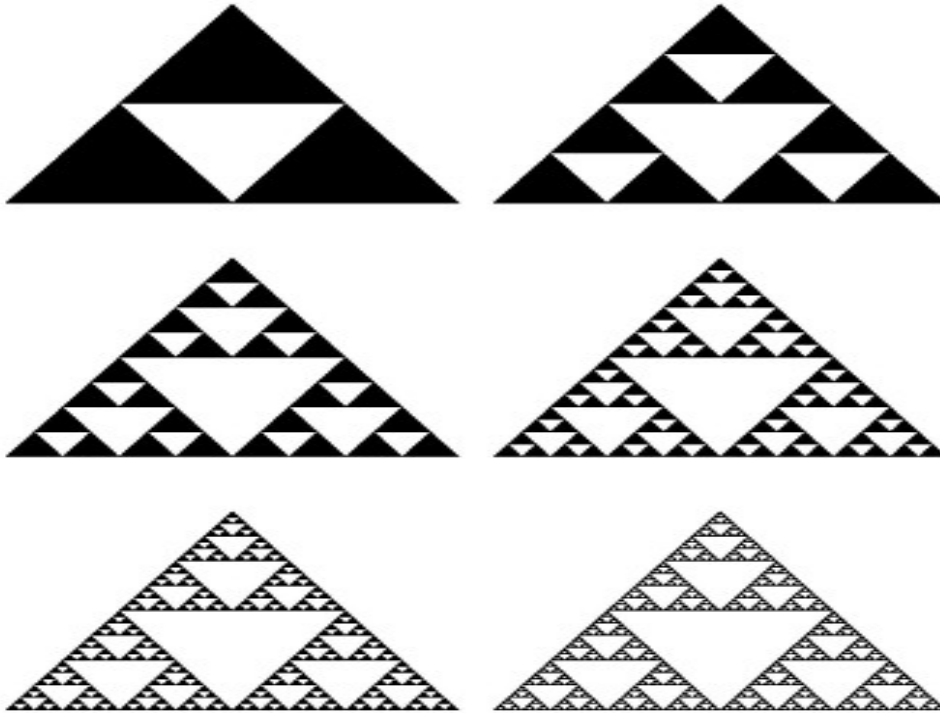


**ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ**  
**Τμήμα Ναυπηγών Μηχανολόγων Μηχανικών**



**«Μοντελοποίηση της Ναυλαγοράς με βάση την θεωρία του Χάους»**

**Διπλωματική Εργασία**

Όνομα σπουδαστή: Κωνσταντίνος Νικολάου

16/9/2010

ΑΘΗΝΑ

Κωνσταντίνος Νικολάου

**Μοντελοποίηση της Ναυλαγοράς με Βάση την Θεωρία του Χάους**

16 Σεπτεμβρίου 2010

**Διπλωματική Εργασία**

**Σχολή Ναυπηγών Μηχανολόγων Μηχανικών**

Συγγραφέας : Κωνσταντίνος Νικολάου

Επιβλέπων : Δημήτριος Β. Λυρίδης      Επ. Καθηγητής ΕΜΠ

ΑΘΗΝΑ

## Σύνοψη

Σε αυτήν την εργασία εφαρμόζονται μέθοδοι για την πρόβλεψη χρονοσειρών σε περίπλοκα πραγματικά δεδομένα που έχουν παρθεί από την ναυλαγορά. Συγκεκριμένα εξετάζονται οι spot τιμές των suezmax tankers της τελευταίας δεκαετίας. Εξετάζεται κατά πόσο παρουσιάζει η χρονοσειρά μη γραμμική ή χαοτική συμπεριφορά. Αναζητούνται επίσης οι βέλτιστες παράμετροι εμβύθισης και γίνεται προσπάθεια εντοπισμού του θορύβου που παρουσιάζεται στα δεδομένα και μείωσής του. Στην έρευνα που πραγματοποιήθηκε καταλήξαμε στο συμπέρασμα ότι η υπό μελέτη χρονοσειρά παρουσιάζει χαοτική συμπεριφορά. Στο ίδιο συμπέρασμα είχε καταλήξει και η διδακτορική διατριβή που πραγματοποίησε η Μαρινίκη Ψιφιά σε συνεργασία με τον καθηγητή Αλέξανδρο Μ. Γουλιέλμο του Πανεπιστημίου Πειραιώς. Στην διδακτορική αυτή διατριβή έγινε η χρήση του BDS test προκειμένου να αποδείξουν ότι 2 χρονοσειρές παρμένες από τις τιμές της ναυλαγοράς παρουσιάζουν μη γραμμική συμπεριφορά και το κατά ποσό αυτές είναι ανεξάρτητες και οι τιμές τους είναι τυχαίες.

Στην διπλωματική αυτή αφού αποδείξαμε ότι η χρονοσειρά μας παρουσιάζει χαοτική συμπεριφορά προχωρήσαμε ένα βήμα παραπέρα. Πραγματοποιήθηκε πρόβλεψη εντός δείγματος με την χρήση 5 μεθόδων σε 7 διαφορετικά σημεία της χρονοσειράς. Οι 5 μέθοδοι προβλέψεων που εφαρμόστηκαν είναι τοπικές γραμμικές, ακτινικές συναρτήσεις βάσης και εκτίμησης πυκνότητας πυρήνα. Τα αποτελέσματα αυτών των προβλέψεων παρατίθενται στο κεφάλαιο 4. Κρίνοντας από τα συμπεράσματα που εξάγονται από αυτές τις προβλέψεις μπορούμε να πούμε πως δεν μπορούμε να τις εφαρμόσουμε στην πράξη μιας και οι τιμές που μας δίνουν δεν είναι ικανοποιητικές.

## Κατάλογος Σχημάτων

- Σχήμα 1: σημεία της λογιστικής εξίσωσης ... σελ. 9
- Σχήμα 2: Δέντρο διπλασιασμού περιόδου ... σελ. 11
- Σχήμα 3: Διάγραμμα διακλάδωσης ... σελ. 11
- Σχήμα 4: σημεία της εξίσωσης Mackey-Glass στην πάροδο του χρόνου ...σελ. 12
- Σχήμα 5: εμπυθισμένη εκδοχή της συνάρτησης Mackey-Glass... σελ. 13.
- Σχήμα 6: παρουσίαση του συστήματος εξισώσεων (Lorenz, 63) ... σελ.14
- Σχήμα 7,8,9:παράσταση του συστήματος Lorenz στις διάφορες διαστάσεις ... σελ. 14, 15
- Σχήμα 10 : η έννοια του ελκυστή σχηματικά ... σελ. 20
- Σχήμα 11 : ελκυστής σημείου ... σελ. 21
- Σχήμα 12:ελκυστής οριακού κύκλου ... σελ. 22
- Σχήμα 13: ελκυστής δύο περιόδων... σελ. 22
- Σχήμα 14: Χαοτικός ελκυστής ... σελ. 23
- Σχήμα 15: Πλωτάρισμα χρονοσειρών της λογιστικής εξίσωσης και της Tent map ... σελ. 25
- Σχήμα 16: Χαρακτηριστική απόκλιση στο χρόνο 2 γειτονικών τροχιών του ελκυστή ... σελ 26.
- Σχήμα 17: Διάγραμμα διάστασης συστήματος –διάστασης εμπύθισης ... σελ. 32
- Σχήμα 18: Δισδιάστατη αναπαράσταση των 5 σημείων και τις ακτίνας  $r=\varepsilon$  από την οποία καθορίζεται ποια είναι τα σημεία που θεωρούνται κοντινότεροι γείτονες για τα σημεία a) 1 και b) 2.
- Σχήμα 19: Διάγραμμα χώρου-χρόνου ... σελ 39.
- Σχήμα 20: Παρουσίαση της υπό μελέτη χρονοσειράς ... σελ 52.
- Σχήμα 21: Παρουσίαση του φάσματος ισχύος για τα πειραματικά δεδομένα ... σελ. 53
- Σχήμα 22: Διάγραμμα ολοκληρώματος συσχέτισης έναντι λογαρίθμου της ακμής  $\log r - \log c(r,m)$  ... σελ 5
- Σχήμα 23: Διάγραμμα των κλίσεων έναντι του λογαρίθμου της ακμής  $\log r - \text{slope}$ ... σελ 56
- Σχήμα 24: Διάγραμμα διάστασης συσχέτισης έναντι διάστασης εμπύθισης ...σελ. 57

- Σχήμα 25: Διάγραμμα  $S(a)$ - $a$  και κλίσης ... σελ. 58
- Σχήμα 26: Παρουσίαση των προβλέψεων για την πρώτη κοπή ... σελ. 60
- Σχήμα 27: Παρουσίαση των προβλέψεων για την δεύτερη κοπή ... σελ. 61
- Σχήμα 28: Παρουσίαση των προβλέψεων για την Τρίτη κοπή ...σελ. 62
- Σχήμα 29: Παρουσίαση των προβλέψεων για την τέταρτη κοπή ...σελ. 63
- Σχήμα 30: Παρουσίαση των προβλέψεων για την Πέμπτη κοπή ... σελ. 64
- Σχήμα 31: Παρουσίαση των προβλέψεων για την έκτη κοπή ... σελ. 65
- Σχήμα 32: Παρουσίαση των προβλέψεων για την έβδομη κοπή... σελ. 66

### **Κατάλογος πινάκων**

- Πίνακας 1: Διάσταση εμπύθισης σε σχέση με την διάσταση συσχέτισης...σελ .59
- Πίνακας 2: Πίνακες υπολογισμού του λογαρίθμου του ολοκληρώματος συσχέτισης ...σελ .59
- Πίνακας 3 : Κλίσεις έναντι του λογαρίθμου της ακμής ... σελ. 60
- Πίνακας 4: Πίνακας της διάστασης εμπύθισης σε σχέση με τη διάσταση συσχέτισης... σελ . 61
- Πίνακας 5: Πίνακας τιμών της διάστασης εμπύθισης και τις τιμές των LLE και  $S(a)$ ... σελ. 63
- Πίνακας 6: Αποτελέσματα των προβλέψεων και με τις 5 μεθόδους στην πρώτη κοπή... σελ. 65
- Πίνακας 7: Πίνακας προβλέψεων στην δεύτερη κοπή δεδομένων... σελ. 66
- Πίνακας 8: Πίνακας προβλέψεων στην τρίτη κοπή δεδομένων ... σελ. 67
- Πίνακας 9: Πίνακας προβλέψεων στην τέταρτη κοπή δεδομένων ...σελ. 68
- Πίνακας 10: Πίνακας προβλέψεων στην πέμπτη κοπή δεδομένων ...σελ. 69
- Πίνακας 11: Πίνακας προβλέψεων στην έκτη κοπή δεδομένων ...σελ. 70
- Πίνακας 12: Πίνακας προβλέψεων στην έβδομη κοπή δεδομένων ...σελ. 71
- Πίνακας 13,14,..., 19: Πίνακες μέτρησης της τάσης (trend) για τα 7 πειράματα ... σελ. 72 έως 78

## Εισαγωγή

Σε αυτήν την εργασία πραγματοποιήθηκε έρευνα και χρησιμοποιήθηκαν μέθοδοι για την πρόβλεψη χρονοσειρών πραγματικών και περίπλοκων χρονοσειρών από τον πραγματικό κόσμο χρησιμοποιώντας αρχές της θεωρίας του χάους και τεχνητή νοημοσύνη. Με μια πιο προσεκτική ματιά οι μεθοδολογίες που χρησιμοποιήθηκαν απαιτούσαν μια γκάμα από παραμέτρους που έπρεπε να βρεθούν μέσω της εμπειρίας μιας και δεν υπήρχαν αναλυτικές μέθοδοι για την εύρεση τους. Η επιλογή των τιμών των παραμέτρων αυτών δεν μπορεί να δικαιολογηθεί με βάση καμία θεωρία, ούτε ο ερευνητής μπορεί να ισχυριστεί ότι τα αποτελέσματά του είναι πραγματικά προϊόν τυφλών δοκιμών.

Στο κεφάλαιο 1 γίνεται αναφορά στους τύπους των διαφόρων χρονοσειρών και αναφορά στις αναλυτικές τεχνικές και στις προβλέψεις που χρησιμοποιήθηκαν σε αυτές. Το κεφάλαιο αυτό μπορεί να θεωρηθεί σαν μια εισαγωγή στις βασικές αρχές και έννοιες της θεωρίας του χάους.

Στο κεφάλαιο 2 παρουσιάζονται αναλυτικότερες μέθοδοι για την εύρεση της παρουσίας χάους. Μελετάται η παρουσία μακροχρόνιας μνήμης στις υπό εξέταση χρονοσειρές, και με την μελέτη του φάσματος ισχύος προσπαθούμε να προσδιορίσουμε εάν έχει εμφανή σημάδια χάους. Επίσης μελετώνται τα δυναμικά συστήματα από τα οποία πηγάζουν οι χρονοσειρές. Γίνεται παρουσίαση των ειδών των ελκυστών και οι ιδιότητές τους. Αναλύεται η εμφύθιση σαν μέσο ανάλυσης των χρονοσειρών, οι εκθέτες Lyapunov, ο εκθέτης Hurst για την μέτρηση της προβλεψιμότητας των χρονοσειρών. Στο κεφάλαιο αυτό μελετώνται και διάφορες γεωμετρικές μέθοδοι όπως οι λανθασμένοι κοντινότεροι γείτονες. Τέλος γίνεται αναφορά στις διαστάσεις τα είδη τους, όπου διάσταση εννοούμε την μέτρηση της διάστασης του συστήματος που προήλθε η χρονοσειρά, και παρουσιάζονται μέθοδοι για την μείωση του θορύβου.

Στο 3 κεφάλαιο ασχολούμαστε με τις προβλέψεις και αναλύονται οι μέθοδοι για την πραγματοποίησή τους. Τοπικές γραμμικές προβλέψεις είναι ένα από τα είδη των μεθόδων αυτών με την χρήση της μεθόδου των κοντινότερων γειτόνων. Άλλο είδος είναι οι ακτινικές συναρτήσεις βάσης και Εκτίμηση πυκνότητας πυρήνα. Τέλος γίνεται αναφορά στις χρηματοοικονομικές χρονοσειρές. Είναι παρόμοιες με αυτές της ναυλαγοράς. Κύριο ενδιαφέρον παρουσιάζεται για το εάν ισχύει ο τυχαίος περίπατος. Έγιναν προβλέψεις με διάφορες μεθόδους εκτός δείγματος και με άλλες εντός και εκτός δείγματος.

Στο κεφάλαιο 4 γίνεται παρουσίαση όλων των πειραματικών αποτελεσμάτων που πραγματοποιήθηκαν με τη χρήση του λογισμικού NLTSΑ. Βρέθηκε το φάσμα ισχύος που είναι μια πρώτη προσπάθεια για εύρεση ύπαρξης ή μη χαοτικής συμπεριφοράς. Επίσης προσδιορίστηκε η διάσταση συσχέτισης που είναι μια από τις πιο σημαντικές διαστάσεις. Επίσης προσδιορίσαμε τον μέγιστο εκθέτη Lyapunov με την χρήση μιας

εύρωστης μεθόδου. Τέλος παραθέτονται οι προβλέψεις που πραγματοποιήθηκαν για την χρονοσειρά μας με την χρήση 5 μεθόδων.

Στο τέλος της διπλωματικής παραθέτονται κάποιες ιδέες για περαιτέρω έρευνα.

## Κεφάλαιο 1

### 1.1 Υπόβαθρο της εργασίας

Τα τελευταία χρόνια παρουσιάστηκε έντονο ενδιαφέρον στις εφαρμογές της θεωρίας του χάους σε μια ποικιλία πραγματικών χρονοεξαρτημένων συστημάτων, και η ικανότητα που παρουσιάζει αυτός ο κλάδος των μαθηματικών να παρέχει τάξη από εκεί που φαινομενικά υπάρχει αταξία. Υπάρχουν πολλές χρονοσειρές, τόσο στις φυσικές επιστήμες όσο και στα οικονομικά που είχαν αποδειχθεί δύσκολο να αναλυθούν με τις συμβατικές γραμμικές μεθόδους, οι οποίες τώρα μοντελοποιούνται χρησιμοποιώντας μη γραμμικές και μη παραμετρικές μεθόδους. Ο Mandelbrot ξεκινάει το βιβλίο του *The Fractal Geometry of Nature* (Mandelbrot, 1983) κατακρίνοντας την επιστήμη που προσπαθεί να εντάξει την φύση σε καλούπια τα οποία είναι συμβατά με ήδη έτοιμα μαθηματικά μοντέλα της Ευκλείδειας γεωμετρίας. Χαρακτηριστικά αναφέρει ‘ τα σύννεφα δεν είναι σφαίρες, τα βουνά δεν είναι κώνοι, οι ακτογραμμές δεν είναι κύκλοι, και ο φλοιός δεν είναι ομαλός ούτε η αστραπή ταξιδεύει σε ευθεία γραμμή’. Συνεχίζοντας ισχυρίζεται πως ο πραγματικός κόσμος τείνει να είναι κυριαρχημένος από μορφές και διαδικασίες που δεν είναι ούτε ομαλές ούτε συνεχείς, ο οποίος δεν ενδίδει εύκολα στην ανάλυση ήδη έτοιμων μαθηματικών τεχνικών.

Μαθηματικές μέθοδοι έχουν αναπτυχθεί για να χαρακτηρίσουν και σε κάποιο βαθμό να αναλύσουν αυτές τις χρονοσειρές. Λόγω των πεπερασμένων δεδομένων τα οποία έχουμε για τις εν λόγω χρονοσειρές, η παρουσία θορύβου και η γραμμικότητα που τις χαρακτηρίζει κάνουν την προσπάθεια για ανάλυση ακόμη πιο δύσκολη και μας οδηγούν στην χρήση εμπειρικών μεθόδων.

