνακα 1 * 3 . Οπότε το αποτέλεσμά του θα είναι ένας πίνακας 3 * 3

$$(X - \mu) * (X - \mu)^T = \begin{bmatrix} (X_1-\mu_1)^2 & (X_1-\mu_1)(X_2-\mu_2) & (X_1-\mu_1)(X_3-\mu_3) \\ (X_2-\mu_2)(X_1-\mu_1) & (X_2-\mu_2)^2 & (X_2-\mu_2)(X_3-\mu_3) \\ (X_3-\mu_3)(X_1-\mu_1) & (X_3-\mu_3)(X_2-\mu_2) & (X_3-\mu_3)^3 \end{bmatrix}$$

αν στον παραπάνω πίνακα πάρουμε αναμενόμενες τιμές τότε

$(X - \mu) * (X - \mu)^T = E\{[\ldots]\}$ οπότε θα γίνει ο πίνακας των συντελεστών συνδιασποράς

Ας ορίσουμε $C = (X - \mu) * (X - \mu)^T$

$$C = \begin{bmatrix} C_0 & C_1 & C_2 \\ C_1 & C_0 & C_1 \\ C_2 & C_1 & C_0 \end{bmatrix} \text{ με } C_0 = \sigma^2_x$$

Παρατηρούμε ότι η μήτρα των συντελεστών συνδιασποράς είναι συμμετρική .

Αν διαιρέσουμε κάθε στοιχείο του C με το γινόμενο των τυπικών αποκλίσεων των μεταβλητών δηλαδή με το

$$\frac{1}{\sigma_{Xi}\sigma_{Xi}}$$

τότε θα μας προκύψει η ακόλουθη μήτρα

$$\begin{bmatrix} \frac{E\{(X_1-\mu_1)^2\}}{\sigma_{X1}\sigma_{X1}} \_ \frac{E\{(X_1-\mu_1)(X_2-\mu_2)\}}{\sigma_{X1}\sigma_{X2}} \_ \frac{E\{(X_1-\mu_1)(X_3-\mu_3)\}}{\sigma_{X1}\sigma_{X3}} \\ \frac{E\{(X_1-\mu_1)(X_2-\mu_2)\}}{\sigma_{X1}\sigma_{X2}} \_ \frac{E\{(X_2-\mu_2)^2\}}{\sigma_{X2}\sigma_{X2}} \_ \frac{E\{(X_2-\mu_2)(X_3-\mu_3)\}}{\sigma_{X2}\sigma_{X3}} \\ \frac{E\{(X_1-\mu_1)(X_3-\mu_3)\}}{\sigma_{X1}\sigma_{X3}} \_ \frac{E\{(X_2-\mu_2)(X_3-\mu_3)\}}{\sigma_{X2}\sigma_{X3}} \_ \frac{E\{(X_3-\mu_3)^2\}}{\sigma_{X3}\sigma_{X3}} \end{bmatrix}$$

η οποία είναι η μήτρα των συντελεστών συσχέτισης

Μια προσεκτικότερη μελέτη της παραπάνω μήτρας θα μας οδηγήσει στο συμπέρασμα ότι τα στοιχεία της κύριας διαγωνίου είναι οι διασπορές των τυχαίων μεταβλητών X_i και πως ο πίνακας είναι συμμετρικός .

Οπότε

$$\rho_{Xi} = \begin{bmatrix} \rho_0 \_ \rho_1 \_ \rho_2 \\ \rho_1 \_ \rho_0 \_ \rho_1 \\ \rho_2 \_ \rho_1 \_ \rho_0 \end{bmatrix}$$ αλλά ξέρουμε ότι $\rho_\kappa = \frac{C_\kappa}{C_0}$ οπότε για $\kappa = 0$, $\rho_0 = 1$

άρα ο πίνακας συσχέτισης μπορεί να γραφεί ως

$$\rho_{Xi} = \begin{bmatrix} 1 \_ \rho_1 \_ \rho_2 \\ \rho_1 \_ 1 \_ \rho_1 \\ \rho_2 \_ \rho_1 \_ 1 \end{bmatrix}$$

Για 3 τυχαίες μεταβλητές δημιουργήθηκε ένας πίνακας συντελεστών συσχέτισης 3 * 3. Αν είχαμε n μεταβλητές θα είχαμε ένα πίνακα n * n .

5. Μοντέλα ARMA

1. Συνάρτηση αυτοσυσχετίσεων Autocorrelation Function (ACF)

Ένας σπουδαίος οδηγός για την ανάλυση των χρονολογικών σειρών είναι οι λεγόμενοι δειγματικοί συντελεστές αυτοσυσχέτισης. Οι συντελεστές αυτοί δίνουν κατά κανόνα πολύτιμη πληροφορία για το μοντέλο το οποί μπορεί να προσαρμοστεί στα δεδομένα και κατά συνέπεια χρησιμεύουν πολύ στη διαδικασία της ταυτοποίησης.

Πολλά΄ δεδομένα στη Οικονομία λαμβάνονται με τη μορφή χρονολογικών σειρών και οι παρατηρήσεις είναι εξαρτημένες μεταξύ τους η εξάρτησή ων οποίων παρουσιάζει ιδιαίτερο ενδιαφέρον. Για παραδειγμα η σημερινή τιμή κλεισίματος μίας μετοχής θα εξαρτάται εν μέρει από τη χθεσινή, τη προχθεσινή κτλ. Η συσχέτιση μεταξύ διαδοχικών τιμών μιας μεταβλητής καλείται αυτοσυσχέτιση.

Έστω ότι έχουμε ένα πλήθος N παρατηρήσεων $x_1, x_2, \ldots x_N$ σε ισαπέχουσες χρονικές στιγμές μιας τυχαίς μεταβλητής x . Από τα δεδομένα αυτά μπορούμε να σχηματίσουμε N – 1 ζευγάρια παρατηρήσεων .

$(x_1, x_2)(x_2, x_3) \ldots (x_{N-1}, x_N)$

Θεωρώντας την πρώτη παρατήρηση σε κάθε ζευγάρι σαν μια τυχαία μεταβλητή και τη δεύτερη σαν παρατήρηση σαν δεύτερη μεταβλητή, ο δειγματικός συντελεστής αυτοσυσχέτισης μεταξύ των μεταβλητών x_t και x_{t+1} $t = 1,2\ldots N-1$

Ορίζεται από την

$$\boxed{r_1 = \frac{\sum_{t=1}^{N-1}(X_t - \bar{X})(X_{t+1} - \bar{X})}{\sum_{t=1}^{N}(X_t - \bar{X})^2}}$$

Όπου $\bar{X} = \frac{1}{N}\sum_{t=1}^{N} X_t$ ο μέσος της σειράς .

Ανάλογα ορίζεται και ο δειγματικός συντελεστής αυτοσυσχέτισης μεταξύ παρατηρήσεων σε «απόσταση κ » ο οποίος δίνεται από τη σχέση

$$r_\kappa = \frac{\sum_{t=1}^{N-\kappa}(X_t - \bar{X})(X_{t+\kappa} - \bar{X})}{\sum_{t=1}^{N}(X_t - \bar{X})^2}$$

Ο αριθμητής θεωρείται η αυτοσυνδιακύμαση γ_κ και ο παρονομαστής η διασπορά γ_o της διαδικασίας.

Το διάγραμμα των σιγμαικών συντελεστών αυτοσυσχέτισης προκύπτει από τα ζεύγη (κ , r_κ) αν αυτά εμφανιστούν σε ένα ορθογώνιο σύστημα αξόνων.

2. Η διαδικασία αυτοπαλινδρόμησης τάξης p

Έστω ε_t $t \in T$ μια τελείως τυχαία διαδικασία καταλοίπων με μέσο 0 και διακύμανση σ^2. Μια διαδικασία x_t $t \in T$ θα καλείται διαδικασία αυτοπαλινδρόμησης τάξης p και θα συμβολίζεται ως AR (p) αν :

$$\boxed{x_t = \xi + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \phi_3 x_{t-3} + \ldots + \phi_p x_{t-p} + e_t} \quad (1)$$

Όπου φ' = ($\varphi_1, \varphi_2, \ldots, \varphi_p$,) ένα διάνυσμα παραμέτρων.

Αν μια χρονολογική σειρά έχει προκύψει από την παραπάνω διαδικασία (1) η τιμή x_t θα εξαρτάται από τις t = p προηγούμενες τιμές συν το τυχαίο σφάλμα ε_t συν μια σταθερά ξ .

Το υπόδειγμα (1) θυμίζει υπόδειγμα γραμμικής παλινδρόμησης με εξαρτημένη μεταβλητή την x_t και ανεξάρτητες την x_{t-1} , x_{t-2} , , x_{t-p} .

Χρησιμοποιώντας το τελεστή οπισθοδρόμησης Β που ορίζεται από την

$B^\kappa x_t = x_{t-\kappa}$

Η (1) μπορεί να γραφτεί ως εξής

$$\boxed{x_t = \xi + (\phi_1 B + \phi_2 B^2 + \phi_3 B^3 + \ldots + \phi_p B^p)x_t + e_t}$$

$$\boxed{(1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3 - \ldots - \phi_p B^p)x_t = \xi + e_t}$$

ή αλλιώς

$\Phi(B)\, x_t = \xi + \varepsilon_t$

με

$$\boxed{\Phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3 - \ldots - \phi_p B^p}$$

να καλείται χαρακτηριστικό πολυώνυμο της διαδικασίας (1)

Η διαδικασία αυτοπαλινδρόμησης θεωρείται ευσταθής αν οι ρίζες του χαρακτηριστικού της πολυωνύμου βρίσκονται έξω από τον μοναδιαίο κύκλο. Δηλαδή αν G_i $i = 1, 2, \ldots p$ οι ρίζες του χαρακτηριστικού πολυωνύμου τότε θα πρέπει

$$| G_i | > 1$$

3. Η διαδικασία αυτοπαλινδρόμησης πρώτης τάξης

Το πιο απλό παράδειγμα αυτοπαλινδρόμησης είναι η διαδικασία αυτοπαλινδρόμησης πρώτης τάξης $p = 1$ AR (1)

$$\boxed{x_t = \xi + \phi_1 x_{t-1} + e_t}$$

και $\xi = 0$

$$\boxed{x_t = \phi_1 x_{t-1} + e_t} \qquad (1)$$

Η διαδικασία αυτή μπορεί να γραφτεί ως εξής

Για $t = t - 1$ έχουμε

$$\boxed{x_{t-1} = \phi_1 x_{t-2} + e_{t-1}} \qquad (2)$$

και αντικαθιστώντας στην (1) την (2) έχουμε

$$\boxed{x_t = \phi_1(\phi_1 x_{t-2} + e_{t-1}) + e_t}$$

$$\boxed{x_t = \phi_1^2 x_{t-2} + \phi_1 e_{t-1} + e_t} \qquad (3)$$

Ομοίως, στην (1) για $t = t - 2$ έχουμε

$$\boxed{x_{t-2} = \phi_1 x_{t-3} + e_{t-2}} \qquad (4)$$

και αντικαθιστώντας στην (3) την (4) έχουμε

$$\boxed{x_t = \phi_1^2(\phi_1 x_{t-3} + e_{t-2}) + \phi_1 e_{t-1} + e_t}$$

$$\boxed{x_t = \phi_1^3 x_{t-3} + \phi_1^2 e_{t-2} + \phi_1 e_{t-1} + e_t}$$

και αν συνεχίσουμε θα φτάσουμε στην ακόλουθη σχέση

$$\boxed{x_t = e_t + \phi_1 e_{t-1} + \phi_1^2 e_{t-2} + \phi_1^3 e_{t-3} + \ldots}$$

δηλαδή η x_t να γράφεται σαν ένα άθροισμα υπολοίπων μέχρι και την χρονική στιγμή t .

Το χαρακτηριστικό πολυώνυμο της διαδικασίας είναι

$\Phi(B) = (1 - \varphi_1 B)$

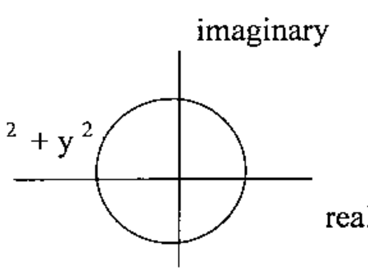
Συνθήκες που πρέπει να ικανοποιούν οι συντελεστές και οι οποίες θα εξασφαλίζουν τη στασιμότητα της διαδικασίας

$\underline{1-\varphi B} = 0$

Φ(Β)

$\varphi B = 1$

Μέτρο του μιγαδικού αριθμού $x^2 + y^2$



$B = \dfrac{1}{\phi}$

i) Αν φ > 1 τότε B < 1 οπότε η χρονοσειρά θα είναι **μη στάσιμη** διότι η ρίζα του χαρακτηριστικού πολυωνύμου **Φ(Β)** θα πέφτει μέσα στον unit cycle (μοναδιαίο κύκλο)

ii) Αν φ < 1 τότε B > 1 οπότε η χρονοσειρά θα είναι **<u>στάσιμη</u>** διότι η ρίζα του χαρακτηριστικού πολυωνύμου **Φ (Β) θα** πέφτει έξω από τον μοναδιαίο κύκλο (unit cycle)

Συνεπώς η συνθήκη για την εξασφάλιση της στασιμότητας είναι

$$\boldsymbol{\varphi < 1}$$

Μια ενδιαφέρουσα σχέση είναι η συνάρτηση των συντελεστών αυτοσυσχέτισης για μια διαδικασία AR (1)

z_n : η τιμή της χρονοσειράς

w_n : η τιμή του τυχαίου σφάλματος

$z_n - \varphi z_{n-1} = w_n$

πολλαπλασιάζουμε με $z_{n-\kappa}$

$z_{n-\kappa} z_n - z_{n-\kappa} \varphi z_{n-1} = z_{n-\kappa} w_n$

παίρνουμε αναμενόμενες τιμές (Expectation)

$E(z_{n-\kappa} z_n) - \varphi E(z_{n-\kappa} z_{n-1}) = E(z_{n-\kappa} w_n)$

<u>Για κ >1</u> $\gamma_\kappa - \varphi\gamma_{\kappa-1} = 0$ (1)

όπου $\gamma_\kappa = Cov(z_n z_{n-\kappa}) = E(z_n - \mu)(z_{n-\kappa} - \mu)$ και

$\gamma_{\kappa-1} = \gamma_{(n-1)-(n-\kappa)} = Cov(z_{n-\kappa} z_{n-1}) = E(z_{n-\kappa} - \mu)(z_{n-1} - \mu)$

και $E(z_{n-\kappa} w_n) = E(z_{n-\kappa}) E(w_n) = E(z_{n-\kappa})\, 0 = 0$ διότι η αναμενόμενη τιμή του λευκού θορύβου είναι 0

Στη σχέση (1) διαιρούμε και με γ_o (διασπορά) και έχουμε

$\gamma_\kappa - \varphi\gamma_{\kappa-1} = 0$

$$\frac{\gamma_\kappa}{\gamma_o} - \varphi\frac{\gamma_{\kappa-1}}{\gamma_o} = 0$$

$\rho_\kappa - \varphi\rho_{\kappa-1} = 0$

$\rho_\kappa = \varphi\, \rho_{\kappa-1}$

για κ = 1 έχουμε

$\rho_1 = \varphi\, \rho_0$

$\rho_1 = \varphi$ (A)

για κ = 2 έχουμε

$\rho_2 = \varphi\, \rho_1$

$\rho_2 = \varphi\varphi$ (από την Α)

$\rho_2 = \varphi^2$

και γενικά

$\rho_\kappa = \varphi^\kappa$

<u>Διασπορά της διαδικασίας</u>

$z_n - \varphi z_{n-1} = w_n$

Η z_n μπορεί να γραφτεί ως εξής

$z_n = w_n + \varphi_1 w_{n-1} + \varphi_1^2 w_{n-2}$ ... (1)

$z_n - \varphi z_{n-1} = w_n$

πολλαπλασιάζουμε με το z_n

$z_n\, z_n - \varphi\, z_n\, z_{n-1} = z_n\, w_n$

παίρνουμε αναμενόμενες τιμές (Expectation)

$E\,(z_n\, z_n) - \varphi E(z_n\, z_{n-1}) = E\,(z_n\, w_n)$ (2)

όπου

$E\,(z_n\, z_n) = Cov(z_n\, z_n) = E\,(z_n - \mu)\,(z_n - \mu) = E\,(z_n - \mu)^2 = \gamma_o = \sigma^2$

$E(z_n\, z_{n-1}) = Cov(z_n\, z_{n-1}) = E\,(z_n - \mu)\,(z_{n-1} - \mu) = \gamma_1$

και

$E\,(z_n\, w_n) = E\,(w_n\, z_n) = E\,(w_n)(w_n + \varphi_1 w_{n-1} + \varphi_1^2\, w_{n-2}\ \ldots\ ) = E\,(w_n\, w_n) + \varphi_1 E(w_n\, w_{n-1}$

$)$

$+ \varphi_1^2 E(w_n\, w_{n-2})$ με

$E\,(w_n\, w_n) = E\,(w_n - \mu)\,(w_n - \mu) = E\,(w_n - \mu)^2 = E\,(w_n)^2 = \sigma^2_w$ αφού ο μ του λευκού θορύβου είναι 0

$E(w_n\, w_{n-1}) = E\,(w_n)(w_{n-1}) = 0$ διότι ο λευκός θόρυβος είναι ένα σύνολο από ασυσχέτιστες τ.μ.

$E(w_n\, w_{n-2}) = 0$ για τον παραπάνω λόγο

Συνεπώς έχουμε

$E\,(w_n)(w_n + \varphi_1 w_{n-1} + \varphi_1^2\, w_{n-2}\ \ldots\ ) = \sigma^2_w$

Άρα στην (2) έχουμε

$E\,(z_n\, z_n) - \varphi E(z_n\, z_{n-1}) = E\,(z_n\, w_n)$

$\sigma^2 - \varphi\, \gamma_1 = \sigma^2_w$

αλλά $\rho_1 = \dfrac{\gamma_1}{\gamma_o}$ οπότε $\gamma_1 = \rho_1\, \gamma_o$ ή $\gamma_1 = \rho_1\, \sigma^2$

Συνεπώς

$\sigma^2 - \varphi\rho_1\sigma^2 = \sigma^2_w$

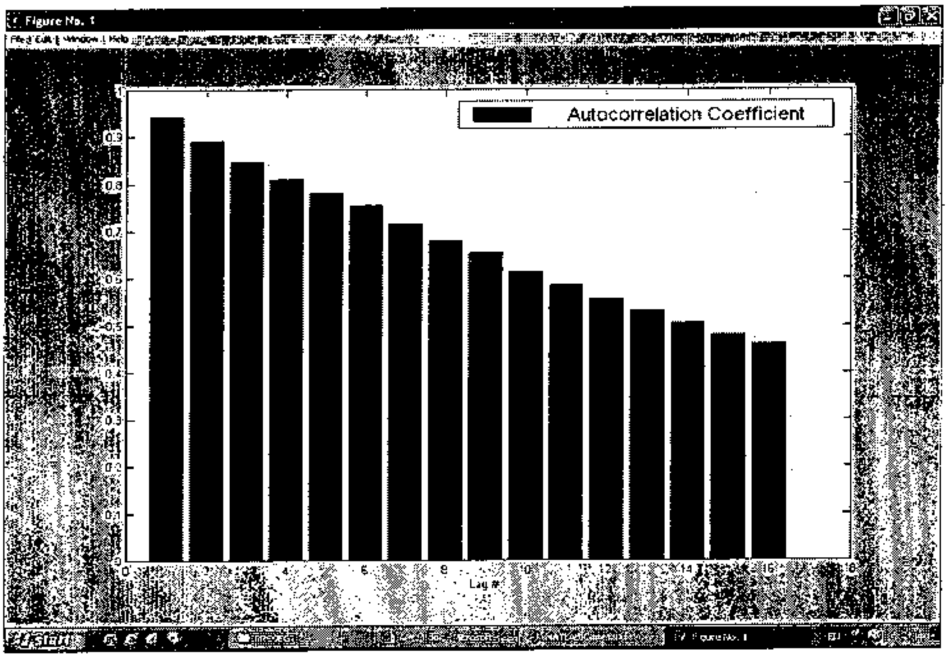
όμως $\rho_\kappa = \varphi^\kappa$ άρα για κ =1 έχουμε $\rho_1 = \varphi^1$

$\sigma^2 - \varphi\varphi\sigma^2 = \sigma^2_w$

$\sigma^2 - \varphi^2\sigma^2 = \sigma^2_w$

$\sigma^2\,(1 - \varphi^2) = \sigma^2_w$ $\sigma^2 = \dfrac{\sigma_w^2}{1-\phi^2}$

Να σημειωθεί ότι σε ευσταθείς διαδικασίες AR (1) επειδή $\varphi_1 < 1$ η συνάρτηση ρ_κ φθίνει κατά απόλυτη τιμή εκθετικά όταν το κ αυξάνει (βλέπε παρακάτω διάγραμμα)



Για παράδειγμα

$z_n = 0.8z_{n-1} + w_n$

$\rho_\kappa = (0.8)^\kappa$

για $\kappa = 1$

$\rho_1 = 0.8$

για $\kappa = 2$

$\rho_2 = 0.64$

για $\kappa = 3$

$\rho_3 = 0.512$

$0.8 > 0.64 > 0.512$

4. Η διαδικασία αυτοπαλινδρόμησης δεύτερης τάξης

Η διαδικασία αυτοπαλινδρόμησης τάξης $p = 2$, AR (2)

$$\boxed{x_t = \xi + \phi_1 x_{t-1} + \phi_2 x_{t-2} + e_t}$$

και $\xi = 0$

$$\boxed{x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + e_t} \qquad (A)$$

Το χαρακτηριστικό πολυώνυμο της διαδικασίας είναι

$\Phi(B) = (1 - \varphi_1 B - \varphi_2 B^2)$

AR (2)

$z_n - \varphi_1 z_{n-1} - \varphi_2 z_{n-2} = w_n$

πολλαπλασιάζουμε με $z_{n-\kappa}$

$z_n z_{n-\kappa} - \varphi_1 z_{n-1} z_{n-\kappa} - \varphi_2 z_{n-2} z_{n-\kappa} = z_{n-\kappa} w_n$

παίρνουμε αναμενόμενες τιμές (Expectation)

$E(z_n z_{n-\kappa}) - \varphi_1 E(z_{n-1} z_{n-\kappa}) - \varphi_2 E(z_{n-2} z_{n-\kappa}) = E(z_{n-\kappa} w_n)$

<u>Για κ >1</u> $\gamma_\kappa - \varphi_1 \gamma_{\kappa-1} - \varphi_2 \gamma_{\kappa-2} = 0$ (1)

όπου $\gamma_\kappa = Cov(z_n z_{n-\kappa}) = E(z_n - \mu)(z_{n-\kappa} - \mu)$ και

$\gamma_{(n-1)-(n-\kappa)} = \gamma_{\kappa-1} = Cov(z_{n-1} z_{n-\kappa}) = E(z_{n-1} - \mu)(z_{n-\kappa} - \mu)$

$\gamma_{(n-2)-(n-\kappa)} = \gamma_{\kappa-2} = Cov(z_{n-2} z_{n-\kappa}) = E(z_{n-2} - \mu)(z_{n-\kappa} - \mu)$

και $E(z_{n-\kappa} w_n) = E(z_{n-\kappa}) E(w_n) = E(z_{n-\kappa})\, 0 = 0$ διότι η αναμενόμενη τιμή του λευκού θορύβου είναι 0

Στη σχέση (1) διαιρούμε και με γ_o (διασπορά) και έχουμε

$\gamma_\kappa - \varphi \gamma_{\kappa-1} - \varphi_2 \gamma_{\kappa-2} = 0$

$$\frac{\gamma_\kappa}{\gamma_o} - \varphi_1 \frac{\gamma_{\kappa-1}}{\gamma_o} - \varphi_2 \frac{\gamma_{\kappa-2}}{\gamma_o} = 0$$

$\rho_\kappa - \varphi_1 \rho_{\kappa-1} - \varphi_2 \rho_{\kappa-2} = 0$

$\rho_\kappa - \varphi_1 B \rho_\kappa - \varphi_2 B^2 \rho_\kappa = 0$

$(1 - \varphi_1 B - \varphi_2 B^2)\, \rho_\kappa = 0$

$\Phi(B)\ \rho_\kappa = 0$

<u>Διασπορά της διαδικασίας</u>

$z_n - \varphi_1 z_{n-1} - \varphi_2 z_{n-2} = w_n$

$z_n = \varphi_1 z_{n-1} + \varphi_2 z_{n-2} + w_n$

$$
\begin{aligned}
\sigma^2 &= E\,(z_n z_n) = E\,[(\varphi_1 z_{n-1} + \varphi_2 z_{n-2} + w_n)(\varphi_1 z_{n-1} + \varphi_2 z_{n-2} + w_n)] \\
&= E\,(\varphi_1^2 (z_{n-1})^2 + \varphi_1\varphi_2 z_{n-1}z_{n-2} + \varphi_1 z_{n-1}w_n + \varphi_2\varphi_1 z_{n-2}z_{n-1} + \varphi_2^2 (z_{n-2})^2 + \varphi_2 z_{n-2}w_n \\
&\quad + \varphi_1 w_n z_{n-1} + \varphi_2 w_n z_{n-2} + (w_n)^2) \\
&= \varphi_1^2 E\,[(z_{n-1})^2] + \varphi_1\varphi_2 E\,(z_{n-1}z_{n-2}) + \varphi_1 E\,(z_{n-1}w_n) + \varphi_2\varphi_1 E\,(z_{n-2}z_{n-1}) + \\
&\quad \varphi_2^2 E\,[(z_{n-2})^2] + \varphi_2 E\,(z_{n-2}w_n) + \varphi_1 E\,(w_n z_{n-1}) + \varphi_2 E\,(w_n z_{n-2}) + E\,[(w_n)^2] \\
&= \varphi_1^2 \gamma_o + \varphi_1\varphi_2 \gamma_1 + 0 + \varphi_1\varphi_2 \gamma_1 + \varphi_2^2 \gamma_o + 0 + 0 + 0 + \sigma^2_w \\
&= \varphi_1^2 \gamma_o + \varphi_1\varphi_2 \gamma_1 + \varphi_1\varphi_2 \gamma_1 + \varphi_2^2 \gamma_o + \sigma^2_w \\
&= \varphi_1^2 \sigma^2 + 2\varphi_1\varphi_2 \gamma_1 + \varphi_2^2 \sigma^2 + \sigma^2_w
\end{aligned}
$$

Άρα $\sigma^2 = \varphi_1^2 \sigma^2 + 2\varphi_1\varphi_2\gamma_1 + \varphi_2^2 \sigma^2 + \sigma^2_w$

$(\varphi_1^2 + \varphi_2^2 - 1)\,\sigma^2 + 2\varphi_1\varphi_2\gamma_1 + \sigma^2_w = 0$

$(\varphi_1^2 + \varphi_2^2 - 1)\,\sigma^2 = 2\varphi_1\varphi_2\gamma_1 + \sigma^2_w$

$$\sigma^2 = \frac{2\phi_1\phi_2\gamma_1 + \sigma_w^2}{\phi_1^2 + \phi_2^2 - 1}$$

Συνθήκες που πρέπει να ικανοποιούν οι συντελεστές και οι οποίες θα εξασφαλίζουν τη στασιμότητα της διαδικασίας

$z_n - \varphi_1 z_{n-1} - \varphi_2 z_{n-2} = w_n$

$\underline{(1 - \varphi_1 B - \varphi_2 B^2)}\ z_n = 0$

$\Phi(B) = 0$

$1 - \varphi_1 B - \varphi_2 B^2 = 0$

$\varphi_2 B^2 + \varphi_1 B - 1 = 0$

Πρέπει οι ρίζες του $\Phi(B)$ να είναι έξω από τον μοναδιαίο κύκλο

Δηλαδή θα πρέπει $G_{1,2} > 1$

Η παρακάτω σχέση

$\rho_\kappa - \varphi_1 B\rho_\kappa - \varphi_2 B^2\rho_\kappa = 0$

μπορεί να γραφτεί

$\rho_\kappa = \varphi_1 B\rho_\kappa + \varphi_2 B^2\rho_\kappa$

$\varphi_1 B\rho_\kappa + \varphi_2 B^2\rho_\kappa = \rho_\kappa$

$\varphi_1\rho_{\kappa-1} + \varphi_2\rho_{\kappa-2} = \rho_\kappa \qquad (B)$

Αν στη (Β) θέσουμε κ = 1 , 2 έχουμε

Για κ = 1 γίνεται $\varphi_1 + \varphi_2\rho_1 = \rho_1$

Για κ = 2 γίνεται $\varphi_1 \rho_1 + \varphi_2 = \rho_2$

Το παραπάνω ζεύγος εξισώσεων είναι γνωστό ως **εξισώσεις Yule – Walker**

Η λύση των εξισώσεων αυτών ως προς φ_1 και φ_2 δίνει

$\varphi_1 = \rho_1 (1 - \rho_2) / (1 - \rho_1)^2$

$\varphi_2 = (\rho_2 - \rho_1^2) / (1 - \rho_1)^2$

5. Η διαδικασία κινητού μέσου όρου τάξης q

Έστω ε_t $t \in T$ μια τελείως τυχαία διαδικασία καταλοίπων με μέσο 0 και διακύμανση σ^2 . Μια διαδικασία x_t $t \in T$ θα καλείται διαδικασία κινητού μέσου όρου q (moving average order q) και θα συμβολίζεται ως MA (q) αν :

$$\boxed{x_t = \xi + \sum_{m=0}^{q} \theta_m e_{t-m}}$$

$$\boxed{x_t = \xi + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \theta_3 e_{t-3} + \dots + \theta_q e_{t-q} + e_t} \qquad (1)$$

Όπου θ' = (θ_1, θ_2, …, θ_q ,) ένα διάνυσμα παραμέτρων.

Αν μια χρονολογική σειρά έχει προκύψει από την παραπάνω διαδικασία (1) η τιμή x_t θα εξαρτάται από την τιμή του τυχαίου σφάλματος ε_t την προηγούμενη χρονική στιγμή συν την τιμή του τυχαίου σφάλματος ε_t πριν 2 χρονικές στιγμές συν την τιμή του τυχαίου σφάλματος ε_t πριν q χρονικές στιγμές συν μια σταθερά ξ .