Σε αυτό το σημείο πρέπει να αναφερθούμε στην διδακτορική διατριβή της Μαρινίκης Ψηφιά υπό την καθοδήγηση του κ. Αλέξανδρου Γουλιέλμου του Πανεπιστημίου Πειραιώς του τμήματος Ναυτιλιακών σπουδών. Στην διατριβή αυτή τίθεται το ερώτημα εάν οι χρονοσειρές των τιμών των ναύλων του ξηρού φορτίου για χρονοναύλωση και για ναύλωση ανά ταξίδι είναι πανομοιότυπα και ανεξάρτητα διαμοιρασμένες και το εάν παρουσιάζουν μη γραμμική εξάρτηση. Για να το επιτύχουν αυτό, βασίστηκαν στο έργο των Brock, Dechert και Scheinkman (Brock W. et al, 1987), στο λεγόμενο BDS τεστ. Τα αποτελέσματα του τεστ αυτού έδειξαν πως οι τιμές των χρονοσειρών δεν είναι τυχαίες και ότι πράγματι υπάρχει μη γραμμική εξάρτηση. Με βάση αυτήν την εργασία τους αλλά και προγενέστερες κατάφεραν να αποδείξουν ότι στην αγορά της χρονοναύλωσης αλλά και των ναύλων ανά ταξίδι υπάρχει μακροχρόνια μνήμη.



## 1.2 Ανασκόπηση των βασικών αρχών

Οι ιδέες που υποστηρίζουν αυτήν την εργασία πηγάζουν από διαφορετικές αρχές. Μερικές περιοχές μελέτης μπορούν να αποφύγουν αναφορά στα μαθηματικά, αλλά πολλές τεχνικές που περιγράφονται πήγασαν όχι από μαθηματικούς αλλά από βιολόγους και φυσικούς προσπαθώντας να αποκτήσουν πληροφορίες από δειγματοληπτούμενες χρονοσειρές οι οποίες ήταν μικρές σε μήκος και θορυβώδεις.

### 1.2.1 Χρονοσειρές

Η έννοια της χρονοσειράς είναι ευρύτατα διαδεδομένη στον επιστημονικό χώρο της στατιστικής, στην ανάλυση σημάτων, στην οικονομετρία και στις οικονομικές επιστήμες. Με τον όρο χρονοσειρά εννοούμε τις τιμές ενός μεγέθους που το μετρήσαμε σε τακτά χρονικά διαστήματα. Παραδείγματα χρονοσειρών αποτελούν η καθημερινή τιμή κλεισίματος του δείκτη Dow Jones και ο ετήσιος όγκος ροής του ποταμού Νείλου στο φράγμα του Aswan. Τα στοιχεία μιας χρονοσειράς έχουν μια φυσική χρονική τάξη. Αυτό καθιστά την ανάλυση χρονοσειρών ευδιάκριτη από άλλα κοινά προβλήματα ανάλυσης στοιχείων, στα οποία δεν υπάρχει καμία φυσική τάξη των παρατηρήσεων. Ένα μοντέλο χρονοσειρών θα αντικατοπτρίζει το γεγονός ότι οι παρατηρήσεις οι οποίες βρίσκονται σε κοντινά χρονικά διαστήματα θα είναι περισσότερο συνδεδεμένες μεταξύ τους (καλύτερη συσχέτιση) από ότι παρατηρήσεις που βρίσκονται πιο μακριά στον χρόνο. Οι μέθοδοι για τις αναλύσεις χρονοσειρών μπορούν να διαιρεθούν σε δύο κατηγορίες: μέθοδοι πεδίου συχνότητας (frequency domain) και μέθοδοι χρονοπεριοχών (time domain). Τα πρώτα περιλαμβάνουν τη φασματική ανάλυση και τα τελευταία περιλαμβάνουν την ανάλυση αυτοσυσχέτισης και συσχέτισης.

Τα μοντέλα των χρονοσειρών έχουν πολλές μορφές και αντιπροσωπεύουν διαφορετικές στοχαστικές διαδικασίες. Τα μοντέλα αυτά χωρίζονται σε 3 βασικές κατηγορίες: στα αυτοανάδρομα μοντέλα (AR), τα ενσωματωμένα μοντέλα (I), και τα μοντέλα κινούμενου μέσου όρου (MA). Αυτές οι τρεις κατηγορίες εξαρτώνται γραμμικά από τα προηγούμενα χρονικά στοιχεία. Οι συνδυασμοί αυτών των μοντέλων παράγουν τον αυτοανάδρομο κινούμενο μέσο όρο (ARMA) και τα αυτοανάδρομα ενσωματωμένα πρότυπα κινούμενου μέσου όρου (ARIMA). Το αυτοανάδρομο fractionally ενσωματωμένο πρότυπο κινούμενου μέσου όρου (ARFIMA) γενικεύει τα προηγούμενα τρία. Ενδιαφέρον παρουσιάζει η ύπαρξη μη γραμμικής εξάρτησης των δεδομένων μιας χρονοσειράς από προγενέστερα στο χρόνο σημεία της, εν μέρη διότι υπάρχει περίπτωση να παραχθεί χαοτική χρονοσειρά.

### 1.2.2 Χαοτικές χρονοσειρές

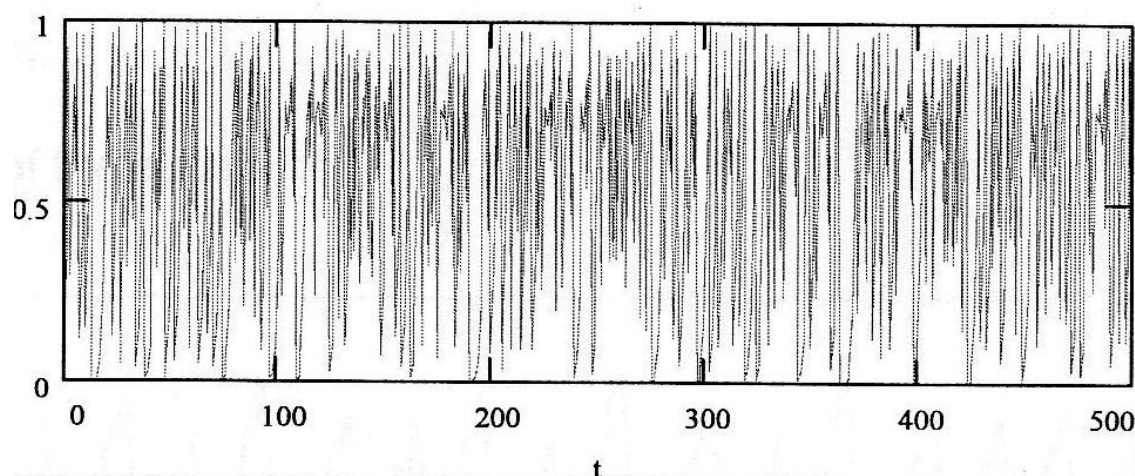
Πολλά συστήματα στον φυσικό κόσμο είναι γνωστό ότι παρουσιάζουν χάος η μη γραμμική συμπεριφορά, η πολυπλοκότητα των οποίων ήταν τόσο σημαντική που τα έκανε να θεωρούνται τυχαία. Απλά πειράματα στα οποία εντοπίστηκε χάος είναι η παρατήρηση μια στρόφιγγας νερού που στάζει ή ακόμη και σε απλά ηλεκτρικά

κυκλώματα. Εκτός εργαστηρίου χαοτική συμπεριφορά έχει εμφανιστεί σε χρονοσειρές κλιματικών αλλαγών στην υδροδυναμική, στην αστροφυσική, στην ιατρική και σε πολλούς άλλους τομείς της επιστήμης. Η διαλεύκανση αυτών των συστημάτων έχει βοηθηθεί από την ανακάλυψη, κυρίως αυτόν τον αιώνα, των μαθηματικών εκφράσεων που παρουσιάζουν παρόμοιες τάσεις(trend). Οι πιο διάσημες και παλαιότερες από αυτές είναι η λογιστική συνάρτηση, η οποία αρχικά χρησιμοποιήθηκε σαν ένα μοντέλο της αύξησης του πληθυσμού.

Η λογιστική εξίσωση:

$$X_{n+1} = \alpha x_n(1 - x_n), n = 0, 1, 2, \dots \quad (1)$$

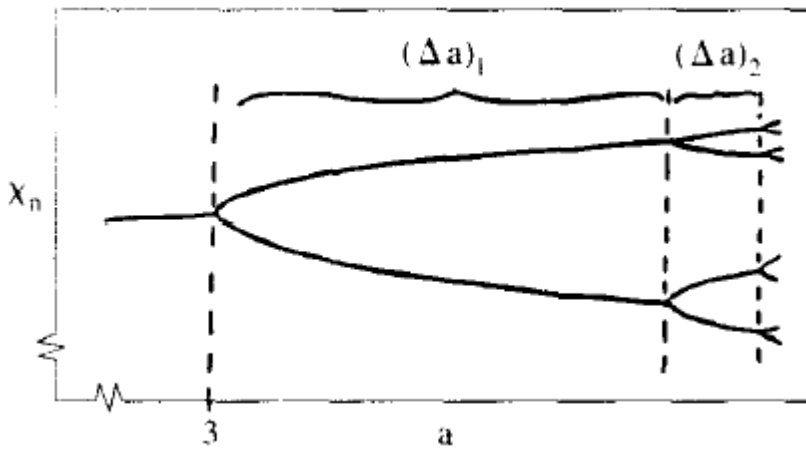
Η τιμή  $\alpha$  μπορεί να πάρει τιμές από 1 μέχρι 4 και το  $X_n$  τιμές μεταξύ 0 και 1. Αυτή η συνάρτηση έχει απασχολήσει ιδιαίτερα την βιβλιογραφία του χάους επειδή παρέχει στοιχεία τα οποία παρουσιάζουν ευθέως την μετάβαση από την τάξη στο χάος.



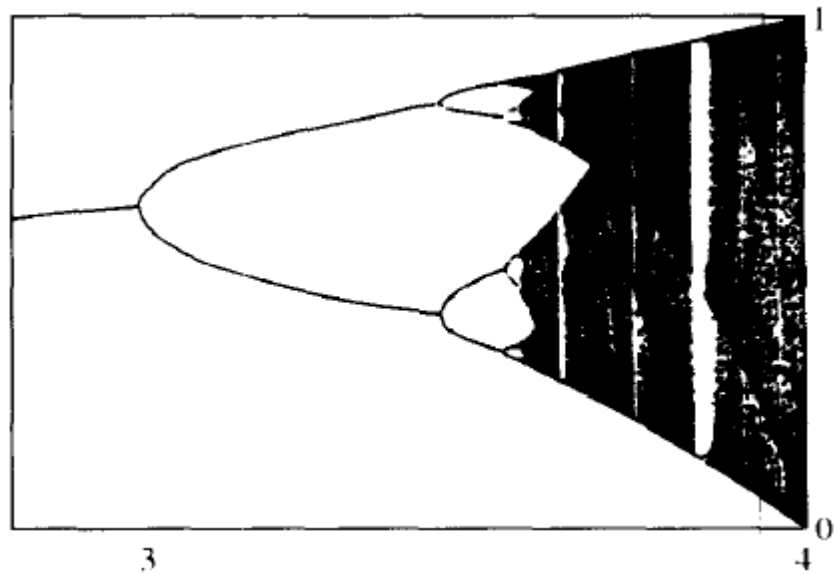
**Σχήμα 1:** Τα πρώτα 500 σημεία της λογιστικής εξίσωσης για  $\alpha=1$

Εάν παραστήσουμε την γραφική παράσταση που παρουσιάζει την σχέση μεταξύ του  $x_n$  και του  $\alpha$  θα δημιουργηθεί το ακόλουθο σχήμα το οποίο ονομάζεται δέντρο διπλασιασμού περιόδου ή αλλιώς bifurcation diagram (σχήμα 2). Ένα αξιοπρόσεκτο χαρακτηριστικό του διαγράμματος αυτού είναι ότι αν πάρουμε το λόγο του  $(\Delta\alpha)_1$  ως προς το  $(\Delta\alpha)_2$  είναι ίσος με 4,6692... του  $(\Delta\alpha)_2$  ως προς το  $(\Delta\alpha)_3$  είναι πάλι 4,6692... και αυτό συνεχίζεται επ' άπειρο. Το φαινόμενο αυτό παρατηρήθηκε για πρώτη φορά από τον Feigenbaum το 1976. Αυτή η αναλογία αναφέρεται ως η σταθερά του Feigenbaum. Η σταθερά Feigenbaum αποτελεί την βάση για την πρόβλεψη της συμπεριφοράς της λογιστικής εξίσωσης σε σχέση με οποιαδήποτε τιμή παίρνει το  $\alpha$ . Ακόμα ο Feigenbaum έκανε μια σημαντική ανακάλυψη. Όταν το  $\alpha$  παίρνει την τιμή 3,5699456... η συνάρτηση γίνεται χαοτική και από κει και πέρα δεν παρουσιάζει προβλεψιμότητα (σχήμα 3). Η τιμή αυτή έχει ονομαστεί σημείο του Feigenbaum και η σημασία της φαίνεται καθαρά στο σχήμα 3. Το σημείο του Feigenbaum χωρίζει το διάγραμμα σε 2 πολύ χαρακτηριστικά τμήματα, στο τμήμα διπλασιασμού περιόδου στα αριστερά και στο τμήμα που κυριαρχείται από χάος, στα δεξιά (Peitgen et al.1992,

σελ 588). Αυτός ο συγκερασμός χάους και τάξης μέσα στο ίδιο ντετερμινιστικό σύστημα είναι καθαρά σε αντίθεση με την Νευτώνεια θεώρηση των πραγμάτων. Καταρρίπτει την θεώρηση πως ντετερμινισμός απαραίτητα σημαίνει και προβλεψιμότητα. Αλλά ακόμη και στην περιοχή που έχουμε χάος παρουσιάζονται μικρά 'παράθυρα προβλεψιμότητας' με πολύ μικρές αυξομειώσεις του  $a$ . Αυτά τα παράθυρα είναι οι άσπρες κατακόρυφες γραμμές που φαίνονται στο διάγραμμα στην δεξιά πλευρά(σχήμα 3), οι οποίες οφείλονται στην εμφάνιση περιοδικής κυκλικής συμπεριφοράς. Τέλος πρέπει να αναφέρουμε πως το διάγραμμα αυτό παρουσιάζει αυτοομοιότητα.



Σχήμα 2: Δέντρο διπλασιασμού περιόδου



Σχήμα 3: Διάγραμμα διακλάδωσης

Άλλες χασοτικές χρονοσειρές παράγονται εξετάζοντας την συνάρτηση Mackey-Glass (Mackey,77)

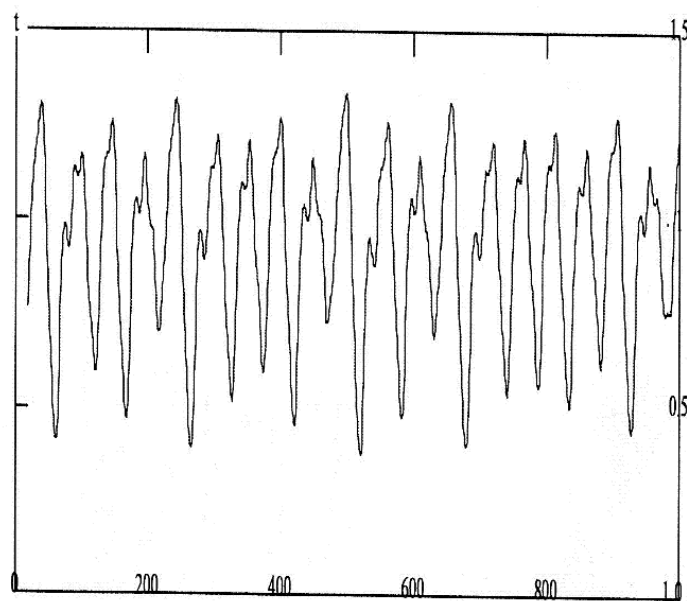
Η συνάρτηση Mackey-Glass στην διακριτική της μορφή :

Οι ακόλουθες επιλογές μεταβλητών και αρχικών συνθηκών προκαλούν χασοτική συμπεριφορά.

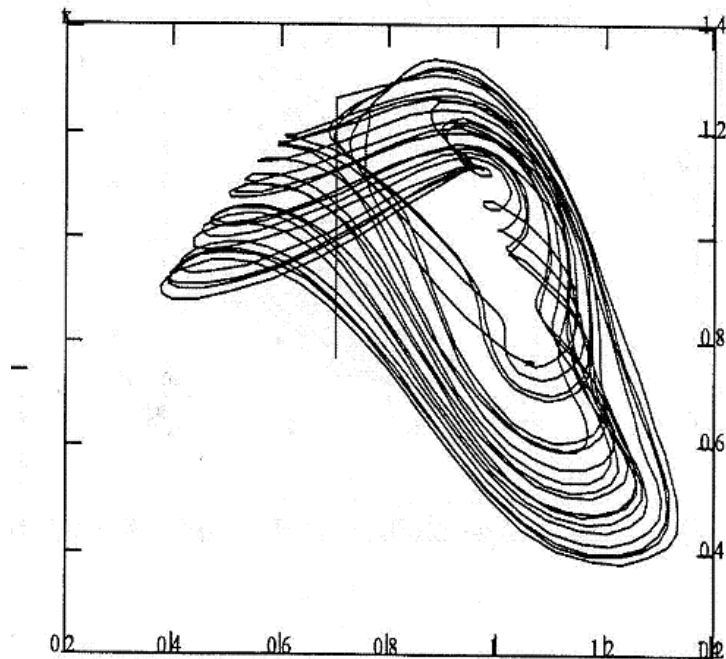
$$a = 0.2, b = 0.9, c = 10, s = 18. \quad x_0 \text{ to } x_{18} = 0.7 \quad (2)$$

Η συνάρτηση :

$$x_t = bx_{t-1} + a \frac{x_{t-s}}{1 + x_{t-s}^c} \quad (3)$$



**Σχήμα 4:** Η εξέλιξη στον χρόνο της συνάρτησης Mackey-Glass

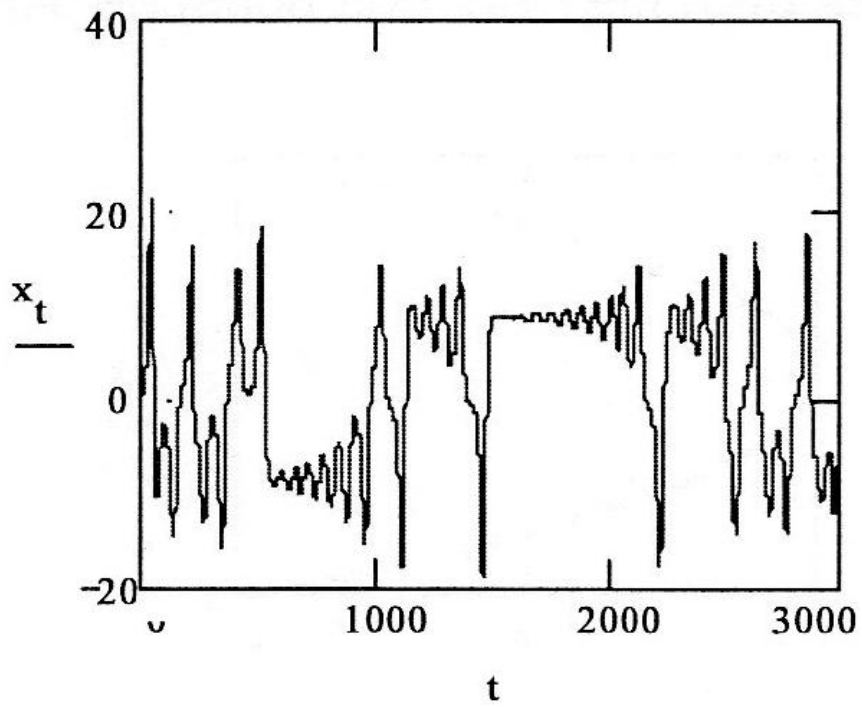


**Σχήμα 5:** Μία εμβυθισμένη εκδοχή της συνάρτησης Mackey-Glass

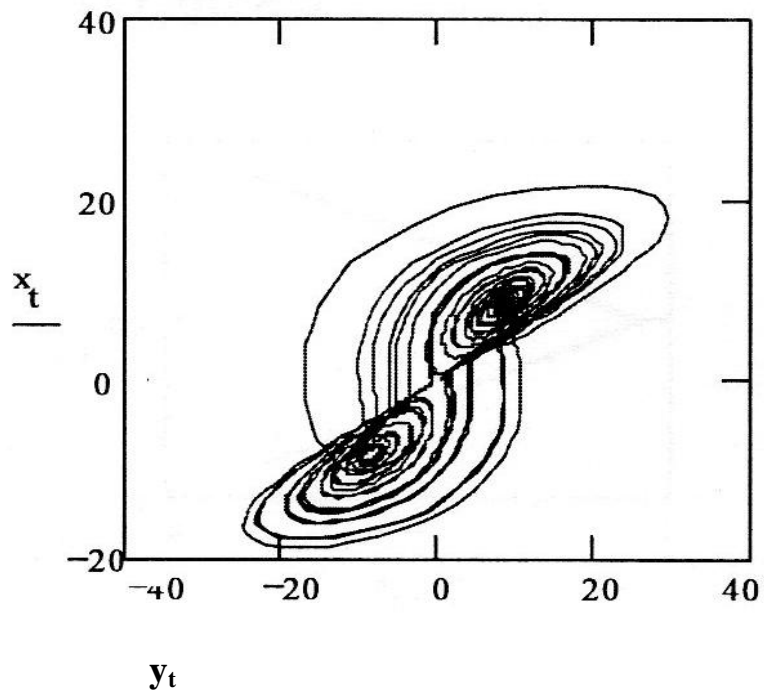
Σαν τελευταίο παράδειγμα υπάρχει ένα σύστημα από εξισώσεις που πρωτοχρησιμοποιήθηκαν από τον Lorenz (Lorenz,63) στην προσπάθεια του να μοντελοποιήσει τις απλές ατμοσφαιρικές αλληλεπιδράσεις :

Αρχικές συνθήκες :

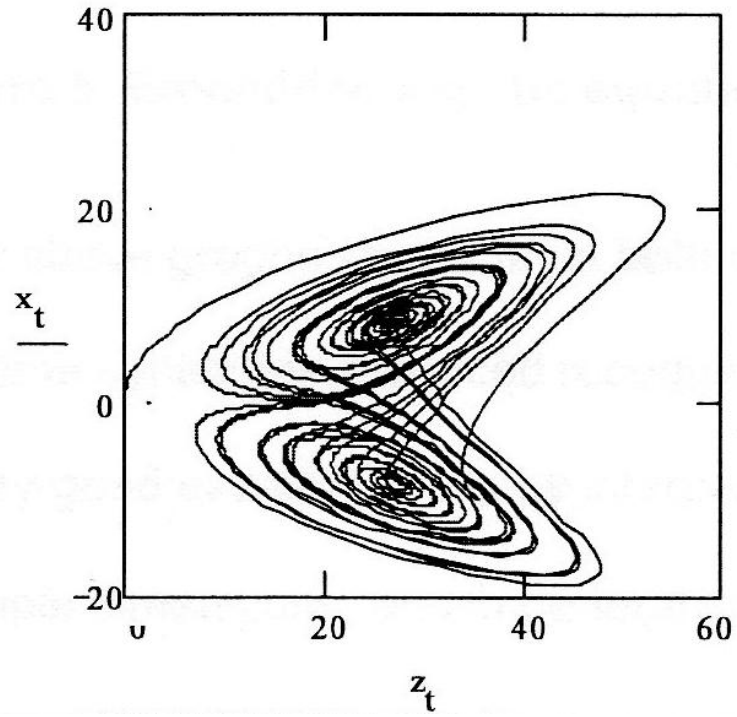
$$c = \frac{8}{3} \quad \text{delta} = 0.01 \quad t = 0,3000 \quad b = 28 \quad a = 10 \quad (4)$$



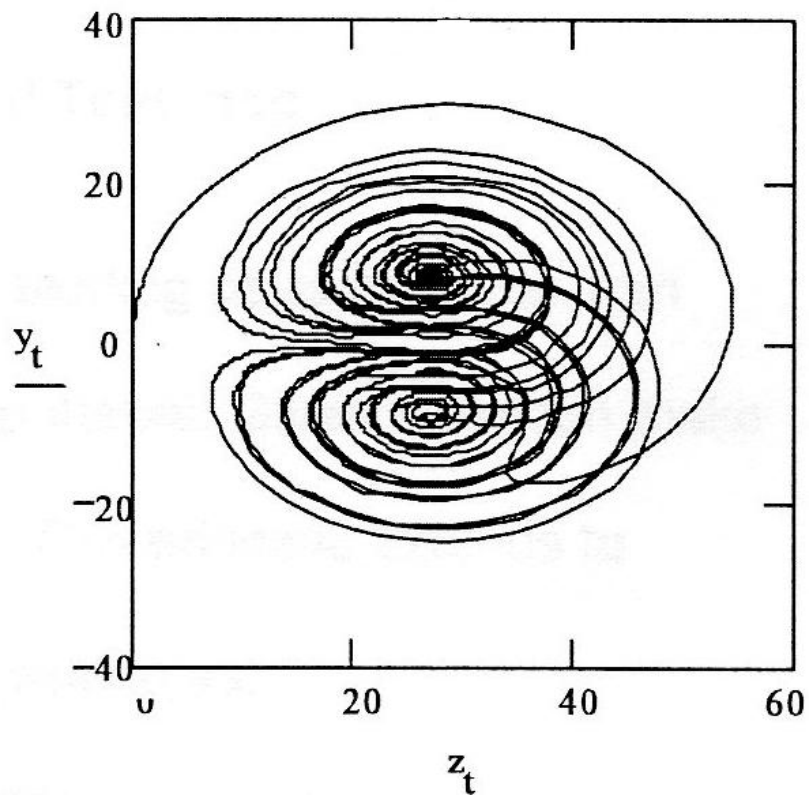
Σχήμα 6: Η συνάρτηση Lorenz με τις δεδομένες αρχικές συνθήκες



Σχήμα 7: Εμβυθισμένη εκδοχή της συνάρτησης Lorenz στους άξονες  $x_t$ ,  $y_t$



Σχήμα 8: Εμβυθισμένη εκδοχή της συνάρτησης Lorenz στους άξονες  $x_t$ ,  $z_t$



Σχήμα 9: Εμβυθισμένη εκδοχή της συνάρτησης Lorenz στους άξονες  $y_t$ ,  $z_t$

Οι παραπάνω χρονοσειρές παρουσιάστηκαν γιατί δείχνουν πως η πολύπλοκη συμπεριφορά μπορεί εύκολα να παραχθεί από απλές συναρτήσεις με μη γραμμικά στοιχεία, αλλά και γιατί αργότερα θα χρησιμοποιηθούν σαν σειρές αναφοράς .

## **Κεφάλαιο 2**

### **Εις βάθος ανάλυση της θεωρίας του χάους**

#### **2.1.1 Εύρεση μακροχρόνιας μνήμης**

Με τον όρο μακροχρόνια μνήμη εννοούμε συνοπτικά την εξής ιδιότητα, που έχουν ορισμένες χρονοσειρές: η τιμή της μεταβλητής για μια ορισμένη χρονική στιγμή προσδιορίζεται από τις τιμές των μεταβλητών του συστήματος για προηγούμενες τιμές του χρόνου, σε ένα βάθος χρόνου. Στην ανάλυση χρονοσειρών ένα από τα βασικά ερωτήματα είναι εάν η συγκεκριμένη χρονοσειρά παρουσιάζει μακροχρόνια μνήμη. Απαντήσεις στο ερώτημα αυτό ήρθαν να δώσουν νέες αριθμητικές μέθοδοι. Εδώ θα εξετάσουμε την Ανάλυση Κανονικοποιημένης Μεταβλητότητας (Rescaled Range Analysis ,R/S analysis ). Ο κλασικός τύπος (Hurst) της ανάλυσης έγινε ευρύτερα γνωστός στο ευρύ κοινό από τις εργασίες του Peters.

Η ανάλυση αυτή επινοήθηκε από τον υδρολόγο Harold Edwin Hurst το 1951, ο οποίος έκανε μια μακροχρόνια μελέτη σχετικά με το ύψος των νερών του Νείλου. Σε μια μεγάλη χρονοσειρά δουλεύουμε χονδρικά ως εξής: χωρίζουμε τη χρονοσειρά σε υποπεριόδους και μετράμε το εύρος της αθροιστικής απόκλισης για κάθε υποπερίοδο με την μέθοδο που αναφέραμε προηγουμένως. Η υποπερίοδος που θα παρουσιάσει το μεγαλύτερο εύρος θα είναι αυτή με την ισχυρότερη τάση.

Όταν έχουμε πολλές χρονοσειρές (ή υποπεριόδους μιας χρονοσειράς ) τις οποίες θέλουμε να συγκρίνουμε ως προς το βαθμό μακροχρόνιας μνήμης, τότε για να βγάλουμε συγκρίσιμα αποτελέσματα θα πρέπει να κάνουμε μια ακόμη στάθμιση που εξασφαλίζει ότι τα δεδομένα όλων των χρονοσειρών είναι συγκρίσιμα στο μέγεθος. Η κλασική μέθοδος είναι να διαιρέσουμε το κάθε εύρος με την τυπική απόκλιση της υποπεριόδου. Η μέθοδος αυτή χρησιμοποιείται στα πλαίσια μιας σημαντικής μεθόδου διάγνωσης μακροχρόνιας μνήμης που προτάθηκε από το Manderbrot, στα τέλη της δεκαετίας του '60 και είναι γνωστή με το όνομα R/S. Στις παραλλαγές δεν διαιρούμε με την τυπική απόκλιση αλλά με αριθμό κοντά στην τυπική απόκλιση.



Στην ανάλυση R/S τελικός στόχος είναι η εύρεση ενός αριθμού γνωστού ως εκθέτη Hurst. Είναι ένας αριθμός στο διάστημα  $[0,1]$ , που δείχνει την πιθανότητα μια χρονοσειρά να συνεχίσει την ίδια πορεία και στην επόμενη περίοδο.

Το φάσμα ισχύος αποτελεί έναν άλλο τρόπο για να υπολογίσουμε εάν η υπό εξέταση χρονοσειρά παρουσιάζει μακροχρόνια μνήμη. Θεωρούμε πως στη χρονοσειρά υπάρχουν μικροί και μεγάλοι κύκλοι με μορφή ημιτόνου (ή συνημίτονου). Πρέπει να διακρίνουμε τους κύκλους που συμβάλλουν σε μεγάλο ποσοστό στην διακύμανση της χρονοσειράς, και να μην λάβουμε υπόψη αυτούς που συμβάλλουν κατά μικρό ποσοστό. Έχουμε μια χρονοσειρά με μέσο όρο 0 και την αναλύουμε σε ένα άθροισμα ημιτόνων και συνημίτονων. Από μια χρονοσειρά  $N$  παρατηρήσεων διακρίνουμε το πολύ  $h = \frac{N-1}{2}$  συχνότητες, όπου το  $h$  είναι ακέραιος. (Αυτό σημαίνει ότι αν το  $N$  είναι ζυγός αριθμός από το  $h$  θα κρατήσουμε μόνο το ακέραιο μέρος του). Η χρονοσειρά, λοιπόν, αναλύεται σε

$$x_n = \sum_{f=1}^h (a_f \cos\left(\frac{2\pi f n}{N}\right) + \beta_f \sin\left(\frac{2\pi f n}{N}\right)) \quad (5)$$

Από την παραπάνω σχέση θα πρέπει να υπολογίσουμε τους συντελεστές  $a_f$  και  $\beta_f$ . Η συχνότητα  $f=1$  ονομάζεται θεμελιώδης αρμονική (fundamental harmonic). Είναι η συχνότητα που θα συναντήσουμε μία μόνο φορά σε όλη την χρονοσειρά. Η συχνότητα  $f=2$  ονομάζεται δεύτερη συχνότητα και θα την συναντήσουμε 2 φορές στην χρονοσειρά μας κοκ. Οι συντελεστές  $a_f$  και  $\beta_f$  ονομάζονται συντελεστές Fourier. Οι συντελεστές  $a_f$  ονομάζονται συντελεστές συνημιτόνων και οι συντελεστές  $\beta_f$  ονομάζονται συντελεστές ημιτόνων. Ο υπολογισμός τους δίνεται από τις ακόλουθες σχέσεις

$$a_f = \frac{2}{N} \sum_{n=1}^N x_n \cos\left(\frac{2\pi f(n-1)}{N}\right) \quad (6)$$

$$\beta_f = \frac{2}{N} \sum_{n=1}^N x_n \sin\left(\frac{2\pi f(n-1)}{N}\right) \quad (7)$$

Από τους συντελεστές αυτούς μπορούμε εύκολα να υπολογίσουμε τη διακύμανση που προέρχεται από κάθε συνδυασμό.

$$P(f) = \sqrt{a_f^2 + b_f^2} \quad (8)$$

Το διάγραμμα που σχηματίζουμε από τους συνδυασμούς  $\log P(f) - \log f$  έχει μορφή παρόμοια με το διάγραμμα της ανάλυσης R/S: και οι δύο άξονες έχουν λογαριθμική μορφή.

Εάν το φάσμα ισχύος παρουσιάζει αρνητική κλίση αποτελεί μια επιπλέον ένδειξη πιθανά χαοτικής συμπεριφοράς. Αντίθετα η διαδικασία του λευκού θορύβου έχει λίγο πολύ οριζόντιο φάσμα ισχύος.

### 2.1.2 Δυναμικά Συστήματα

Σύστημα είναι η διάταξη αλληλεπιδρώντων μερών ή αλλιώς ένα σύνολο στοιχείων σε αλληλουχία, κυρίως χρονολογικά τοποθετημένα. Το ενδιαφέρον μας θα επικεντρωθεί στα δυναμικά συστήματα. Δυναμικό σύστημα (dynamical system) είναι το σύστημα που εξελίσσεται με το χρόνο. Στη συνέχεια με τον όρο σύστημα θα αναφερόμαστε σε δυναμικό σύστημα.

Κάθε σύστημα δίνει μια ή περισσότερες χρονοσειρές. Οι χρονοσειρές δημιουργούνται από την χρονική εξέλιξη των μεταβλητών του συστήματος. Οι εξισώσεις που συνδέουν τις μεταβλητές ενός συστήματος μπορεί να αντιστοιχούν στην περιγραφή ενός φυσικού, βιολογικού, ή οικονομικού υποδείγματος. Συνήθως, το υπόδειγμα είναι γνωστό, μπορεί όμως να παρατηρηθεί η χρονική εξέλιξη (χρονοσειρά) μίας ή περισσότερων από τις μεταβλητές που το αποτελούν. Εμείς θα επικεντρωθούμε στο να εξάγουμε συμπεράσματα για το σύστημα παρατηρώντας μόνο τη μια χρονοσειρά από όσες είναι ικανό να μας δώσει.

Τα συστήματα χωρίζονται σε δύο βασικές κατηγορίες ανάλογα με το εάν χάνουν ή όχι ενέργεια. Τα **διατηρητικά** (ή συντηρητικά) στα οποία παρατηρείται πως δεν έχουν απόσβεση, δηλαδή δεν χάνουν ενέργεια με την πάροδο του χρόνου. Τα συστήματα με **απώλειες** (ή καταναλίσκοντα) στα οποία παρατηρείται πως έχουν απόσβεση και συνεπώς χάνουν ενέργεια με την πάροδο του χρόνου, ενώ ασυμπτωτικά πέφτουν πάνω στον ελκυστή που δημιουργείται. Τα συστήματα αυτά αποτελούν τη συνηθέστερη περίπτωση και θα μας απασχολήσουν στη συνέχεια.