Χρησιμοποιώντας το τελεστή οπισθοδρόμησης Β που ορίζεται από την

$B^\kappa x_t = x_{t-\kappa}$

Η (1) μπορεί να γραφτεί ως εξής

$$\boxed{x_t = \xi + (\theta_1 B + \theta_2 B^2 + \theta_3 B^3 + ... + \theta_q B^q) e_t + e_t}$$

$$\boxed{x_t = (1 + \theta_1 B + \theta_2 B^2 + \theta_3 B^3 + ... + \theta_q B^q) e_t + \xi}$$

ή αλλιώς

$x_t = \Theta(B)\, \varepsilon_t + \xi$

με

$$\boxed{\Theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \theta_3 B^3 - ... - \theta_q B^q}$$

να καλείται χαρακτηριστικό πολυώνυμο της διαδικασίας (1)

Ο μέσος της διαδικασίας μπορεί να υπολογιστεί ως εξής : (για $\xi = 0$)

$$\boxed{x_t = \sum_{m=0}^{q} \theta_m e_{t-m}}$$

$$\boxed{\mu = \mathrm{E}\{x_t\} = \mathrm{E}\{\sum_{m=0}^{q} \theta_m e_{t-m}\} = \sum_{m=0}^{q} \theta_m \mathrm{E}\{e_{t-m}\} = \sum_{m=0}^{q} \theta_m 0 = 0}$$

γιατί ο μέσος του τυχαίου σφάλματος είναι ο0 για κάθε χρονική στιγμή

$\mathrm{E}\{\varepsilon_t\} = 0$

Η διασπορά της διαδικασίας *(10^ο^ στη Βιβλιογραφία)*

μπορεί να υπολογιστεί ως εξής : (για $\xi = 0$)

$\sigma^2 = \mathrm{E}\{x_t - \mu\}^2 = \mathrm{E}\{x_t - 0\}^2 = \mathrm{E}\{x_t\}^2 =$

$$\mathrm{E}\{\sum_{m=0}^{q} \theta_m e_{t-m}\}^2$$

Με τη μορφή διανυσμάτων η παραπάνω σχέση μπορεί να γραφτεί ως εξής

$$\sum_{m=0}^{q} e_{t-m}^2 = \mathbf{e}_t^T \mathbf{e}_t \quad \text{όπου } \mathbf{e}_t = \begin{bmatrix} e_t \\ e_{t-1} \\ . \\ . \\ . \\ e_{t-q} \end{bmatrix} \quad \& \quad \mathbf{e}_t^T = [e_t \; e_{t-1} \; \ldots \; e_{t-q}]$$

$$\sum_{m=0}^{q} e_{t-m}^2 = [\,(e_t)^2 + (e_{t-2})^2 + \ldots + (e_{t-q})^2\,] = [e_t\ e_{t-1} \ldots e_{t-q}] \begin{bmatrix} e_t \\ e_{t-1} \\ . \\ . \\ . \\ e_{t-q} \end{bmatrix}$$

δηλαδή

$$\sum_{m=0}^{q} e_{t-m}^2 = \mathbf{e}_t^{\mathbf{T}}\ \mathbf{e}_t \quad \text{(A)}$$

$$\text{όπου } e_t = \begin{bmatrix} e_t \\ e_{t-1} \\ . \\ . \\ . \\ e_{t-q} \end{bmatrix} \quad \& \quad e_t^{T} = [e_t\ e_{t-1} \ldots e_{t-q}]$$

με T = transpose (ανάστροφος, δηλαδή οι γραμμές γίνονται στήλες και οι στήλες γραμμές)

αλλά και

$$\sum_{m=0}^{q} \theta_m^2 = [\,(\theta_o)^2 + (\theta_1)^2 + \ldots + (\theta_q)^2\,] = [\theta_o\ \theta_1 \ldots \theta_q] \begin{bmatrix} \theta_0 \\ \theta_1 \\ . \\ . \\ . \\ \theta_q \end{bmatrix}$$

$$\sum_{m=0}^{q} \theta_m^2 = \boldsymbol{\theta}_m^{\mathbf{T}}\ \boldsymbol{\theta}_m \quad \text{(B)}$$

όπου $\theta_m = \begin{bmatrix} \theta_0 \\ \theta_1 \\ . \\ . \\ . \\ \theta_q \end{bmatrix}$ & $\theta_m^T = [\theta_o\ \theta_1\ \dots\ \theta_q]$

χρησιμοποιώντας τις Α και Β έχουμε

$E\{\sum_{m=0}^{q} \theta_m e_{t-m}^2\}^2 =$

$E\{ \boldsymbol{\theta}_m^T \mathbf{e}_t \mathbf{e}_t^T \boldsymbol{\theta}_m \}=$

$E\{ \boldsymbol{\theta}_m^T \mathbf{e}_t \mathbf{e}_t^T \boldsymbol{\theta}_m \}=$

$\boldsymbol{\theta}_m^T E\{ \mathbf{e}_t \mathbf{e}_t^T \} \boldsymbol{\theta}_m =$

όμως

$$E\{ \mathbf{e}_t \mathbf{e}_t^T \} = E\{ \begin{bmatrix} e_t \\ e_{t-1} \\ . \\ . \\ . \\ e_{t-q} \end{bmatrix} [e_t\ e_{t-1}\ \dots\ e_{t-q}] \} = E\{ \begin{bmatrix} e_t^2 \_ e_t e_{t-1} \_ \dots \_ e_t e_{t-q} \\ e_{t-1} e_t \_ e_{t-1}^2 \_ \dots \_ e_{t-1} e_{t-q} \\ . \\ . \\ . \\ e_{t-q} e_t \_ e_{t-q}^2 e_{t-1} \_ \dots \_ e_{t-q}^2 \end{bmatrix} \}=$$

$$\begin{bmatrix} \sigma_{et}^2 \_ 0 \_ \dots \_ 0 \\ 0 \_ \sigma_{et-1}^2 \_ \dots \_ 0 \\ . \\ . \\ . \\ 0 \_ 0 \_ \dots \_ \sigma_{et-q}^2 \end{bmatrix} = \boldsymbol{\sigma}^2{}_e \begin{bmatrix} 1 \_ 0 \_ \dots \_ 0 \\ 0 \_ 1 \_ \dots \_ 0 \\ . \\ . \\ . \\ 0 \_ 0 \_ \dots \_ 1 \end{bmatrix} = \boldsymbol{\sigma}^2{}_e \mathbf{I}$$

Άρα

$$\boldsymbol{\sigma}^2 = \boldsymbol{\sigma}^2{}_e \mathbf{I} \sum_{m=0}^{q} \theta_m^2$$

6. Η διαδικασία κινητού μέσου όρου πρώτης τάξης

Το πιο απλό παράδειγμα κινητού μέσου όρου τάξης q είναι η διαδικασία q = 1 MA (1)

$$\boxed{x_t = \xi + \theta_1 e_{t-1} + e_t}$$

και $\xi = 0$

$$\boxed{x_t = \theta_1 e_{t-1} + e_t} \quad (1)$$

Η συνάρτηση αυτοσυσχέτισης αυτής έχει ως ακολούθως

MA (1)

$z_n = w_n + \theta_1 w_{n-1}$

Για κ = 1

Πολλαπλασιάζουμε με z_{n+1}

$z_n z_{n+1} = w_n z_{n+1} + \theta_1 w_{n-1} z_{n+1}$

Παίρνουμε αναμενόμενες τιμές

$E(z_n z_{n+1}) = E(w_n z_{n+1}) + \theta_1 E(w_{n-1} z_{n+1})$ (1)

Μπορούμε να πούμε $z_{n+1} = w_{n+1} + \theta_1 w_n$

Αντικαθιστούμε στην (1) όπου $z_{n+1} = w_{n+1} + \theta_1 w_n$

$\gamma_1 = E\ (w_n (w_{n+1} + \theta_1 w_n)) + \theta_1 E\ (w_{n-1})(w_{n+1} + \theta_1 w_n))$

$\gamma_1 = E\ (w_n w_{n+1}) + \theta_1 E\ ((w_n)^2) + \theta_1 E\ (w_{n-1} w_{n+1}) + \theta_1^{\ 2}\ E\ (w_{n-1} w_n)$

$\gamma_1 = 0 + \theta_1 E\ (w_n)^2) + \theta_1 * 0 + \theta_1^{\ 2} * 0$

$\gamma_1 = \theta_1 \sigma^2_w$

διαιρούμε με γ_o (διασπορά – variance)

$$\frac{\gamma_1}{\gamma_o} = \frac{\theta_1 \sigma_w^2}{\gamma_o}$$

$$\rho_1 = \frac{-\theta_1 \sigma_w^2}{\gamma_o} \quad (2)$$

Πρέπει να βρούμε τη διασπορά γ_o

$z_n = w_n + \theta_1 w_{n-1}$

$$\gamma_0 = E(z_n\ z_n) = E[(w_n + \theta_1 w_{n-1})(w_n + \theta_1 w_{n-1})]$$
$$= E[(w_n)^2 + 2\,\theta_1\, w_n\, w_{n-1} + \theta_1^{\ 2} (w_{n-1})^2]$$
$$= E(w_n)^2 + 2\,\theta_1\, E(w_n\, w_{n-1}) + \theta_1^{\ 2} E((w_{n-1})^2)$$
$$= \sigma^2_w + 2 * 0 + \theta_1^{\ 2}\sigma^2_w \qquad \text{διότι } \sigma^2_{w1} = \sigma^2_{w2} = \sigma^2_{w3} = \ldots = \sigma^2_{wn}$$
$$= \sigma^2_w (1 + \theta_1^{\ 2})$$

Συνεπώς πηγαίνουμε στην (2) και αντικαθιστούμε όπου $\gamma_0 = \sigma^2_w (1 + \theta_1^{\ 2})$

$$\rho_1 = \frac{\theta_1 \sigma_w^2}{(1+\theta_1^2)\sigma_w^2}$$

$$\rho_1 = \frac{\theta_1}{1+\theta_1^2}$$

Για κ > 1 η ακολουθία των συντελεστών αυτοσυσχέτισης (ACF) είναι 0 .

Απόδειξη

Για κ = 2

Πολλαπλασιάζουμε με z_{n+2}

$z_n\, z_{n+2} = w_n\, z_{n+2} + \theta_1 w_{n-1}\, z_{n+2}$

Παίρνουμε αναμενόμενες τιμές

$E(z_n\, z_{n+2}) = E(w_n\, z_{n+2}) + \theta_1 E(w_{n-1}\, z_{n+2})$ (1)

Μπορούμε να πούμε $z_{n+2} = w_{n+2} + \theta_1 w_{n+1}$

Αντικαθιστούμε στην (1) όπου $z_{n+2} = w_{n+2} + \theta_1 w_{n+1}$

$\gamma_2 = E(w_n (w_{n+2} + \theta_1 w_{n+1})) + \theta_1 E(w_{n-1})(w_{n+2} + \theta_1 w_{n+1}))$

$\gamma_2 = E(w_n\, w_{n+2}) + \theta_1 E((w_n\, w_{n+1}) + \theta_1 E(w_{n-1} w_{n+2}) + \theta_1^{\ 2}\ E(w_{n-1}\, w_{n+1})$

$\gamma_2 = 0 + \theta_1 * 0 + \theta_1 * 0 + \theta_1^{\ 2} * 0$

$\gamma_2 = 0$

διαιρούμε με γ_0 (διασπορά – variance)

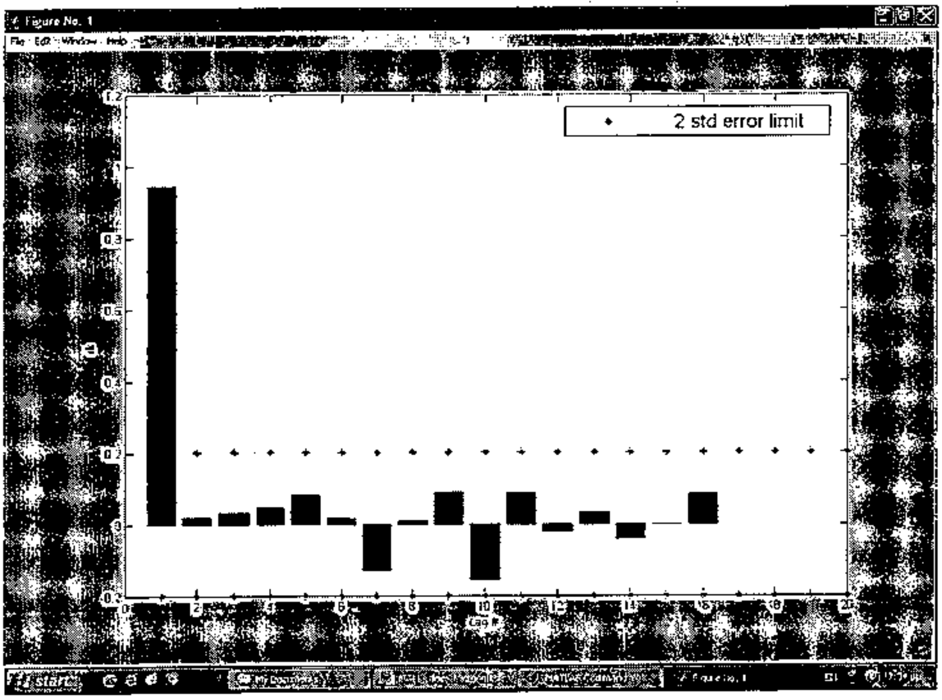
$$\frac{\gamma_2}{\gamma_o} = \frac{0}{\gamma_o}$$

$$\rho_2 = 0 \quad (2)$$

Συνεπώς έχουμε

$$\rho_\kappa \begin{cases} \dfrac{-\theta_1}{1+\theta_1^2} & \text{για } \kappa = 1 \\ 0 & \text{για } \kappa > 1 \end{cases}$$

<u>Διαγραμματικά η ACF της MA(1) θα πρέπει να είναι σύμφωνα με το ακόλουθο διάγραμμα</u>



Συνθήκες που επιτρέπουν την αντιστρεψιμότητά της MA (1) διαδικασίας

$z_n = w_n + \theta_1 w_{n-1}$

$z_n = w_n + \theta_1 B w_n$

$z_n = (1 + \theta_1 B)\, w_n$

$z_n = \Theta(B)\, w_n$

$$\frac{z_n}{\Theta(B)} = \frac{\Theta(B)}{\Theta(B)} w_n$$

$$\frac{z_n}{\Theta(B)} = w_n$$

$\Theta(B)^{-1}\; z_n = w_n$

$(1 + \theta_1 B)^{-1}\; z_n = w_n$

$\Theta(B)^{-1} = 0$

$$1 + \frac{1}{\theta_1 B} = 0$$

$$-\frac{1}{\theta_1 B} = 1$$

$\theta_1 B = -1$

$$B = -\frac{1}{\theta_1}$$

Πρέπει $|B| > 1$ άρα $|\frac{1}{\theta_1}| > 1$ δηλαδή θα πρέπει $\theta_1 < 1$

7. Η διαδικασία κινητού μέσου όρου δεύτερης τάξης

Ένα απλό παράδειγμα κινητού μέσου όρου τάξης q είναι η διαδικασία q = 2 MA (2)

$$\boxed{x_t = \xi + \theta_1 e_{t-1} + \theta_2 e_{t-2} + e_t}$$

και $\xi = 0$

$$\boxed{x_t = \theta_1 e_{t-1} + \theta_2 e_{t-2} + e_t} \quad (1)$$

Η συνάρτηση αυτοσυσχέτισης αυτής έχει ως ακολούθως

MA (2)
$z_n = w_n + \theta_1 w_{n-1} + \theta_2 w_{n-2}$

Για κ = 1

Πολλαπλασιάζουμε με z_{n+1}

$z_n z_{n+1} = w_n z_{n+1} + \theta_1 w_{n-1} z_{n+1} + \theta_2 w_{n-2} z_{n+1}$
Παίρνουμε αναμενόμενες τιμές

$$E(z_n z_{n+1}) = E(w_n z_{n+1}) + \theta_1 E(w_{n-1} z_{n+1}) + \theta_2 E(w_{n-2} z_{n+1}) \quad (1)$$

Μπορούμε να πούμε $z_{n+1} = w_{n+1} + \theta_1 w_n + \theta_2 w_{n-1}$

Αντικαθιστούμε στην (1) όπου $z_{n+1} = w_{n+1} + \theta_1 w_n + \theta_2 w_{n-1}$

$$E(w_n (w_{n+1} + \theta_1 w_n + \theta_2 w_{n-1})) + \theta_1 E(w_{n-1} (w_{n+1} + \theta_1 w_n - \theta_2 w_{n-1})) + \theta_1 E(w_{n-1} (w_{n+1} + \theta_1 w_n + \theta_2 w_{n-1}))$$

(A) (B) (Γ)

$$E(w_n w_{n+1}) + \theta_1 E((w_n)^2) + \theta_2 E(w_n w_{n-1}) = 0 + \theta_1 \sigma^2_w + \theta_2 * 0 = + \theta_1 \sigma^2_w$$

$$\theta_1 E(w_{n-1} (w_{n+1} + \theta_1 w_n + \theta_2 w_{n-1})) = \theta_1 E(w_{n-1} w_{n+1}) + \theta_1^2 E(w_{n-1} w_n) + \theta_1 \theta_2 E((w_{n-1})^2) = \theta_1 * 0 +$$
$$0 + \theta_1 \theta_2 \sigma^2_w = \theta_1 \theta_2 \sigma^2_w$$

$$\theta_1 E(w_{n-1} (w_{n+1} + \theta_1 w_n + \theta_2 w_{n-1})) = \theta_1 E(w_{n-1} w_{n+1}) + \theta_1^2 E(w_{n-1} w_n) + \theta_1 \theta_2 E((w_{n-1})^2) = \theta_1 * 0$$
$$0 + \theta_1 \theta_2 \sigma^2_w = \theta_1 \theta_2 \sigma^2_w$$

Συνεπώς στην **(1)** έχουμε

$\gamma_1 = \theta_1 \sigma^2_w + \theta_1 \theta_2 \sigma^2_w + \theta_1 \theta_2 \sigma^2_w$

$\gamma_1 = 2\theta_1 \theta_2 \sigma^2_w + \theta_1 \sigma^2_w$

$\gamma_1 = \theta_1 \sigma^2_w (2\theta_2 + 1)$

διαιρούμε με γ_o (διασπορά – variance)

$$\frac{\gamma_1}{\gamma_o} = \frac{\theta_1 \sigma_w^2 (2\theta_2 + 1)}{\gamma_o}$$

$$\rho_1 = \frac{\theta_1 \sigma_w^2 (2\theta_2 + 1)}{\gamma_o} \quad \textbf{(2)}$$

Πρέπει να βρούμε τη διασπορά γ_o

$z_n = w_n + \theta_1 w_{n-1} + \theta_2 w_{n-2}$

$\gamma_o = E(z_n \ z_n) = E[(w_n + \theta_1 w_{n-1} + \theta_2 w_{n-2})(w_n + \theta_1 w_{n-1} + \theta_2 w_{n-2})]$
$= E[(w_n)^2 + \theta_1 w_n w_{n-1} + \theta_2 w_n w_{n-2} + \theta_1 w_{n-1} w_n + \theta_1{}^2 (w_{n-1})^2 + \theta_1\theta_2 w_{n-1} w_{n-2}$
$+ \theta_2 w_{n-2} w_n + \theta_2\theta_1 w_{n-2} w_{n-1} + \theta_2{}^2 (w_{n-2})^2$

$\gamma_o = E[(w_n)^2] + \theta_1 E(w_n w_{n-1}) + \theta_2 E(w_n w_{n-2}) + \theta_1 E(w_{n-1} w_n) + \theta_1{}^2 E[(w_{n-1})^2]$
$+ \theta_1\theta_2 E(w_{n-1} w_{n-2}) + \theta_2 E(w_{n-2} w_n) + \theta_2\theta_1 E(w_{n-2} w_{n-1}) + \theta_2{}^2 E[(w_{n-2})^2]$

$\gamma_o = \sigma^2_w + \theta_1 * 0 + \theta_2 * 0 + \theta_1 * 0 + \theta_1{}^2 \sigma^2_w + \theta_1\theta_2 * 0 + \theta_2 * 0 + \theta_2\theta_1 * 0 + \theta_2{}^2$
σ^2_w

$\gamma_o = \sigma^2_w + \theta_1{}^2 \sigma^2_w + \theta_2{}^2 \sigma^2_w$

$\gamma_o = \sigma^2_w (1 + \theta_1{}^2 + \theta_2{}^2)$ **(3)**

Συνεπώς πηγαίνουμε στην (2) και αντικαθιστούμε όπου $\gamma_o = \sigma^2_w (1 + \theta_1{}^2)$

$$\rho_1 = \frac{\theta_1 \sigma_w^2 (2\theta_2 + 1)}{\sigma_w^2 (1 + \theta_1^2 + \theta_2^2)}$$

8. Συνάρτηση μερικών αυτοσυσχετίσεων Partial Autocorrelation Function (PACF)

Έστω ότι μια χρονολογική σειρά έχει προέλθει από μια διαδικασία αυτοπαλινδρόμησης άγνωστης όμως τάξης έστω κ με $1 \leq \kappa \leq p$.

Ας συμβολίσουμε με $\varphi_{\kappa j}$ τον j συντελεστή σε μια διαδικασία AR (κ) έτσι ώστε ο $\Phi_{\kappa\kappa}$ να είναι ο τελευταίος συντελεστής της διαδικασίας. Το σύστημα των εξισώσεων των **Yule – Walker** σε μορφή μήτρων μπορεί να γραφτεί

$$\rho_\kappa = \varphi_{\kappa 1}\rho_{\kappa-1} + \varphi_{\kappa 2}\rho_{\kappa-2} + \ldots + \varphi_{\kappa\kappa}\rho_{\kappa-p} \quad \text{ή}$$

$$\varphi_{\kappa 1}\rho_{\kappa-1} + \varphi_{\kappa 2}\rho_{\kappa-2} + \ldots + \varphi_{\kappa\kappa}\rho_{\kappa-p} = \rho_\kappa$$

Για κ = 1 $\quad \varphi_{\kappa 1} + \varphi_{\kappa 2}\rho_1 + \ldots + \varphi_{\kappa\kappa}\rho_{1-p} = \rho_1$

Για κ = 2 $\quad \varphi_{\kappa 1}\rho_1 + \varphi_{\kappa 2} + \ldots + \varphi_{\kappa\kappa}\rho_{2-p} = \rho_2$

Για κ = p $\quad \varphi_{\kappa 1}\rho_{p-1} + \varphi_{p-2} + \ldots + \varphi_{\kappa\kappa} = \rho_3$

$$\begin{bmatrix} 1\_\rho_1\_\rho_2\_\cdots\_\rho_{\kappa-1} \\ \rho_1\_1\_\rho_1\_\cdots\_\rho_{\kappa-2} \\ . \\ . \\ . \\ \rho_{\kappa-1}\_\rho_{\kappa-2}\_\rho_{\kappa-3}\cdots\_1 \end{bmatrix} \begin{bmatrix} \phi_{\kappa 1} \\ \phi_{\kappa 2} \\ . \\ . \\ . \\ \phi_{\kappa\kappa} \end{bmatrix} = \begin{bmatrix} \rho_1 \\ \rho_2 \\ . \\ . \\ . \\ \rho_{\kappa\kappa} \end{bmatrix}$$ *(Σημείωση : Η _ είναι το κενό)*

Λύνοντας το σύστημα για κ = 1, 2 , 3 … παίρνουμε διαδοχικά :

$$\varphi_{11} = \rho_1$$

Η ορίζουσά του είναι

$$D = \begin{vmatrix} 1\_\_\rho_1 \\ \rho_1\_\_1 \end{vmatrix}$$

Και για λύσουμε ως προς φ_2

$$Df = \begin{vmatrix} 1\_\_\rho_1 \\ \rho_1\_\rho_2 \end{vmatrix}$$

$\varphi_{22} = Df / D$

Για μια **AR (3)** δηλαδή για p = 3 έχουμε με τον ίδιο τρόπο

$$\rho_\kappa - \varphi_1\rho_{\kappa-1} - \varphi_2\rho_{\kappa-2} - \varphi_3\rho_{\kappa-3} = 0$$

$$\rho_\kappa = \varphi_1\rho_{\kappa-1} + \varphi_2\rho_{\kappa-2} + \varphi_3\rho_{\kappa-3}$$

$$\varphi_1\rho_{\kappa-1} + \varphi_2\rho_{\kappa-2} + \varphi_3\rho_{\kappa-3} = \rho_\kappa \quad (3)$$

Για $\kappa = 1$ η (3) γίνεται $\varphi_1 + \varphi_2\rho_1 + \varphi_3\rho_2 = \rho_1$

Για $\kappa = 2$ η (3) γίνεται $\varphi_1\rho_1 + \varphi_2 + \varphi_3\rho_1 = \rho_2$

Για $\kappa = 3$ η (3) γίνεται $\varphi_1\rho_2 + \varphi_2\rho_1 + \varphi_3 = \rho_3$

Η ορίζουσά του είναι

$$D = \begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}$$

Και για λύσουμε ως προς φ_3

$$Df = \begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}$$

$\varphi_{33} = Df / D$

9. Η μικτή διαδικασία τάξης p, q

Έστω ε_t $t \in T$ μια τελείως τυχαία διαδικασία καταλοίπων με μέσο 0 και διακύμανση σ^2. Μια διαδικασία x_t $t \in T$ θα καλείται μικτή διαδικασία τάξης p, q και θα συμβολίζεται ως ARMA (p, q) αν :

(1)

$$\boxed{x_t - \phi_1 x_{t-1} - \phi_2 x_{t-2} - \ldots - \phi_p x_{t-p} = e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \ldots + \theta_q e_{t-q} + \xi}$$

$$\boxed{(1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p)x_t = (1 + \theta_1 B + \theta_2 B^2 + \ldots + \theta_q B^q)e_t + \xi}$$

$$\boxed{\phi(B)x_t = \theta(B)e_t + \xi}$$

$$x_t = \frac{\theta(B)}{\phi(B)} e_t + \xi$$

Το πηλίκο $\frac{\theta(B)}{\phi(B)}$ ονομάζεται συνάρτηση μεταφοράς

Για q = 0 η (1) ταυτίζεται με μια AR (p) ενώ για p = 0 με μια MA (q) . Με άλλα λόγια τα υποδείγματα ARMA (p, q) περιέχουν σαν υπομοντέλα τις διαδικασίες αυτοπαλινδρόμησης και κινητού μέσου .

10. Η μικτή διαδικασία τάξης (1, 1)

Η διαδικασία ARMA (1, 1) στην περίπτωση μηδενικού ξ έχει την παρακάτω μορφή

$$\boxed{x_t - \phi_1 x_{t-1} = e_t + \theta_1 e_{t-1}}$$

$$\boxed{(1-\phi_1 B)x_t = (1+\theta_1 \mathrm{B})e_t}$$

$$\boxed{x_t = \phi_1 x_{t-1} + e_t + \theta_1 e_{t-1}}$$

Η συνάρτηση αυτοσυσχέτισης έχει ως εξής :

z_n : η τιμή της χρονοσειράς

w_n : η τιμή του τυχαίου σφάλματος

$z_n - \varphi z_{n-1} = w_n + \theta_1 w_{n-1}$

$z_n = \varphi z_{n-1} + w_n + \theta_1 w_{n-1}$

πολλαπλασιάζουμε με $z_{n-\kappa}$

$z_{n-\kappa} z_n = z_{n-\kappa} \varphi z_{n-1} + z_{n-\kappa} w_n + z_{n-\kappa} \theta_1 w_{n-1}$

παίρνουμε αναμενόμενες τιμές (Expectation)

$$E (z_{n-\kappa} z_n) = E (z_{n-\kappa} \varphi z_{n-1}) + \underbrace{E (z_{n-\kappa} w_n)}_{A} + \underbrace{E (z_{n-\kappa} \theta_1 w_{n-1})}_{B}$$

$$\boxed{z_n = w_n + \phi_1 w_{n-1} + \phi_1^2 w_{n-2} + \phi_1^3 w_{n-3} + \ldots}$$

$$\boxed{z_{n-\kappa} = w_{n-\kappa} + \phi_1 w_{n-1-\kappa} + \phi_1^2 w_{n-2-\kappa} + \phi_1^3 w_{n-3-\kappa} + \ldots}$$

$$\boxed{z_{n-1} = w_{n-1} + \phi_1 w_{n-2} + \phi_1^2 w_{n-3} + \phi_1^3 w_{n-4} + \ldots}$$

<u>Για κ >1</u>

A : E ($z_{n-\kappa}$ w_n) = E (z_{n-1} w_n) = E (($w_{n-\kappa} + \phi_1 w_{n-1-\kappa} + \phi_1^2 w_{n-2-\kappa} + \phi_1^3 w_{n-3-\kappa} + \ldots$)

w_n) = 0 *καθώς* $E(\, w_n \; w_{n-\kappa}\,) = 0$

B : E ($z_{n-\kappa}$ $\theta_1 w_{n-1}$) = E (z_{n-1} $\theta_1 w_{n-1}$) =

E (($w_{n-\kappa} + \phi_1 w_{n-1-\kappa} + \phi_1^2 w_{n-2-\kappa} + \phi_1^3 w_{n-3-\kappa} + \ldots$)$\theta_1 w_{n-1}$) = 0

καθώς $E(\, w_{n-1} \; w_{n-\kappa}\,) = 0$ *για* $\kappa > 1$

Άρα

$\gamma_\kappa = \varphi\gamma_{\kappa-1} + 0 + 0$ (1)

Στη σχέση (1) διαιρούμε και με γ_o (διασπορά) και έχουμε

$\gamma_\kappa - \varphi\gamma_{\kappa-1} = 0$

$$\frac{\gamma_\kappa}{\gamma_o} - \varphi\frac{\gamma_{\kappa-1}}{\gamma_o} = 0$$

$\rho_\kappa - \varphi\rho_{\kappa-1} = 0$

$\rho_\kappa = \varphi\, \rho_{\kappa-1}$

Δηλαδή για $\kappa > 1$ ταυτίζεται με την ACF μιας AR (1)

Η συνάρτηση μερικής αυτοσυσχέτισης της διαδικασίας συμπεριφέρεται όπως και η συνάρτηση μερικής αυτοσυσχέτισης της διαδικασίας MA (1) δηλαδή φθίνει κατά απόλυτη τιμή εκθετικά όσο αυξάνει το κ με αρχική τιμή το $\varphi_{11} = \rho_1$

11. Υπόδειγμα εξομάλυνσης ασταθών σειρών

Μέχρι τώρα έχουμε ασχοληθεί μόνο με στάσιμες (ευσταθείς) διαδικασίες . Δεν είναι όμως όλες οι διαδικασίες ευσταθείς οπότε τα υποδείγματα ARMA (p , q) δεν είναι κατάλληλα για να περιγράψουν μια χρονοσειρά .

Πολλές χρονοσειρές συμπεριφέρονται συνήθως κατά τρόπο που δείχνει φανερά που ότι έχουν κάποιο σταθερό μέσο κατά τη διαχρονική τους εξέλιξη όπως για παράδειγμα η τιμή κλεισίματος μιας μετοχής. Μια ομοιογένεια όμως παροσιάζεται στο είδος της αστάθειας που συναντάμε σε οικονομικές χρονολογικές σειρές . Αυτό σημαίνει ότι παρόλο που η σειρά μεταβάλλεται ανεξάρτητα από κάποιο μέσο η συμπεριφορά της σε διαφορετικές χρονικές περιόδους είναι περίπου η ίδια. Με άλλοα λόγια αν μια ασταθής χρονική σειρά x_t $t = 1, 2, 3 \ldots$ παρουσιάζει ομοιογένεια , οι πρώτες (ή διαδοχικές) διαφορές αυτής , δηλαδή η σειρά :

$$z_t = x_t - x_{t-1}, \quad t = 1, 2, 3 \ldots$$

ή αλλιώς

$$\nabla x_t = x_t - x_{t-1} = x_t - B x_t = (1 - B) x_t$$

όπου ∇ο τελεστής διαφόρισης (difference operator)

και $\nabla^2 x_t = \nabla\nabla x_t = (1-B)(1-B) x_t = (1-B)^2 x_t$

γενικά $\nabla^d x_t = (1-B)^d x_t$

Σε κάποιες περιπτώσεις όμως ενδεχομένως να χρειαστούν οι δεύτερες διαφορές δηλαδή ,

$$w_t = z_t - z_{t-1} = (x_t - x_{t-1}) - (x_{t-1} - x_{t-2}) = x_t - 2 x_{t-1} + x_{t-2}$$

, $t = 1, 2, 3 \ldots$

ή αλλιώς

$$\nabla^2 x_t = \nabla\nabla x_t = (1-B)(1-B) x_t = (1-B)(x_t - x_{t-1}) =$$

$$x_t - x_{t-1} - x_{t-1} + x_{t-2} = x_t - 2 x_{t-1} + x_{t-2}$$

Οπότε το υπόδειγμα ARMA (p , q)

$$(1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p) x_t = (1 + \theta_1 B + \theta_2 B^2 + \ldots + \theta_q B^q) e_t + \xi$$

$$\phi(B) x_t = \theta(B) e_t + \xi$$

σε περίπτωση χρήσης των d διαφορών θα πρέπει να γραφτεί σαν

$\phi(B)(1-B)^d x_t = \theta(B)e_t + \xi$

το υπόδειγμα αυτό είναι γνωστό σαν **ARIMA (p , d , q)**

Για παράδειγμα η διαδικασία **ARIMA (1 , 1 , 1)** έχει ως ακολούθως

$(1-\phi_1 B)(1-B)x_t = (1+\theta_1 B)e_t + \xi$

Ακολουθούν κάποια παραδείγματα ταυτοποίησης με διαφορές

1) $z_n - 2z_{n-1} + z_{n-2} = w_n + \theta_1 w_{n-1} + \theta_2 w_{n-2}$

$z_n - 2Bz_n + B^2 z_{n-2} = w_n + \theta_1 B w_n + \theta_2 B^2 w_n$

$(1-2B+B^2)z_n = (1+\theta_1 B + \theta_2 B^2) w_n$

$(1-B)^2 z_n = (1+\theta_1 B + \theta_2 B^2) w_n$

MA (2) όπου 2 είναι ο
αριθμός των παραμέτρων θ_i

Συνεπώς ARIMA (0, 2, 2)

2) $z_n - 3z_{n-1} + 3z_{n-2} - z_{n-3} = w_n$

$z_n - 3Bz_n + 3B^2 z_n - B^3 z_n = w_n$

$(1-3B+3B^2-B^3)z_n = w_n$

$(1-3B+3B^2-B^3)z_n = w_n$

$(1-B)^3 = w_n$

Συνεπώς ARIMA (0, 3, 0)

12. Μετασχηματισμοί χρονικών σειρών

Στις περισσότερες περιπτώσεις που αφορούν μη εποχικά δεδομένα , η λήψη των πρώτων ή δεύτερων διαφορών είναι αρκετή για να μας οδηγήσει σε ευστάθεια. Μερικές φορές όμως η λήψη των διαφορών δεν είναι αρκετή . Για παράδειγμα η διαχρονική εξέλιξη του ΑΕΠ μιας χώρας όπου εκτός από μακροχρόνια τάση παρουσιάζει και μεγαλύτερη μεταβλητότητα χρόνο με το χρόνο. Οπότε ένας άλλος μετασχηματισμός, οι λογαριθμικές μεταβολές, μπορούν να επιφέρουν την επιθυμητή ευστάθεια (στασιμότητα)

Πράγματι αν x_t είναι τα αρχικά δεδομένα και L τα λογαριθμισμένα έχουμε

$$L_t = \log x_t ,\ t = 1, 2, 3 \ldots$$

Και οι λογαριθμικές διαφορές θα είναι

$$L_t - L_{t-1} = \log x_t - \log x_{t-1} = \log \frac{x_t}{x_{t-1}}$$

13. Εκτίμηση των παραμέτρων

Αφού βρεθούν τα p , q το επόμενο βήμα είναι να εκτιμηθούνε οι παράμετροι φ_p και θ_q .