Τα συστήματα ανάλογα με το αν παρουσιάζουν συνεχείς ή διακριτές διαδικασίες στην εξέλιξη τους διακρίνονται σε δύο κατηγορίες :

1. Τα συνεχή συστήματα. Στα συστήματα αυτά κάθε μεταβλητή θεωρείται συνεχής συνάρτηση του χρόνου. Οι χρονοσειρές των συστημάτων αυτών μετρούνται με μικρό χρονικό βήμα. Επίσης, έχουμε το δικαίωμα να αυξομειώσουμε τη συχνότητα δειγματοληψίας. Τα υποδείγματα που φτιάχνουμε για τα συστήματα αυτά αποτελούνται από διαφορικές εξισώσεις. Το σύνολο των διαφορικών εξισώσεων που απαρτίζουν ένα συνεχές σύστημα ορίζει μια ροή στο χώρο των φάσεων. Παράδειγμα ροής είναι η μεταβολή πίεσης και θερμοκρασίας μέσα σε ένα δωμάτιο, η ροή των ποταμών κ.α. Τα δύο αυτά μεγέθη μεταβάλλονται και αλληλεπιδρούν σε συνεχή χρόνο.
2. Τα διακριτά συστήματα. Στα συστήματα αυτά οι μεταβλητές μετρούνται σε συγκεκριμένες χρονικές στιγμές και ενδιάμεσα από δύο μετρήσεις δεν είναι γνωστή η εξέλιξη των μεγεθών. Τα υποδείγματα που φτιάχνουμε από αυτά αποτελούνται από εξισώσεις διαφορών. Αυτές λύνονται με

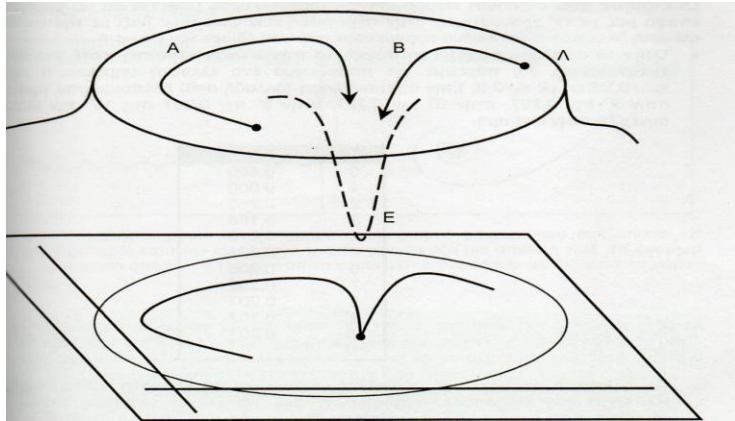
επαναλήψεις(iterations). Συνήθως, τα φαινόμενα στα οποία αναφέρονται τα διακριτά συστήματα είναι συνεχή. Παραδείγματα αποτελούν οι σεισμοί, οι καταιγίδες και οι ηφαιστειακές εκρήξεις.

Ένας από τους καλύτερους τρόπους να κατανοήσουμε τα δυναμικά συστήματα είναι να τα παραστήσουμε οπτικά. Για να γίνει αυτό πρέπει να κάνουμε την γραφική τους παράσταση. Ο χώρος των φάσεων (phase space) είναι ο μαθηματικός χώρος του οποίου οι συντεταγμένες είναι ανεξάρτητες μεταβλητές που χρειάζονται για να καθορίσουν τη φάση του δυναμικού συστήματος κάθε χρονική στιγμή.

Σαν συμπλήρωμα στο συνηθισμένο πλωτάρισμα των χρονοσειρών, το πλωτάρισμα στον χώρο των φάσεων παρέχει μια διαφορετική οπτική της εξέλιξης του συστήματος. Επίσης όταν έχουμε χρονοσειρές με πολλά δεδομένα τα οποία είναι δύσκολο να παρασταθούν στο απλό διάγραμμα χρόνου, στο χώρο των φάσεων μπορούν να παρασταθούν πιο εύκολα. Τέλος με το πλωτάρισμα των τιμών στο χώρο των φάσεων μπορεί να μας δώσει πολύ διαφορετικές δομές από αυτές που συναντάμε στο απλό πλωτάρισμα.

## **Ο ελκυστής**

Πιθανότατα η πιο βασική έννοια στην χαοτική ανάλυση είναι αυτή του ελκυστή (attractor). Ο ελκυστής είναι ένα σύνολο σημείων στο χώρο των φάσεων, προς το οποίο τείνουν ασυμπτωτικά στο χρόνο οι τροχιές για ένα εύρος αρχικών συνθηκών. Στον χώρο των φάσεων ο ελκυστής δείχνει την μακροχρόνια συμπεριφορά του συστήματος. Ο ελκυστής φαίνεται στο σύστημα αφού αυτό σταθεροποιηθεί και αρχίσει να επαναλαμβάνει κάποια συμπεριφορά. Συνεπώς, οι αρχικές συνθήκες και η συμπεριφορά στις πρώτες επαναλήψεις δε συμπεριλαμβάνονται στον ελκυστή. Ο ελκυστής είναι κατά μια έννοια ένα περιληπτικό μαθηματικό αντικείμενο, που μπορούμε να το σχεδιάσουμε στο χαρτί και συνήθως μας δίνει μια συναρπαστική εικόνα.



**Σχήμα 10:** Έχουμε δύο σφαίρες την A και την B. Οι αρχικές θέσεις των σφαιρών αυτών είναι οι αρχικές συνθήκες του συστήματος. Στην πάνω εικόνα βλέπουμε πως αυτές οι σφαίρες βρίσκονται μέσα στην λεκάνη έλξης Λ. Η λεκάνη έλξης είναι όλες οι δυνατές θέσεις που θα οδηγήσουν τις σφαίρες στη λακκούβα E. Έξω από την λεκάνη έλξης οι σφαίρες θα οδηγηθούν σε θέσεις μακριά από την λακκούβα. Η λακκούβα στο έδαφος δρα ως ελκυστής για τις δύο σφαίρες, δηλαδή έλκει όλες τις αρχικές συνθήκες οι οποίες ανήκουν στην λεκάνη έλξης. Στην κάτω εικόνα βλέπουμε σε ένα σύστημα συντεταγμένων. Έχουμε σχεδιάσει την προβολή της κίνηση των δύο σφαιρών μέχρι να καταλήξουν στη λακκούβα-ελκυστή. Παρατηρούμε επίσης το γεωμετρικό τόπο των σημείων της λεκάνης έλξης.

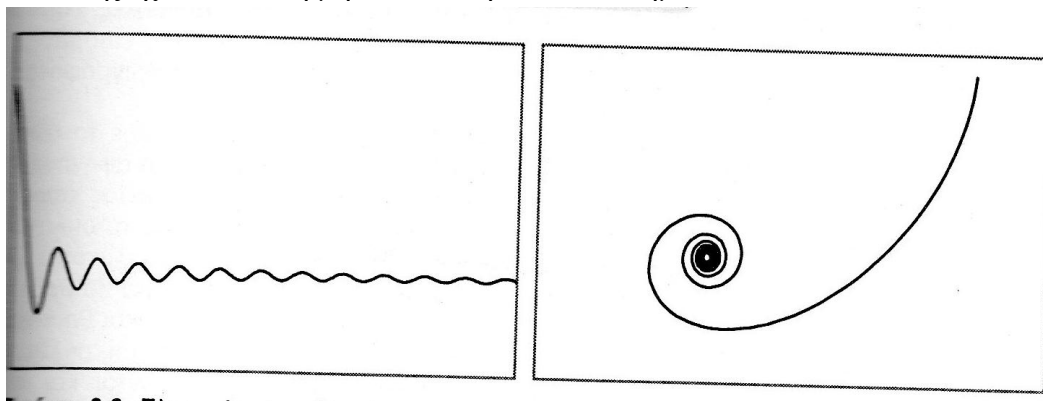
Ας δούμε στην πραγματικότητα πως δουλεύει η παραπάνω έννοια. Σκάβουμε μια λακκούβα στο έδαφος όπως δείχνει σχήμα 10 πάνω. Από δύο κοντινές περιοχές της λακκούβας αυτής αφήνουμε να κυλήσουν δύο σφαίρες, η A και η B. Το σύστημα αυτό είναι ένα σύστημα με απώλειες αφού το μήκος μεταξύ των δύο σφαιρών τελικά θα μηδενιστεί. Το σύστημα είναι συνεχές αφού οι σφαίρες κυλούν σε συνεχή χρόνο. Και οι δύο σφαίρες θα καταλήξουν στο βαθύτερο σημείο της λακκούβας. Τα αντίστοιχα βέλη A και B δείχνουν την πορεία που θα ακολουθήσουν οι δύο σφαίρες μέχρι να σταθεροποιήσουν την συμπεριφορά τους στο βάθος της λακκούβας. Η σταθεροποίηση της συμπεριφοράς σε αυτή την περίπτωση είναι η ακινησία στην πάροδο του χρόνου. Ο ελκυστής (E) σε αυτήν την περίπτωση είναι η λακκούβα. Η περιφέρεια Λ ονομάζεται λεκάνη έλξης. Η λεκάνη έλξης (basin of attraction) ενός ελκυστή είναι το σύνολο αρχικών συνθηκών που η εξέλιξή τους οδηγεί στον ελκυστή. Πέρα από τη λεκάνη έλξης το σύστημα μπορεί να κατευθυνθεί σε άλλο ελκυστή ή να φύγει στο  $\infty$ . Στο σχήμα 10 στο κάτω βλέπουμε την προβολή της εξέλιξης σε ένα διδιάστατο διάγραμμα. Επειδή η έλξη γίνεται από ένα σημείο ο ελκυστής αυτός ονομάζεται ελκυστής σημείου (fixed-point attractor). Είναι η πιο απλή μορφή ελκυστή. Τις υπόλοιπες κατηγορίες ελκυστών θα τις δούμε παρακάτω.

## Οι τύποι των ελκυστών

Για να μην χάνουμε την επαφή μας με τις χρονοσειρές, στην περιγραφή των ελκυστών μαζί με την εικόνα του ελκυστή δίνουμε και την εικόνα της χρονοσειράς που έδωσε τον ελκυστή.

Όλοι οι ελκυστές χωρίζονται σε **χαοτικούς** και **μη χαοτικούς**. Οι μη χαοτικοί ελκυστές είναι 3 τύπων, ελκυστές σημείου, περιοδικοί ελκυστές, και οι τοροειδείς ελκυστές. Οι μη χαοτικοί ελκυστές αποτελούνται από κανονικές, προβλέψιμες τροχιές. Προβλέψεις μακροχρόνιας εξέλιξης ή επαναληπτικής συμπεριφοράς μπορεί να είναι αρκετά ακριβής, σε αρκετά μακρή διάστημα στο μέλλον. Δυο γειτονικές τροχιές μόλις βρεθούν πάνω στον ελκυστή, μένουν σε κοντινή απόσταση μεταξύ τους, καθώς εξελίσσονται στο χρόνο. Οι χαοτικοί ελκυστές, αναπτύσσουν πολλά ενδιαφέροντα και σύνθετα σχήματα στον χώρο των φάσεων. Δεν μπορείς να προβλέψεις την μακροχρόνια εξέλιξη αυτών των ελκυστών με αξιοπιστία. Περίεργως, οι περιοδικοί ελκυστές παρουσιάζονται όχι μόνο σε μη χαοτικούς τομείς αλλά και εκεί που παρουσιάζεται χάος.

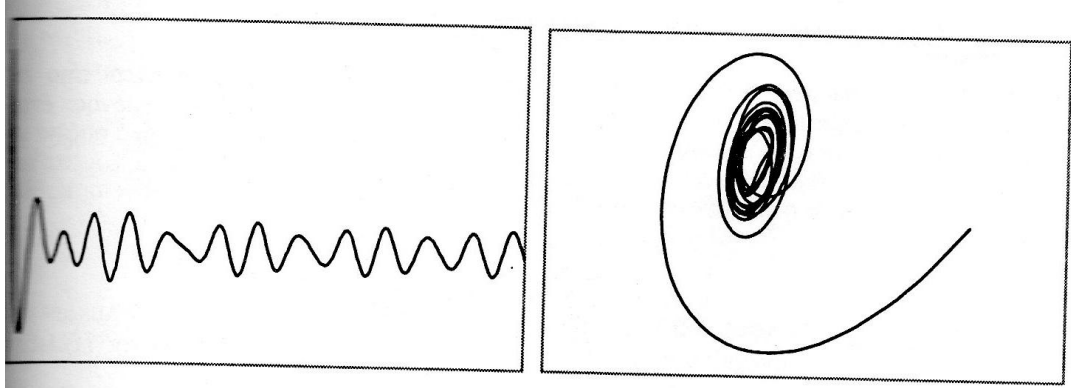
1. Ελκυστής σημείου: τα σημεία στο χώρο των φάσεων έλκονται από ένα σημείο του (σχήμα 11). Είναι ο απλούστερος τύπος ελκυστή, ο οποίος είναι ένα συγκεκριμένο σημείο στον χώρο των φάσεων. Αντιπροσωπεύει όλα τα συστήματα τα οποία καταλήγουν σε κατάσταση ξεκούρασης με την πάροδο του χρόνου ( Crutchfield et al, 1986) ή τα οποία εξελίσσονται σε μια κατάσταση που δεν αλλάζει με τον χρόνο. Μόλις η σταθερή κατάσταση επιτευχθεί, ο ελκυστής σημείου είναι ανεξάρτητος του χρόνου. Στο χώρο των φάσεων είναι πλέον στατικό. Οι ελκυστές σημείου περιλαμβάνουν 2 μεταβλητές, επομένως ο χώρος των φάσεων είναι διδιάστατος. Εντούτοις, ο ελκυστής έχει διάσταση μηδέν, επειδή είναι ένα σημείο.



**Σχήμα 11:** Ελκυστής σημείου

Μπορούμε εύκολα να φανταστούμε ένα τέτοιο σύστημα: ένα εκκρεμές που το εκτρέπουμε και τελικά ηρεμεί. Ένα σύστημα λοιπόν που το εκτρέπουμε ενώ αυτό έχει την τάση να ακινητοποιηθεί, είναι ένα σύστημα που παρουσιάζει ελκυστή σημείου.

2. Ελκυστής οριακού κύκλου (limit circle attractor) ή περιοδικός ελκυστής: τα σημεία στο χώρο των φάσεων έλκονται από ένα κύκλο (Σχήμα 12). Είναι ο αμέσως πιο περίπλοκος τύπος ελκυστή από τον ελκυστή σημείου.

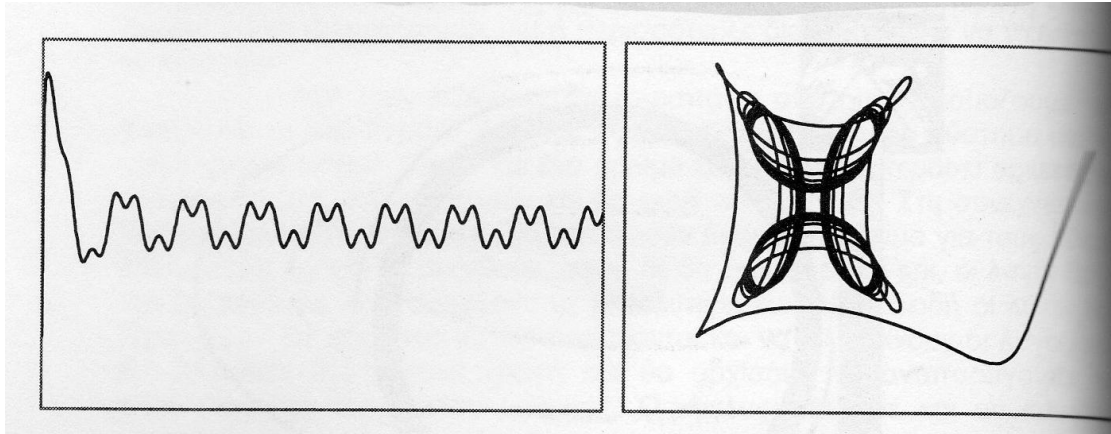


**Σχήμα 12:** Ελκυστής οριακού κύκλου

Ο κύκλος που σχηματίζεται έχει περίοδο ίση με τον αριθμό των επαναλήψεων που χρειάζεται για να εκτελέσει μια πλήρη περιστροφή.

Αποτελείται από δύο ή περισσότερες τιμές οι οποίες επαναλαμβάνονται με την ίδια σειρά. Ένα καλό παράδειγμα είναι ένας οποιοσδήποτε ελκυστής της περιοχής διπλασιασμού περιόδου της λογιστικής συνάρτησης. Παρατηρούμε τέτοιου τύπου ελκυστές επίσης στο ρολόι με εκκρεμές. Η δυναμική ενέργεια είναι συσσωρευμένη στο ελατήριο του ρολογιού αναπληρώνει την απώλεια από τις τριβές κατά την κίνηση των γραναζιών. Στον μετρονόμο που χρησιμοποιούν οι μουσικοί συμβαίνει κάτι ανάλογο με το ρολόι, η ενέργεια από το κούρδισμα αναπληρώνει την απώλεια από τις τριβές.

3. Ελκυστής δύο περιόδων ή τοροειδής ελκυστής (toroid attractor): τα σημεία στο χώρο των φάσεων έλκονται από ένα σχήμα που μοιάζει με δύο κύκλους, ο ένας πάνω στον άλλο (εικόνα 5). Ο τόρος ορίζεται ως το καρτεσιανό γινόμενο 2 κύκλων. (N-τόρος ονομάζεται το καρτεσιανό γινόμενο N κύκλων.) Είναι ο αμέσως πιο περίπλοκος τύπος ελκυστή μετά από τον ελκυστή οριακού κύκλου.

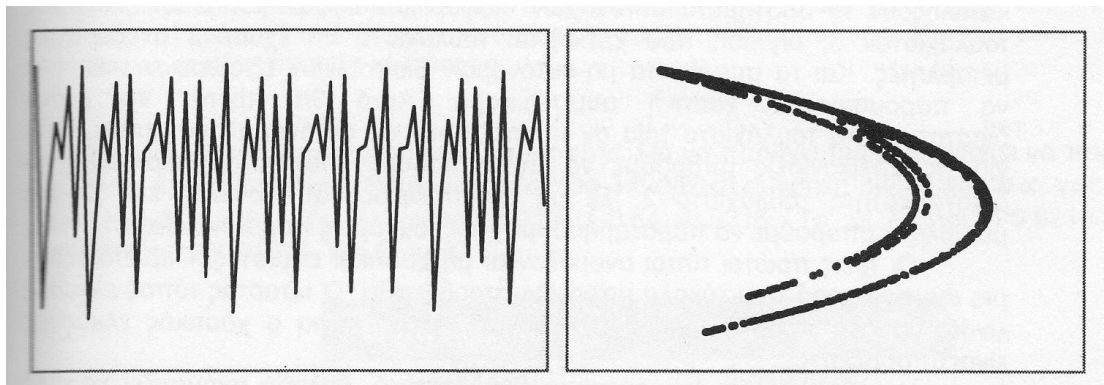


**Σχήμα 13:** τοροειδής ελκυστής

Θα δούμε τώρα τι γίνεται όταν προσθέσουμε σε μια χρονοσειρά μιας περιόδου μια χρονοσειρά με διαφορετική περίοδο. Στο χώρο των φάσεων θα δούμε ένα σχήμα που επαναλαμβάνεται περιοδικά και επιπλέον μέσα του περιέχει δύο περιόδους. Σαν παράδειγμα τέτοιας συμπεριφοράς μπορούμε να φέρουμε τη διακύμανση της θερμοκρασίας σε ένα συγκεκριμένο τόπο. Ας υποθέσουμε ότι οι διακυμάνσεις αυτές εξαρτώνται μόνο από δύο παράγοντες: την εποχή και την ώρα της ημέρας. Η εποχή επηρεάζει ως εξής: το καλοκαίρι ανεβαίνει η θερμοκρασία και το χειμώνα πέφτει. Η ώρα της ημέρας επηρεάζει ως εξής: τις πρωινές ώρες ο ήλιος θερμαίνει την ατμόσφαιρα άρα ανεβαίνει η θερμοκρασία ενώ τη νύχτα η θερμοκρασία πέφτει. Μπορούμε εύκολα να φανταστούμε μια τέτοια χρονοσειρά σε μακροχρόνια εξέλιξη. Μεγάλοι κύκλοι ξεχωρίζουν καθαρά(εποχικότητα ,ενώ πάνω τους ξεχωρίζουν μικρές περιοδικότητες (ώρα της ημέρας ). Συνεπώς, αν αφαιρέσουμε από τη χρονοσειρά την μεγάλη περίοδο τα κατάλοιπα θα είναι μόνο η μικρή περίοδος. Αντίστροφα ,αν από τη χρονοσειρά αφαιρέσουμε τη μικρή περίοδο τα κατάλοιπα θα είναι η μεγάλη περίοδος.

4. Χαοτικό (ή παράξενο) ελκυστή(chaotic attractor): τα σημεία στο χώρο των φάσεων έλκονται από ένα αρκετά παράξενο αντικείμενο, δηλαδή δεν έχει ένα συγκεκριμένο σχήμα σαν αυτά που γνωρίζουμε (σχήμα 14). Παράλληλα ,δεν είναι και ένα σύννεφο από ισοκατανεμημένα σημεία. Ο χαοτικός ελκυστής είναι μια περίπλοκη επιφάνεια στον χώρο των φάσεων στην οποία η τροχιά είναι ασυμπτωτική με τον χρόνο και πάνω στην οποία περιφέρεται χαοτικά.

(Grebogi et al.1982)



**Σχήμα 14:** Χαστικός ελκυστής

Τα χαρακτηριστικά του παράξενου ελκυστή είναι τα εξής:

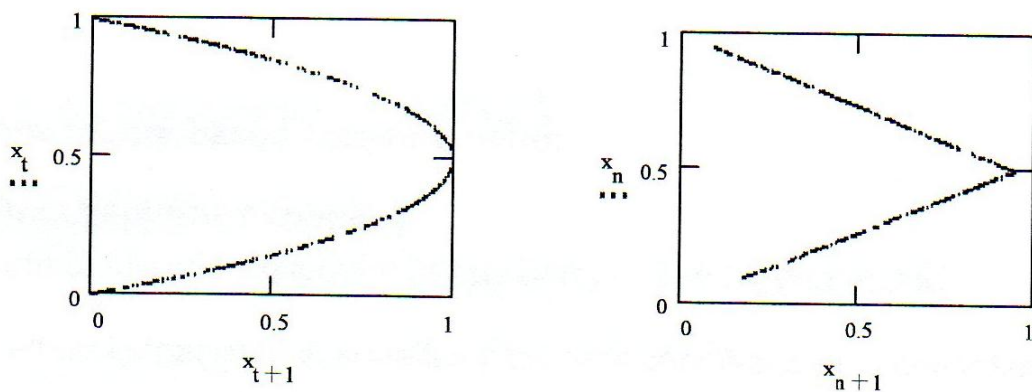
- Μια τροχιά μέσα στην χαστική περιοχή είναι συνήθως πιο περίπλοκη από ένα απλό επαναλαμβανόμενο βρόχο. Για κάποιες τιμές παραμέτρων ελέγχου, η τροχιά δεν επαναλαμβάνεται δηλαδή είναι απεριοδική η μη περιοδική. Πολλές άλλες τιμές των παραμέτρων φαίνεται να δίνουν περιοδικές τροχιές με μεγάλη περιοδικότητα. Στην πράξη είναι δύσκολο να ξεχωρίσεις τροχιές με μεγάλη περιοδικότητα από τις μη περιοδικές. Η κυρίαρχη περιοδικότητα, αν υπάρχει, του παράξενου ελκυστή καθορίζεται από τη μελέτη του φάσματος των συχνοτήτων του.
- Οι τροχιές στον χαστικό ελκυστή δεν διασταυρώνονται. Εάν διασταυρώνονταν, τότε το σύστημα θα συμπεριφερόταν με πολλούς διαφορετικούς τρόπους οπότε οι συνθήκες στην διασταύρωση ξαναπραγματοποιώντουσαν.
- Δύο τροχιές οι οποίες κάποια στιγμή είναι κοντά αποκλίνουν και τελικά ακολουθούν πολύ διαφορετικά μονοπάτια. Αυτό οφείλεται στην ευαισθησία στις αρχικές συνθήκες. Αυτή η απόκλιση των τροχιών καλείται από τους τοπολόγους ως τέντωμα.
- Το μονοπάτι μιας χαστικής τροχιάς στον χώρο των φάσεων παρουσιάζει το χαρακτηριστικό του διπλώματος. Αυτό πραγματοποιείται όταν η τροχιά φτάσει στα όρια του χώρου των φάσεων (στα όρια μιας ή περισσότερων μεταβλητών) και ανακλάται στο πλωταρισμένο της σχέδιο.
- Ένας χαστικός ελκυστής έχει πολύπλοκη, πολλών στρώσεων εσωτερική δομή. Η δομή αυτή είναι συνήθως κλασματική.
- Η διάσταση του χαστικού ελκυστή δεν χρειάζεται να είναι ακέραια. Η μη ακέραια και συνήθως κλασματική φύση των χαστικών ελκυστών οδηγεί τον Mandelbrot(1983:197) να συνιστά να ονομάζουμε τους εκλυστές αυτούς κλασματικούς ελκυστές.
- Μη κανονική κατανομή συχνοτήτων. Το σύστημα προσεγγίζοντας τον ελκυστή ακολουθεί συγκεκριμένες τροχιές και δεν περιφέρεται εντελώς



τυχαία. Αφού με την πάροδο του χρόνου το σύστημα τείνει να αγγίξει τον ελκυστή, η κατανομή συχνοτήτων τείνει σε κάποια συγκεκριμένη κατανομή.

### 2.1.3 Η εμβύθιση σαν ένα μέσο ανάλυσης των χαοτικών χρονοσειρών.

Οι (Ruelle,81) και (Takens,81) πρώτοι περιέγραψαν μια απλή μέθοδο για να αναλύουν τις χρονοσειρές που την ονόμασαν εμβύθιση των χρονοσειρών. Μπορεί να παρουσιαστεί απλά παρατηρώντας τις συνέπειες του πλοταρίσματος ζευγαριών σημείων  $X_t$  και  $X_{t+1}$  και για την λογιστική συνάρτηση και για την Tent map όπως φαίνεται στο παρακάτω σχήμα.



Σχήμα 15: Εμβυθισμένη παράσταση της tent map

Οι παραπάνω γραφικές παραστάσεις δείχνουν ότι και στις δυο περιπτώσεις τα πλοταρισμένα ζεύγη σημείων από τις χρονοσειρές δημιούργησαν αναγνωρίσιμες πατέντες. Αν μας δίνεται το  $X_t$  μπορούμε να κάνουμε μια πολύ καλή εκτίμηση του  $X_{t+1}$  με παρεμβολή. Αυτή η αρχή επεκτείνεται σε πολλαπλές διαστάσεις και γενικά μπορεί να γραφεί ως εξής :

$$X_t = X_t, X_{d+t}, X_{2d+t}, X_{3d+t}, \dots, X_{nd+t} \quad (9)$$

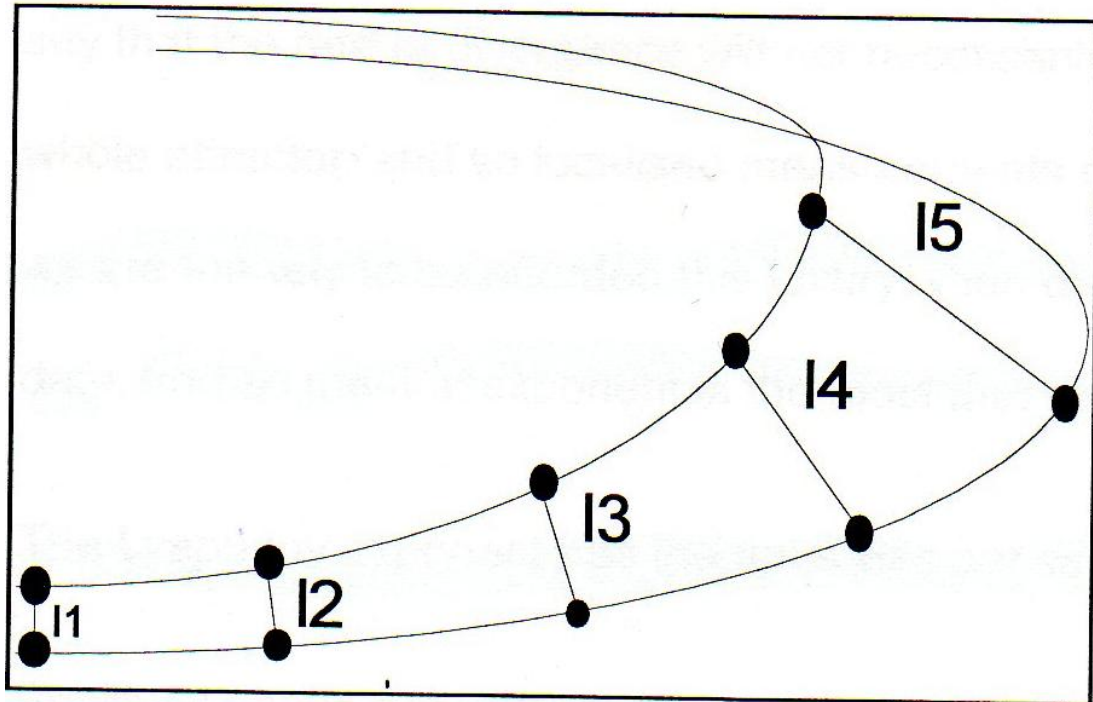
Όπου  $X$  είναι το εμβυθισμένο διάνυσμα,  $d$  είναι η διαχώριση και  $n$  διάσταση εμβύθισης.

Ο Takens (takens,81) έδειξε ότι αυτή η αρχή γενικεύεται και ότι εμβυθίζοντας σωστά την χαοτική χρονοσειρά μια ομαλή συνάρτηση την μοντελοποιεί τέλεια. Και οι δυο όμως, η διάσταση εμβύθισης και η ομαλή συνάρτηση πρέπει να βρεθούν εμπειρικά.

## 2.2 Οι βασικές μετρήσεις της θεωρίας του χάους

### 2.2.1 Ο εκθέτης Lyapunov

Ένα από τα χαρακτηριστικά των χαοτικών συστημάτων που παίζουν πολύ σημαντικό ρόλο είναι η ευαισθησία στις αρχικές συνθήκες. Ένας εκθέτης Lyapunov είναι ένας αριθμός που περιγράφει την δυναμική της εξέλιξης των τροχιών. Οι τιμές που πέρνει μπορεί να είναι αρνητικές, μηδέν, ή θετικές. Αρνητικές τιμές σημαίνουν ότι δύο γειτονικές τροχιές πλησιάζουν η μια στην άλλη. Θετικοί εκθέτες, από την άλλη μεριά, σημαίνουν πως δύο γειτονικές τροχιές αποκλίνουν η μια από την άλλη και παρουσιάζονται στις περιοχές που παρατηρείται χάος. Θετικός εκθέτης Lyapunov είναι ένδειξη χάους. Ένας θετικός εκθέτης Lyapunov ποσοτικοποιεί την εξάρτηση ευαισθησίας στις αρχικές συνθήκες δείχνοντας τον μέσο όρο στον οποίο δυο γειτονικά σημεία χωρίζονται στο χρόνο. η τιμή 0 του εκθέτη Lyapunov είναι χαρακτηριστικό κυκλικής συμπεριφοράς. Θεωρείται πως το σύστημα μας παρουσιάζει χαοτική συμπεριφορά αν αυτό το χαρακτηριστικό μπορεί να ανιχνευτεί σε αυτό(Wolf, 85). Η στάνταρ μέθοδος ανίχνευσης αυτού του χαρακτηριστικού είναι ο υπολογισμός του εκθέτη Lyapunov (wolf, 85).



**Σχήμα 16:** Χαρακτηριστική απόκλιση στο χρόνο 2 γειτονικών τροχιών του ελκυστή

Στο παραπάνω σχήμα βλέπουμε 2 γειτονικές τροχιές του ελκυστή και την χαρακτηριστική απόκλιση στο χρόνο που σχετίζεται με το χάος .

Οι τροχιές στον ελκυστή είναι εμβυθισμένες σε ένα πολυδιάστατο χώρο, επομένως η απόκλιση είναι καλά παρουσιασμένη σαν την διαφορά μεταξύ 2 n-tuples. Ο κυρίαρχος μέσος εκθέτης Lyapunov καθορίζεται ως :

$$L = \log_2 \frac{(\sum_{n-1}^{l_{n+1}} l_n)}{n-1} \quad (10)$$

Όπου το n δείχνει τον αριθμό των δηγμάτων και το 1 την ευκλείδεια απόσταση μεταξύ των τροχιών και των κοντινότερων γειτόνων.

Με ένα πολύ μεγάλο αριθμό από δεδομένα η πυκνότητα των σημείων σε μια περιοχή του ελκυστή θα επιτρέψει τον υπολογισμό των τοπικών εκθετών Lyapunov. Ένα παράδειγμα του παραπάνω είναι ο ελκυστής ενός ημιτονοειδούς συστήματος όπως είναι ένα διεγερμένο εκκρεμές που είναι ένας κύκλος. Ο κυρίαρχος εκθέτης Lyapunov επειδή υπολογίζει τους μέσους όρους των συγκλίσεων και των αποκλίσεων σε πολλά σημεία σε όλο τον ελκυστή, είναι μια σφαιρική ποσότητα και όχι τοπική. Αυτό μπορούμε να το μεταφράσουμε με τρεις τρόπους.

- Σε n διαστάσεις , ο L ποσοτικοποιεί, σε ένα μόνο αριθμό, τον μέσο όρο κατά τον οποίο αναπτύσσεται η διάσταση με την μεγαλύτερη ανάπτυξη στον χώρο των φάσεων.
- Ποσοτικοποιεί την μέση προβλεψιμότητα πάνω στον ελκυστή.

- Επειδή οι κοντινές τροχιές αντιπροσωπεύουν αλλαγές στις αρχικές συνθήκες ενός συστήματος, ο L αποτελεί την μέτρηση του πόσο ευαίσθητο είναι το σύστημα σε μικρές αλλαγές ή διαταραχές.

### 2.2.2 Ο εκθέτης Hurst

Ο Η.Ε. Hurst (Hurst,65) είναι υπεύθυνος για μια μέτρηση προβλεψιμότητας των χρονοσειρών που παρουσιάζει ενδιαφέροντα χαρακτηριστικά. Ο εκθέτης πηγάζει από την χρήση της ανάλυσης R/S. Όπου R/S(rescaled range) είναι μια μέθοδος που αναπτύχθηκε από τον Hurst για να καθοριστούν τα μακροχρόνιας μνήμης αποτελέσματα και η Brownian κλασματική κίνηση. Η ανάλυση R/S μετράει πόση απόσταση διανύει ένα σωματίδιο αυξάνει καθώς κοιτάει σε μεγαλύτερες και μεγαλύτερες χρονικές κλίμακες. Δοθείσας μιας χρονοσειράς X που περιέχει ένα αριθμό σημείων, n, και διαλέγοντας ένα ακέραιο διαιρέτη p, όπου για ευκολία  $10 \leq p < n/2$ , τα δεδομένα μπορούν να διαχωριστούν σε n/p μπλοκς .

Για κάθε μπλοκ η μέση τιμή υπολογίζεται, μετά η μέγιστη εμβέλεια του κάθε μπλοκ και η στάνταρ απόκλιση του καθενός. Η τιμή (εμβέλεια)/(στάνταρ απόκλιση) υπολογίζεται για κάθε μπλοκ και ο μέσος όρος σε όλα τα μπλοκ. Αυτή η μέση τιμή rs σχετίζεται με τον εκθέτη Hurst με την ακόλουθη φόρμουλα.

$$rs = \left(\frac{p}{2}\right)^H \quad (11)$$

Όπου H είναι ο εκθέτης Hurst. Προκειμένου να αποκτήσουμε μια πιο αξιόπιστη εκτίμηση της τιμής του rs υπολογίζεται για όλες τις πιθανές τιμές του p, και τα αποτελέσματα εισάγονται και γραμμική οπισθοδρόμηση πραγματοποιείται σε αυτά. Είναι η κλίση της οπισθοδρόμησης που χρησιμοποιείται σαν ο εκθέτης Hurst.