Η διαδικασία εκτίμησης των παραμέτρων συνίσταται στην εύρεση των τιμών εκείνων των παραμέτρων οι οποίες ελαχιστοποιούν το άθροισμα $\sum_{i=1}^{n} e_i^2$

με $e_t = x_t - \hat{x}_t$ με $\hat{x}_t$ η εκτίμηση του x_t

$$S_e(\varphi) = \sum_{i=1}^{n} e_i^2 = e_t^T e_t \quad \text{όπου } e_t = \begin{bmatrix} e_1 \\ e_2 \\ . \\ . \\ . \\ . \\ e_n \end{bmatrix} \quad \& \quad e_t^T = [e_1\ e_2 \ldots e_n]$$

$$[e_1\ e_2 \ldots e_n] \begin{bmatrix} e_1 \\ e_2 \\ . \\ . \\ . \\ e_n \end{bmatrix} = [\ (e_1)^2 + (e_2)^2 + \ \ldots + (e_n)^2\ ] = \sum_{i=1}^{n} e_i^2$$

Αν υποθέσουμε ότι έχουμε σε μορφή διανύσματος

1. X τις τιμές της χρονοσειράς, π.χ. $\begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ x_n \end{bmatrix}$
2. Z το πλήθος των μεταβλητών που επηρεάζουν τις τιμές της χρονοσειράς (x_{t-p} ή e_{t-q})
3. φ το πλήθος των παραμέτρων
4. e_t τα σφάλματα

τότε

$X = Z\varphi + e_t$ ή

$e_t = X - Z\varphi$

$$S_e(\varphi) = \sum_{i=1}^{n} e_i^2 = e_t^T\ e_t = (X - Z\varphi)^T\ (X - Z\varphi)$$

$$S_e(\varphi) = X^2 - 2\,XZ\varphi + (Z\varphi)^2$$

$$S_e(\varphi) = X^TX - 2\,X^TX\varphi + \varphi^T(Z^TZ)\varphi$$

$$S_e(\varphi) = X^2 - 2\,Z^TX\varphi + \varphi^2(Z^TZ)$$

Και επειδή θέλουμε εκείνος τις παραμέτρους για τις οποίες ελαχιστοποιείται το άθροισμα χρειαζόμαστε να βρούμε που μηδενίζεται η μερική παράγωγος της συνάρτησης S_e.

$$\frac{dS_e}{d\phi} = 0$$

$$-2\,Z^T X + 2\,\varphi(Z^TZ) = 0$$

$$2\,\varphi(Z^TZ) - 2Z^T X = 0$$

$$\varphi(Z^TZ) = Z^T X$$

$$\varphi = (Z^T Z)^{-1} Z^T X \quad (2)$$

14. Κριτική στα μοντέλα ARMA

Η χρήση των ARIMA υποδειγμάτων είναι μια ισχυρή προσέγγιση στην επίλυση πολλών προβλημάτων πρόβλεψης. Μπορεί να παρέχει εξαιρετικά ακριβείς προβλέψεις των χρονολογικών σειρών και προσφέρει μια επίσημη και καλά δομημένη προσέγγιση στην ανάλυση και κατασκευή υποδειγμάτων. Ωστόσο, δε λείπουν σε αυτά τα υποδείγματα και αρκετοί σοβαροί περιορισμοί.

Γενικά, απαιτούμε τουλάχιστον 50 και κυρίως 100 παρατηρήσεις για να αναπτύξουμε ένα ικανοποιητικό υπόδειγμα ARIMA. Αυτό είναι μια σχετικά μεγάλη ποσότητα δεδομένων και υπάρχουν αρκετοί τύποι προβλημάτων στα οποία αυτά τα ιστορικά δεδομένα δε θα είναι διαθέσιμα. Για παράδειγμα στην παραγωγή εποχιακών ειδών, η ιστορία των προηγούμενων περιόδων είναι συχνά μικρής αξίας στην πρόβλεψη της τρέχουσας περιόδου εξαιτίας των αλλαγών που γίνονται στη μόδα και στις στρατηγικές μάρκετινγκ . Ακόμη και στην πρόβλεψη ζήτησης για ένα σχετικά σταθερό προϊόν σε μηνιαία βάση , μπορεί να μην είμαστε σε θέση να συλλέξουμε δεδομένα από 50 έως 100 μήνες που να είναι τελείως αξιόπιστα καθώς οι αλλαγές στο πληροφοριακό σύστημα ή στο σύστημα ελέγχου παραγωγής μπορεί να έχουν τροποποιήσει τον τρόπο με τον οποίο έχουν αποθηκευτεί τα δεδομένα. Για τις εποχικές χρονολογικές σειρές, το πρόβλημα είναι ακόμη πιο σύνθετο. Μέχρι τώρα, όσο αφορά την ταυτοποίηση υποδείγματος και την εκτίμηση, κάθε πλήρης περίοδος είναι περίπου σαν μία ξεχωριστή παρατήρηση. Αυτό σημαίνει ότι σε πέντε χρόνια μηνιαίων δεδομένων, θα έχουμε μόνο 5 παρατηρήσεις για τον Ιανουάριο, 5 για τον Φεβρουάριο, κτλ. Συνεπώς, ίσως να απαιτηθούν πολλές περίοδοι ιστορικών δεδομένων για να κατασκευαστεί ένα κατάλληλο εποχιακό υπόδειγμα. Ωστόσο, πρέπει να επισημάνουμε ότι εξαιτίας ανεπαρκών ιστορικών δεδομένων , κάποιος μπορεί να αναλύσει τα δεδομένα που είναι διαθέσιμα για να αποκτήσει ένα προκαταρκτικό υπόδειγμα. Αυτό το υπόδειγμα μπορεί στη συνέχεια να επαναλαμβάνεται καθώς όλο και πιο πρόσφατα δεδομένα θα είναι διαθέσιμα.

Παράλληλα, ένα μειονέκτημα των υποδειγμάτων ARIMA είναι ότι δεν υπάρχει, προς το παρόν, ένας βολικός τρόπος για να ενημερώσουμε ή να τροποποιήσουμε τις εκτιμήσεις των παραμέτρων του υποδείγματος κάθε φορά που μια νέα παρατήρηση θα γίνεται διαθέσιμη ανάλογο με αυτό που υπάρχει στη άμεση εξομάλυνση. Κάποιος θα πρέπει, περιοδικά, να κατασκευάζει το υπόδειγμα από την αρχή. Πιθανώς, ακόμη πιο σοβαρό μειονέκτημα είναι ότι θα πρέπει να υποθέσουμε ότι η εξέλιξη των χρονοσειρών θα είναι ίδια με αυτήν του παρελθόντος, γεγονός που σημαίνει ότι η μορφή του υποδείγματος δε θα αλλάξει μέσα στο χρόνο. Δεν υπάρχουν συστηματικοί μέθοδοι για την παρακολούθηση της απόδοσης του υποδείγματος ή της αυτόματης προσαρμογής των παραμέτρων στις αλλαγές που συντελούνται στην εν λόγω διαδικασία.

Ταυτόχρονα, ένα βασικό μειονέκτημα των ARIMA υποδειγμάτων είναι η «επένδυση» σε χρόνο αλλά και άλλες πηγές που απαιτούντια για την κατασκευή ενός ικανοποιητικού υποδείγματος. Αυτό μπορεί να μην είναι σημαντικό όταν μελετούνται λίγες χρονοσειρές. Εντούτοις, μια τυπική εφαρμογή πρόβλεψης στα συστήματα παραγωγής – αποθήκης μπορεί να περιλαμβάνει από μερικές εκατοντάδες έως αρκετές χιλιάδες διαφορετικές χρονοσειρές. Είναι αμφίβολο αν οι βελτιώσεις στην ακρίβεια των προβλέψεων – πιθανότατα μέσω της μεθοδολογίας ARIMA – θα μπορούσαν να δικαιολογήσουν το κόστος της κατασκευής υποδειγμάτων σε μια τέτοια περίπτωση.

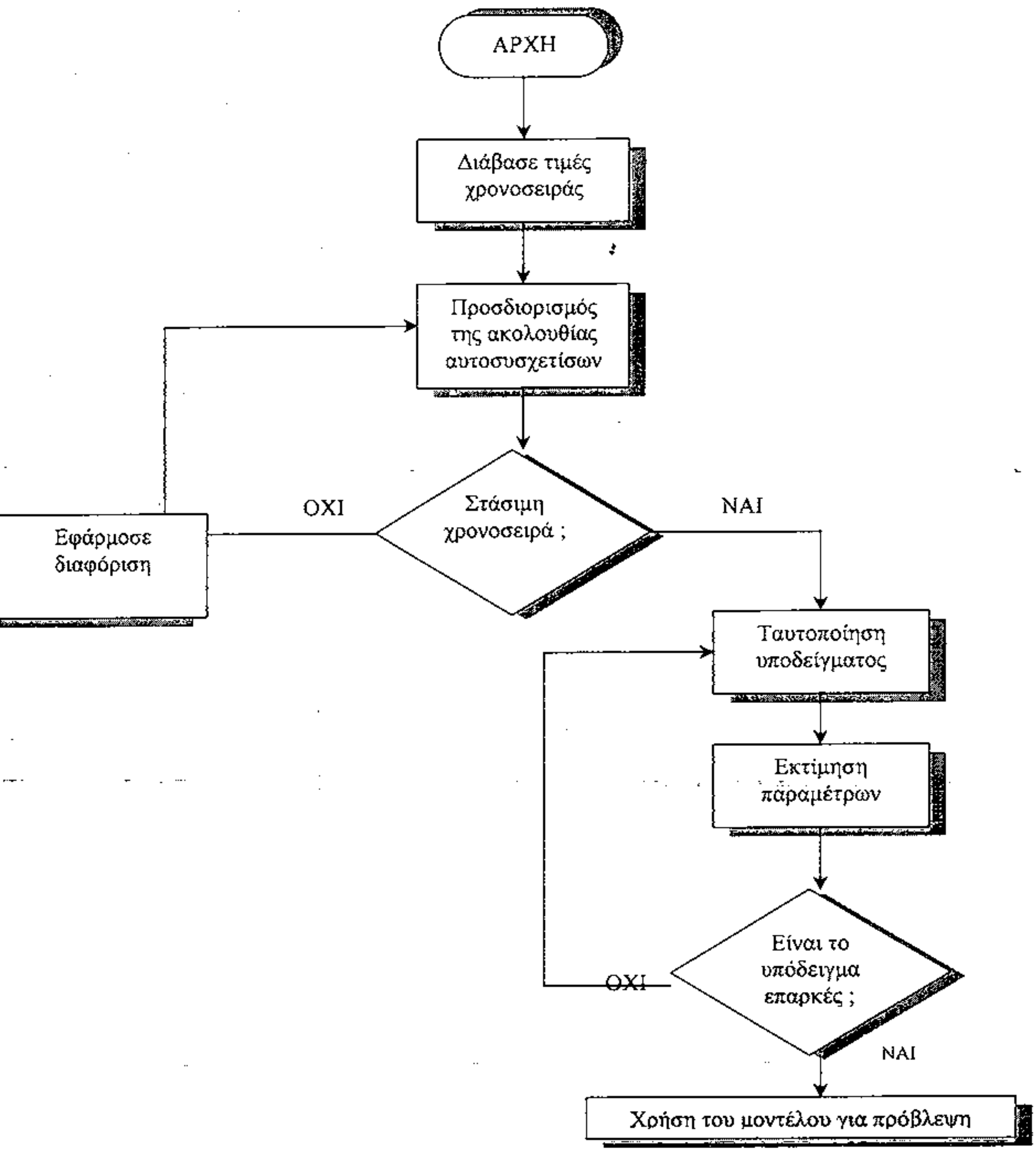
Ανεξαρτήτως αυτών των μειονεκτημάτων , τα υποδείγματα ARIMA είναι πιθανώς η πιο ακριβής ομάδα μοντέλων πρόβλεψης που είναι διαθέσιμη σήμερα. Είναι ιδιαίτερα κατάλληλα για χρονοσειρές στις οποίες το διάστημα του δείγματος είναι πολύ μικρό αφού σχετικά μεγάλη ποσότητα ιστορικών δεδομένων μπορεί να ανακτηθεί εύκολα. Για αυτόν το λόγο, έχουν εφαρμοστεί ευρέως σε χρονολογικές σειρές όπου έχουν ενδιαφέρον ωριαίες , καθημερινές ή εβδομαδιαίες παρατηρήσεις. Τα υποδείγματα ARIMA προτιμούνται ιδιαίτερα σε καταστάσεις όπου ένας μικρός αριθμός χρονοσειρών είναι κάποιου ενδιαφέροντος και η διοίκηση είναι πρόθυμη να δαπανήσει όλους εκείνους τους πόρους που είναι απαραίτητοι για να επιτύχει υψηλό βαθμό ακρίβειας πρόβλεψης.

6. Ανάλυση και σχεδιασμός συστημάτων για πρόβλεψη

1. Βήματα ανάπτυξης Πληροφοριακών συστημάτων

Το Πληροφοριακού Συστήματος διαχείρισης Οικονομικού Χαρτοφυλακίου με έμφαση στη σχεδίαση ευφυών αλγορίθμων για εκτίμηση και πρόβλεψη για να αναπτυχθεί έχει εισέλθει ή θα εισέλθει σε μια σειρά από φάσεις που είναι γνωστές ως «Βήματα Ανάπτυξης Συστημάτων».

Στην προκαταρκτική φάση αρκούσε να καθορίσουμε τις παραμέτρους του προβλήματος . Στη φάση της ανάλυση του υπάρχοντος συστήματος προσδιορίζουμε τις απαιτήσεις του συστήματος. Αυτές μπορεί να προέλθουν με τη χρήση διάφορων μεθόδων όπως είναι το διάγραμμα ροής δεδομένων. Ένα τέτοιο απεικονίζει τον αλγόριθμο των ερευνητών Box – Jenkins για την πρόβλεψης χρονολογικών σειρών.



Ακολούθως, στο βήμα του σχεδιασμού του νέου πλέον συστήματος καθορίζουμε τις αναλυτικές του προδιαγραφές σε επίπεδο software. Το μέσο που προτείνεται για την υλοποίηση των αλγορίθμων εκτίμησης και πρόβλεψης είναι η γλώσσα Matlab έκδοση 5.2 .Ο ισχυρός συνδυασμός μεταξύ μήτρων, προγραμματιστικών χαρακτηριστικών, γραφικών και εργαλείων ανάπτυξης εφαρμογών δικαιολογεί την παραπάνω επιλογή . Το Matlab ενσωματώνει στην υπολογιστική του διαδικασία τη σύντομη δημιουργία και απεικόνιση γραφικών και τον προγραμματισμό σε ένα ευέλικτο και ανοικτό υπολογιστικό περιβάλλον.

Στη Matlab υλοποιήθηκε το παρόν σύστημα , εκτελούνται δοκιμές και παρατηρούμε κατά πόσο το σύστημα ανταποκρίνεται στις προδιαγραφές του σχεδιασμού του . Το προσωπικό που θα κάνει χρήση του νέου συστήματος χρειάζεται ένα εγχειρίδιο που θα περιέχει μηχανισμούς παροχής βοήθειας.

Στην τελική φάση, στην εκμετάλλευση του συστήματος θα πρέπει η ομάδα εργασίας να υποβάλλει προτάσεις αναπροσαρμογής και να επεμβαίνει προληπτικά για να συντηρεί το σύστημα. Η αξιολόγησή του θα συντελείται με συνεργασία του προσωπικού που το χρησιμοποιεί.

2. Ο αλγόριθμος Box - Jenkins

ΤΟ ΠΡΟΒΛΗΜΑ

Να γίνει πρόγραμμα στο οποίο ο χρήστης θα θέλει να δίνει τιμές μιας χρονοσειράς και στη συνέχεια να βλέπει τη γραφική παράσταση της σειράς και μετά να παίρνει την πρόβλεψη για ένα χρονικό διάστημα που να το προσδιορίζει εκείνος.

ΣΤΟΧΟΣ : Μέθοδος κατασκευής υποδείγματος για προβλέψεις

1. Διάβασμα t στοιχείων t = 1, 2 ,3 … (βλέπε βρόχος σελίδα 40 Ανάπτυξη εφαρμογών)

$$\boxed{r_\kappa = \frac{\sum_{1}^{n-\kappa}(X_t - \overline{X}_t)(X_{t-\kappa} - \overline{X}_t)}{\sum_{1}^{n}(X_t - \overline{X}_t)^2}}$$

2. Προσδιορισμός της ακολουθίας συντελεστών αυτοσυσχέτισης

   1. Εμφάνισή τους στον πίνακα Α
   2. Προσδιορισμός του μέσου όρου των τιμών $\overline{X}_t$
   3. Προσδιορισμός της διασποράς του x
   4. Δημιουργία ενός νέου πίνακα που θα έχει για κάθε στοιχείο τη διαφορά $x_t - \overline{X}_t$
   5. Δημιουργία ανάστροφου πίνακα Β
   6. Πολλαπλασιασμός των ανάστροφων πινάκων όπου θα δημιουργηθεί ένας t x t πίνακας με το όνομα Rκ
   7. Να διαιρεθεί το κάθε στοιχείο του πίνακα με τη διασπορά της διαδικασίας
   8. Να εμφανίζονται τα στοιχεία της 1ης γραμμής του Rκ
   9. Να βρεθεί η διασπορά του συντελεστή συσχέτισης $\sigma^2 = \frac{1}{N}$ και η τυπική απόκλιση $\sigma = \sqrt{\frac{1}{N}}$
   10. Το τυπικό σφάλμα θα είναι se = 2 * s
   11. Αν $| r_\kappa | <$ se τότε το r_κ θα είναι 0 δεν είναι σημαντικά διάφορο του 0.

12. Αν $|r_κ| = 0$ ύστερα από το 1 ή το 2 κ τότε η χρονοσειρά θα καλείται στάσιμη

13. Αν $|r_κ| <> 0$ ύστερα από το 1 ή το 2 κ τότε η χρονοσειρά θα καλείται μη στάσιμη και χρειάζεται διαφόριση

$$\nabla x_t = (1 - B) x_t = x_t - x_{t-1}$$

και ξανακάνουμε την ίδια διαδικασία (βήματα 1 έως 11)

3. Προσδιορισμός τάξης p και q

   $p + 1 = κ$ & $q + 1 = κ$ όπου κ η χρονική υστέρηση στην οποία η ακολουθία των συντελεστών αυτοσυσχέτισης ξεκινά να μηδενίζεται

4. Προσδιορισμός θεωρητικής μορφή μοντέλου

$$(1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_p B^p)(1 - B)^d z_n = (1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q) w_n$$

ARIMA (p , d , q)

Πίνακας Αποφάσεων

$\bar{3}\,\bar{3}\,\bar{3}$

$$(1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_p B^p) (1 - B)^d z_n = (1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q) w_n$$

Σενάρια																											
p	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2
d	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2
q	0	0	0	1	1	1	2	2	2	0	0	0	1	1	1	2	2	2	0	0	0	1	1	1	2	2	2
Υπόδειγμα	Δ	A	b	c	d	e	f	g	h	i	j	k	l	M	n	o	p	q	r	s	t	u	v	w	x	y	z

Υποδείγματα

a) $(1 - B)z_n = w_n$

b) $(1 - B)^2 z_n = w_n$

c) $z_n = (1 + \theta_1 B) w_n$

d) $(1 - B)z_n = (1 + \theta_1 B) w_n$

e) $(1 - B)^2 z_n = (1 + \theta_1 B) w_n$

f) $z_n = (1 + \theta_1 B + \theta_2 B^2) w_n$

g) $(1 - B) z_n = (1 + \theta_1 B + \theta_2 B^2)$

h) $(1 - B)^2 z_n = (1 + \theta_1 B + \theta_2 B^2)$

i) $(1 - \varphi_1 B)z_n = w_n$

j) $(1- \varphi_1 B)(1 - B)z_n = w_n$

k) $(1 - \varphi_1 B) (1 - B)^2 z_n = w_n$

l) $(1- \varphi_1 B) (1 - B)z_n = w_n$

m) $(1 - \varphi_1 B) (1 - B)z_n = (1 + \theta_1 B)w_n$

n) $(1- \varphi_1 B) (1- B)^2 z_n = (1 + \theta_1 B)w_n$

o) $(1 - \varphi_1 B)z_n = (1 + \theta_1 B+ \theta_2 B^2)w_n$

p) $(1- \varphi_1 B) (1 -B)z_n = (1 + \theta_1 B+ \theta_2 B^2)w_n$

q) $(1- \varphi_1 B) (1- B)^2 z_n = (1 + \theta_1 B+ \theta_2 B^2)w_n$

r) $(1 - \varphi_1 B - \varphi_2 B^2) z_n = w_n$

s) $(1 - \varphi_1 B - \varphi_2 B^2) (1 -B) z_n = w_n$

t) $(1 - \varphi_1 B - \varphi_2 B^2) (1 -B)^2 z_n = w_n$

u) $(1 - \varphi_1 B - \varphi_2 B^2) z_n = (1 + \theta_1 B) w_n$

v) $(1 - \varphi_1 B - \varphi_2 B^2) (1 - B) z_n = (1 + \theta_1 B) w_n$

w) $(1 - \varphi_1 B - \varphi_2 B^2) (1 - B)^2 z_n = (1 + \theta_1 B) w_n$

x) $(1 - \varphi_1 B - \varphi_2 B^2)\, z_n = (1 + \theta_1 B + \theta_2 B^2)\, w_n$

y) $(1 - \varphi_1 B - \varphi_2 B^2)(1 - B)\, z_n = (1 + \theta_1 B + \theta_2 B^2)\, w_n$

z) $(1 - \varphi_1 B - \varphi_2 B^2)(1 - B)^2\, z_n = (1 + \theta_1 B + \theta_2 B^2)\, w_n$

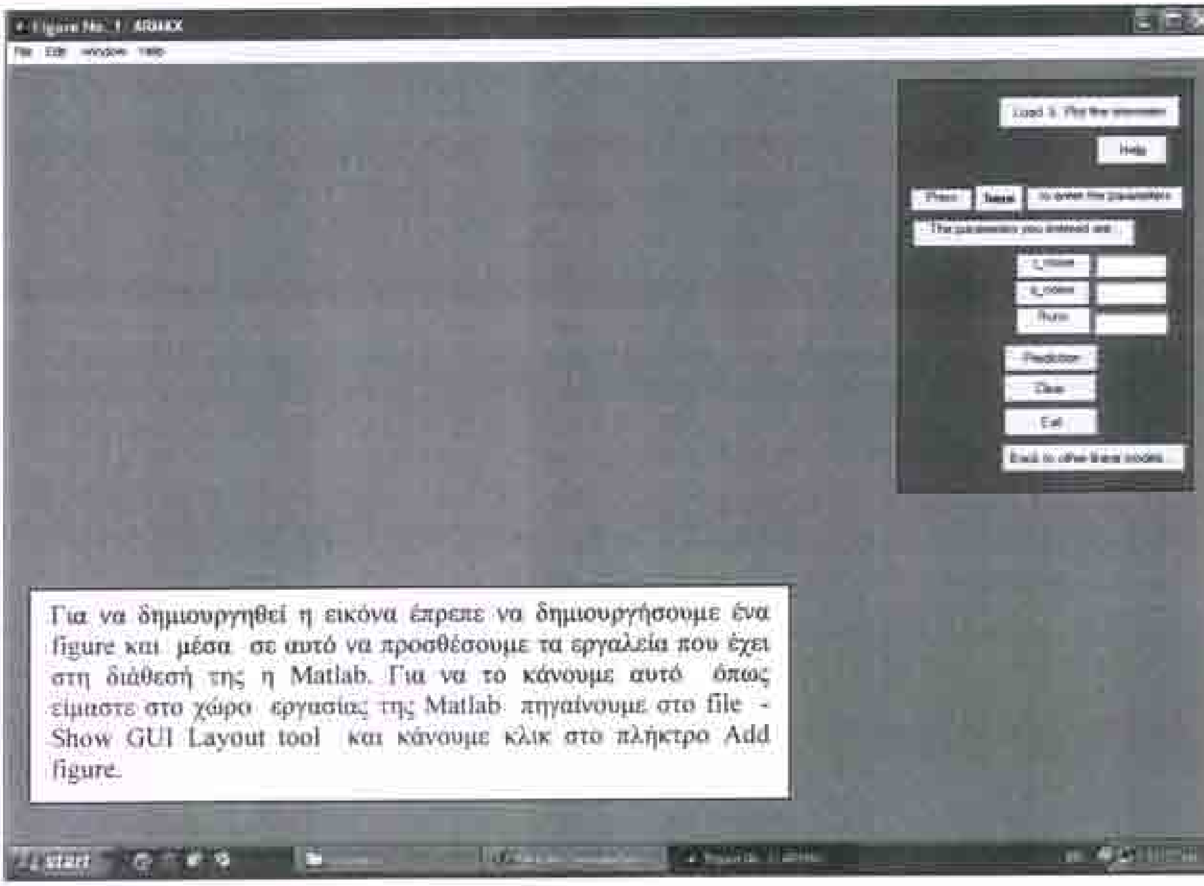
Σημείωση : Δ Δεν υπάρχει υπόδειγμα

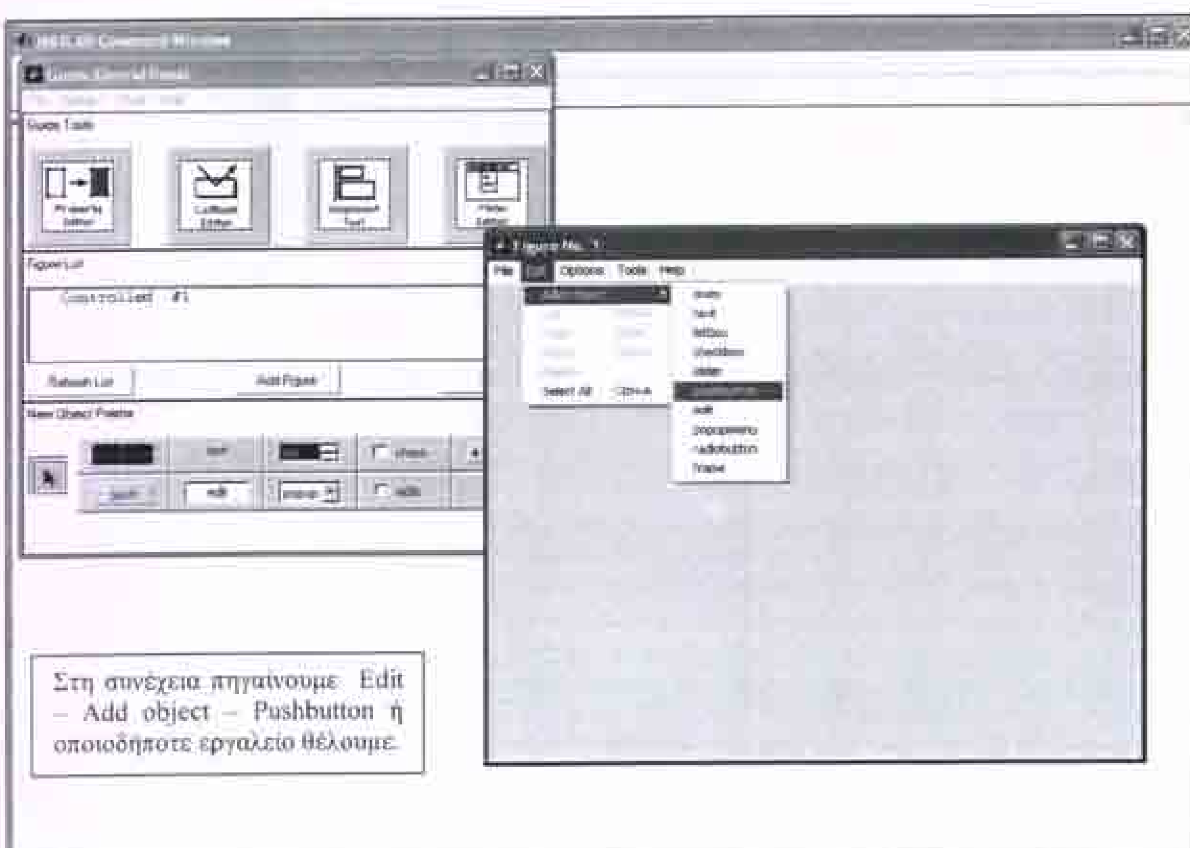
5. Παράμετροι φ_i και θ_i
6. Χρήση του μοντέλου για πρόβλεψη

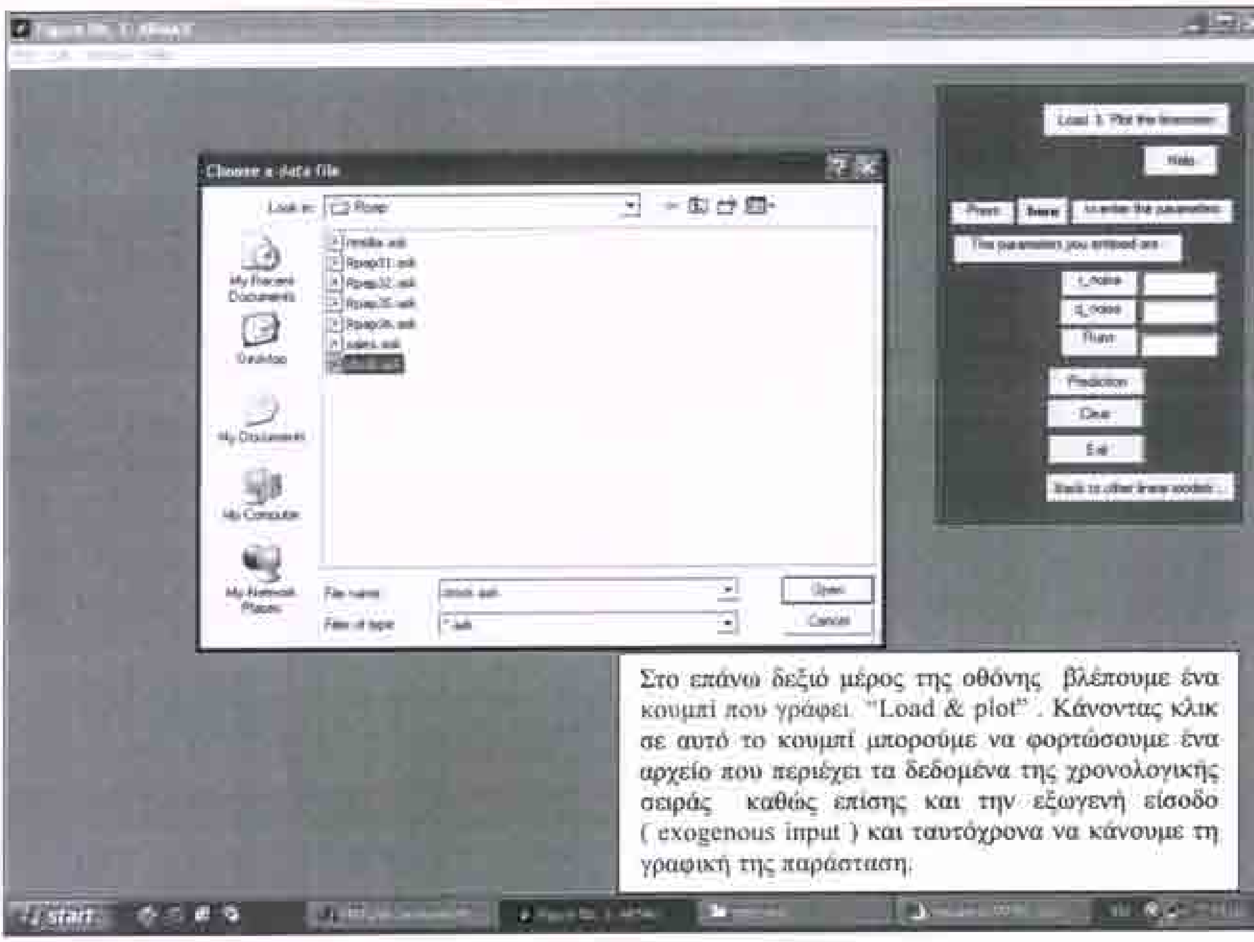
7. Δημιουργία του γραφικού περιβάλλοντος χρήστη

1. ARMAX.

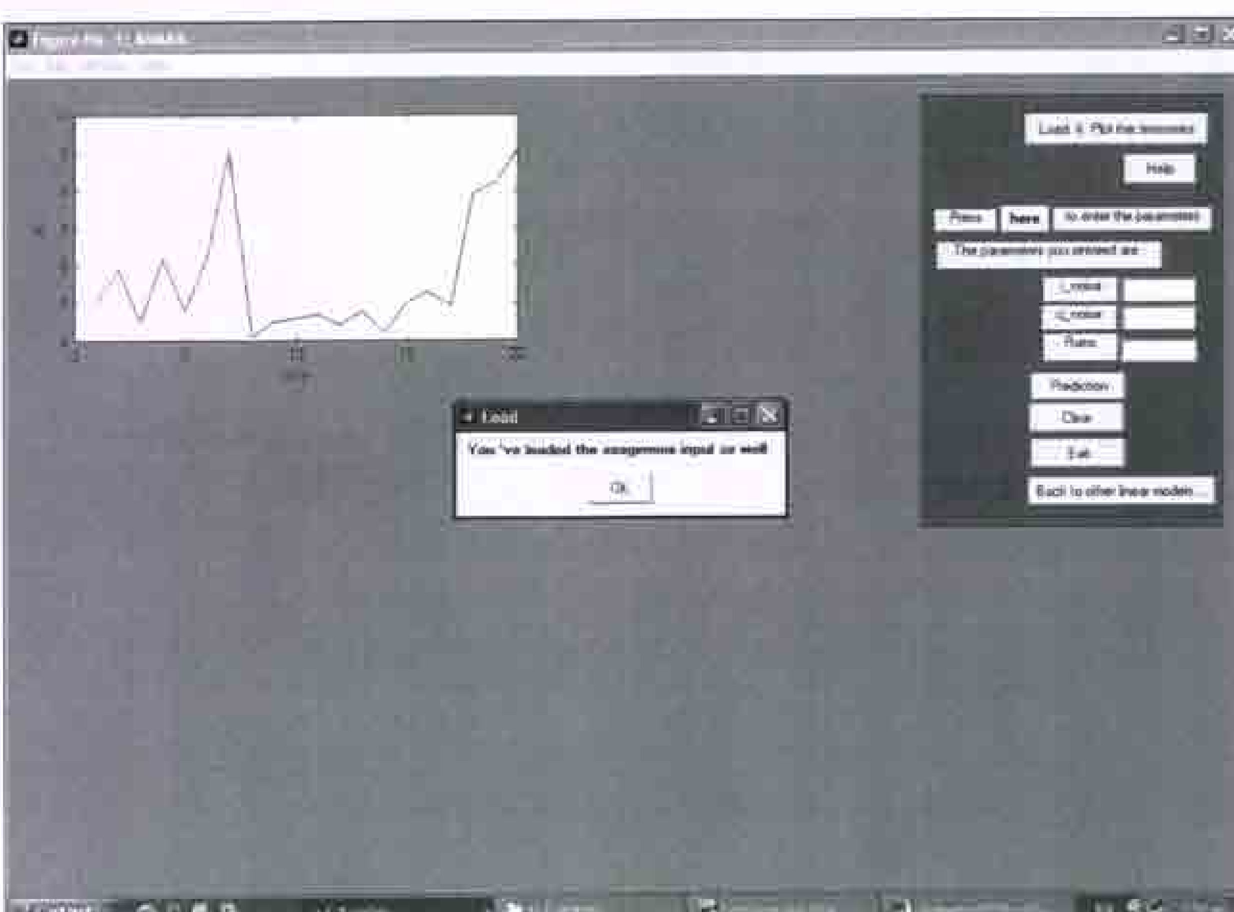
(ptixiaki)

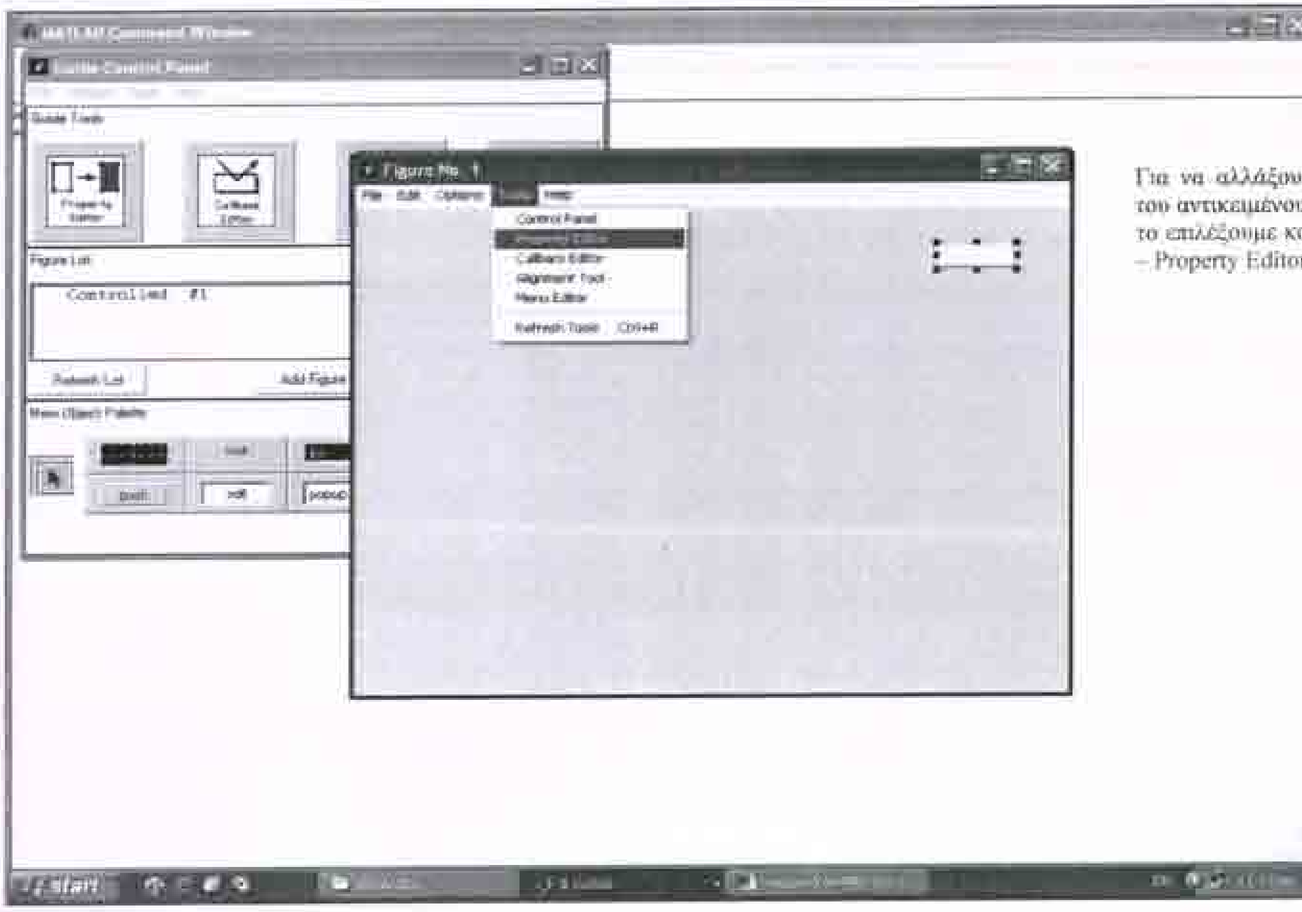






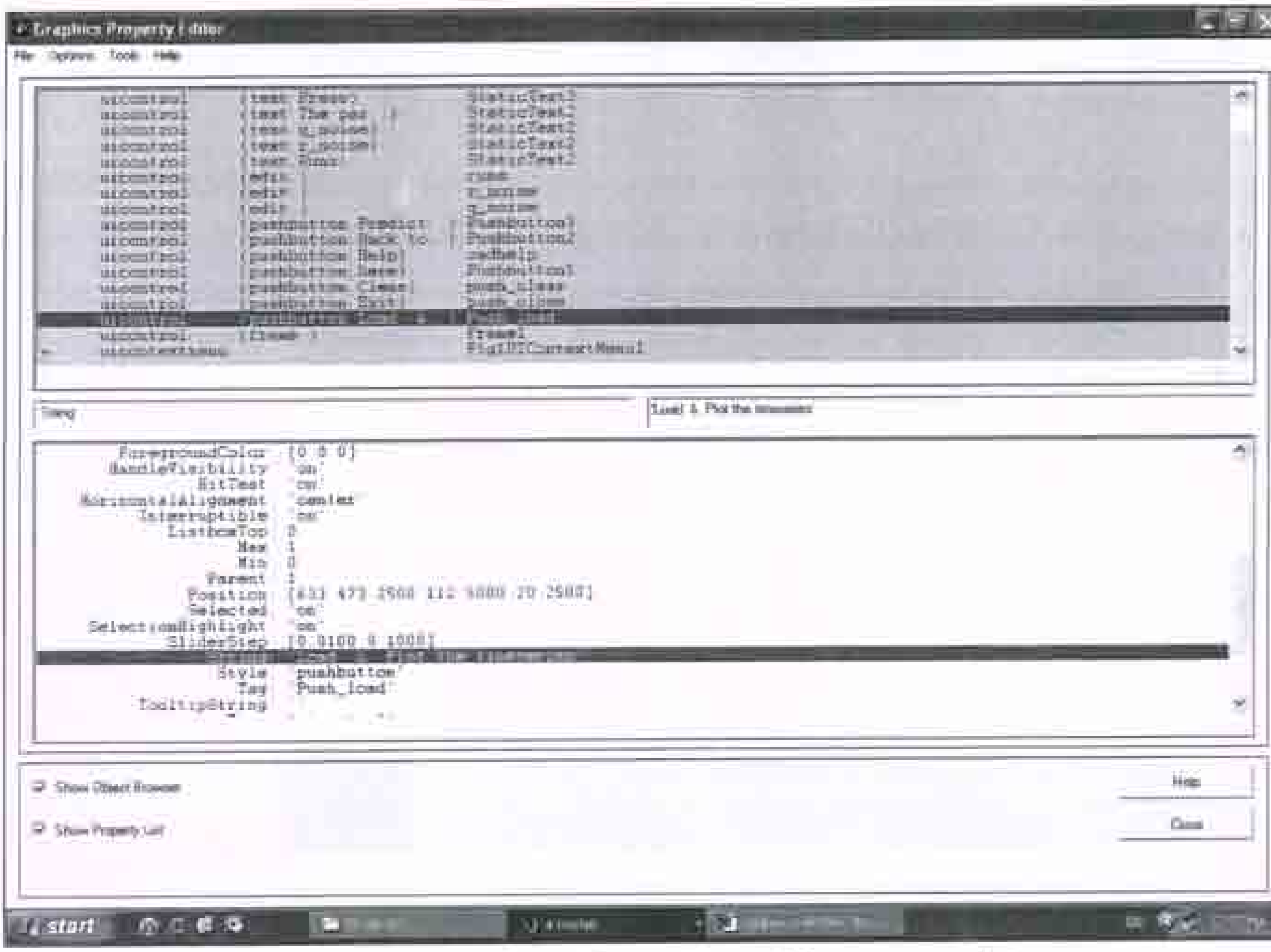
Στο επάνω δεξιό μέρος της οθόνης βλέπουμε ένα κουμπί που γράφει "Load & plot". Κάνοντας κλικ σε αυτό το κουμπί μπορούμε να φορτώσουμε ένα αρχείο που περιέχει τα δεδομένα της χρονολογικής σειράς καθώς επίσης και την εξωγενή είσοδο (exogenous input) και ταυτόχρονα να κάνουμε τη γραφική της παράσταση.



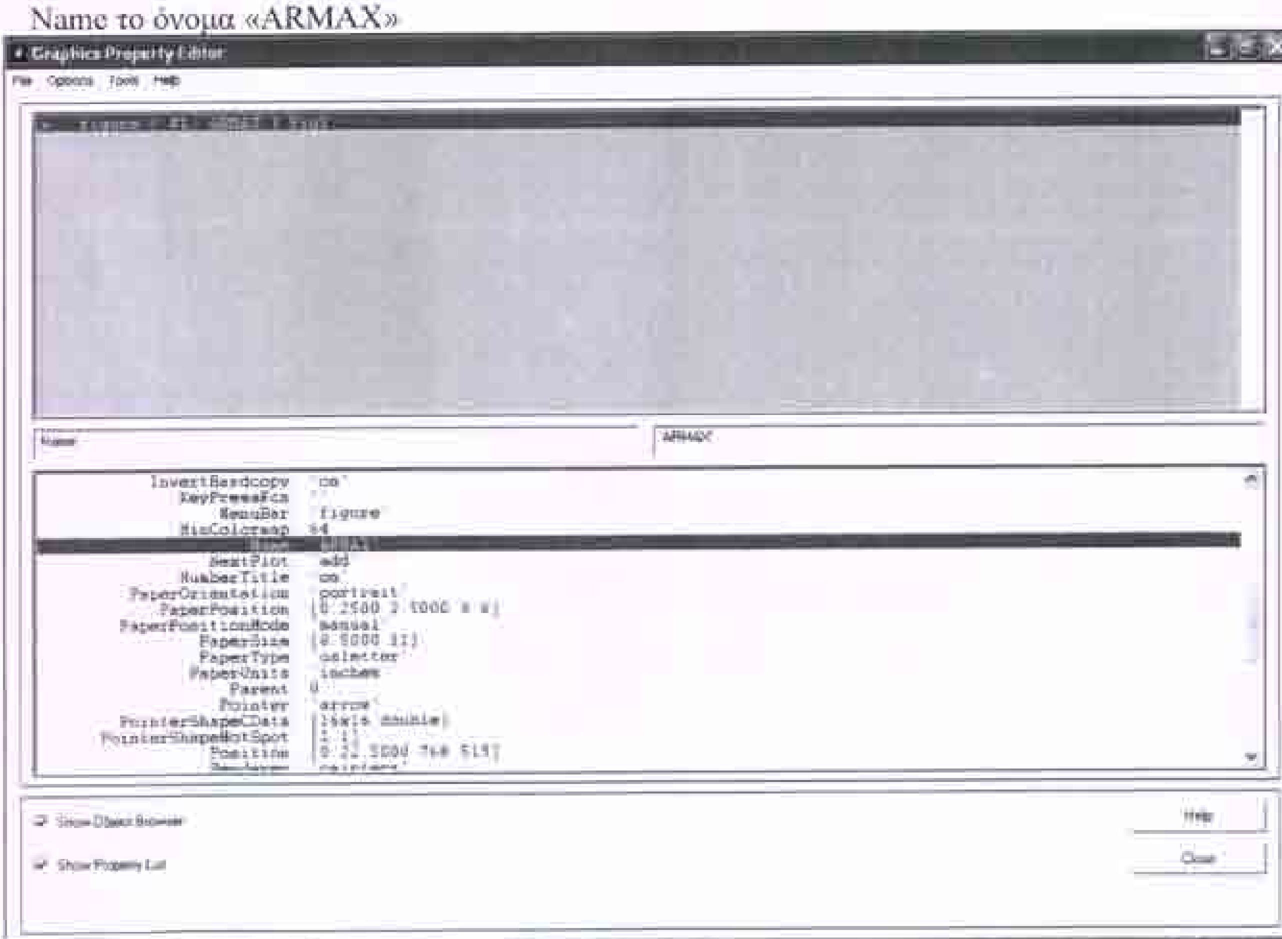


Για να αλλάξουμε τις ιδιότη του αντικειμένου αυτού αρκεί το επιλέξουμε και να πάμε Too – Property Editor

Οι ιδιότητες του είναι το χρώμα, ο τύπος της γραμματοσειράς ,το μέγεθος της γραμματοσειράς. Στην ιδιότητα 'string' θέτουμε την ετικέτα του κουμπιού και στην ιδιότητα 'tag' το όνομα του κουμπιού για να το ξεχωρίζουμε από τα υπόλοιπα όπως φαίνεται και από την ακόλουθη εικόνα.

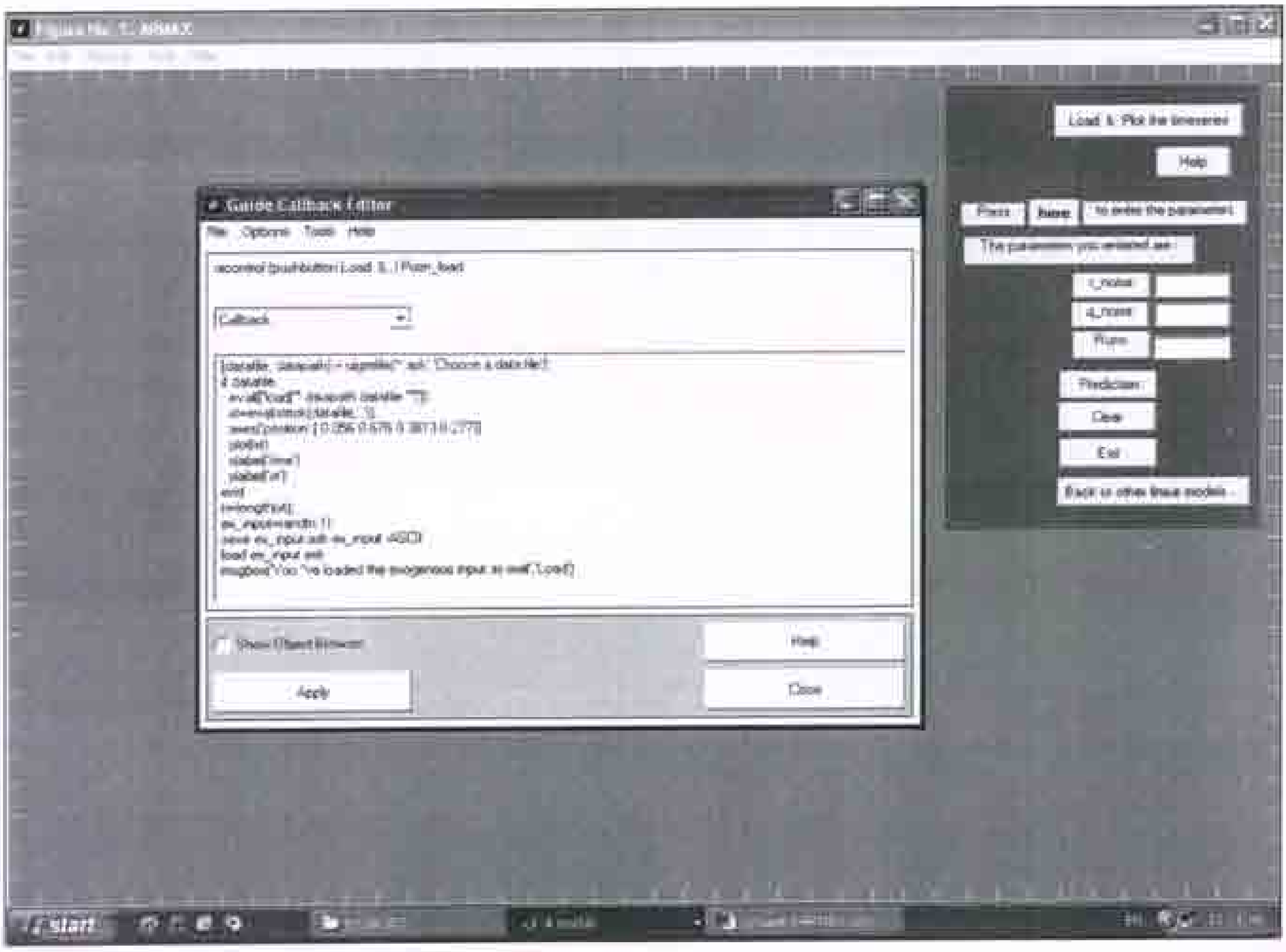


Ιδιότητες όμως έχουμε και για το ίδιο το figure . Σε αυτό δίνουμε στην ιδιότητα Name το όνομα «ARMAX»



Για να αποθηκεύσουμε τις αλλαγές αρκεί να πατήσουμε το κλείσιμο από το figure και να το αποθηκεύσουμε σε ένα αρχείο με την προέκταση .m Στην εργασία, το ARMAX είναι αποθηκευμένο στο αρχείο με το όνομα ptixiaki.m

Ο κώδικας που μας επιτρέπει να φορτώσουμε ένα αρχείο της επιλογής που είναι αποθηκευμένο σε κάποιο φάκελο του υπολογιστή μας και να κάνουμε τη γραφική του παράσταση προστίθεται αν πάμε Tools – Callback Editor



Η συνάρτηση που χρησιμοποιείται για να επιλέξουμε ένα αρχείο είναι η uigetfile η οποία εκχωρεί στη μεταβλητή datafile το όνομα του αρχείου και τη διαδρομή του στο δίσκο στη μεταβλητή datapath και μετά φορτώνει το αρχείο που βρ΄σικεται στη συγκεκριμένη διαδρομή . Στη συνέχεια του δίνει το όνομα xt και του ζητά στην θέση με τις συντεταγμένες [0.056 0.676 0.3613 0.277] να εμφανίσει το γράφημα της χρονοσειράς στους άξονες με την εντολή plot και να βάλει ετικέτες στον κάθε άξονα.

```
[datafile, datapath] = uigetfile('*.ask','Choose a data file');
if datafile,
 eval(['load('" datapath datafile '")']);
```

xt=eval(strtok(datafile,'.'));

axes('position',[0.056 0.676 0.3613 0.277]) {Οι συντεταγμένες βρέθηκαν κάνοντας Edit – Add object – Axes και τοποθετώντας τους άξονες στο σημείο που θέλαμε πήγαμε στις ιδιοτητές του Tools – Property Editor και είδαμε πως οι συντεταγμένες 0.056 0.676 0.3613 0.277 αντιστοιχούν στη θέση που βάλαμε τους άξονες. Μετά κρατήσαμε τις συντεταγμένες και σβήσαμε τους άξονες αφού με την εντολή axes(position,'[ ]) δημιουργούμε άξονες στη θέση που θέλουμε }

plot(xt)

xlabel('time')

ylabel('xt')

end

n=length(xt); *{Το πλήθος των παρατηρήσεων της χρονοσειράς εκχωρείται στη μεταβλητή n }*

ex_input=rand(n,1) *{Μια χρονοσειρά με τυχαίους αριθμούς που έχει για γραμμές το ίδιο πλήθος παρατηρήσεων n και μία στήλη εκχωρείται στη μεταβλητή με το όνομα ex_input }*

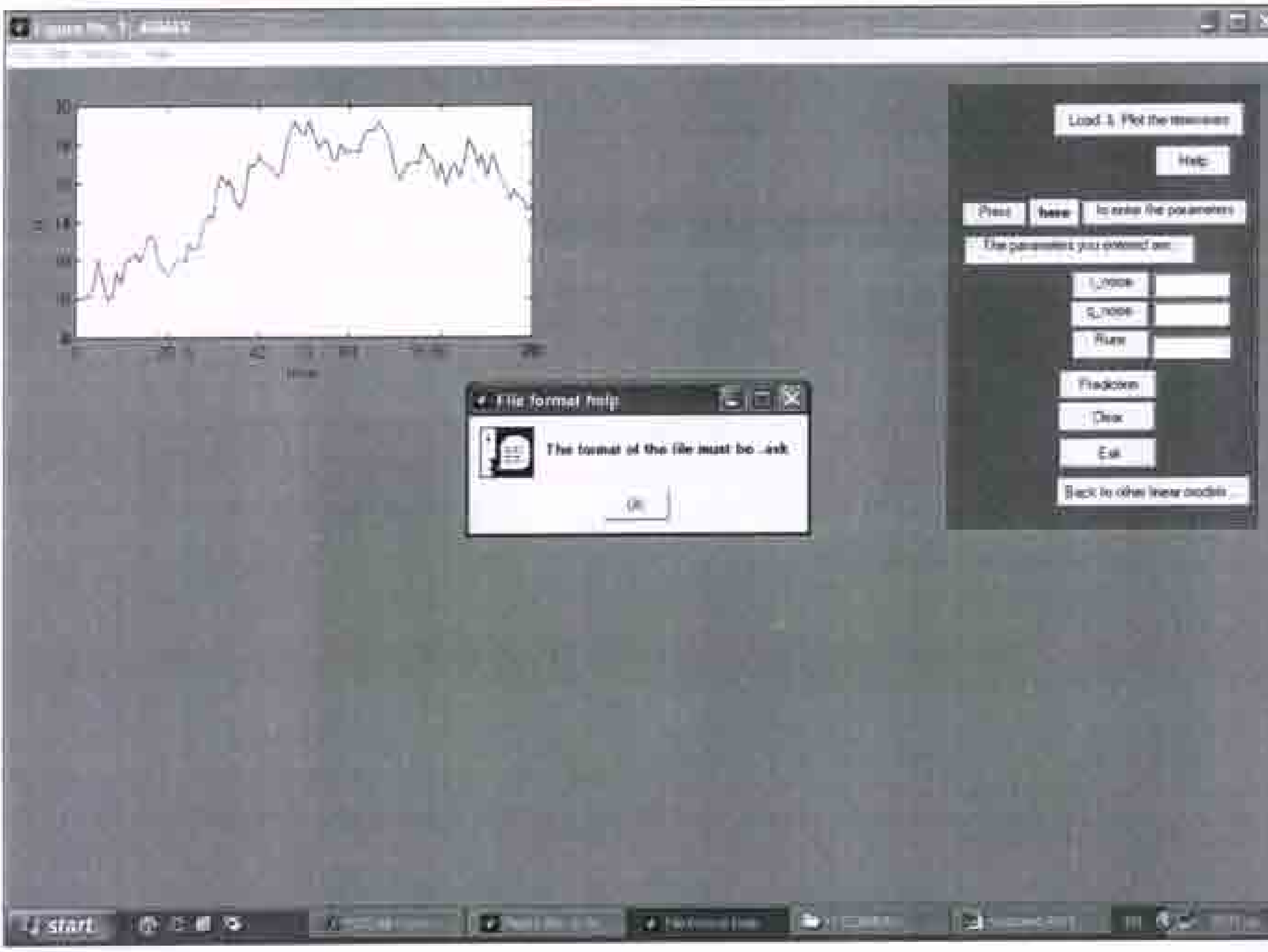
save ex_input.ask ex_input –ASCII *{Αποθηκεύουμε στο δίσκο, στο τρέχοντα φάκελο, την μεταβλητή ex_input ως ένα αρχείο με το όνομα ex_input και την προέκταση .ask . Η προέκταση αυτή σημαίνει πως το αρχείο περιέχει μόνο αριθμούς σαν απλούς χαρακτήρες ASCII . }*

load ex_input.ask *{ φορτώνουμε το αρχείο με το όνομα ex_input.ask }*

msgbox('You ''ve loaded the exogenous input as well','Load') *{ εμφανίζεται μήνυμα που πληροφορεί το χρήστη ότι παράλληλα με το αρχείο που περιέχει τη χρονοσειρά φόρτωσε και την εξωγενή είσοδο στο σύστημα }*

Το επόμενο κουμπί (βλέπε 1^{η} εικόνα) που γράφει Help εμφανίζει ένα μήνυμα βοηθητικού χαρακτήρα που πληροφορεί το χρήστη ότι τα αρχεία που μπορεί να φορτώσει είναι εκείνα με την προέκταση .ask.