Ο εκθέτης Hurst παίρνει τιμές μεταξύ 0 και 1. Η τιμή 0.5 μας δείχνει ότι πρόκειται για πραγματικά τυχαίο περίπατο. Εάν η τιμή είναι  $0.5 < H < 1$  μας δείχνει μια επίμονη συμπεριφορά, που σημαίνει πως κάποιος μπορεί να περιμένει με αυξανόμενη σιγουριά καθώς η τιμή του εκθέτη τείνει στο 1 ότι οποιαδήποτε πορεία αλλαγής υπάρχει τώρα θα συνεχίσει να είναι αυτή. Μια ευθεία γραμμή με μηδενική κλίση θα είχε εκθέτη Hurst 1. Παρόμοια οι τιμές  $0 < H < 0.5$  δείχνουν αντιεπίμονη συμπεριφορά, στην οποία μπορεί κάποιος να περιμένει ότι οποιαδήποτε κατεύθυνση αλλαγής είναι παρούσα είναι απίθανο να συνεχιστεί. Στο όριο 0 οι χρονοσειρές πρέπει να αλλάζουν κατεύθυνση σε κάθε δείγμα. Αυτό δίνει ένα στοιχείο για την σχέση μεταξύ μιας διευκρίνισης της διάστασης fractal και του εκθέτη Hurst.

$$D = 2 - H \quad (12)$$

Επομένως εκθέτης Hurst 1 δίνει διάσταση fractal 1, όπως θα περίμενε κανείς από μια ευθεία γραμμή. Η τιμή 0 για τον εκθέτη Hurst πρέπει να ανήκει σε μια χρονοσειρά που πρέπει να κινείται στον δισδιάστατο χώρο και έτσι να έχει διάσταση 2.

## 2.3 Η παραγωγή των παραμέτρων εμβύθισης

### 2.3.1 Αναλυτικές μέθοδοι

Ο Takens (Takens, 81) καθόρισε τα άνω όρια για μια επιτυχημένη εμβύθιση εάν ξέρεις την διάσταση fractal του ελκυστή  $d_a$ . Δηλαδή :

$$d_e \geq 2d_a + 1 \quad (13)$$

Όπου  $d_e$  είναι η διάσταση εμβύθισης, ακέραια τιμή.

Όμως, όπως έδειξε ο Medio (Medio, 92) η διάσταση fractal του ελκυστή του Lorenz (Lorenz, 63) είναι 2.06, έτσι ο Takens μας οδηγεί σε μια διάσταση εμβύθισης 5, όπου είναι γνωστό

ότι η πρακτική διάσταση εμβύθισης αυτών των σειρών είναι 3. Όταν αντιμετωπίζουμε δεδομένα χωρίς θόρυβο, μια επιλογή μεγάλης διάστασης εμβύθισης δεν θα επηρεάσει τον υπολογισμό των εκθετών Lyapunov, αλλά μπορεί να είναι μοιραία για την μοντελοποίηση.

Το πάνω όριο του Takens εξακολουθεί να απαιτεί τον υπολογισμό της διάστασης του ελκυστή, και στα πειραματικά δεδομένα όπου οι συναρτήσεις των δυναμικών συστημάτων δεν είναι γνωστές απαιτούνται δοκιμαστικές εμβυθίσεις, επομένως η όλη διαδικασία είναι κατά κάποιον τρόπο κυκλική. Κάποια άλλη μέθοδος είναι απαραίτητη για τον υπολογισμό της διάστασης εμβύθισης. Υπάρχουν αρκετές αναλυτικές μέθοδοι για την επιλογή της παραπάνω. Αυτή με την παλαιότερη γενιά (Farmer, 89) είναι να κοιτάξεις τον κορεσμό κάποιας αμετάβλητης τιμής του συστήματος, όπως είναι ο εκθέτης Lyapunov, αυξάνοντας την διάσταση εμβύθισης. Μια άλλη είναι ο αλγόριθμος του λανθασμένου κοντινότητας γείτονα και μια άλλη η ανάλυση της στατιστικής Kaplan-Glass (Kaplan, 91) με αυξανόμενη διάσταση εμβύθισης. Από αυτές η πιο απλή να εφαρμόσεις και η πιο εύρωστη είναι ο αλγόριθμος του λανθασμένου κοντινότητας γείτονα. Έχει ένα γεωμετρικό σχηματισμό που τον κάνει να ξεχωρίζει από τις άλλες 2, οι οποίες θεωρούνται με τους τοπικούς παράγοντες του ελκυστή, και έτσι δίνει μια πιο ανεξάρτητη οπτική της διάστασης του ελκυστή. Η επιλογή μιας μεθοδολογίας για την επιλογή της διαχώρισης είναι πιο περίπλοκη. Ο Medio ισχυρίζεται πως η διαχώριση είναι άσχετη εάν είναι μικρή.

#### 2.3.1.1 Λανθασμένοι κοντινοί γείτονες

Οι λανθασμένοι κοντινοί γείτονες (Kennel, 92) είναι μια μεθοδολογία για να παράγουμε μια διάσταση εμβύθισης για μια δοθείσα χρονοσειρά.

Μια δοκιμαστική εμβύθιση σχηματίζεται με διαχώριση που διαλέγουμε με όποια τεχνική θέλουμε και διάσταση εμβύθισης  $m$ . Το κοντινό γειτονικό διάστημα για

κάθε εμβυθισμένο διάνυσμα βρίσκεται και η ευκλείδεια απόσταση μεταξύ των 2 διανυσμάτων μετρείται. Μετά η διάσταση εμβύθισης αυξάνεται κατά 1 και για τα δύο, για το παρόν διάνυσμα και για το γειτονικό του προσάπτοντας την ανάλογη τιμή δεδομένων. Εάν η απόσταση μεταξύ των νέων διανυσμάτων είναι δραματικά διαφορετική μετά από αυτήν την αλλαγή ο γείτονας θεωρείται εσφαλμένος. Ο αριθμός αυτός μετρείται και εκφράζεται σαν κλάσμα του αριθμού των σειρών. Μια καμπύλη από αυτές τις τιμές έναντι του m πλοτάρονται και μια διάσταση εμβύθισης με μικρό αριθμό από λανθασμένους γείτονες επιλέγεται. Για κάθε διάνυσμα  $x_i$  και τον κ-οστο κοντινότερο γείτονα του ( $x_k$ ) στη διάσταση m, την απόσταση που δίνεται από τον τύπο

$$R(x_i, k, m) = \sqrt{\sum_{j=1}^m (x_{i+(j-1)\tau} - x_{k+(j-1)\tau})^2} \quad (14)$$

Καθώς πηγαίνουμε από την διάσταση m στην διάσταση m+1 η απόσταση γίνεται

$$R(x_i, k, m+1) = \sqrt{\sum_{j=1}^{m+1} (x_{i+(j-1)\tau} - x_{k+(j-1)\tau})^2} = \sqrt{(R(x_i, k, m))^2 + (x_{i+m\tau} - x_{k+m\tau})^2} \quad (15)$$

Για το εάν ένα σημείο θεωρείται κοντινότερος γείτονας ή όχι έχουμε 2 κριτήρια. Το πρώτο κριτήριο δείχνει εάν είναι σημαντική η αύξηση της απόστασης του σημείου και του κοντινότερου γείτονα όταν αυξάνει η διάσταση εμβύθισης.

$$\sqrt{\frac{R^2(x_i, k, m+1) - R^2(x_i, k, m)}{R^2(x_i, k, m)}} = \frac{|x_{i+m\tau} - x_{k+m\tau}|}{R(x_i, k, m)} > R_{tol} \quad (16)$$

Εάν ισχύει το παραπάνω κριτήριο τότε έχουμε λανθασμένο κοντινότερο γείτονα. Το  $R_{tol} > 10$  με συνήθη τιμή το 15.

Το δεύτερο κριτήριο συγκρίνουμε την απόσταση του σημείου από τον κοντινότερο γείτονα με το μέγεθος του ελκυστή

$$\frac{R(x_i, 1, m+1)}{R_A} > A_{tol} \quad (17)$$

$$R_A = \sqrt{\frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N}} \quad (18)$$

Η συνήθης τιμή του  $A_{tol}$  είναι 2. Αν ισχύει το κριτήριο πάλι έχουμε λανθασμένο κοντινότερο γείτονα.

Μια ικανοποιητική διάσταση εμβύθισης έχουμε όταν το ποσοστό το λανθασμένων κοντινότερων γειτόνων βρεθεί κάτω από το 1% και σταθεροποιηθεί εκεί και για τις επόμενες διαστάσεις εμβύθισης.

### 2.3.1.2 αμοιβαία πληροφορία

Η αμοιβαία πληροφορία είναι μια ιδέα που βρέθηκε από τον Claude Shannon (Shannon, 49). Η αμοιβαία πληροφορία προσπαθεί να μετρήσει σε bits το ποσό της πληροφορίας που μπορεί να συμπεραθεί για μια σειρά από σύμβολα από μια άλλη. Γενικά δοθέντων 2 σειρών  $x, y$  με δείκτες  $i$  και  $j$  αντίστοιχα, η μέση αμοιβαία πληροφορία μπορεί να υπολογιστεί ως εξής :

$$I(x, y) = \sum_{i,j} P(x_i, y_j) \log \left( \frac{P(x_i, y_j)}{P(x_i)P(y_j)} \right) \quad (19)$$

Υπόψην ότι η αμοιβαία πληροφορία είναι θετική και συμμετρική, δηλαδή  $I(x, y) > 0$ ,  $I(x, y) = I(y, x)$ .

Η αμοιβαία πληροφορία μετριέται χρησιμοποιώντας 2 αντίγραφα των ίδιων σειρών, η μια με καθυστέρηση  $t$ . Καθώς το  $t$  αυξάνεται ελπίζουμε πως τα διάφορα μηνύματα θα παρατηρηθούν στην καμπύλη που παράγεται και η τιμή του  $t$  στο πρώτο ελάχιστο επιλέγεται. Θέλουμε να διαλέξουμε διαχώριση εμπύθισης έτσι ώστε κάθε στήλη στις εμπυθιζόμενες προμήθειες το μέγιστο ποσό νέας πληροφορίας .

## 2.4 Διαστάσεις

Οι διαστάσεις μας βοηθάνε να υπολογίσουμε πόσο περίπλοκο είναι ένα σύστημα. Η μέτρηση των διαστάσεων μας επιτρέπει να εκτιμήσουμε ή να συγκρίνουμε την γεωμετρική πολυπλοκότητα αντικειμένων ή συστημάτων, διαφορετικού μεγέθους, σχήματος και κατασκευής. Οι διαστάσεις είναι χρήσιμες στην θεωρία του χάους για δύο λόγους κυρίως:

- Μετρώντας την διάσταση μπορούμε να διακρίνουμε μεταξύ χάους και τυχαιότητας. Η διάσταση ενός αντικειμένου χαρακτηρίζει την γεωμετρία του και την φυσική του οντότητα. Ο χασοτικός ελκυστής έχει ένα σχήμα στον χώρο των φάσεων. Αυτό το σχήμα έχει μια πεπερασμένη και μετρήσιμη διάσταση. Αντίθετα εάν ένα σύστημα λειτουργεί τυχαία τότε η διάσταση εμπύθισης του ταυτίζεται με την διάσταση του συστήματος η οποία γίνεται άπειρη εάν συνεχίζουμε να αυξάνουμε την διάσταση εμπύθισης του. Στην πράξη, δεδομένα με θόρυβο μας εμποδίζουν να αναγνωρίσουμε την χασοτική συμπεριφορά.
- Η διάσταση του ελκυστή μας επιτρέπει να υπολογίσουμε τον αριθμό των μεταβλητών που είναι αναγκαίες για να περιγράψουμε το σύστημα απ' το οποίο προήλθε. Για να μοντελοποιήσουμε επομένως ένα σύστημα χρειάζεται να έχουμε τουλάχιστον τόσες μεταβλητές όσο είναι η διάσταση του.

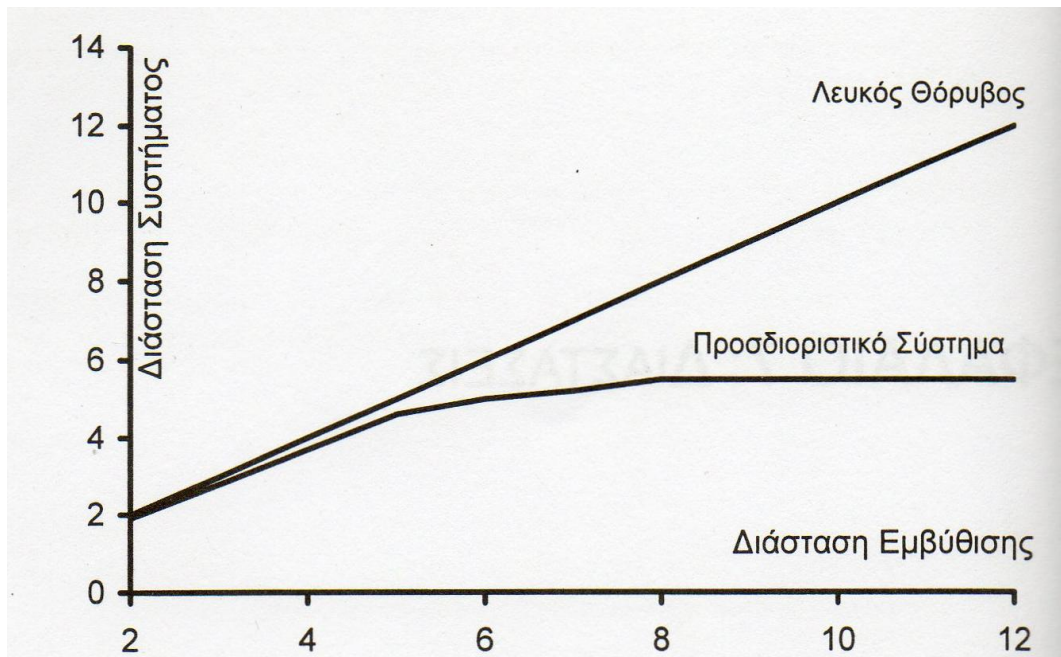
Υπάρχουν πολλές διαφορετικοί τύποι διαστάσεων. Οι πιο γνωστοί είναι:

- Η διάσταση ομοιότητας
- Η διάσταση χωρητικότητας
- Η διάσταση Hausdorff-Besicovich
- Η διάσταση πληροφορίας
- Η διάσταση συσχέτισης
- Η κλασματική διάσταση
- Η διάσταση Lyapunov

Από τις παραπάνω διαστάσεις, θα αναλύσουμε πέντε από αυτές. Οι διάφορες αυτές διαστάσεις είναι αλληλοσυσχετιζόμενες. Μερικές μάλιστα έχουν την ίδια αριθμητική τιμή υπό κατάλληλες συνθήκες. Οι περισσότερες από τις παραπάνω διαστάσεις εντάσσονται σε 2 κατηγορίες. Η πρώτη κατηγορία μετράει μόνο την γεωμετρία του ελκυστή. Τέτοιες είναι η διάσταση ομοιότητας, η διάσταση χωρητικότητας και η διάσταση Hausdorff-Besicovich. Η άλλη κατηγορία εξετάζει όχι μόνο την γεωμετρία αλλά και τις πιθανοθεωρητικές και πληροφοριακές πλευρές του ελκυστή. Λαμβάνει δηλαδή υπόψη της ότι η τροχιά μπορεί να επισκεφτεί κάποιες ‘γειτονίες’ στον χώρο των φάσεων πιο συχνά από τις άλλες. Παραδείγματα αυτής της κατηγορίας αποτελούν η διάσταση πληροφορίας και η διάσταση συσχέτισης. Η τελευταία φαίνεται να είναι και η πιο διάσημη στην θεωρία του χάους.

Το σκεπτικό, συνοπτικά, για τη μέτρηση των διαστάσεων είναι το ακόλουθο: εμβυθίζουμε τη χρονοσειρά σε διαδοχικές διαστάσεις και προσπαθούμε να διακρίνουμε κάποια σταθεροποίηση στη σύνθεση του αντικειμένου που σχηματίζεται. Το αντικείμενο που σχηματίζεται είναι ο ελκυστής όπως τον βλέπουμε μέσα από τη μέθοδο εμβύθισης που χρησιμοποιούμε. Βέβαια σε 4 και άνω διαστάσεις δε μπορούμε να έχουμε μια συνολική οπτική εικόνα του ελκυστή. Ο ελκυστής αυτός, ακόμη και όταν έχουμε χρονοσειρά που προέρχεται από συνεχές σύστημα, δεν παύει να είναι ένα σύνολο σημείων. Το σύνολο αυτό των σημείων σε άλλες περιοχές του χώρου των φάσεων έχει υψηλή πυκνότητα και σε άλλες χαμηλή πυκνότητα. Έχουμε, λοιπόν, ένα αντικείμενο με διαφορετική σύνθεση και, για κάποια συγκεκριμένη σύνθεση προσπαθούμε να διακρίνουμε το ελάχιστο αριθμό των διαστάσεων. Επιπλέον, δεν έχουμε χρονοσειρά λευκού θορύβου ώστε διάσταση να μη σταθεροποιείται ποτέ (Σχήμα 17).





Σχήμα 17: Διάγραμμα διάστασης συστήματος –διάστασης εμβύθισης

## Η διάσταση Fractal

Η διάσταση fractal η αλλιώς κλασματική διάσταση είναι μια στατιστική ποσότητα που δίνει μια ένδειξη κατά το πόσο γεμίζει το fractal χώρο, καθώς κάποιος μεγεθύνει σε μικρότερες κλίμακες. Το fractal είναι γενικά ένα κομματιασμένο γεωμετρικό σχήμα το οποίο μπορεί να χωριστεί σε τμήματα το καθένα από τα οποία είναι (τουλάχιστον περίπου) ένα μειωμένο σε μέγεθος αντίγραφο του όλου, μια ιδιότητα που ονομάζεται αυτοομοιότητα. Παράδειγμα ενός fractal αποτελεί το τρίγωνο Sierpinski. Τα fractals έχουν τα ακόλουθα χαρακτηριστικά:

1. Έχουν καλή δομή σε αυθαίρετα μικρές κλίμακες .
2. Είναι πολύ ακανόνιστα για να περιγραφούν με την παραδοσιακή Ευκλείδεια γλώσσα .
3. Παρουσιάζει αυτοομοιότητα.
4. Έχει διάσταση Hausdorff μεγαλύτερη από την τοπολογική του διάσταση .
5. Έχει απλό και επαναληπτικό προσδιορισμό.