(Proektasi)



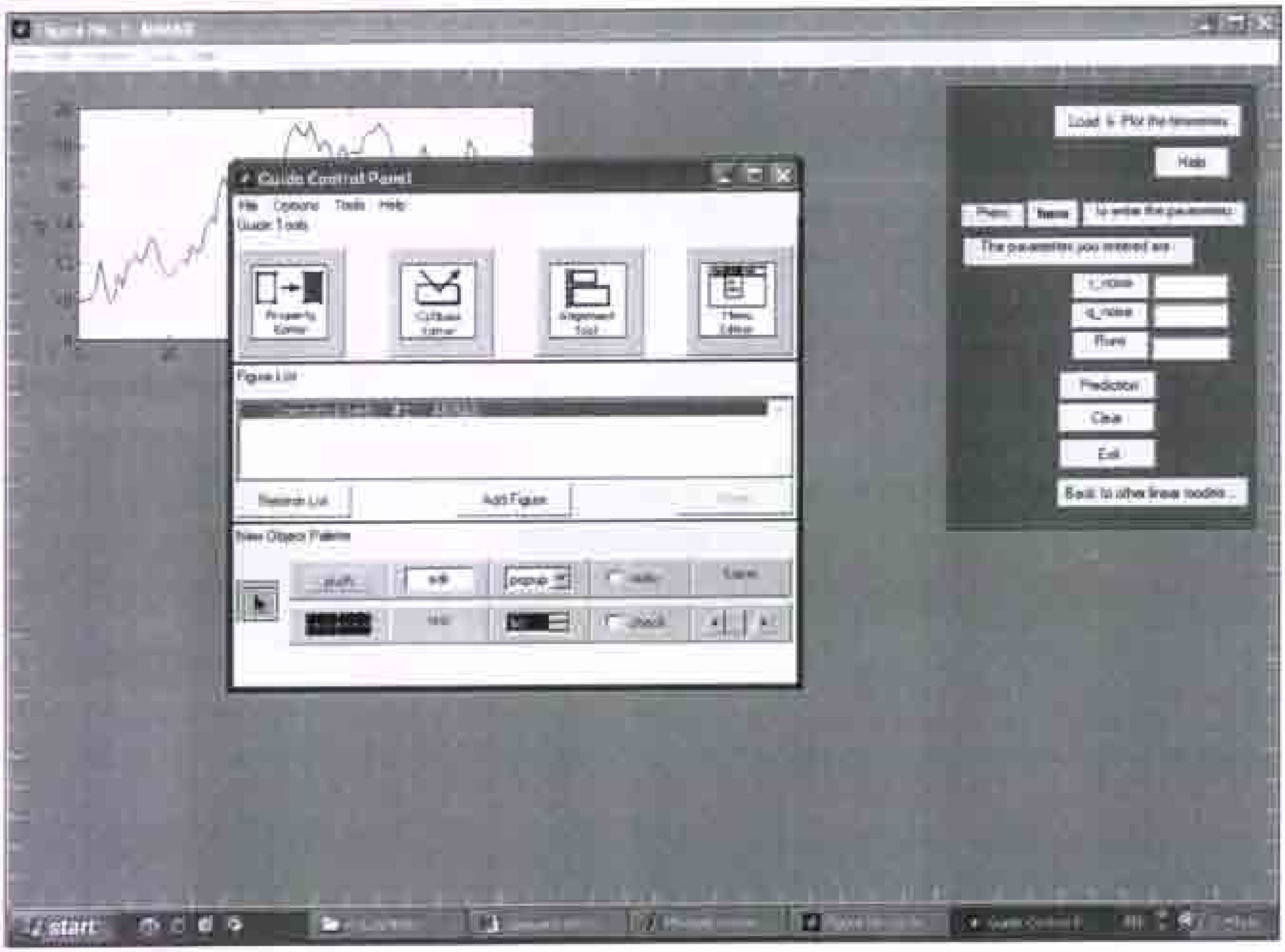
Να σημειωθεί πως για να δημιουργήσουμε ένα τέτοιο αρχείο θα πρέπει να πάμε στα Windows, Εργαλεία – Επιλογές Φακέλων – στην καρτέλα View και να απενεργοποιήσουμε την επιλογή «Απόκρυψη προεκτάσεων για γνωστούς τύπους αρχείων». Αφού κάνουμε αυτό, παρατηρούμε ότι εμφανίζονται η προέκταση σε όλα τα αχρεία που είναι αποθηκευμένα στον υπολογιστή μας. Οπότε, αρκεί να δημιουργήσουμε ένα έγγραφο κειμένου και να το μετονομάσουμε προσέχοντας όμως το όνομα που θα δώσουμε να τελειώνει σε .ask.

Για να εμφανιστεί το μήνυμα θα πρέπει να επιλέξουμε το κουμπί και να πάμε Tools – Callback Editor και μέσα να γράψουμε την ακόλουθη εντολή :

h=helpdlg('The format of the file must be .ask ','File format help')

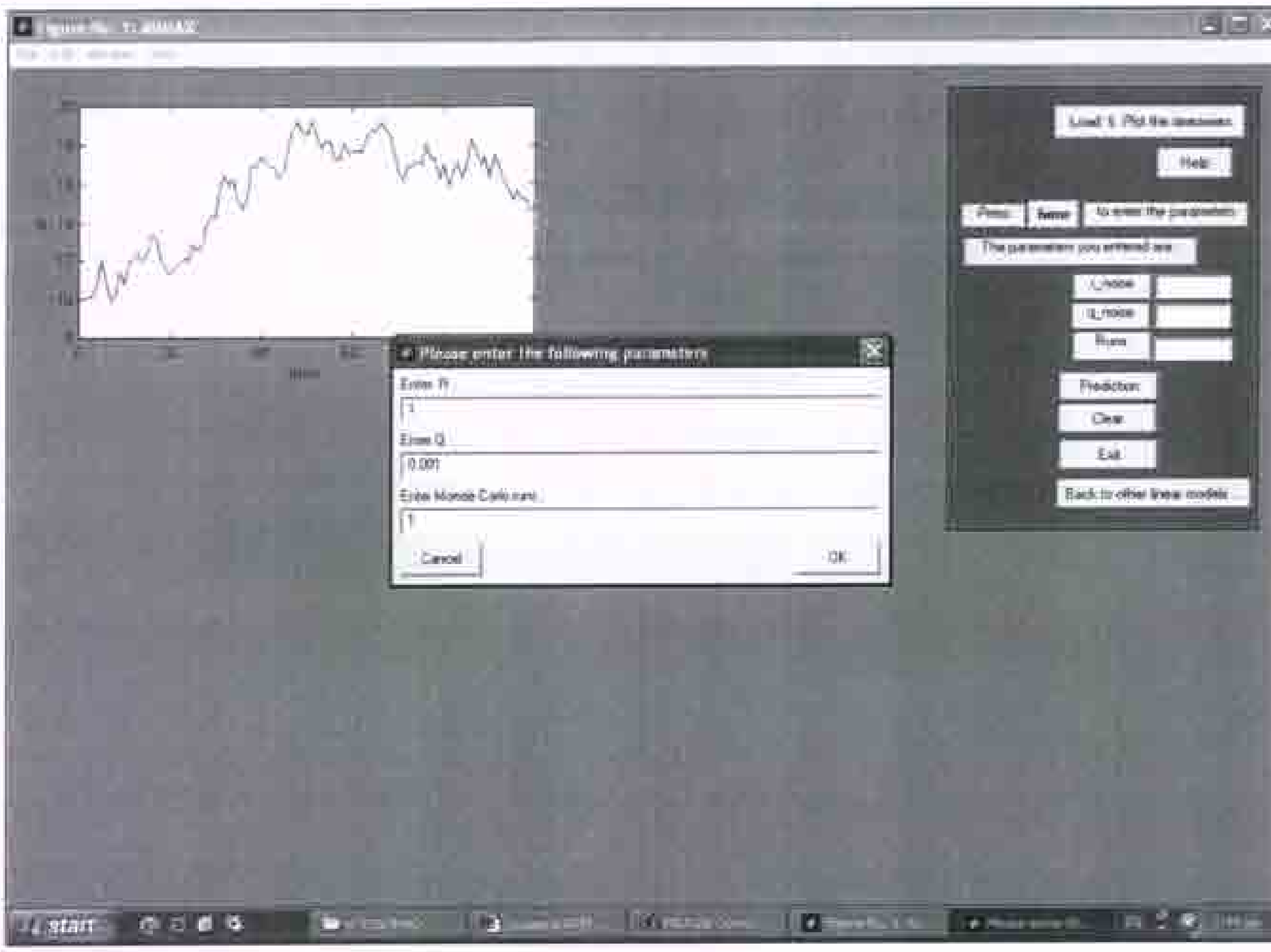
Να τονιστεί πως για να κάνουμε επεμβάσεις σε ένα figure θα πρέπει να πάμε στη γραμμή εντολών της Matlab και μετά File – Show GUI Layout tool και μετά στο παράθυρο που θα εμφανιστεί , Guide control Panel , να κάνουμε διπλό κλικ στο figure που βρίσκεται στη λίστα προκειμένου από Active να μετατραπεί σε Controlled

(energo parathiro)



Ακολούθως, κάνουμε κλικ στο κουμπί **here** για να εισάγουμε τις παραμέτρους του συστήματος . Πριν εξηγήσουμε τον κώδικα απλώς να σημειώσουμε πως οι ετικέτες που γράφουν Press, to enter the parameters, The parameters you entered are : , r_noise, q_noise και runs δημιουργήθηκαν από την κατάσταση Controlled πηγαίνοντας Edit – Add object – Text και στη συνέχεια στην ιδιότητα "String" γράφουμε τα παραπάνω.

Κάνοντας κλικ στο here εμφανίζεται το παρακάτω παράθυρο το οποίο ζητά από το χρήστη να εισαγάγει κάποιες τιμές στις μεταβλητές r_noise, q_noise και runs *(parametroi)*



Ο κώδικας για το κουμπί **here**

prompstr={'Enter R : ','Enter Q : ','Enter Monde Carlo runs : '};

inistr={'1','0.001','1'}; { Αρχικές τιμές που ορίζονται κατά τη σχεδίαση}

titlestr='Please enter the following parameters'; { Ο τίτλος του παραθύρου }

nlines=1; { Το πλήθος των γραμμών για κάθε τιμή }

result=inputdlg(prompstr,titlestr,nlines,inistr); { Οι τιμές αποθηκεύονται σε μια μεταβλητή με το όνομα result και αποθηκεύονται ως συμβολοσειρές (string) }

r_noise_c=[result{1}] { Το πρώτο κελί της μεταβλητής εκχωρείται σε μεταβλητή τύπου string με το όνομα r_noise_c }

q_noise_c=[result{2}] { Ομοίως για το 2° κελί } runs_c=[result{3}] { Ομοίως για το 3° κελί }

r_noise = str2num(r_noise_c) {Με τη str2num μετατρέπουμε τη μεταβλητή r_noise_c που είναι string σε αριθμητικού τύπου (number) και της δίνουμε πλέον ντο όνομα r_noise }

q_noise = str2num(q_noise_c) {Ομοίως }

runs= str2num(runs_c) {Ομοίως }

Άρα μέχρι στιγμής έχουμε εκχωρήσει τιμές αριθμητικές στις μεταβλητές r_noise, q_noise και runs. Η μετατροπή από συμβολοσειρά σε αριθμητικό τύπο ήταν απαραίτητη διότι μας ενδιαφέρει να επεξεργαστούμε τις μεταβλητές ως αριθμούς για κάνουμε τους διάφορους υπολογισμούς.

Οι τιμές που εκχωρήθηκαν όμως υπόκεινται σε περιορισμούς. Δηλαδή η τιμή του r_noise πρέπει να είναι μικρότερη του 1. Για αυτό το λόγο υλοποιήθηκε και ένας βρόγχος

while r_noise >1, { όσο η τιμή είναι μεγαλύτερη του 1 κάνε }

waitfor(msgbox({'Sorry, the r you entered should not be greater than 1 '})) { εμφάνισε το μήνυμα «Sorry ...» }

{Στη συνέχεια εμφανίζεται πάλι το παράθυρο εισαγωγής παραμέτρων }

```
prompstr={'Enter R : ','Enter Q : ','Enter Monde Carlo runs : '};
inistr={'1','0.001','1'};
titlestr='Please enter the following parameters'
   nlines=1;
   result=inputdlg(prompstr,titlestr,nlines,inistr);
   r_noise_c=[result{1}]
   q_noise_c=[result{2}]
   runs_c=[result{3}]
   r_noise = str2num(r_noise_c)
   q_noise = str2num(q_noise_c)
   runs= str2num(runs_c)
 end
```

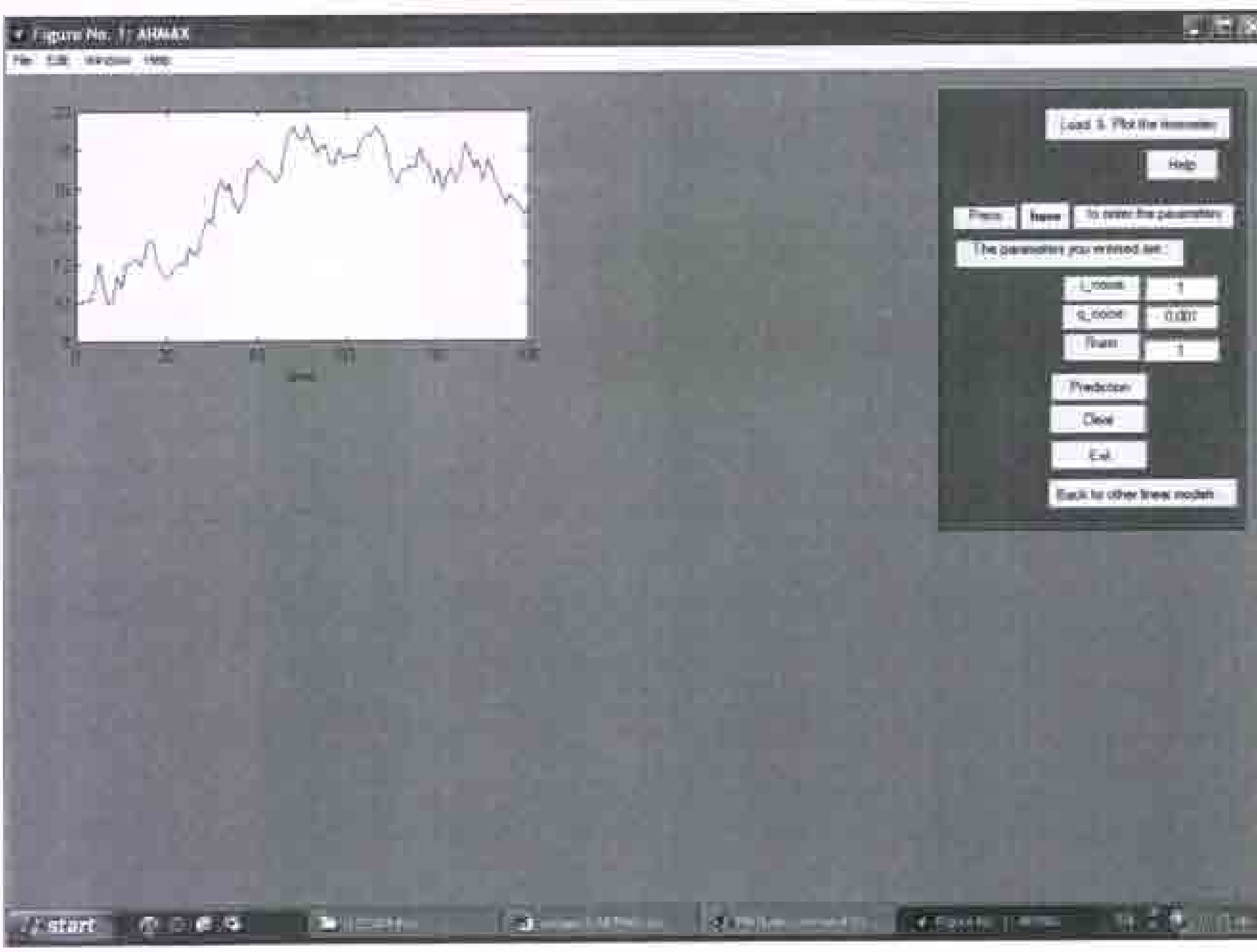
Όσο ισχύει η συνθήκη r_noise >1 θα εξακολουθεί να εμφανίζεται το παράθυρο της εισαγωγής των παραμέτρων.

Όταν το r_noise <1 τότε θα μπορούμε να συνεχίσουμε εμφανίζοντας τις τιμές που εκχώρησε ο χρήστης στα πλαίσια κειμένου (Edit – Add object – edit). Για να το πετύχουμε αυτό χρησιμοποιούμε την εντολή findobj και set

h = findobj('tag','r_noise') ; {Βρίσκει το αντικείμενο που έχει στην ιδιότητα tag το όνομα r_noise

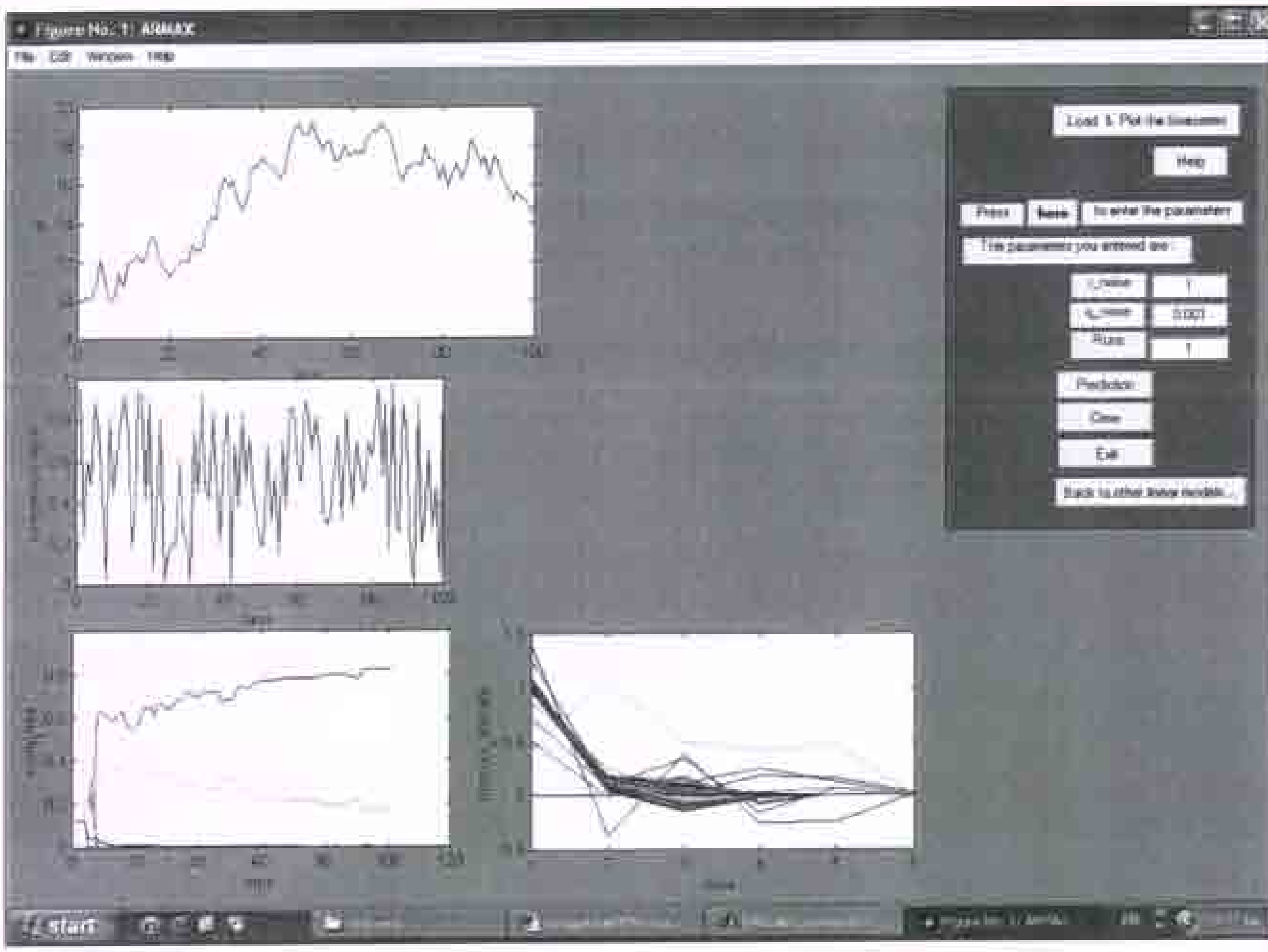
set(h,'string',r_noise); {Στο ίδιο αντικείμενο, του ζητάμε στην ιδιότητα string να προσθέσει τη μεταβλητή r_noise

Με τον ίδιο τρόπο ενεργούμε και για τα άλλα πλαίσια κειμένου για τις μεταβλητές q_noise και runs.



Στη συνέχεια, στο κουμπί Prediction ζητάμε να εκτελεστεί ο αλγόριθμος εκτίμησης και πρόβλεψης τα αποτελέσματα του οποίου εμφανίζονται στην παρακάτω εικόνα.

(apotelemata)



```
prompstr={'From t = :', 'Up to  t = :'};
titlestr= 'Enter prediction period ';
nlines=1;
result=inputdlg(prompstr,titlestr,nlines);
start_t=[result{1}]
end_t=[result{2}]
start_t_n = str2num(start_t)
end_t_n = str2num(end_t)
eval  petrosbat2
```

{ Εκτελούμε το αρχείο που περιέχει τον αλγόριθμο επιλογής του kalman και έχει το όνομα petrosbat2. Θα πρέπει να είναι όμως στο ίδιο φάκελο που είναι αποθηκευμένο και το figure ARMAX } load petrosout.mat { φορτώνουμε τα αποτελέσματα του αλγορίθμου που είναι αποθηκευμένα σε ένα αρχείο με το όνομα petrosout.mat . Τα αρχεία με την προέκταση .mat περιέχουν τις μεταβλητές }

axes('position',[0.056 0.676 0.3613 0.277]) { βάζουμε άξονες στις συντεταγμένες που έχουμε ορίσει έτσι ώστε στους άξονες να εμφανίσουμε κάποια γραφήματα}

plot(xt) {κάνουμε το γράφημα της χρονοσειράς }

xlabel('time') { Στον οριζόντιο άξονα βάζουμε τη λέξη time }

ylabel('xt') { Στον κάθετο άξονα βάζουμε τη λέξη xt}

axes('position',[0.055 0.382 0.291 0.247])

plot(ex_input) {Ομοίως με το πρώτο axes }

xlabel('time')

ylabel('exogenous input')

axes('position',[0.052 0.069 0.300 0.260])

plot(monp_thita') {κάνουμε το γράφημα της μεταβλητής monp_thita' (ανάστροφος) η οποία περιλαμβάνει την πιθανότητα ένα μοντέλο να περιγράφει καλύτερα την χρονοσειρά }

xlabel('time')

ylabel('monp_thita')

axes('position',[0.415 0.063 0.3018 0.2601])

plot(monxkk_estimate) { κάνουμε το γράφημα της μεταβλητής monxkk_estimate η οποία μας δείχνει την τιμή που παίρνει η εκτιμήτρια κάθε παραμέτρου του συστήματος, π.χ. των συντελεστών του AR (p)

xlabel('time')

ylabel('monxkk_estimate')

Ακολούθως, στην 1η εικόνα παρατηρούμε πως υπάρχει ένα κουμπί "Clear" με το οποίο καθαρίζουμε το figure από τα αποτελέσματα του τρέχοντος πειράματος και το επαναφέρουμε στην αρχική του κατάσταση έτοιμο για τη διενέργεια νέων πειραμάτων.

Ο κώδικας που μας επιτρέπει τα παραπάνω είναι :

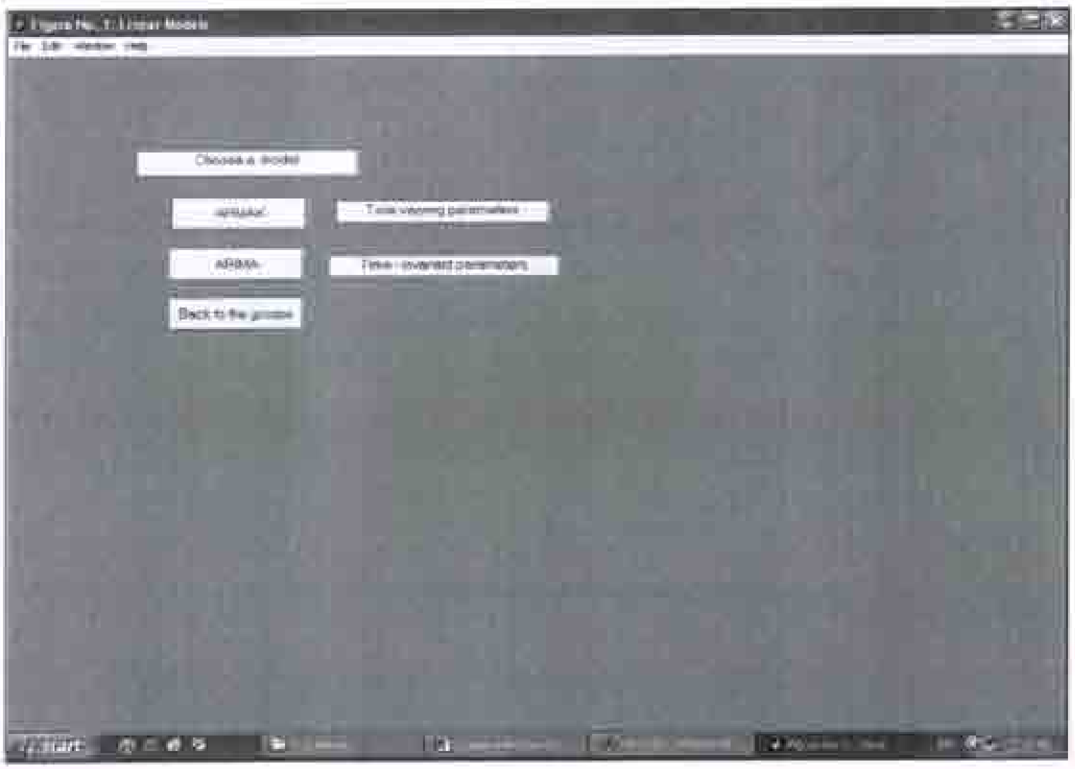
Close {κλείσε το τρέχον αρχείο δηλαδή το ptixiaki.m}

eval ptixiaki{άνοιξε το αρχείο με το όνομα ptixiaki }

Το κουμπί "Exit" έχει μια παραπλήσια λειτουργία δηλαδή κλείνει το ptixiaki.m με την εντολή close.

Το κουμπί με το όνομα "Back to other linear models" μας δίνει τη δυνατότητα να πάμε σε ένα άλλο figure από το οποίο μπορούμε να διαλέξουμε μια διαφορετική προσέγγιση για την ταυτοποίηση (αναγνώριση) του συστήματος – system identification, το μοντέλο ARIMA.

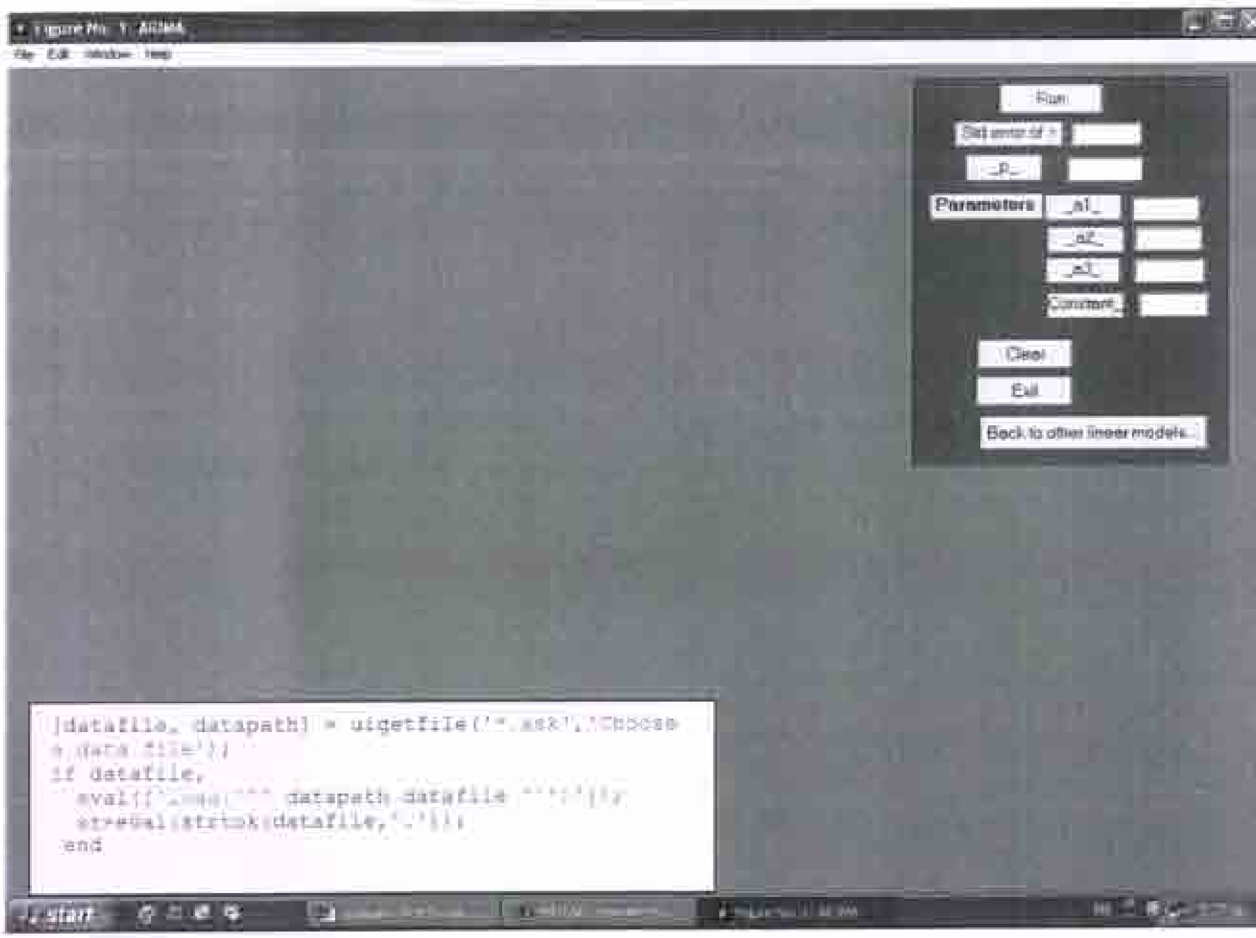
Για να το πετύχουμε χρησιμοποιούμε την εντολή close με την οποία κλείνει το τρέχον figure και την eval ανοίγει το αρχείο με το όνομα linear1 το οποίο φαίνεται στην παρακάτω εικόνα.



2. AR(p) (sales.ask)

Για να ξεκινήσουμε την ταυτοποίηση του μοντέλου χρειαζόμαστε όπως άλλωστε και στην προηγούμενη περίπτωση να φορτώσουμε το αρχείο που περιέχει τη χρονοσειρά. Ο κώδικας που περιέχει τον αλγόριθμο ταυτοποίησης και εκτίμησης περιέχεται σε ένα αρχείο που ονομάζεται arma5.m Στη παρακάτω εικόνα βλέπουμε το πρώτο κουμπί πάνω δεξιά με το όνομα Run . το κουμπί αυτό ενεργοποιεί το αρχείο arma5.m με την εντολή eval.

κόνα 1



Για να ξεκινήσει η ανάλυση της χρονολογικής σειράς θα πρέπει αυτή πρώτα να φορτωθεί. Είδαμε στην προηγούμενη υλοποίηση, αναφορικά με τα φίλτρα Kalman, τον κώδικα που μας επιτρέπει να φορτώσουμε τη χρονοσειρά ή οποία είναι και η είσοδός μας στο σύστημα. Τον υπενθυμίζουμε στο πλαίσιο της 1ης εικόνας.

Υπολογισμός ακολουθίας των συντελεστών αυτοσυσχέτισης και μερικής αυτοσυσχέτισης

Στη συνέχεια θα πρέπει να υπολογίσουμε την ακολουθία των συντελεστών αυτοσυσχέτισης και μερικής αυτοσυσχέτισης. Για να το κάνουμε αυτό χρειαζόμαστε την μήτρα των συντελεστών αυτοσυνδιασποράς Αυτή τη μήτρα την έχουμε ορίσει σαν C. Για να τη προσδιορίσουμε θα πρέπει να αφαιρέσουμε από κάθε τιμή της χρονικής σειράς το μέσο της και στη συνέχεια η διαφορά τους να πολλαπλασιαστεί με τη διαφορά του ανάστροφου πίνακα της xt μείον τον μέσο της . Να σημειώσουμε ότι ο ανάστροφος ενός πίνακα A είναι εκείνος ο πίνακας στον οι γραμμές του πρώτου γίνονται στήλες στο δεύτερο και οι στήλες του πρώτου γραμμές στο δεύτερο. Για παράδειγμα :

A =

1

4

5

0

7

4

Με την εντολή B = A' πετυχαίνουμε το ακόλουθο

= 1 4 5 0 7 4

Στον πίνακα A έχουμε 6 γραμμές και 1 στήλη. Στον ανάστροφό του δηλαδή στον B έχουμε τώρα 1 γραμμή και 6 στήλες.

Η εντολή που μας δίνει την μήτρα των συντελεστών αυτοσυνδιασποράς είναι η

```
C=(xt-mean(xt))*(xt'-mean(xt))
```

Και τα αποτελέσματά της έστω για

x t =

1
4
5
0
7
4

είναι τα ακόλουθα

C =

6.2500	-1.2500	-3.7500	8.7500	-8.7500	-1.2500
-1.2500	**0.2500**	0.7500	-1.7500	1.7500	0.2500
-3.7500	0.7500	**2.2500**	-5.2500	5.2500	0.7500
8.7500	-1.7500	-5.2500	**12.2500**	-12.2500	-1.7500
-8.7500	1.7500	5.2500	-12.2500	**12.2500**	1.7500
-1.2500	0.2500	0.7500	-1.7500	1.7500	**0.2500**

Παρατηρούμε ότι ο πίνακας των συντελεστών αυτοσυνδιασποράς είναι ένας πίνακας συμμετρικός, 6 * 6 διότι 6 είναι και το πλήθος των τιμών της xt. Αν η xt είχε 100 τιμές τότε ο πίνακας αυτοσυνδιαπορών θα ήταν 100 * 100 .

Με την ακόλουθη for υπολογίζουμε την sum((xt - mean(xt))^2) της διαδικασίας Το Co είναι το πηλίκο του **δειγματικού συντελεστή αυτοσυσχέτισης** που είναι και το εργαλείο μας για την ταυτοποίηση της χρονοσειράς μας

$$\boxed{r_\kappa = \frac{\sum_{1}^{n-\kappa}(X_t - \bar{X}_t)(X_{t-\kappa} - \bar{X}_t)}{\sum_{1}^{n}(X_t - \bar{X}_t)^2}}$$

m = length(xt) { η μεταβλητή m κρατά το πλήθος των τιμών της xt }

n=length(xt) { ομοίως και για τη n }

```
Co=0
for i=1:n { για κάθε γραμμή του C }
    for j=1:m { για κάθε στήλη του C }
        if i==j { αν η γραμμή είναι ίση με τη στήλη δηλαδή 1η γραμμή = 1η
στήλη )
     Co=Co+C(i,j) { {πρόσθεσε το στοιχείο του πίνακα στο Co }
         end
    end
end
```

Η παραπάνω for θα προσθέσει τα στοιχεία C_{11}, C_{22}, ... C_{nn}

Για να υπολογίσουμε τη αυτοσυνδιασπορά για κ εργαζόμαστε με τον ίδιο τρόπο

```
A=(1:16)'
B=A-A

for i=1:n
    for j=1:m
        for k=1:16
```

```
    if  j-i==k
       B(k)=B(k)+C(i,j)
       end
     end
   end
end
r=B./Co
```

Για κ = 1, η παραπάνω for θα προσθέσει τα στοιχεία C_{12}, C_{23}, C_{35} ...

Για κ = 2, η παραπάνω for θα προσθέσει τα στοιχεία C_{13}, C_{24}, C_{46} ...

.

.

.

Για κ = 16, η παραπάνω for θα προσθέσει τα στοιχεία C_{116}, C_{217}, C_{319} ...

Κάθε συντελεστή αυτοσυνδιασποράς τον αποθηκεύουμε σε ένα πίνακα που ονομάζουμε B και είναι 1 * 16 . Επειδή θα πρέπει να έχουμε μια αρχικοποίηση των συντελεστών δηλαδή B(1) B(2), ... B(16) = 0 ο πίνακας B θα πρέπει να έχει 16 θέσεις και σε κάθε θέση να έχουμε την τιμή 0. Για να το πετύχουμε αυτό δημιουργήσαμε πρώτα ένα πίνακα A (1 στήλη και 16 γραμμές) με αριθμούς από 1 έως το 16

```
A=(1:16)'
```

και μετά αφαιρέσαμε κάθε αριθμό με τον εαυτό του.

```
B=A-A
```

Για αυτό το λόγο οι `A=(1:16)'` και `B=A-A` προηγούνται του βρόγχου

Οπότε αφού έχουμε τα B και το Co μπορούμε να βρούμε το r = B/Co, δηλαδή το δειγματικό συντελεστή αυτοσυσχέτισης για κάθε υστέρηση κ .

Στο αλγόριθμο που υλοποιήθηκε το κ δηλαδή η υστέρηση (lag) φτάνει ως και το 16.

Για την ακολουθία **συντελεστών μερικής αυτοσυσχέτισης (Partial Autocorrelation Function – PACF) $\Phi_{\kappa\kappa}$** πριν αναπτύξουμε τον αλγόριθμο θα πρέπει να θυμηθούμε το θεωρητικό υπόβαθρο.

Έστω ότι έχουμε μια **AR (1)** της οποίας τη συνάρτηση αυτοσυσχέτισης θα υπολογίσουμε

$z_n = \varphi z_{n-1} + w_n$

$z_n - \varphi z_{n-1} = w_n$

πολλαπλασιάζουμε με $z_{n-\kappa}$

$z_{n-\kappa} z_n - z_{n-\kappa} \varphi z_{n-1} = z_{n-\kappa} w_n$

παίρνουμε αναμενόμενες τιμές (Expectation)

$E(z_{n-\kappa} z_n) - \varphi E(z_{n-\kappa} z_{n-1}) = E(z_{n-\kappa} w_n)$

<u>Για κ >1</u> $\gamma_\kappa - \varphi\gamma_{\kappa-1} = 0 \quad (1)$

Στη σχέση (1) διαιρούμε με γ_o (διασπορά) και έχουμε

$\gamma_\kappa - \varphi\gamma_{\kappa-1} = 0$

$$\frac{\gamma_\kappa}{\gamma_o} - \varphi\frac{\gamma_{\kappa-1}}{\gamma_o} = 0$$

$\rho_\kappa - \varphi\rho_{\kappa-1} = 0$

$\rho_\kappa = \varphi\rho_{\kappa-1}$ Θυμίζουμε πως $\rho_o = 1$

Για $\kappa = 1$

$\varphi = \rho_1$

Έστω τώρα ότι έχουμε μια **AR (2) .**

$z_n - \varphi_1 z_{n-1} - \varphi_2 z_{n-2} = w_n$

πολλαπλασιάζουμε με $z_{n-\kappa}$

$z_n z_{n-\kappa} - \varphi_1 z_{n-1} z_{n-\kappa} - \varphi_2 z_{n-2} z_{n-\kappa} = z_{n-\kappa} w_n$

παίρνουμε αναμενόμενες τιμές (Expectation)

$E(z_n z_{n-\kappa}) - \varphi_1 E(z_{n-1} z_{n-\kappa}) - \varphi_2 E(z_{n-2} z_{n-\kappa}) = E(z_{n-\kappa} w_n)$

<u>Για κ >1</u> $\gamma_\kappa - \varphi_1 \gamma_{\kappa-1} - \varphi_2 \gamma_{\kappa-2} = 0 \quad (1)$

Στη σχέση (1) διαιρούμε με γ_o (διασπορά) και έχουμε

$\gamma_\kappa - \varphi\gamma_{\kappa-1} - \varphi_2 \gamma_{\kappa-2} = 0$

$$\frac{\gamma_\kappa}{\gamma_o} - \varphi_1\frac{\gamma_{\kappa-1}}{\gamma_o} - \varphi_2\frac{\gamma_{\kappa-2}}{\gamma_o} = 0$$

$$\rho_\kappa - \varphi_1\rho_{\kappa-1} - \varphi_2\rho_{\kappa-2} = 0$$

$$\rho_\kappa = \varphi_1\rho_{\kappa-1} + \varphi_2\rho_{\kappa-2}$$

$$\varphi_1\rho_{\kappa-1} + \varphi_2\rho_{\kappa-2} = \rho_\kappa \quad (2)$$

Για $\kappa = 1$ η (2) γίνεται $\varphi_1 + \varphi_2\rho_1 = \rho_1$

Για κ = 2 η (2) γίνεται $\varphi_1 \rho_1 + \varphi_2 = \rho_2$

Το παραπάνω ζεύγος εξισώσεων είναι γνωστό ως εξισώσεις Yule – Walker

Η ορίζουσά του είναι

$$D = \begin{vmatrix} 1 \_\_ \rho_1 \\ \rho_1 \_\_ 1 \end{vmatrix}$$ *(Η _ είναι το κενό)*

Και για λύσουμε ως προς φ_2

$$Df = \begin{vmatrix} 1 \_\_ \rho_1 \\ \rho_1 \_ \rho_2 \end{vmatrix}$$

$\varphi_2 = Df / D$

Για μια **AR (3)** δηλαδή για p = 3 έχουμε με τον ίδιο τρόπο

$$\rho_\kappa - \varphi_1\rho_{\kappa-1} - \varphi_2\rho_{\kappa-2} - \varphi_3\rho_{\kappa-3} = 0$$

$$\rho_\kappa = \varphi_1\rho_{\kappa-1} + \varphi_2\rho_{\kappa-2} + \varphi_3\rho_{\kappa-3}$$

$$\varphi_1\rho_{\kappa-1} + \varphi_2\rho_{\kappa-2} + \varphi_3\rho_{\kappa-3} = \rho_\kappa \quad (3)$$

Για κ = 1 η (3) γίνεται $\varphi_1 + \varphi_2\rho_1 + \varphi_3\rho_2 = \rho_1$

Για κ = 2 η (3) γίνεται $\varphi_1 \rho_1 + \varphi_2 + \varphi_3\rho_1 = \rho_2$

Για κ = 3 η (3) γίνεται $\varphi_1 \rho_2 + \varphi_2 \rho_1 + \varphi_3 = \rho_3$

Η ορίζουσά του είναι

$$D = \begin{vmatrix} 1 \_\_ \rho_1 \_\_ \rho_2 \\ \rho_1 \_\_ 1 \_\_ \rho_1 \\ \rho_2 \_\_ \rho_1 \_\_ 1 \end{vmatrix}$$

Και για λύσουμε ως προς φ_3

$$Df = \begin{vmatrix} 1 \_\_ \rho_1 \_\_ \rho_1 \\ \rho_1 \_\_ 1 \_\_ \rho_2 \\ \rho_2 \_\_ \rho_1 \_ \rho_3 \end{vmatrix}$$

$\varphi_3 = Df / D$

Για μια **AR (4)** δηλαδή για p = 4 έχουμε με τον ίδιο τρόπο

$$\rho_\kappa - \varphi_1\rho_{\kappa-1} - \varphi_2\rho_{\kappa-2} - \varphi_3\rho_{\kappa-3} - \varphi_4\rho_{\kappa-4} = 0$$

$$\rho_\kappa = \varphi_1\rho_{\kappa-1} + \varphi_2\rho_{\kappa-2} + \varphi_3\rho_{\kappa-3} + \varphi_4\rho_{\kappa-4}$$
$$\varphi_1\rho_{\kappa-1} + \varphi_2\rho_{\kappa-2} + \varphi_3\rho_{\kappa-3} + \varphi_4\rho_{\kappa-4} = \rho_\kappa \quad (4)$$

Για κ = 1 η (4) γίνεται $\varphi_1 + \varphi_2\rho_1 + \varphi_3\rho_2 + \varphi_4\rho_3 = \rho_1$

Για κ = 2 η (4) γίνεται $\varphi_1\rho_1 + \varphi_2 + \varphi_3\rho_1 + \varphi_4\rho_2 = \rho_2$

Για κ = 3 η (4) γίνεται $\varphi_1\rho_2 + \varphi_2\rho_1 + \varphi_3 + \varphi_4\rho_1 = \rho_3$

Για κ = 4 η (4) γίνεται $\varphi_1\rho_3 + \varphi_2\rho_2 + \varphi_3\rho_1 + \varphi_4 = \rho_4$

Η ορίζουσά του είναι

$$D = \begin{vmatrix} 1 & \rho_1 & \rho_2 & \rho_3 \\ \rho_1 & 1 & \rho_1 & \rho_2 \\ \rho_2 & \rho_1 & 1 & \rho_1 \\ \rho_3 & \rho_2 & \rho_1 & 1 \end{vmatrix}$$

Και για λύσουμε ως προς φ_4

$$Df = \begin{vmatrix} 1 & \rho_1 & \rho_2 & \rho_1 \\ \rho_1 & 1 & \rho_1 & \rho_2 \\ \rho_2 & \rho_1 & 1 & \rho_3 \\ \rho_3 & \rho_2 & \rho_1 & \rho_4 \end{vmatrix} \qquad \varphi_4 = Df / D$$

Με γνώμονα τα παραπάνω και για r = ρ θα υλοποιήσουμε τον αλγόριθμο

`A=[1:16]'` Εργαζόμαστε όπως στην ACF δημιουργώντας ένα

`F=A-A` πίνακα 16 θέσεων με τιμή 0 προκειμένου να εισάγουμε σε αυτόν

τους $\mathbf{\Phi}_{\kappa\kappa}$

`F(1)=r(1)` { Ο $\mathbf{\Phi}_{11}$ είναι ο πρώτος συντελεστής μερικής αυτοσυσχέτισης }

`D=[1 r(1);r(1) 1]` { $D = \begin{vmatrix} 1 & r_1 \\ r_1 & 1 \end{vmatrix}$ Η ορίζουσα αυτή προκύπτει από

τις

εξισώσεις Yule – Walker αν για ρ = r

}

`Df=[1 r(1);r(1) r(2)]{` και $Df = \begin{vmatrix} 1 & r_1 \\ r_1 & r_2 \end{vmatrix}$ }

`f2=det(Df)/det(D)` { Φ_{22} = Df / D με τη det να μας υπολογίζει την ορίζουσα }

`F(2)=f2`

Κάθε φορά ο συντελεστής θα προκύπτει από τη σχέση $\Phi_{\kappa\kappa}$ = Df / D . Κάθε νέα D που προκύπτει για p =3 ..16 θα παρατηρήσουμε πως είναι ίδια με την προηγούμενή της με τη διαφορά ότι έχει μια καινούρια στήλη και γραμμή. Η νέα στήλη θα έχει τους συντελεστές αυτοσυσχέτισης r(p – 1) έως r(1) .

p = 2 $\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}$

p = 3 $\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}$

p = 4 $\begin{vmatrix} 1 & \rho_1 & \rho_2 & \rho_3 \\ \rho_1 & 1 & \rho_1 & \rho_2 \\ \rho_2 & \rho_1 & 1 & \rho_1 \\ \rho_3 & \rho_2 & \rho_1 & 1 \end{vmatrix}$

Για παράδειγμα αν θέλουμε να υπολογίσουμε την D για p = 3 αυτή θα είναι

$\begin{vmatrix} 1 & r_1 & r_2 \\ r_1 & 1 & r_1 \\ r_2 & r_1 & 1 \end{vmatrix}$ που είναι ίδια με αυτήν για p = 2 $\begin{vmatrix} 1 & r_1 \\ r_1 & 1 \end{vmatrix}$ μόνο που έχει μία επιπλέον στήλη και μία επιπλέον γραμμή. Η στήλη και η γραμμή έχουν r(2), r(1) και το 1.

```
K=[16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1]'
```

```
r_c=r(K(1:16))
```
{ Αντιμεταθέτει το διάνυσμα r }

Δηλαδή αν

r =	Τότε r_c =
0.9437	0.4594
0.8925	0.4788
0.8475	0.5036
0.8095	0.5309
0.7816	0.5568
0.7555	0.5852
0.7154	0.6122
0.6784	0.6528
0.6528	0.6784
0.6122	0.7154
0.5852	0.7555
0.5568	0.7816
0.5309	0.8095
0.5036	0.8475
0.4788	0.8925
0.4594	0.9437

Η αντιμετάθεση μας χρειάζεται διότι θα πρέπει να θέσουμε τους συντελεστές αυτοσυσχέτισης r(p – 1) έως r(1) . Αν έχουμε p = 3 χρειαζόμαστε r(2), r(1) και το 1.

Στη Matlab μπορούμε με την εντολή A (1: 3) να δείξουμε τα 3 πρώτα στοιχεία ενός πίνακα A 1 * μ . Δε γίνεται το ίδιο όμως και με την A (3: 1).

Για αυτό το λόγο αντιμεταθέσαμε το r σε r_c οπότε το r(1) έγινε r_c (16) , το r(2) έγινε r_c(15), το r(3) έγινε r_c(14) κτλ. Ταυτόχρονα επειδή στην εισαγωγή στοιχείων σε κάθε στήλη θα χρειαστούμε και την μονάδα, θέσαμε ως 17^o στοιχείο του r_c το 1.

Στην περίπτωση του p = 3 (δηλαδή για $Φ_{33}$) χρειαζόμαστε r(2), r(1) και το 1. Δηλαδή το r_c(15), r_c(16) και r_c(17). Πιο απλά χρειαζόμαστε το r_c(15 :17).

Στην περίπτωση τώρα του p = 16 (δηλαδή για $Φ_{1616}$) χρειαζόμαστε r(16), r(15), r(14) … r(1) και το 1. Δηλαδή το r_c(1), r_c(1) και r_c(3) … r_c(17) . Πιο απλά χρειαζόμαστε το r_c(1 :17).

Παραμετρικά χρειαζόμαστε το [r_c((18-i):16);] όπου i ο – ιοστός συντελεστής Φ που θέλουμε να προσδιορίσουμε.

```
r_c(17)=1
for i=3:16
```

`D(:,i)=[r_c((18-i):16);]` *{Προσθέτουμε στη D. i–οστή στήλη με τα στοιχεία του r_c}*

`D(i,:)=[r_c((18-i):17)]'` *{Προσθέτουμε στη D. i–οστή γραμμή με τα στοιχεία του r_c}*

`Df=D` *{ Η Df είναι ίδια με την D μόνο που ...*

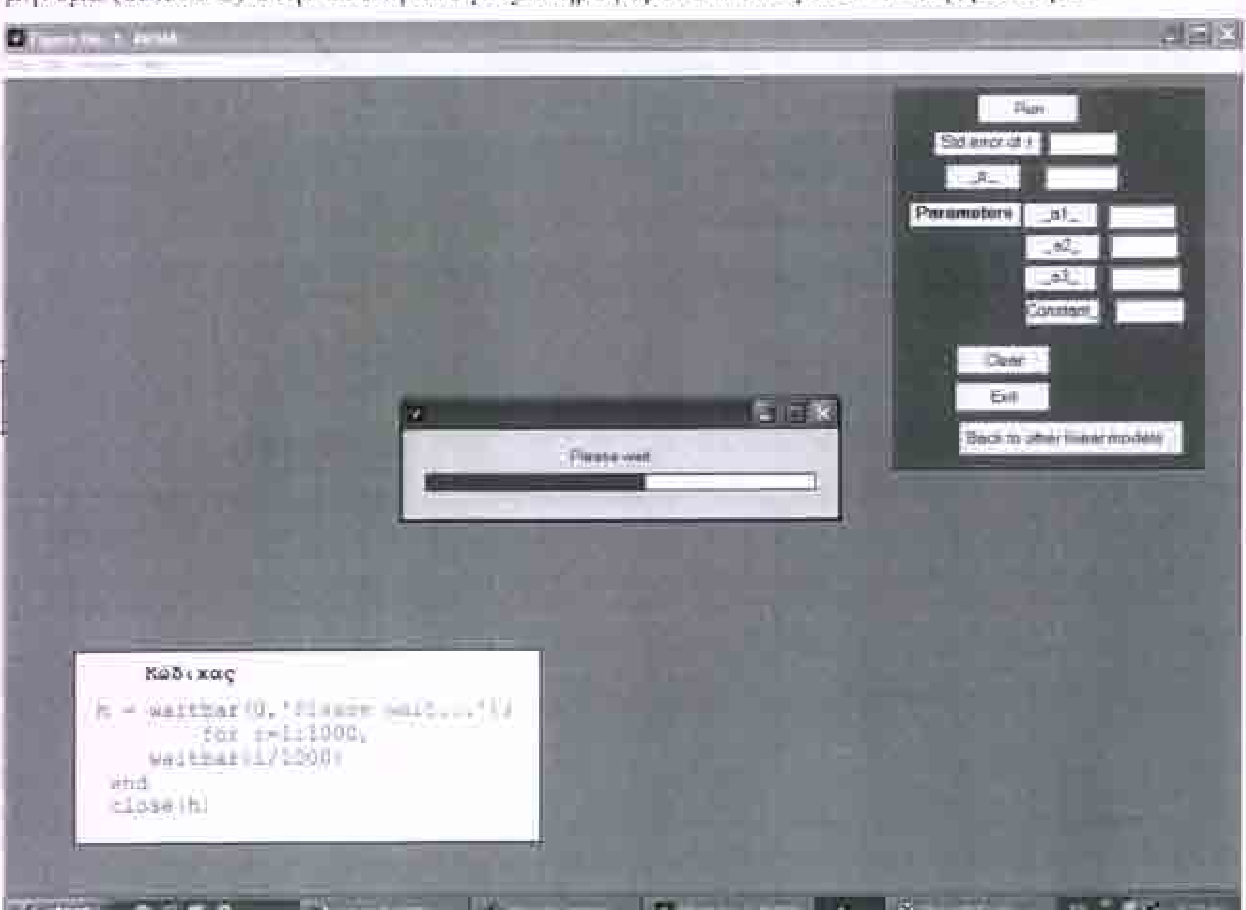
`Df(:,i)=[r(1:i);]` *... η i–οστή της στήλη έχει τα στοιχεία του r }*

`F(i)=F(i)+ (det(Df)/det(D))` *{ Κάθε στοιχείο του πίνακα συντελεστών μερικής αυτοσυσχέτισης προκύπτει αν κάθε φορά διαιρούμε την ορίζουσα D με την ορίζουσα της Df }*

```
End
```

Επειδή ο προσδιορισμός τη μήτρας των συντελεστών συσχέτισης και μερικής αυτοσυσχέτισης είναι μια χρονοβόρα διαδικασία εμφανίζουμε και ένα βοηθητικό μήνυμα (εικόνα 2) στην οθόνη που μας πληροφορεί ότι θα πρέπει να περιμένουμε .

Εικόνα 2



Για το συντελεστή αυτοσυσχέτισης θα πρέπει να βρούμε το τυπικό του σφάλμα που προσδιορίζεται από τη σχέση $\pm 2\sqrt{\frac{1}{n}}$

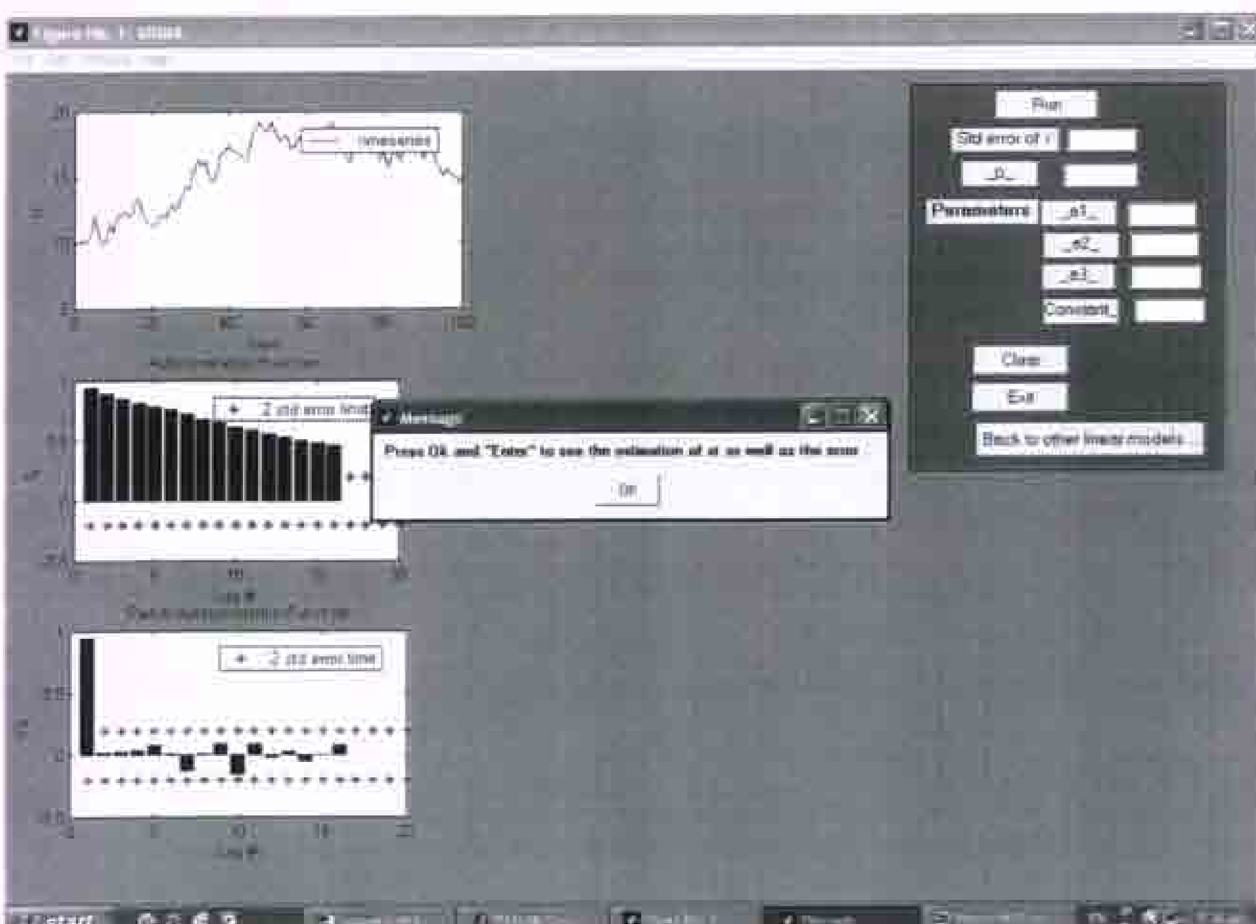
`var_r=1/length(xt);` Διασπορά

`s=sqrt(var_r);` Τυπική απόκλιση

`se=2*s;` Τυπικό σφάλμα

Τα αποτελέσματα του συντελεστή αυτοσυσχέτισης φαίνονται στην ακόλουθη εικόνα

να 3



Άξονες με τις γραφικές παραστάσεις

axes('position',[0.055 0.713 0.312 0.238]) *{Στους άξονες με αυτές τις συντεταγμένες }*

plot(xt) *{...κάνει το γράφημα της xt }*

xlabel('time') *{ στον άξονα των x βάζουμε την ετικέτα time }*

ylabel('xt') *{ στον άξονα των y βάζουμε την ετικέτα xt }*

legend('timeseries') *{ εμφανίζουμε το χρώμα που απεικονίζει τη μεταβλητή }*

axes('position',[0.055 0.410 0.262 0.218])

plot(k,se,'r*'); *(κάνουμε το γράφημα του τυπικού σφάλματος του συντελεστή αυτοσυσχ. με κόκκινο 'red' αστερίσκο *. }*

Hold *(Με αυτήν την εντολή κρατάμε το προηγούμενο γράφημα και είμαστε έτοιμοι το επόμενο να το εμφανίσουμε στους ίδιους άξονες}*

plot(k,-se,'r*');

legend('2 std error limit')

Bar(r) *{δείχνει τη γραφική παράσταση της ACF σε ράβδους 2° διάγραμμα }*

xlabel('Lag #')

ylabel('r_k')

title('Autocorrelation Function') *{ ο τίτλος του γραφήματος }*

Με ίδιες εντολές εργαστήκαμε για να δημιουργήσουμε και το τρίτο διάγραμμα των συντελεστών μερικής αυτοσυσχέτισης αυτή τη φορά.

Αφού ο αλγόριθμος υπολογίσει την ACF και την PACF ξεκινά να προσδιορίσει την τάξη του AR μοντέλου , την τιμή των παραμέτρων $φ_i$ την εκτίμηση και το λάθος. Πατώντας το OK του μηνύματος που εμφανίζεται και μετά το Enter

Προσδιορισμός του p

Το p προσδιορίζεται από τη σχέση p = κ – 1 όπου κ εκείνο το lag στο οποίο η ακολουθία των συντελεστών μερικής αυτοσυσχέτισης ξεκινά να μηδενίζεται ή καλύτερα πέφτει μέσα στα όρια του διαστήματος εμπιστοσύνης που δημιουργείται από την απόλυτη τιμή του τυπικού σφάλματος του συντελεστή αυτοσυσχέτισης.

Από τον παρακάτω αλγόριθμο ζητούμε για κάθε στοιχείο του F αν το απόλυτό του είναι μεγαλύτερο από την απόλυτη τιμή του τυπικού σφάλματος του συντελεστή αυτοσυσχέτισης τότε κράτησέ του τη θέση στη μεταβλητή p στην οποία έχουμε δώσει μια αρχική τιμή p = 0.

```
p=0;
for i=1:16
    if abs(F(i))>abs(se)
        p=i
    end
end
```

Στο παράδειγμά μας βλέπουμε πως p = 1

Προσδιορισμός των παραμέτρων φ_i

Έστω

$x_t = \xi + \varphi_1 x_{t-1} + e_t$

$\hat{x}_t = \hat{\xi} + \alpha_1 x_{t-1}$ όπου $\hat{\xi}$ και α_1 είναι η εκτίμηση των φ_1 και β αντίστοιχα

$e_t = x_t - \hat{x}_t$ με $\hat{x}_t$ η εκτίμηση του xt

$$S_e(\varphi) = \sum_{i=1}^{n} e_i^2 = e_t^T e_t \quad \text{όπου } e_t = \begin{bmatrix} e_1 \\ e_2 \\ . \\ . \\ . \\ e_n \end{bmatrix} \quad \& \quad e_t^T = [e_1\ e_2 \ldots e_n]$$

$$[e_1\ e_2 \ldots e_n] \begin{bmatrix} e_1 \\ e_2 \\ . \\ . \\ . \\ e_n \end{bmatrix} = [\,(e_1)^2 + (e_2)^2 + \ldots + (e_n)^2\,]$$

Η $x_t = \xi + \varphi_1 x_{t-1} + e_t$ σε μορφή διανυσμάτων μπορεί να γραφτεί

$$\begin{bmatrix} x_2 \\ x_3 \\ . \\ . \\ x_t \end{bmatrix} = \begin{bmatrix} 1\_x_1 \\ 1\_x_2 \\ . \\ . \\ 1\_x_{t-1} \end{bmatrix} \begin{bmatrix} \xi \\ \phi_1 \end{bmatrix} + \begin{bmatrix} e_2 \\ e_3 \\ . \\ . \\ e_n \end{bmatrix} (1)$$

$X = Z\varphi + e_1$ ή

$e_1 = X - Z\varphi$

$S_e(\varphi) = \sum_{i=1}^{n} e_i^2 = e_1^T e_1 = (X - Z\varphi)^T (X - Z\varphi)$

$S_e(\varphi) = X^2 - 2XZ\varphi + (Z\varphi)^2$

$S_e(\varphi) = X^T X - 2X^T X\varphi + \varphi^T (Z^T Z)\varphi$

$S_e(\varphi) = X^2 - 2Z^T X\varphi + \varphi^2 (Z^T Z)$

$$\frac{dS_e}{d\phi} = 0$$

$-2Z^T X + 2\varphi(Z^T Z) = 0$

$2\varphi(Z^T Z) - 2Z^T X = 0$

$\varphi(Z^T Z) = Z^T X$

$\varphi = (Z^T Z)^{-1} Z^T X \quad (2)$

Στη Matlab για να φτιάξουμε τον πίνακα Z φτιάξαμε πρώτα ένα διάνυσμα, το s που είχε n γραμμές και 1 στήλη και είχε μόνο τους αριθμούς από 1 ...n όπου n , το πλήθος των παρατηρήσεων της χρονοσειράς. Επειδή όμως θέλουμε το διάνυσμα να έχει μονάδες διαιρέσαμε κάθε αριθμό με τον εαυτό του. Οπότε δημιουργήθηκε ένα νέο διάνυσμα, το s_c που είχε n γραμμές και 1 στήλη και είχε μόνο τον αριθμό 1 .