Επειδή παρουσιάζονται να είναι παρόμοια σε όλα τα επίπεδα μεγέθυνσης, τα fractals συχνά θεωρούνται πως είναι απείρως περίπλοκα. Φυσικά αντικείμενα που προσεγγίζουν σε κάποιο βαθμό τα fractals είναι τα σύννεφα, οι βουνοκορφές, οι αστραπές, οι ακτογραμμές και οι νιφάδες χιονιού.

## Η διάσταση χωρητικότητας

Θα ξεκινήσουμε με την διάσταση χωρητικότητας (capacity dimension) ή διάσταση άθροισης κουτιών (box-counting dimension). Είναι πολύ εύκολη στον υπολογισμό και, συνεπώς μπορεί πιο εύκολα να μας εισάγει στην έννοια των υπολοίπων διαστάσεων.

Με την διάσταση αυτή μετράμε τον όγκο που καταλαμβάνει ο ελκυστής στο χώρο των φάσεων. Βέβαια φαίνεται το ερώτημα: Πως είναι δυνατόν ένα σύνολο από πεπερασμένα σημεία να έχουν όγκο; Καλύπτουμε με ένα πλέγμα (grid) τον ελκυστή και μετράμε πόσα κουτιά (πλευράς  $r$ ) περιέχουν έστω και 1 σημείο. Αυτό που κάνουμε ουσιαστικά είναι να υπολογίζουμε την πιθανότητα να υπάρχει τουλάχιστο 1 σημείο μέσα σε ένα κουτί πλευράς  $r$  και, στη συνέχεια, να υψώσουμε την πιθανότητα αυτή στη μηδενική δύναμη. Ας πούμε ότι είναι  $N(r)$  το όλων των κουτιών που μετρήσαμε. Στη συνέχεια, αυξάνουμε την ανάλυση. Το  $r$  γίνεται  $r'$ , με  $r' < r$ , άρα το πλέγμα γίνεται πιο πυκνό, αφού αποτελείται τώρα από περισσότερα κουτιά. Μετράμε ξανά πόσα κουτιά περιέχουν έστω και 1 σημείο και είναι  $N(r')$ . Επαναλαμβάνουμε τη διαδικασία διαδοχικά σε όλο και μικρότερα μήκη πλευρών, ώσπου το μήκος της πλευράς να τείνει στο 0. Η διάσταση χωρητικότητας είναι η κλίση της παλινδρόμησης  $\log N(r), \log N(r'), \dots$  έναντι  $-\log r, -\log r', \dots$ . Πιο αναλυτικά υπολογίζουμε:

$$\log N(r) = D_0 \cdot \log\left(\frac{1}{r}\right) + a \quad (20)$$

Όπου :  $D_0$  είναι η κλίση της παλινδρόμησης (διάσταση χωρητικότητας)

$a$  είναι η σταθερά της παλινδρόμησης που δε μας ενδιαφέρει

Η σχέση 1 προκύπτει από τον ορισμό της διάστασης χωρητικότητας που είναι

$$D_0 = \lim_{r \rightarrow 0} \frac{\log N(r)}{\log\left(\frac{1}{r}\right)} \quad (21)$$

Η διάσταση χωρητικότητας είναι προσδιορίσιμη μόνο όταν η ακτίνα  $r$  πλησιάζει την τιμή μηδέν. Στην πράξη, υπάρχει προφανώς ένα όριο του πόσο μικρή μπορεί να είναι η ακτίνα αυτή. Επομένως, η διάσταση χωρητικότητας μπορεί μόνο να εκτιμηθεί στο περίπου και όχι ακριβώς.

## Η διάσταση πληροφορίας

Η τροχιά σε ένα χαοτικό ελκυστή μπορεί να ακολουθήσει πολλές διαφορετικές κατευθύνσεις. Η διάσταση ομοιότητας, χωρητικότητας και η Hausdorff λαμβάνουν υπόψη τους μόνο την γεωμετρία του ελκυστή. Οι επόμενες δύο διαστάσεις, η διάσταση πληροφορίας και συσχέτισης, προσπαθούν όχι μόνο να αντικατοπτρίσουν την γεωμετρία του ελκυστή αλλά και την συχνότητα με την οποία το σύστημα βρίσκεται σε διάφορους τομείς του χώρου των φάσεων. Με όλες τις μεθόδους υπολογισμού της διάστασης καλύπτουμε τον ελκυστή με ένα επιλεγμένο πρότυπο μέτρησης (κύκλο, σφαίρα, κύλινδρο κτλ.) και του δίνουμε ένα μέγεθος. Και μετά το

επαναλαμβάνουμε μεγαλώνοντας (ή μικραίνοντας) το μέγεθος αυτού του προτύπου. Στις δύο αυτές μεθόδους υπολογισμού της διάστασης κάνουμε μετρήσεις που αντικατοπτρίζουν πόσο συχνά μια τροχιά πλησιάζει ένα προηγούμενο μονοπάτι της, η συχνότητας με την οποία επισκέπτεται ένα δεδομένο υποχώρο στον ελκυστή. Για να έχουμε τις καλύτερες δυνατές μετρήσεις τα δεδομένα μας πρέπει να είναι πολλά και να καλύπτουν ένα μεγάλο χρονικό περιθώριο.

Αυτό που κάνουμε στη διάσταση πληροφορίας (information dimension) είναι να μετρήσουμε από κάθε κουτί την εντροπία πληροφορίας. Υπολογίζουμε το  $I(r)$  το οποίο δίνεται από την σχέση:

$$I(r) = \sum_{i=1}^{N_s} P_i \log_2 \left( \frac{1}{P_i} \right) \quad (22)$$

Όπου  $N_s$  είναι ο αριθμός των κουτιών ακτίνας  $r$  που έχουν μη μηδενική πιθανότητα,  $i$  είναι το κουτί  $i$ , και  $P$  η πιθανότητα που καθορίστηκε για το κουτί  $i$ .

Το  $r$ , όπως και στη συνέχεια μειώνεται σε  $r'$ . Υπολογίζουμε ξανά το  $I(r')$  και συνεχίζουμε. Η διάσταση πληροφορίας βρίσκεται στην κλίση της παλινδρόμησης  $I(r)$ ,  $I(r')$ , ... έναντι  $-\log r$ ,  $-\log r'$ , ... . Παρόμοια με τα παραπάνω υπολογίζουμε:

$$I(r) = D_1 \cdot \log \left( \frac{1}{r} \right) + a \quad (23)$$

Όπου  $D_1$  είναι η κλίση της παλινδρόμησης (διάσταση πληροφορίας)

$A$  είναι η σταθερά της παλινδρόμησης που δε μας ενδιαφέρει

Η σχέση 3 προκύπτει από τον ορισμό της διάστασης πληροφορίας που είναι

$$D_1 = \lim_{r \rightarrow 0} \frac{I(r)}{\log \left( \frac{1}{r} \right)} \quad (24)$$

Η πληροφορία συνήθως (όπως έχουμε πει) μετριέται σε bits. Αν θέλουμε να τη μετρήσουμε με αυτόν τον τρόπο χρησιμοποιούμε λογάριθμο με βάση το 2. Με οποιαδήποτε βάση λογαρίθμου μετρήσουμε την πληροφορία θα πρέπει να έχουμε υπόψη μας τα εξής :

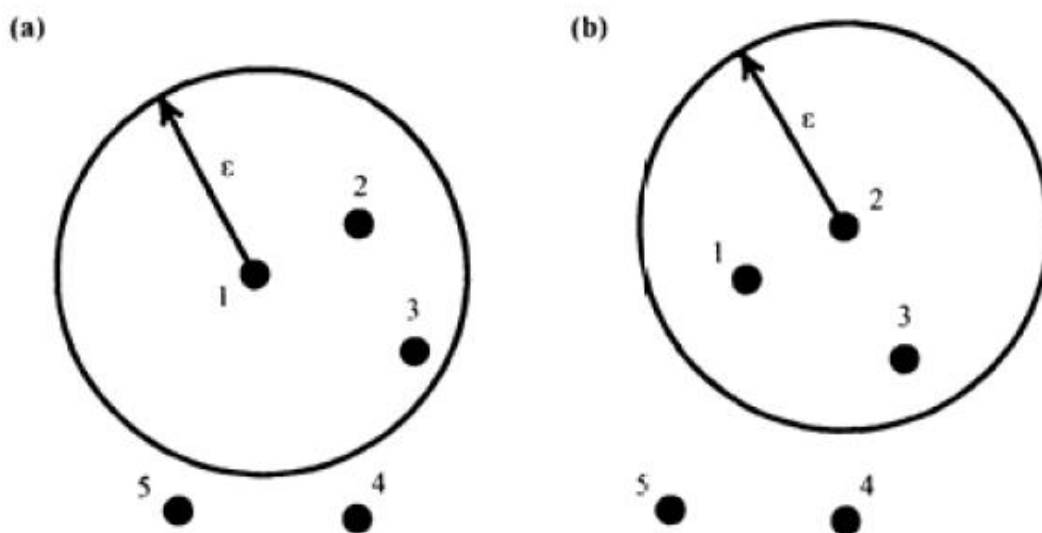
1. Θα πρέπει να χρησιμοποιούμε την ίδια βάση για την μέτρηση της απόστασης  $r$  και του  $I(r)$  .
2. Δεν έχει σημασία η βάση των λογαρίθμων για την μέτρηση της διάστασης πληροφορίας. Η διάσταση πληροφορίας δεν εξαρτάται από τη βάση.  
Τέλος αν δεν συγκλίνει πουθενά η διάσταση πληροφορίας αυτό σημαίνει είτε ότι έχουμε μεγάλο ποσοστό θορύβου, είτε ότι η χρονοσειρά μας είναι πολύ μικρή σε μήκος, ή και τα δύο.

## Η διάσταση συσχέτισης

Η διάσταση συσχέτισης (correlation dimension) είναι αυτή που χρησιμοποιείται συνήθως για να μας δείξει τη διάσταση του συστήματος. Η διάσταση συσχέτισης είναι σαν την διάσταση πληροφορίας αλλά λίγο πιο περίπλοκη. Πολλοί τύποι διαστάσεων είναι αδύνατον να τους υπολογίσεις στην πράξη, διότι είναι υπολογιστικά ανεπαρκής. Η διάσταση συσχέτισης δεν έχει τέτοια προβλήματα για τον υπολογισμό της. Δύο σημεία δεδομένων στο χώρο των φάσεων έχουν μεγάλη χωρική συσχέτιση ( η τιμή του ενός είναι παραπλήσια με του άλλου). Παρόλα αυτά όμως τα δύο αυτά σημεία μπορεί να μην έχουν καμία συσχέτιση στο χρόνο. Η διάσταση συσχέτισης ελέγχει μόνο την χωρική τους συσχέτιση και αγνοεί τον χρόνο.

Η διαδικασία υπολογισμού της διάστασης συσχέτισης περιλαμβάνει τον υπολογισμό πρώτα όχι μόνο της χρονικής υστέρησης αλλά και τον υπολογισμό της διάστασης εμφύθισης. Όμως δεν υπάρχει τρόπος να υπολογίσουμε εξ αρχής την διάσταση εμφύθισης, διότι αυτή εξαρτάται από την πραγματική διάσταση του ελκυστή, και αυτή είναι που ψάχνουμε να βρούμε. Η σωστή διάσταση εμφύθισης προκύπτει από την ανάλυση που θα πραγματοποιηθεί.

Μόλις καθοριστεί η χρονική υστέρηση, ξεκινάμε συνήθως με διάσταση εμφύθισης δύο. Στην συνέχεια τοποθετούμε τα κελιά μέτρησης, τα οποία στην συγκεκριμένη περίπτωση είναι κύκλοι, με τέτοιο τρόπο ώστε το σημείο που εξετάζουμε να είναι στο κέντρο. Μετά μετράμε τον αριθμό των σημείων που βρίσκονται μέσα στο κελί αυτό, και μετά μεταφέρουμε το κέντρο στο επόμενο σημείο της τροχιάς και ακολουθούμε την ίδια διαδικασία. Έστω ότι έχουμε ένα δείγμα από πέντε σημεία μόνο.



Σχήμα 18: διδιάστατη αναπαράσταση των 5 σημείων και τις ακτίνας  $r=\epsilon$  από την οποία καθορίζεται ποια είναι τα σημεία που θεωρούνται κοντινότεροι γείτονες για τα σημεία a) 1 και b) 2.

Στο παραπάνω σχήμα φαίνονται τα πέντε σημεία αυτά και είναι αριθμημένα με βάση με βάση την σειρά που τα επισκέπτεται η τροχιά. Γύρω από το σημείο 1 δημιουργούμε ένα κύκλο με ακτίνα  $r = \varepsilon$  και κέντρο το σημείο αυτό. Παρατηρούμε ότι τα σημεία 2 και 3 βρίσκονται μέσα στην ακτίνα αυτή, ενώ τα σημεία 4 και 5 εκτός (εδώ αξίζει να σημειωθεί ότι ο υπολογιστής υπολογίζει αν τα σημεία είναι εντός ή εκτός υπολογίζοντας την ευκλείδεια απόσταση τους από το σημείο 1).

Στην συνέχεια κεντράρουμε τον κύκλο στο σημείο 2 με την ίδια ακτίνα  $r$ . Τα σημεία 1 και 3 βρίσκονται μέσα στον κύκλο. Αφού ακολουθήσουμε την ίδια διαδικασία για όλα τα σημεία με την ίδια ακτίνα, προσθέτουμε τον αριθμό όλων των σημείων που βρίσκονταν εντός κύκλου για αυτήν την ακτίνα  $r$ . Για το παράδειγμά μας ο ολικός αριθμός ήταν 8 (2 και το σημείο 1, 2 για τα σημεία 2 και 3, και 1 για τα σημεία 4 και 5).

Προφανώς ο ολικός αριθμός των σημείων που καθορίζουν την τροχιά επηρεάζει τον ολικό αριθμό των σημείων που βρίσκονται εντός ακτίνας, για δεδομένη ακτίνα. Κανονικοποιούμε τον αριθμό αυτών των σημείων, με τον αριθμό όλων των διαθέσιμων σημείων στην τροχιά. Το κανονικοποιημένο αποτέλεσμα αυτό καλείται ακέραιος συχέτιση άθροισμα συχέτισης ακτίνας  $r$ .

$$C_r = \frac{\text{ολικός αριθμός σημείων εντός ακτίνας } r}{\text{ο μεγαλύτερος αριθμός πιθανών σημείων}} \quad (25)$$

$$C_r = \frac{\text{ολικός αριθμός σημείων εντός ακτίνας } r}{N(N-1)} \quad (26)$$

Όπου  $N =$  ο ολικός αριθμός των σημείων της τροχιάς.

Μια ειδική εκδοχή της παραπάνω συνάρτησης προκύπτει όταν το  $N$  γίνεται πολύ μεγάλο. Όταν το  $N$  είναι πολύ μεγάλο τότε το  $N-1 = N$  και συνάρτησή μας γίνεται:

$$C_r = \lim_{N \rightarrow \infty} \frac{\text{ολικός αριθμός εντός ακτίνας } r}{N^2} \quad (27)$$

Πρακτικά η παραπάνω σχέση είναι χρησιμοποιήσιμη όταν το  $N$  είναι μεγαλύτερο από μερικές εκατοντάδες.

Αφού καθορίσουμε το  $C_r$  για την ακτίνα  $r$  αυξάνουμε την ακτίνα και επαναλαμβάνουμε την διαδικασία. Συνεχίζουμε αυξάνοντας την ακτίνα και συνεπώς παίρνοντας όλο και μεγαλύτερα αθροίσματα συχέτισης. Στην συνέχεια αυξάνουμε την διάσταση εμβύθισης.

Το 1983 οι Grasberger και Procaccia έκαναν ένα πολύ σημαντικό βήμα στη μέτρηση της διάστασης συχέτισης. Έδειξαν μια καινούργια μεθοδολογία με υπολογιστική διαδικασία πολύ ταχύτερη από την τότε υπάρχουσα. Ο αλγόριθμος αυτός ως προς την ταχύτητα έχει βελτιωθεί ξανά.. Σε μεγάλη διάσταση εμβύθισης, η αύξηση της ακτίνας θα επιφέρει εκθετική αύξηση των σημείων που περιλαμβάνει. Ξεκινάμε με μια μικρή ακτίνα  $r$ . Ο μέσος όρος των πιθανοτήτων όλων των σημείων, να περικλείσουμε κάποιο άλλο σημείο του ελκυστή  $C(r)$ . Μια μεγαλύτερη ακτίνα  $r'$  μας δίνει το  $C(r')$ .

Η διάσταση συσχέτισης βρίσκεται στην κλίση της παλινδρόμησης  $C(r), C(r'), \dots$ , έναντι  $\log r, \log r', \dots$ . Παρόμοια με τα παραπάνω υπολογίζουμε:

$$C(r) = D_2 \cdot \log r + a \quad (28)$$

Όπου:  $D_2$  είναι η κλίση της παλινδρόμησης (διάσταση συσχέτισης)

$a$  είναι η σταθερά της παλινδρόμησης, που δε μας ενδιαφέρει

Η σχέση (28) προκύπτει από τον ορισμό της διάστασης χωρητικότητας που είναι

$$D_2 = \lim_{r \rightarrow 0} \frac{\log C(r)}{\log r} \quad (29)$$

Η εύρεση της διάστασης συσχέτισης γίνεται στο χώρο των φάσεων με τη μέθοδο κλίσης-παλινδρόμησης και για διάφορες διαστάσεις εμβύθισης.