Από αυτό το διάνυσμα αποτελείται αρχικά ο πίνακας Z . Η δεύτερη στήλη του όμως αποτελείται από την $1^η$ παρατήρηση της έως την προτελευταία, δηλαδή xt(1 : n-1). Για αυτό το λόγο και ο Z έχει n –1 . γραμμές.

```
if p==1
  s=(1:n)'  { Δημιουργία του διανύσματος s }
  s_c=s./s  { Δημιουργία του διανύσματος s_c μόνο με μονάδες }
```

```
Z=s_c(2:n)
```
{Δημιουργία του Ζ με τις τιμές του s_c μείον μίας }

```
Z(:,2)=[xt(1:n-1);]
```
{Εισαγωγή 2ης στήλης στον Ζ με στοιχεία του xt }

```
f=(inv(Z'*Z))*Z'*(xt(2:n))
```
{ Υλοποιούμε την σχέση (2) με τη συνάρτηση inv (αντίστροφος) και τα αποτελέσματα εκχωρούνται σε μια μεταβλητή , την f }

```
a=f(1)
```
{ το πρώτο στοιχείο του διανύσματος είναι μια σταθερά }

```
a1=f(2)
```
{ το δεύτερο στοιχείο του διανύσματος είναι ο συντελεστής φ_1 }

```
a2=0
```
{ Σε μια AR(1) δεν υπάρχουν άλλοι συντελεστής φ για αυτό παίρνουν την τιμή 0 }

```
a3=0
end
```

Με τον ίδιο τρόπο ενεργούμε για p = 2

$x_t = \xi + \varphi_1 x_{t-1} + \varphi_2 x_{t-2} + e_t$

$$\begin{bmatrix} x_3 \\ x_4 \\ . \\ . \\ x_t \end{bmatrix} = \begin{bmatrix} 1\_x_2\_\_x_1 \\ 1\_x_3\_\_x_2 \\ .\_\dots\dots \\ \dots\dots\dots\dots \\ 1\_x_{t-1}\_x_{t-2} \end{bmatrix} \begin{bmatrix} \xi \\ \phi_1 \\ \phi_2 \end{bmatrix} + \begin{bmatrix} e_3 \\ e_4 \\ . \\ . \\ e_t \end{bmatrix} (1)$$

```
f=(inv(Z'*Z))*Z'*(xt(3:n))
```

και για p = 3

$x_t = \xi + \varphi_1 x_{t-1} + \varphi_2 x_{t-2} + + \varphi_3 x_{t-3} + e_t$

$$\begin{bmatrix} x_4 \\ x_5 \\ . \\ . \\ x_t \end{bmatrix} = \begin{bmatrix} 1\_\_x_3\_x_2\_\_x_1 \\ 1\_\_x_4\_x_3\_\_x_2 \\ . \\ . \\ 1\_x_{t-1}\_x_{t-2}\_x_{t-3} \end{bmatrix} \begin{bmatrix} \xi \\ \phi_1 \\ \phi_2 \\ \phi_3 \end{bmatrix} + \begin{bmatrix} e_4 \\ e_5 \\ . \\ . \\ e_t \end{bmatrix} (1)$$

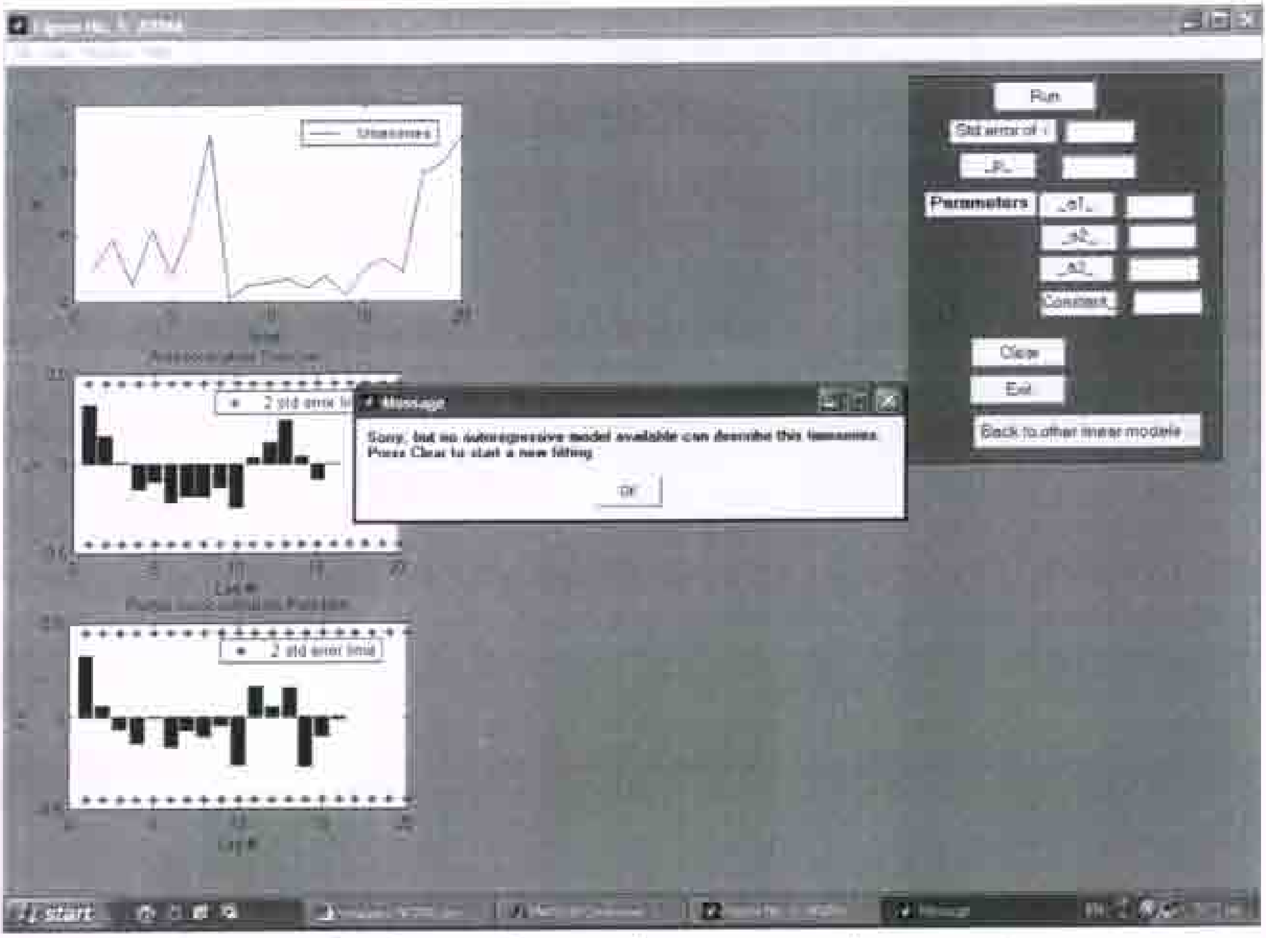
Εκτίμηση και λάθος

Με τον προσδιορισμό των παραμέτρων είμαστε πλέον σε θέση να καθορίσουμε μια εκτίμηση της χρονοσειράς . Ανάλογα με την τάξη του αυτοπαλίνδρομου μοντέλου που καθορίστηκε με τη συνάρτηση μερικών αυτοσυσχετίσεων (PACF) έχουμε και την αντίστοιχη εκτίμηση (xt_estimate) όπως βλέπουμε παρακάτω

```
if p==1
   xt_estimate=[a; a1*(xt(1:n-1))+a;]
end

 if p==2
   xt_estimate= [a; a1*xt(1)+a; a1*xt(2:n-1)+a2*xt(1:n-
2)+a;]
end

if p==3
  xt_estimate=[a; a1*xt(1)+a; a1*xt(2)+a2*xt(1)+a;
a1*xt(3:n-1)+a2*xt(2:n-2)+a3*xt(1:n-3)+ a]
end

if p==0
   msgbox('Sorry, but no autoregressive model available
can describe this timeseries. Press Clear to start a new
fitting','Message')
end
```

Παράλληλα, οφείλουμε να καλύψουμε κα την περίπτωση που p = 0 . Αν δηλαδή η συνάρτηση μερικών αυτοσυσχετίσεων είναι 0 σε όλα τα κ (βλ. εικόνα 4) τότε ζητούμε από το χρήστη του συστήματος να κλείσει τη τρέχουσα ταυτοποίηση και να ξεκινήσει μια άλλη.

κόνα 4



Το λάθος (error_est) δεν είναι παρά μια αφαίρεση του xt μείον την εκτίμηση

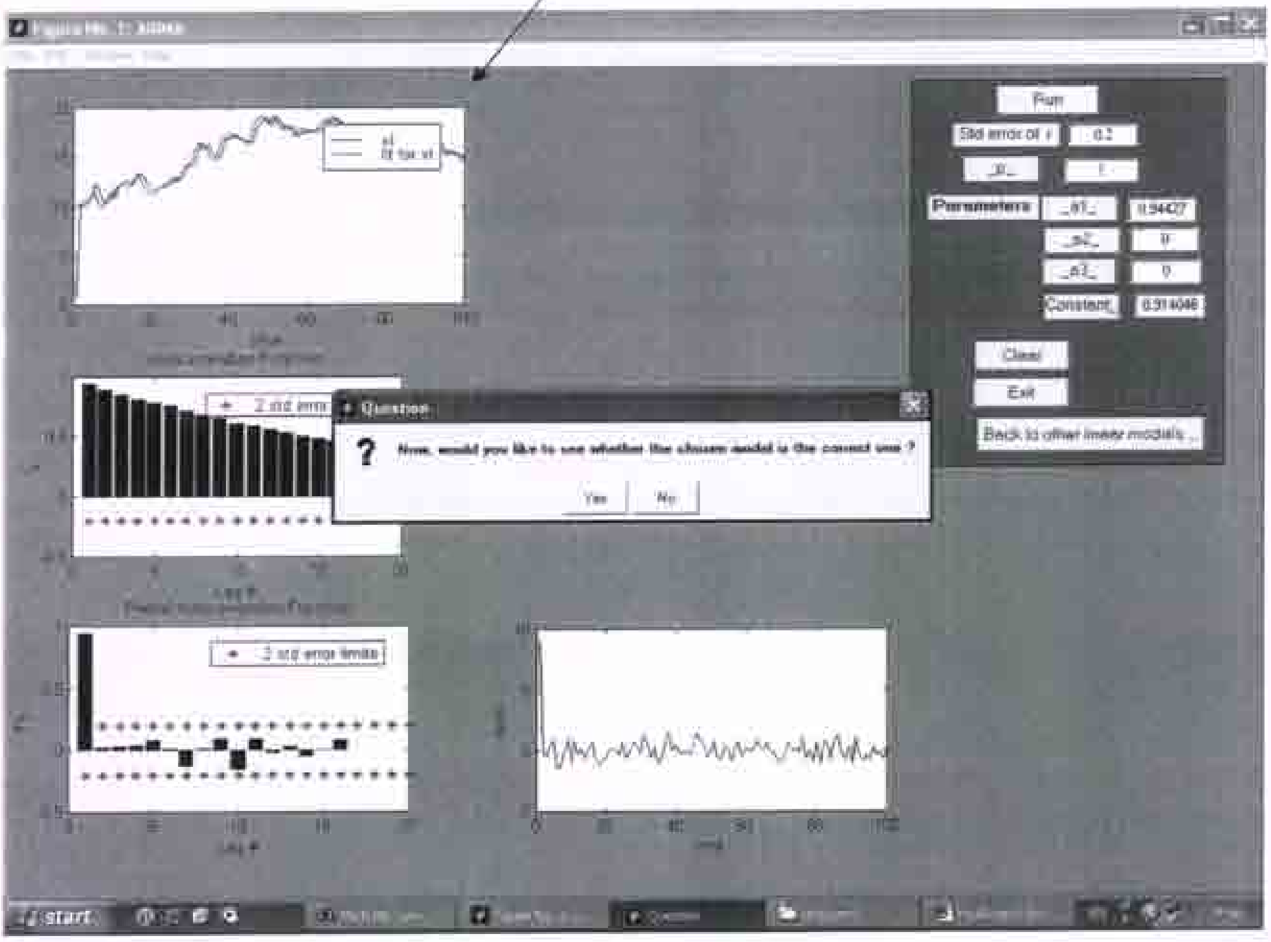
```
axes('position',[ 0.055 0.713 0.312 0.238])
error_est=xt-xt_estimate
```

`plot(xt)` *{ Κάνουμε το γράφημα της xt }*

`hold` *{Κρατάμε τους άξονες}*

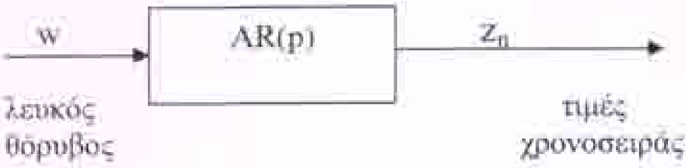
`plot(xt_estimate,'r')` {Στους ίδιους άξονες κάνουμε και το γράφημα του εκτιμηθέντος xt} κόνα 5

legend('xt','fit for xt',-1)



Επαλήθευση επιλογής του σωστού μοντέλου

Για να διαπιστώσουμε αν πράγματι το μοντέλο που επιλέχθηκε από το σύστημα είναι κατάλληλο τότε μπορούμε να υπολογίσουμε τη συνάρτηση αυτοσυσχετίσεων (ACF) του error. Τη χρειαζόμαστε για δούμε αν το σφάλμα έχει την ιδιότητα του λευκού θορύβου περί ασυσχέτιστων τ.μ.

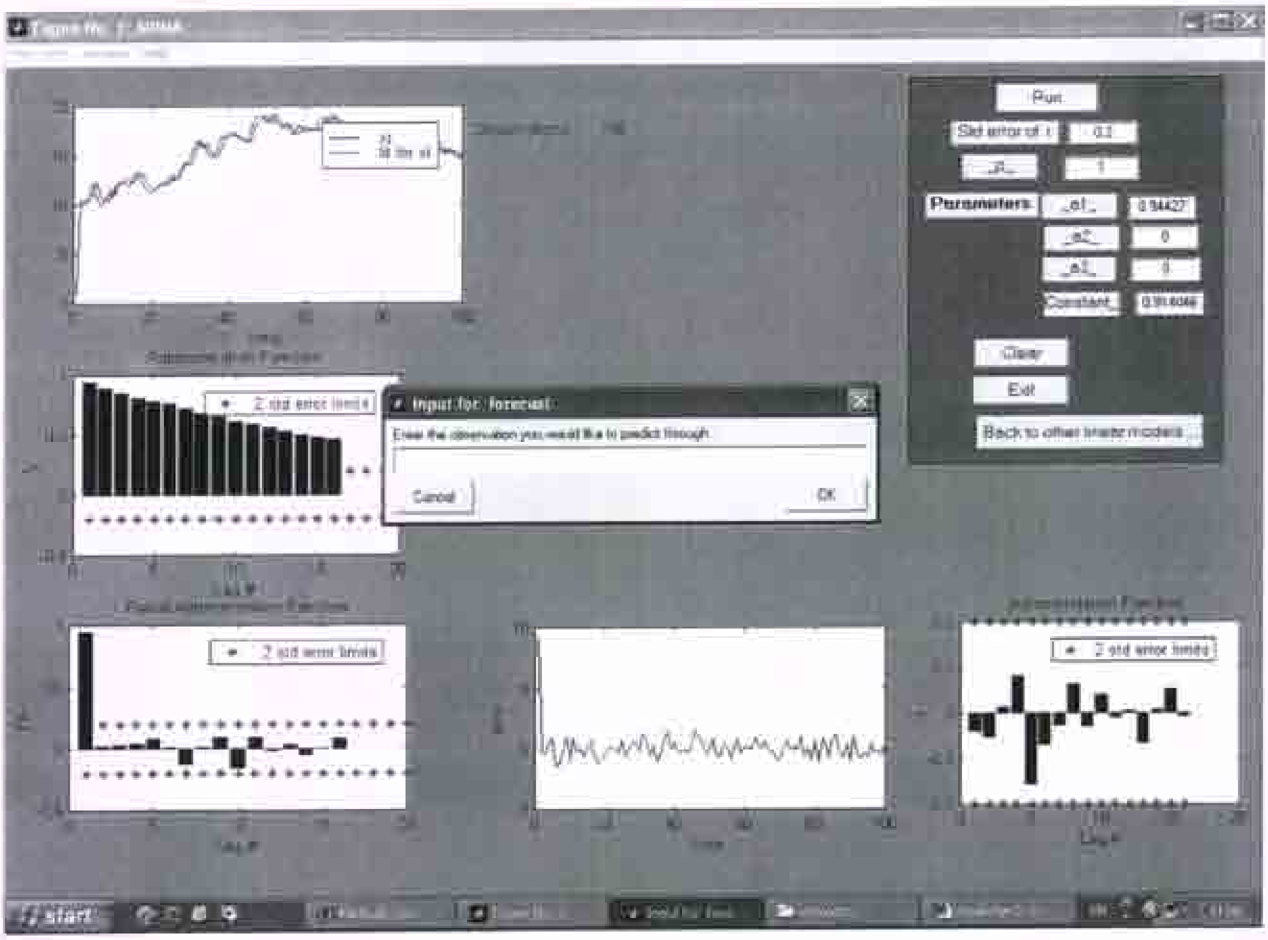


Απόδειξη ότι έχουμε επιλέξει το κατάλληλο υπόδειγμα είναι τα λάθη (errors) να έχουν τις ιδιότητες του λευκού θορύβου, δηλαδή

- Το w_n ακολουθεί την κανονική κατανομή
- Τα w_n ασυσχέτιστα δηλαδή οι συντελεστές αυτοσυσχέτισης να είναι 0

Η δεύτερη ιδιότητα φαίνεται και από το διάγραμμα κάτω δεξιά στην παρακάτω εικόνα

κόνα 6



Για να εμφανιστεί όμως το error ζητούμε την απάντηση του χρήστη. Μόνο αν ο χρήστης επιλέξει Yes στο παράθυρο της εικόνας 5 θα ξεκινήσει ο αλγόριθμος υπολογισμού της ακολουθίας των συντελεστών αυτοσυσχέτισης. Διαφορετικά θα εμφανιστεί μήνυμα που θα ζητά από το χρήστη να αρχίσει μια νέα ταυτοποίηση.

Το παράθυρο της εικόνας 5 στηρίζεται στην εντολή «questdlg» που περιέχει 2 επιλογές, την Yes και την No . Η επιλογή αποθηκεύεται στη μεταβλητή «ButtonName» . Αν η επιλογή είναι Yes τότε υπολογίζεται ο αλγόριθμος της ACF με το ίδιο τρόπο που περιγράψαμε και για τη xl με τη διαφορά βέβαια ότι εδώ η μεταβλητή μας είναι η error_est.

```
ButtonName=questdlg('Now, would you like to see whether
the chosen model is the correct one ?', ...
                              'Question', ...
                              'Yes','No','Yes');
   switch ButtonName,
   case 'Yes',
 m = length(error_est)
 n=length(error_est)

C=(error_est-mean(error_est))*(error_est'-
mean(error_est))
```
{ η μήτρα των συντελεστών αυτοσυνδιασποράς του σφάλματος }

```
h = waitbar(0,'Please wait while ACF of error is
computed...');
        for i=1:1000,
    waitbar(i/1000)
 end
 close(h)
```
{ Το μήνυμα που μας καλεί να περιμένουμε }

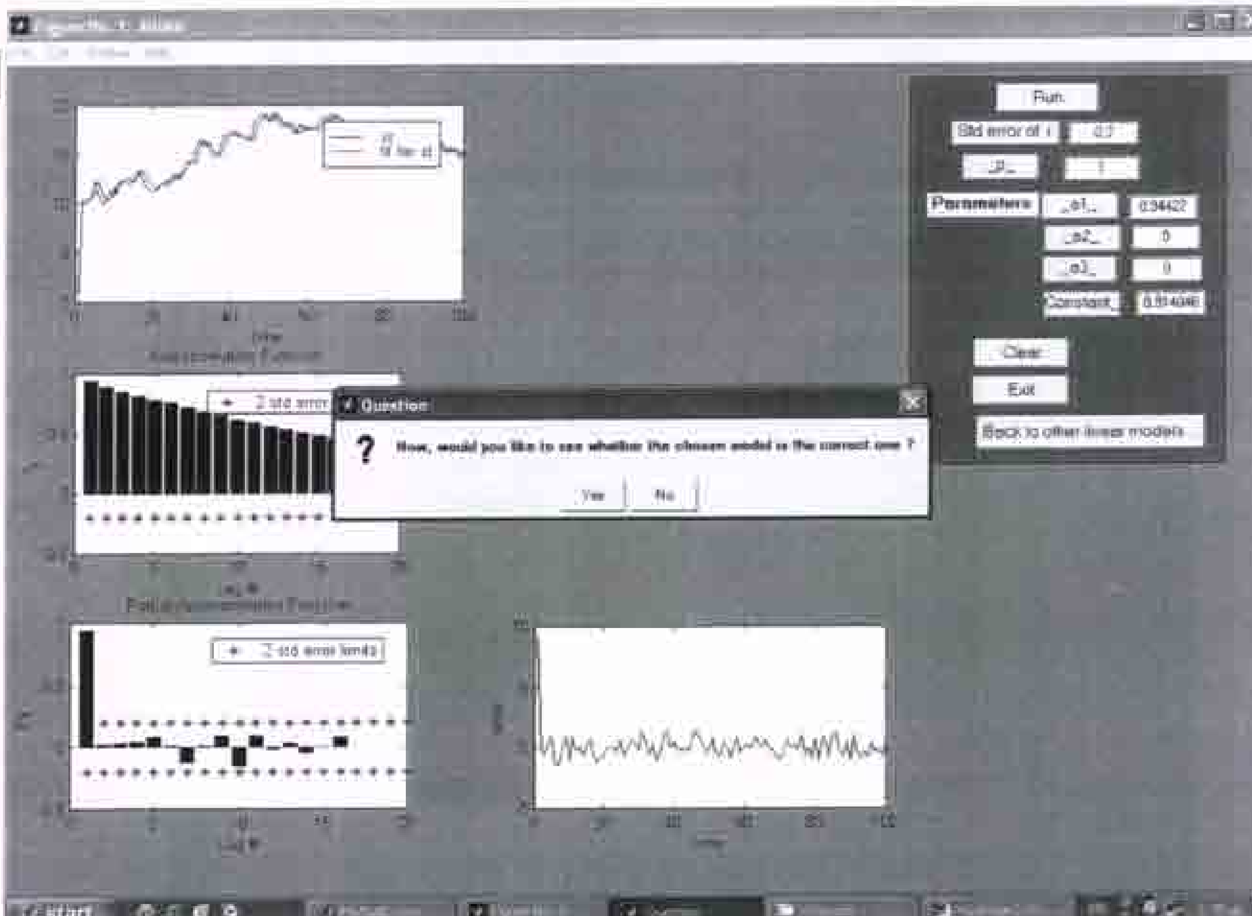
```
Co=0
for i=1:n
   for j=1:m
      if i==j
         Co=Co+C(i,j)
         end
   end
end
```
{ Η διασπορά του σφάλματος }

```
A=(1:16)'
B=A-A
for i=1:n
   for j=1:m
      for k=1:16
         if  j-i==k
            B(k)=B(k)+C(i,j)
         end
      end
   end
end
r_e=B./Co
k=1:16
```
*{ Η ACF του σφάλματος. Οι συντελεστές μπαίνουν σε ένα πίνακα r_e 1 * 16 }*

```
 axes('position', [ 0.761 0.105 0.222 0.221])
plot(k,se,'r*');
 hold
 plot(k,-se,'r*');
 bar(r_e)    xlabel('Lag #')
 ylabel('r_k')
 title('Autocorrelation Function ')
 legend('2 std error limits')
```
{ Η ACF του σφάλματος εμφανίζεται στους ίδιους άξονες με το τυπικό σφάλμα του r_e }

```
        case 'No',
         msgbox('Press clear to start a new
fitting','Message')
      end % switch
```
{ Αν επιλέξει Όχι ο χρήστης εμφανίζεται μήνυμα (βλ. Εικόνα 7) }

Εικόνα 7



Μόλις τελειώσει ο αλγόριθμος τους υπολογισμούς του, εμφανίζονται στα πλαίσια κειμένου ο βαθμός του p , οι τιμές των συντελεστών φ_i που εδώ συμβολίζονται με a1, a2 , a3 καθώς επίσης και η τιμή της σταθεράς a (Εικόνα 7).

Η εμφάνιση τιμών σε πλαίσια κειμένου γίνεται με μια γνωστή εντολή που την είχαμε δει και στο γραμμικό ARMAX, την findobj.

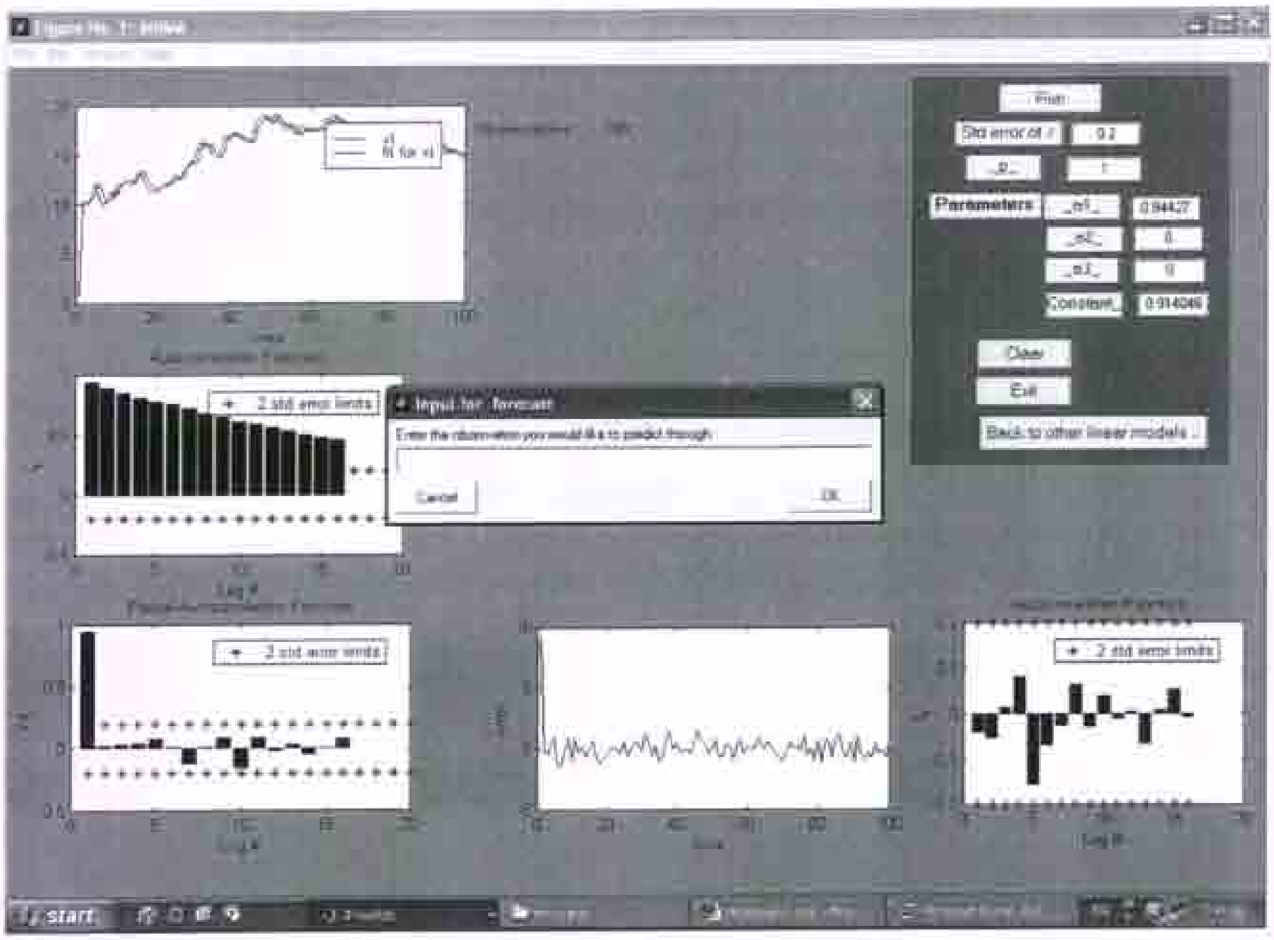
h = findobj('tag','txtp') ; *{Βρίσκει το αντικείμενο που έχει στην ιδιότητα tag το όνομα txtp}*

set(h,'string',p); {Στο ίδιο αντικείμενο, του ζητάμε στην ιδιότητα string να προσθέσει τη μεταβλητή p}. Ομοίως και για τα υπόλοιπα αντικείμενα.

Πρόβλεψη

Στην παρακάτω εικόνα (Εικόνα 8) ο χρήστης βλέπει ένα παράθυρο που τον καλεί να δώσει μέχρι ποια παρατήρηση θα ήθελε να φτάσει η πρόβλεψή του. Εδώ ο χρήστης δίνει την τιμή 105 που σημαίνει ότι θα ήθελε να υπολογιστούν οι προβλεφθείσες τιμές μέχρι και την 105 χρονική στιγμή. Να σημειωθεί πως σε αυτό το παράδειγμα οι παρατηρήσεις της χρονοσειράς είναι 100.

κόνα 8



Για να εμφανιστεί αυτό το μήνυμα απαιτείται ο παρακάτω κώδικας που τον είχαμε συναντήσει και στην περίπτωση του γραμμικού ARMAX όταν θέλαμε ο χρήστης να εισαγάγει τις παραμέτρους του συστήματος. :

prompt={'Enter the observation you would like to predict through :'}; *{ το περιεχόμενο της γραμμής εισαγωγής }*

title='Input for forecast'; *{ ο τίτλος του παραθύρου }*

lineNo=1;

answer=inputdlg(prompt,title,lineNo); *{Η μεταβλητή answer είναι αυτή στην οποία αποθηκεύεται η τιμή που εκχωρήθηκε από το χρήστη }*

obs_c=[answer{1}] *{ Η μεταβλητή στην οποία αποθηκεύεται η τιμή είναι obs_c η οποία όμως δεν είναι αριθμητικού τύπου αλλά χαρακτήρας string}*

obs = str2num(obs_c) *{ Επειδή πρέπει να κάνουμε πράξεις με την obs_c θα πρέπει να τη μετατρέψουμε σε αριθμητικού τύπου με τη συνάρτηση str2num (string to number). Το αποτέλεσμα θα είναι η μεταβλητή obs (**obs**ervation = παρατήρηση }*

Τι θα γίνει όμως στην περίπτωση που ο χρήστης θα εκχωρήσει εκ παραδρομής μια τιμή μικρότερη από το πλήθος των παρατηρήσεων , δηλαδή θα ζητήσει να γίνει πρόβλεψη για τη χρονική στιγμή 97 ενώ οι παρατηρήσεις που έχουν ήδη συλλεχθεί είναι 100 ; Πιο απλά, τι θα γίνει σε περίπτωση που ζητήσει να γίνει πρόβλεψη για το παρελθόν ;

Για να αντιμετωπιστεί αυτή η κατάσταση έχει συνταχθεί μια συνθήκη που λειτουργεί με τη βοήθεια της while. Πιο αναλυτικά, όσο η παρατήρηση πρόβλεψης που εκχώρησε ο χρήστης είναι μικρότερη από το πλήθος των παρατηρήσεων της χρονοσειράς τότε 2 μηνύματα θα εμφανίζονται α) ένα που θα ζητά από το χρήστη να εισάγει μια παρατήρηση μεγαλύτερη από το πλήθος των τιμών της χρονοσειράς και β) το παράθυρο εισαγωγής .

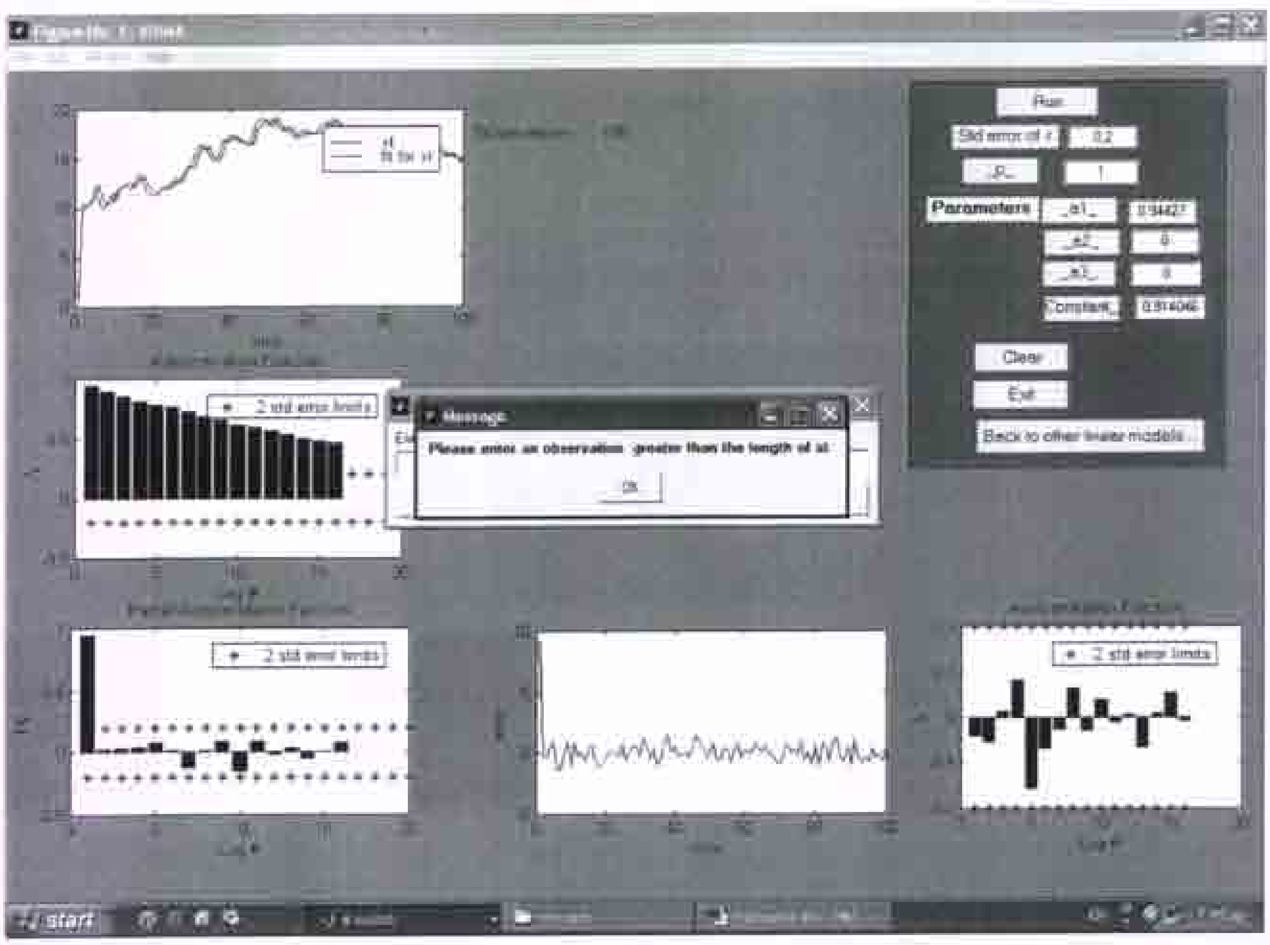
Δηλαδή,

```
while obs<n
```

```
MsgBox('Please enter an observation  greater than the
length of xt','Message')
prompt=('Enter the observation you would like to predict
through :');
title='Input for  forecast';
lineNo=1;
answer=inputdlg(prompt,title,lineNo);
obs_c=[answer(1)]
obs = str2num(obs_c)
end
```

Όσο ο χρήστης εκχωρεί τιμή μικρότερη από το πλήθος των παρατηρήσεων θα εμφανίζονται τα μηνύματα που περιγράφηκαν παραπάνω (δείτε και εικόνα 9).

όνα 9



Επειδή η πρόβλεψη αποτελεί συνέχεια της εκτίμησης θα πρέπει οι προβλεφθείσες τιμές της χρονοσειράς να είναι συνέχεια του πίνακα στον οποίο είναι καταχωρημένες οι τιμές του εκτιμηθέντος xt, δηλαδή του xt _estimate.

Δηλαδή αν ο πίνακας εκτίμησης (n γραμμές * 1 στήλη) είναι ο παρακάτω

xt _estimate =

$$\begin{bmatrix} 12.5 \\ 15.8 \\ 14.7 \\ 17.8 \\ . \\ . \\ . \\ . \\ 15.7 \end{bmatrix} \qquad n = 100$$

τότε θα πρέπει σε αυτόν αν προστεθούν τόσες θέσεις όσες χρειάζονται για να φτάσουν την πρόβλεψη. Αν για παράδειγμα ο χρήστης εκχωρήσει 105 τότε αυτό σημαίνει ότι χρειαζόμαστε την 101^{η}, 102^{η}, 103^{η}, 104^{η} και 105^{η} στιγμή. Άρα θέλουμε 5 θέσεις ακόμα (105 – 100) και μάλιστα αυτές οι 5 νέες θέσεις θα πρέπει αρχικά να έχουν την τιμή 0.

Για να κάνουμε το πρόβλημά πιο παραμετρικό λοιπόν θέσαμε ως πλήθος παρατηρήσεων της χρονοσειράς το **n** (100), την παρατήρηση που θα ήθελε ο χρήστης να φτάσει η πρόβλεψή του **obs** (105) και τις επιπλέον θέσεις ως **time** που προκύπτει από την διαφορά obs-n. Επιπλέον ονομάσαμε ως Forecast (Πρόβλεψη) την xt_estimate.

Αν θέλουμε την πρόβλεψη για την 101^{η} (t = 101) χρονική στιγμή τότε σε ένα μοντέλο AR (3) θα πρέπει να έχουμε

$$x_t = \xi + \varphi_1 x_{t-1} + \varphi_2 x_{t-2} + \varphi_3 x_{t-3} + \varepsilon$$

δηλαδή

$$x_{101} = \xi + \varphi_1 x_{100} + \varphi_2 x_{99} + \varphi_3 x_{98} + \varepsilon$$

Για το ίδιο μοντέλο και για την 102^{η} χρονική στιγμή θα είχαμε

$$x_{102} = \xi + \varphi_1 x_{101} + \varphi_2 x_{100} + \varphi_3 x_{99} + \varepsilon$$

Για το ίδιο μοντέλο και για την 103^{η} χρονική στιγμή θα είχαμε

$x_{103} = \xi + \varphi_1 x_{102} + \varphi_2 x_{101} + \varphi_3 x_{100} + \varepsilon$

Για την 104^{η} χρονική στιγμή θα είχαμε

$x_{104} = \xi + \varphi_1 x_{103} + \varphi_2 x_{102} + \varphi_3 x_{101} + \varepsilon$

Για την 105^{η} χρονική στιγμή θα είχαμε

$x_{105} = \xi + \varphi_1 x_{104} + \varphi_2 x_{103} + \varphi_3 x_{102} + \varepsilon$

Παρατηρούμε ότι η πρόβλεψη είναι μια επανάληψη αν θέσουμε όπου i να παίρνει τιμές από το 1 έως το 5 (5 = time = obs-n.) και όπου n = 100

```
for i=1:time
Forecast(n+i)=Forecast(n+i)+a1*Forecast(n+i-1)+a2*Forecast(n+i-2)+a3*Forecast(n+i-3)+a
End
```

Το Forecast(n+i) είναι η αρχική τιμή της πρόβλεψης που είναι το 0 για καθαρά υπολογιστικούς λόγους με τη χρήση for.

`n=length(xt_estimate)` *{ Θέτουμε ως μεταβλητή n το πλήθος των παρατηρήσεων του εκτιμηθέντος xt }*

`Forecast=xt_estimate` *{Θέτουμε ως μεταβλητή Forecast το εκτιμηθέν xt }*

`time=obs-n` *{ Η διαφορά ανάμεσα στην προβλεφθείσα χρονική στιγμή και το πλήθος των παρατηρήσεων της χρονοσειράς }*

`Forecast_a=[1:time]'` *{Δημιουργούμε ένα πίνακα time * 1 που θα έχει τιμές από το 1 έως το time (105 – 100) }*

`Forecast_b=Forecast_a - Forecast_a'` *{Δημιουργούμε ένα πίνακα time * 1 που θα έχει σε κάθε θέση του την τιμή 0 }*

`Forecast(n+1:obs)=Forecast_b` *{ο προηγούμενος πίνακας με τα 0 προστίθεται πλέον στον πίνακα Forecast , από την 101 θέση και κάτω}*

Forecast =

$$\begin{bmatrix} 12.5 \\ 15.8 \\ 14.7 \\ 17.8 \\ . \\ . \\ . \\ 15.7 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Ο βρόγχος για τον υπολογισμό της πρόβλεψης

```
for i=1:time
Forecast(n+i)=Forecast(n+i)+a1*Forecast(n+i-1)+a2*Forecast(n+i2)+a3*Forecast(n+i-3)+a
End
```

`Forecast_1=Forecast(n+1:obs)` *{Ορίζουμε μια μεταβλητή, την* Forecast_1 *που θα έχει τις προβλεφθείσες τιμές δηλαδή τις τιμές από τη θέση 101 (n + 1) έως και 105(obs) }*

Παράλληλα για να δούμε με πιθανότητα έστω 97,5 % πού θα διακυμανθούν οι τιμές υπολογίζουμε και τα όρια του διαστήματος εμπιστοσύνης της πρόβλεψης.

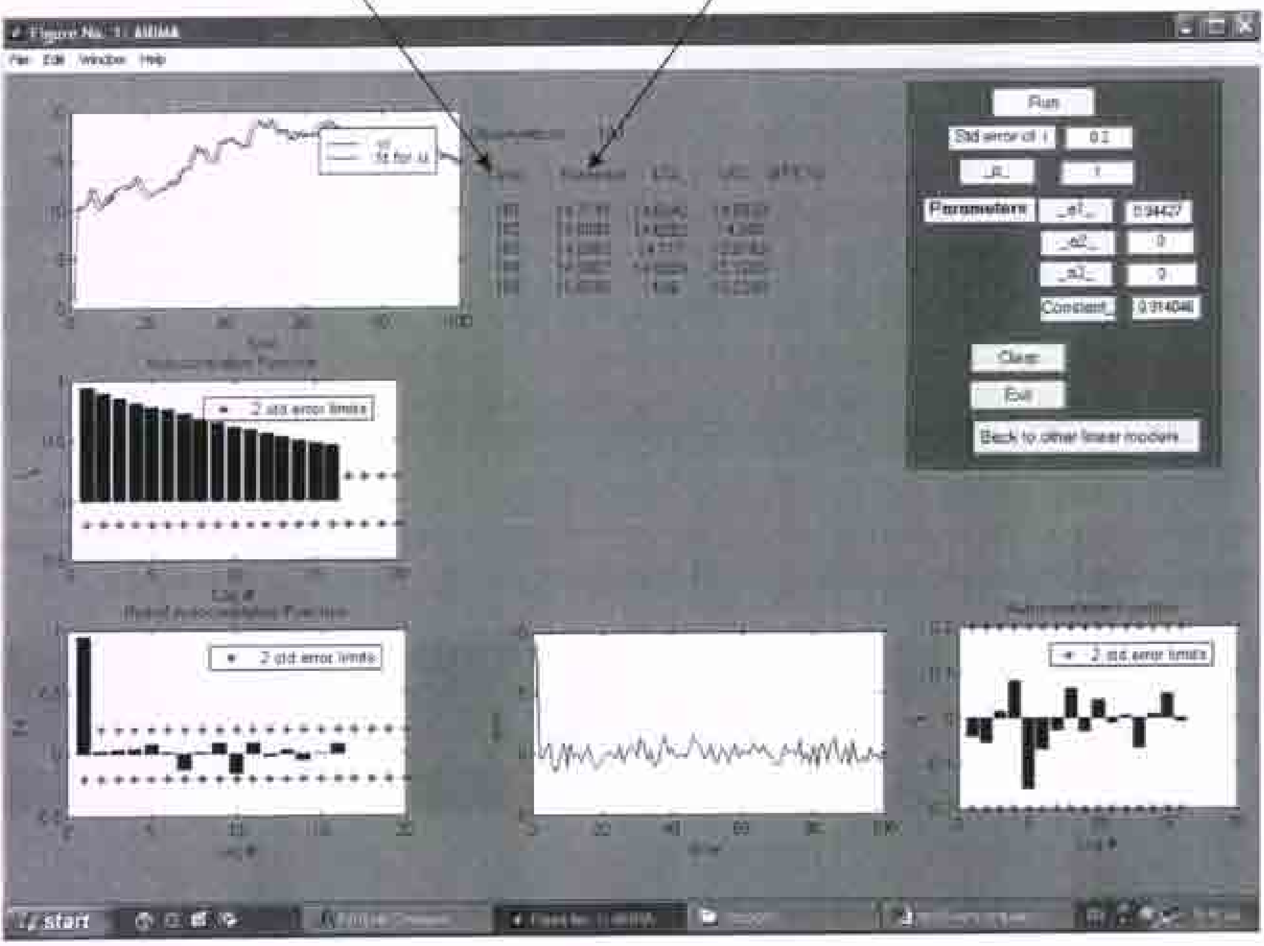
Το υψηλότερο (Upper Confidence Limit – UCL) και χαμηλότερο όριο εμπιστοσύνης (Lower Confidence Limit – LCL) υπολογίζονται από τη σχέση

$X_T \pm 1.96\, s_\varepsilon$ όπου X_T η προβλεφθείσα τιμή και $s_\varepsilon = \sqrt{\dfrac{\sum_{i=1}^{n} \varepsilon_i^2}{r-p}}$

Η μεταβλητή period=(n+1:obs)' Η μεταβλητή Forecast_1

Τα αποτελέσματα της πρόβλεψης εμφανίζονται στην ακόλουθη εικόνα.

10



Βιβλιογραφία

1. Demiris E.N. , Likothanassis S.D. , *(a) ARMAX Model Identification with Unknown Process Order and Time Varying Parameters, (b) Time Varying NonLinear AR Model Identification : Lainiotis's Multi – Model Methodology* (2001)
2. Δημητριάδης Α. *Διοίκηση – Διαχείριση πληροφοριακών Συστημάτων* (1998) εκδ. Νέων Τεχνολογιών Αθήνα.
3. Hanselman Duane, Littlefield Bruce, *Mastering Matlab 5* (1998), εκδ. Prentice Hall
4. Jarret Jeffrey, *Μέθοδοι Προβλέψεων για Οικονομικές – Επιχειρηματικές Αποφάσεις* (1987) εκδ. Guttenberg, Αθήνα, επιμέλεια Ανδρέας Κίντης
5. Κιόχος Πέτρος, *Στατιστική* (1993) εκδ. «INTERBOOKS» , Αθήνα
6. Montgomery Douglas Jonhnson A. Lynwood, Gardiner John, *Forecasting & Time Series Analysis* (1990) 2nd edition, εκδ. McGraw – Hill, Inc.,
7. Μπόρα - Σέντα Ε., Χ Μωυσιάδης, *Εφαρμοσμένη Στατιστική* (1997) εκδ. Ζήτη, Θεσσαλονίκη
8. Παπαδήμας Όθωνας, *Στατιστική ΙΙ*, (1994) Μακεδονικές εκδόσεις.
9. Παπαρρίζος Κωνσταντίνος, *Matlab 6*, εκδ. Ζυγός, Θεσσαλονίκη
10. Ψαράκης Ε.Ζ. *Σημειώσεις μαθήματος Τεχνικές Προβλέψεων και ελέγχου* (2003) .ΤΕΙ Πάτρας, Τμήμα Επιχειρηματικού Σχεδιασμού και Πληροφοριακών Συστημάτων, Ε΄ εξάμηνο.

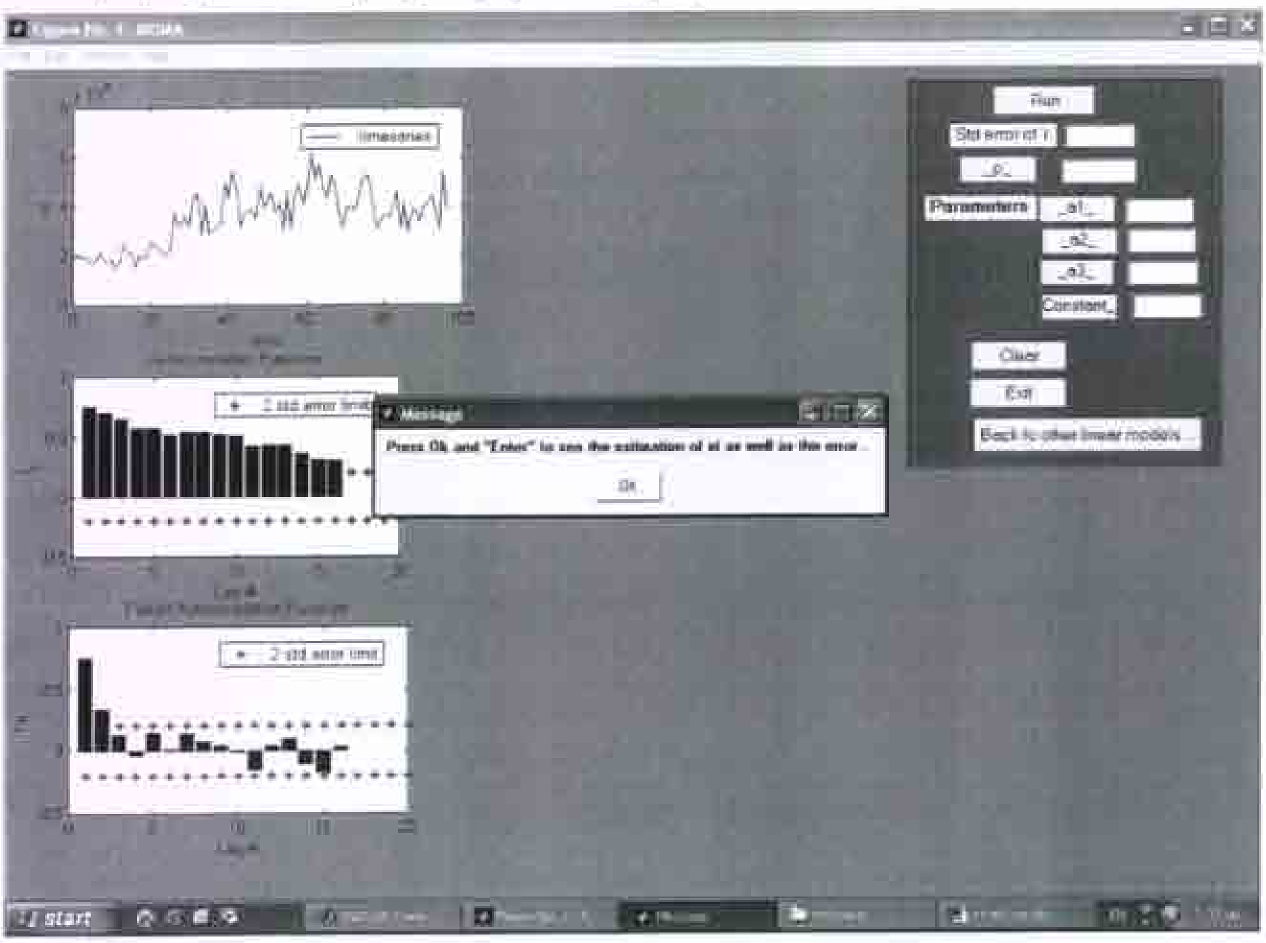
Παράρτημα

Παράδειγμα εκτίμησης και πρόβλεψης

Νοσήλια από το ΙΚΑ του Καραμανδάνειου Νοσοκομείου Παίδων Πάτρας Ιανουάριος 1996 – Δεκέμβριος 2003 (nosila.ask)

	1996	1997	1998	1999	2000	2001	2002	2003
Ιαν	20281,53	25553,51	22272,75	32288,27	39295,23	61256,02	42774,31	41416,43
Φεβ	19423,04	21441,20	37500,54	32729,09	41884,77	51161,00	50746,77	29723,84
Μαρτ	19650,18	21266,91	32336,49	49093,55	45987,15	57073,37	53246,49	39071,40
Απρ	17901,28	15152,19	33120,76	44007,08	37414,84	46155,19	48238,67	38291,75
Μαΐος	16891,77	19297,08	35928,80	54846,06	40918,44	43830,79	37765,32	33827,80
Ιουν	18155,51	19529,27	32220,35	45371,69	30513,31	52492,62	30583,81	40014,78
Ιουλ	21388,85	20012,85	44765,88	31856,60	46167,78	49655,41	35029,78	40876,68
Αυγ	18899,55	25396,68	45165,75	37143,77	31664,42	38400,35	35864,26	44172,95
Σεπτ	17127,42	24303,68	28205,11	38206,05	44923,26	31335,76	36018,21	37522,66
Οκτωβ	18500,22	23426,36	41309,46	38699,45	49213,21	40171,47	41124,50	30555,70
Νοεμβ	15321,20	20522,57	29403,49	35037,77	41176,79	39958,74	45674,70	54125,75
Δεκ	16861,42	20721,07	30321,69	48655,16	47628,91	38776,96	31175,66	38370,68

Γραφική απεικόνιση της σειράς και διαγράμματα ACF & PACF



Σχ. 1

Από το ACF plot (Σχ. 2 Β) βλέπουμε ότι φθίνει αργά και από το PACF (Σχ. 2 Γ) παρατηρούμε πως για το 1^o και το 2^o lag το $r \neq 0$ και μετά η ακολουθία των συντελεστών μερικής αυτοσυσχέτισης πέφτει γρήγορα μέσα στα όρια του διαστήματος εμπιστοσύνης γεγονός που σημαίνει ότι μιλάμε για μια διαδικασία AR (2).

Από το παρακάτω σχήμα βλέπουμε ότι το σύστημα αναγνώρισε την τάξη του υποδείγματος p = 2 και εκτιμήθηκαν οι παράμετροι (Σχ. 2 Γ)

$a1(\varphi_1) = 0{,}454$

$a2(\varphi_2) = 0{,}378$

constant(ξ) = 6261,95

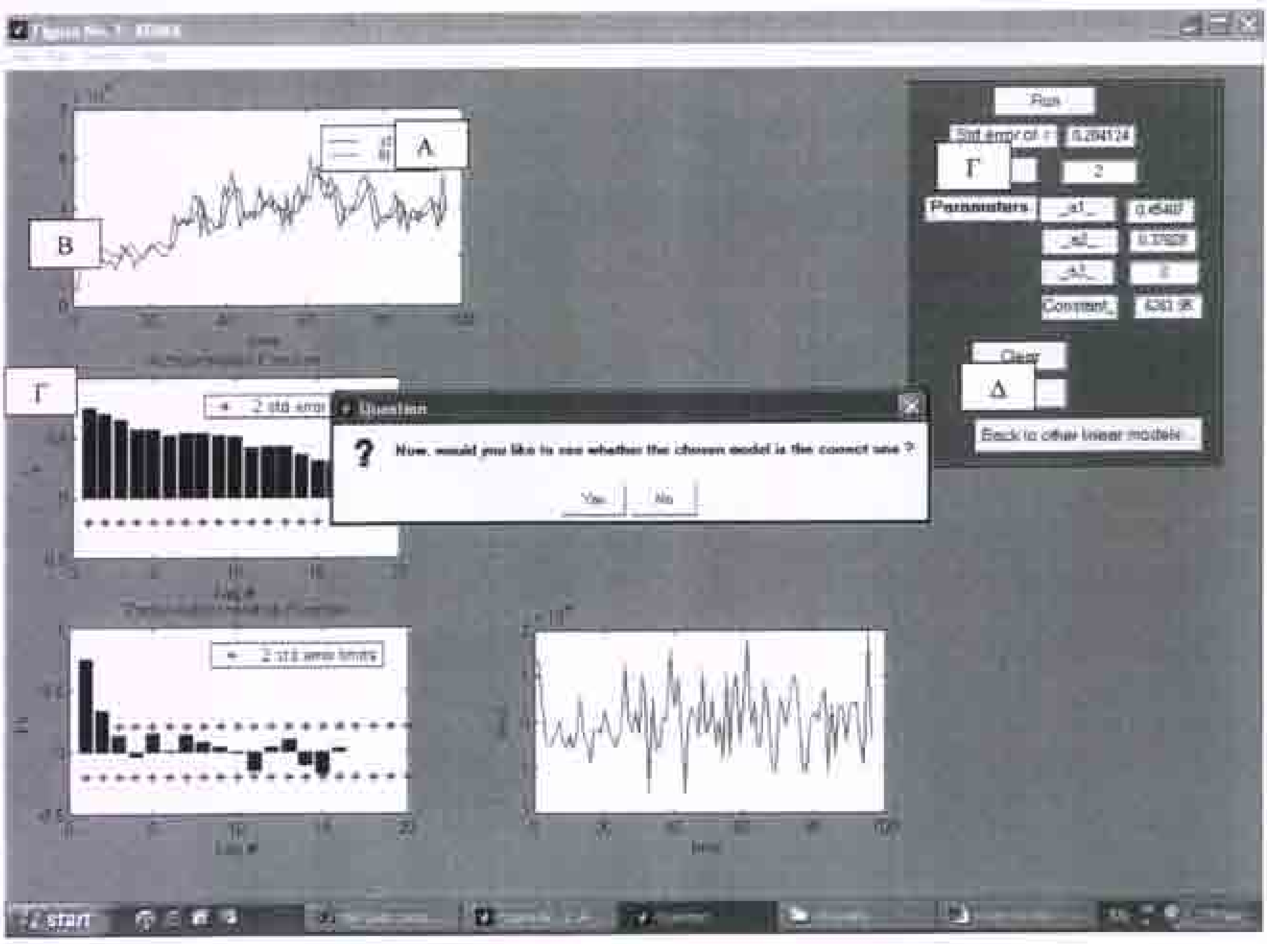
Συνεπώς, AR (2)

$(1 - \varphi_1 B - \varphi_2 B^2)\, x_t = e_t$ ή

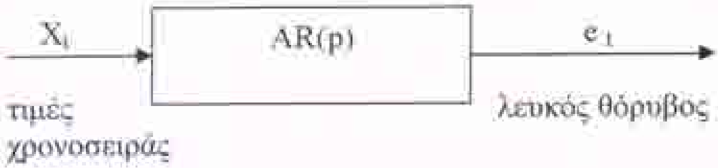
$(1 - 0{,}454\, B - 0.378\, B^2)\, x_t = e_t + 6261{,}95$ ή

$x_t = 1 + 0{,}454\, x_{t-1} + 0.378\, x_{t-2} + e_t + 6261{,}95$

2

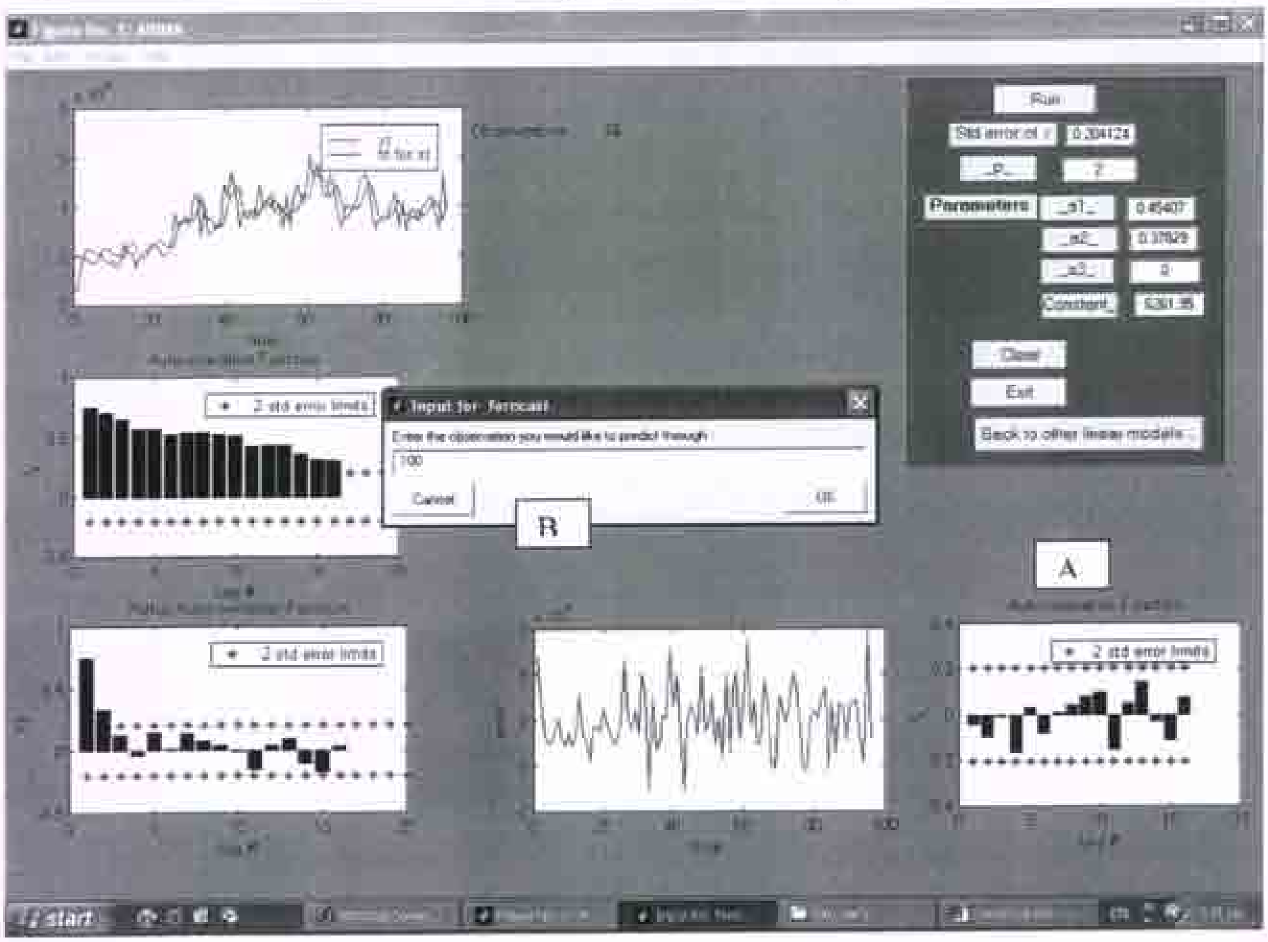


Αμέσως μετά εμφανίζεται ένα μήνυμα που μας ζητά να δούμε αν το μοντέλο που επιλέχθηκε είναι το σωστό . Επιλέγουμε «Yes» και υπολογίζεται η ακολουθία των συντελεστών αυτοσυσχέτισης του σφάλματος.



Απόδειξη ότι έχουμε επιλέξει το κατάλληλο υπόδειγμα είναι τα λάθη (errors) να έχουν τις ιδιότητες του λευκού θορύβου, δηλαδή τα e_t ασυσχέτιστα οπότε οι συντελεστές αυτοσυσχέτισης να είναι 0

Η ιδιότητα φαίνεται και από το σχήμα 3 A (Οι αυτοσυσχετίσεις των σφαλμάτων δεν είναι έξω από τα όρια του τυπικού σφάλματος οπότε θεωρούνται 0.)



..3

Ακολούθως, εκχωρούμε μέχρι ποια χρονική στιγμή θέλουμε να γίνει η πρόβλεψη δηλαδή μέχρι και την 100 χρονική στιγμή (σχήμα 3 B) και τα αποτελέσματα της πρόβλεψης είναι στο σχήμα 4 A όπου μάλιστα δίνεται και το διάστημα εμπιστοσύνης δηλαδή τα όρια μέσα στα οποία με πολύ μεγάλη πιθανότητα (97,5 %) θα διακυμανθεί η προβλεφθείσα τιμή

Μήνας	Έτος	Πρόβλεψη για Νοσήλια (σε €)
JAN	2004	38501,6
FEB	2004	39783,8
MARC	2004	38891,9
APRIL	2004	38972