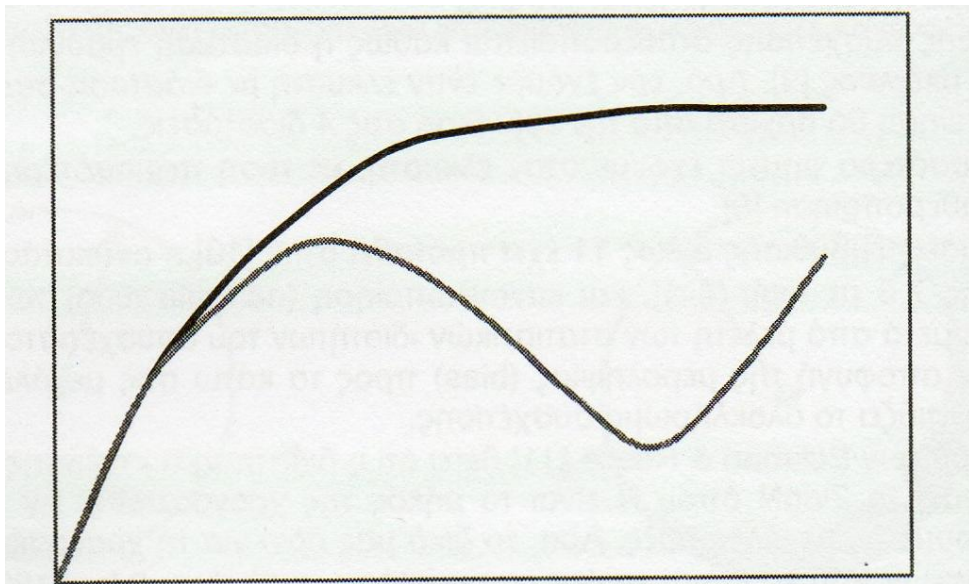
Μερικές άλλες παρατηρήσεις πάνω στο θέμα της διάστασης συσχέτισης παραθέτονται αμέσως παρακάτω, ταξινομημένες.

- Είναι δυνατό να ξεχωρίσουμε μεταξύ περιοδικής, χαοτικής, και τυχαίας συμπεριφοράς ακόμη και με λίγα γεμάτα θόρυβο δεδομένα.
- Ενδιαφέρον έχει, επίσης, η μέτρηση μέσα από το φίλτρο της PCA με σκοπό τη μείωση της επίδρασης του θορύβου.
- Σε περίπτωση που η παρουσία του θορύβου είναι ισχυρή θα δούμε το ολοκλήρωμα συσχέτισης να παρουσιάζει δύο καμπύλες: μια καμπύλη στα μικρά  $r$ , ίσο με τη διάσταση εμβύθισης και, μια δεύτερη, ίση με τη διάσταση συσχέτισης χωρίς θόρυβο. Η αναγνώριση των δύο κλίσεων είναι εξαιρετικά δύσκολη λόγω των περιορισμένων δεδομένων. Όσο μεγαλώνει το  $r$  και τείνει να εξισωθεί με τη διάμετρο του ελκυστή, η κλίση του ολοκληρώματος συσχέτισης τείνει στο 0.
- Η διάσταση συσχέτισης σταθεροποιείται καθώς η διάσταση εμβύθισης είναι όσο ο επόμενος ακέραιος. Άρα, εάν έχουμε έναν ελκυστή με διάσταση συσχέτισης 3, 4 η σταθεροποίηση θα αρχίσει από την εμβύθιση στις 4 διαστάσεις.
- Όσο περισσότερα σημεία έχουμε στον ελκυστή, με τόση περισσότερη σαφήνεια θα γίνει η σταθεροποίηση.
- Για διαστάσεις εμβύθισης 5 ως 11 έχει προταθεί η αντικατάσταση του  $\log r$  της σχέσης 5 με  $\log[r(2-r)]$ , και κοινωνικοποίηση (normalization) των δεδομένων. Αυτό έγινε μετά από μελέτη των στατιστικών ιδιοτήτων του ασυσχέτιστου θορύβου με σκοπό την αποφυγή της μεροληψίας προς τα κάτω στις μεγάλες διαστάσεις, που παρουσιάζει το ολοκλήρωμα συσχέτισης.
- Το θεώρημα των Eckman & Ruelle θέτει ότι η διάσταση συσχέτισης δε θα πρέπει να ξεπερνάει το  $2 \log N$  όπου  $N$  είναι το μήκος της χρονοσειράς. Αν το ξεπερνάει, απορρίπτουμε τις μεγάλες τιμές.
- Δύο μεγάλα θέματα που απασχολούν τους ερευνητές είναι ο περιορισμένος αριθμός των παρατηρήσεων και η μη στασιμότητα των δεδομένων.

- Μια ακόμη εκτενής ανάλυση για τη διάκριση μεταξύ χαμηλοδιάστατης συμπεριφοράς και τυχαίας (συσχετισμένης ή ασυσχέτιστης). Είναι γνωστό ότι αυτοσυσχετισμένος θόρυβος μπορεί να σταθεροποιήσει τη διάσταση συσχέτισης σε κάποια τιμή.

Χρησιμοποιούνται οι έλεγχοι:

1. Διάγραμμα απόστασης χώρου-χρόνου. Έστω ότι, έχουμε δύο τυχαία σημεία στο χώρο των φάσεων της χρονοσειράς.



**Σχήμα 19:** Διάγραμμα χώρου-χρόνου. Στον οριζόντιο άξονα είναι η χρονική διαφορά δύο αρχικά κοντινών διανυσμάτων. Με τη μαύρη γραμμή βλέπουμε τα αναμενόμενα αποτελέσματα για μια τυχαία χρονοσειρά. Η απόσταση μεγαλώνει καθώς αυξάνεται ο χρόνος. Με τη γκριζα γραμμή βλέπουμε τα αποτελέσματα για μια ντετερμινιστική χρονοσειρά. Τα σημεία του χώρου των φάσεων αρχικά απομακρύνονται, άρα η απόσταση μεγαλώνει, ενώ στη συνέχεια τα σημεία επιστρέφουν στις ίδιες περίπου περιοχές, άρα η απόσταση μειώνεται ξανά.

Σε ντετερμινιστική χρονοσειρά παρατηρούμε μια απομάκρυνση των σημείων που στη συνέχεια μειώνεται και αυξάνεται ξανά. Σε τυχαία χρονοσειρά η απομάκρυνση στο χρόνο είναι και απομάκρυνση στο χώρο(βλέπε σχήμα 19).

2. Αναδιαμορφωμένα δεδομένα (surrogate data)
3. Πρώτες διαφορές. Σε τυχαία χρονοσειρά οι πρώτες διαφορές μπορούν να διασκορπίσουν συσχετίσεις που σταθεροποιούν την διάσταση.
4. Ανεξάρτητες θεωρήσεις (independent realizations): παίρνουμε διαφορετικές αρχικές συνθήκες και πραγματοποιούμε το πείραμα. Σε τυχαία χρονοσειρά αν κατά τύχη την πρώτη φορά θα δούμε σταθεροποίηση. Για να πούμε ότι επαναλαμβάνουμε το πείραμα με

διαφορετικές αρχικές συνθήκες, αρκεί να έχουμε μια μεγάλη χρονοσειρά και να τη χωρίσουμε στα δύο. Οι δύο χρονοσειρές που θα πάρουμε με αυτό τον τρόπο θα πρέπει να έχουν τα ίδια χαρακτηριστικά(διάσταση, εκθέτες Lyapunov κτλ). Θεωρούμε έτσι ότι, έχουμε δύο διαφορετικές αρχικές συνθήκες που είναι ανεξάρτητες μεταξύ τους. Αν πάρουμε όμως διαφορετικά αποτελέσματα αυτό σημαίνει ότι έχουμε τυχαιότητα.

5. Συνάρτηση δομής (structure function): Δίνεται από τον τύπο  $S(n) = \sum_{j=1}^{N-n} (X_{i+n} - X_i)^2$ . Για χαμηλής διάστασης χρονοσειρά περιμένουμε  $S(n) \propto n^{2H}$ . Αυτός ο έλεγχος θα πρέπει οπωσδήποτε να συνδυαστεί με τους 3 και 4.

## Η διάσταση Lyapunov

Είναι ένα μέγεθος που έχει ορισθεί από τους Kaplan και Yorke ως εξής:

$$D_L = j + \frac{\sum_{i=1}^j \lambda_i}{|\lambda_{j+1}|} \quad (30)$$

Όπου το  $j$  καθορίζεται ως εξής:

$$\sum_{i=1}^j \lambda_i > 0 \wedge \sum_{i=1}^{j+1} \lambda_i < 0 \quad (31)$$

Το σύμβολο  $\wedge$  είναι το λογικό και, δηλαδή θα πρέπει να ισχύουν και οι δύο σχέσεις.

Η διάσταση Lyapunov είναι ίση με τη διάσταση πληροφορίας. Η μέτρηση των διαστάσεων ουσιαστικά μας φανερώνει την πολυπλοκότητα της χρονοσειράς. Θα πρέπει να γίνει με υπομονή και προσοχή, γιατί τα αποτελέσματα θα ενισχύσουν την άποψη ότι έχουμε να κάνουμε με χάος και όχι με θόρυβο. Προτείνουμε τη διάσταση συσχέτισης σαν ένα ικανό εκτιμητή της διάστασης του συστήματος που δίνει την χρονοσειρά. Είναι πραγματικά πολύ δύσκολο να ξεχωρίσουμε μεταξύ χαμηλοδιάστατου χάους και αυτοσυσχετιζόμενου θορύβου. Επίσης, είναι πολύ εύκολο να αναγνωρίσουμε χάος σε κυκλική διαδικασία με θόρυβο

## Έλεγχος στατικής σημαντικότητας

Μέχρι στιγμής, αυτό που κάναμε είναι να μετράμε τη διάσταση από τη χρονοσειρά και να θεωρούμε ότι αυτό είναι το σωστό αποτέλεσμα. Σε ένα αποτέλεσμα όμως θα πρέπει πάντα να υπολογίζουμε και τη στατική του σημαντικότητα. Δεν μπορούμε για παράδειγμα να είμαστε βέβαιοι ότι η διάσταση συσχέτισης είναι 5 όταν οι κλίσεις που σχηματίζει το ολοκλήρωμα συσχέτισης κυμαίνονται ομοιόμορφα από 1 μέχρι 9. Από την άλλη πλευρά, έχει αποδειχθεί ότι, ο φιλτραρισμένος θόρυβος μπορεί να μιμηθεί τη χαμηλοδιάστατη χαοτική συμπεριφορά.

Η μέθοδος για τον έλεγχο στατιστικής σημαντικότητας που έχει επινοηθεί και έχει εφαρμοστεί σε χρηματοοικονομικά δεδομένα με επιτυχία ονομάζεται



αναδιαμορφωμένα δεδομένα (surrogate data). Τα αναδιαμορφωμένα δεδομένα είναι χρονοσειρές οι οποίες έχουν τα ίδια χαρακτηριστικά με τη χρονοσειρά που έχουμε. Τα χαρακτηριστικά αυτά είναι ο μέσος όρος, η τυπική απόκλιση, η κατανομή συχνοτήτων, η συνάρτηση αυτοσυσχέτισης κ.α. Η χρονοσειρά μας είναι μία από τις πολλές που μπορούν να σχηματιστούν με τα χαρακτηριστικά αυτά. Γνωρίζουμε επίσης ότι η χρονοσειρά μας έχει κάτι το ιδιαίτερο που οι άλλες χρονοσειρές δεν έχουν: παράγεται από μια χαμηλοδιάστατη ντετερμινιστική διαδικασία (ενώ οι άλλες είναι τυχαίοι αριθμοί με τα ίδια χαρακτηριστικά). Είναι πράγματι έτσι, ή οποιαδήποτε χρονοσειρά με τα ίδια χαρακτηριστικά μπορεί να δώσει τα ίδια αποτελέσματα όσον αφορά τη διάσταση;

Ο έλεγχος γίνεται ως εξής : μετράμε τη διάσταση των αναδιαμορφωμένων δεδομένων, υπολογίζουμε το μέσο όρο ( $D_{surr}$ ), και την τυπική απόκλιση ( $STD_{surr}$ ) των αποτελεσμάτων. Στη συνέχεια, υπολογίζουμε το

$$G = \frac{|D - D_{surr}|}{STD_{surr}} \quad (32)$$

Όπου  $D$  είναι η διάσταση της χρονοσειράς μας

Εάν  $G \geq 1,96$ , τότε η χρονοσειρά μας διαφέρει σημαντικά από άλλες με τα ίδια χαρακτηριστικά, διαφορετικά δεν μπορούμε να είμαστε βέβαιοι ότι έχουμε να κάνουμε με χαμηλοδιάστατο χάος.

Με τα αναδιαμορφωμένα δεδομένα μπορούμε να διακρίνουμε εάν η διαδικασία μας είναι ARMA(q,p), κυκλική με θόρυβο, χαμηλοδιάστατη χασοτική με θόρυβο κτλ. Αν για οποιονδήποτε λόγο χρειαστεί να μετασχηματίσουμε τα αρχικά μας δεδομένα (πχ πρώτες διαφορές), τότες θα πρέπει να μετασχηματίσουμε τα αναδιαμορφωμένα με τον ίδιο τρόπο. Το να φτιάξουμε αναδιαμορφωμένα από τα μετασχηματισμένα δεδομένα θα μας οδηγήσει σε λάθος συμπεράσματα.

Στον τομέα αυτό, του ελέγχου δηλαδή της στατικής σημαντικότητας των αποτελεσμάτων, έχουν γίνει πολλές έρευνες και γίνονται ακόμα. Όχι τόσο στην εκτίμηση των διαστημάτων εμπιστοσύνης της διάστασης συσχέτισης, αλλά περισσότερο στην παραγωγή των αναδιαμορφωμένων δεδομένων.

## Μείωση θορύβου

Ακόμη και αν καταφέρουμε να λύσουμε το πρόβλημα των λίγων δεδομένων (π.χ. έχουμε πάνω από 10.000 παρατηρήσεις) μένει να αντιμετωπίσουμε το πρόβλημα του θορύβου. Ο θόρυβος, όχι μόνο εμποδίζει να διακρίνουμε τη δυναμική του συστήματος, αλλά περιορίζει και την ικανότητα για έγκυρες προβλέψεις. Ο θόρυβος και οι προβλέψεις δεν είναι δύο πράγματα ξεχωριστά. Κάθε φορά που θέλουμε να κάνουμε μια πρόβλεψη θα πρέπει εκτός των άλλων να λάβουμε υπόψη και το θόρυβο που επικρατεί. Κάθε φορά που θέλουμε να μειώσουμε το θόρυβο θα πρέπει να υποθέσουμε ότι δεν καταστρέφουμε τη δυναμική της χρονοσειράς, η οποία δυναμική μας οδηγεί σε σωστές προβλέψεις.

Τα ερωτήματα είναι πολλά. Πόσο ικανά είναι τα φίλτρα στο να μειώσουν το θόρυβο; μπορούμε με βάση τη δυναμική του συστήματος να μειώσουμε το θόρυβο, τουλάχιστο όταν η δυναμική είναι χαμηλής διάστασης; Μπορεί η δυναμική να υπερισχύσει του θορύβου;

Διακρίνουμε τις ακόλουθες καταστάσεις :

- Γνωρίζουμε το σύστημα που δίνει τη χρονοσειρά. Μπορούμε τότε να μειώσουμε το θόρυβο μέχρι την ακρίβεια μηχανής.
- Γνωρίσαμε τη χρονοσειρά χωρίς θόρυβο. Ας πούμε ότι μετρήσαμε στο παρελθόν το μέγεθος και υποθέτουμε ότι, αυτό που μετρήσαμε ήταν χωρίς θόρυβο. Τώρα έχουμε ακόμα μια μέτρηση και θεωρούμε ότι περιέχει κάποιο ποσοστό θορύβου. Καθαρίζουμε το καινούργιο σήμα πιθανοκρατικά. Βρίσκουμε τη μεγαλύτερη πιθανότητα της δεδομένης χωρίς θόρυβο χρονοσειράς για να υπολογίσουμε την καθαρή χρονοσειρά.
- Δεν γνωρίζουμε τίποτα για την δυναμική του συστήματος. Καθαρίζουμε το θόρυβο βαδίζοντας τυφλά! Αυτή είναι και η προσέγγιση που θα κάνουμε. Ο πειραματιστής, ο οικονομολόγος, που δεν γνωρίζει σχεδόν τίποτα για τη δυναμική του συστήματος θα πρέπει να επιλέξει ανάμεσα σε ένα πλήθος μεθόδων που δεν θα πρέπει να είναι ευαίσθητες στις παραμέτρους που θα θέσει. Οι τρεις πρώτες μέθοδοι που θα περιγράψουμε είναι βασισμένες στη γεωμετρία του παράξενου ελκυστή(geometry oriented), ενώ η τελευταία στη δυναμική της τροχιάς του(trajjectory oriented).

## Ορισμοί

Διακρίνουμε 2 είδη θορύβου ανάλογα με το αν έχει ή όχι επιδράσεις στις μελλοντικές τιμές της χρονοσειράς:

- Θόρυβος στις παρατηρήσεις(observational noise): Προέρχεται από τα όργανα των παρατηρήσεων, την έλλειψη ακρίβειας. Τα πεπερασμένα δεκαδικά ψηφία που πάντα διαθέτουμε για μια παρατήρηση αποτελούν θόρυβο. Το χαρακτηριστικό αυτού του είδους είναι ότι, αφού επιδράσει στη χρονοσειρά δεν έχει επιπτώσεις στο μέλλον. Με αυτό το είδος θα ασχοληθούμε παρακάτω.

- Δυναμικός θόρυβος (dynamical noise): Έχει το χαρακτηριστικό ότι, ενσωματώνεται στη δυναμική του συστήματος. Από τη στιγμή που θα ενσωματωθεί αποτελεί αναπόσπαστο μέρος της τιμής της χρονοσειράς και στο μέλλον θα χρησιμοποιείται στην εξέλιξη του συστήματος.

Άλλες δύο κατηγορίες θορύβου ανάλογα με το πώς επιδρά είναι οι ακόλουθες:

- Προσθετικός (addictive): Οι τυχαίες παρατηρήσεις προστίθενται στη χρονοσειρά. Όλες οι μέθοδοι με τις οποίες θα ασχοληθούμε παρακάτω θεωρούν προσθετικό θόρυβο.
- Πολλαπλασιαστικός (multiplicative): Οι τυχαίες παρατηρήσεις πολλαπλασιάζονται στις τιμές της χρονοσειράς.

Επίσης, θόρυβο θα μπορούσαμε να πούμε το λάθος κατά τη μεταφορά των αριθμών στους τόπους αποθήκευσης. Αλλά είναι πολύ περιορισμένο το είδος αυτού του θορύβου και το αφήνουμε έξω από την ανάλυση. Μπορούμε να το εντάξουμε στο πρώτο είδος που αναφέραμε.

Θόρυβος στις χρηματοοικονομικές χρονοσειρές είναι οι μικρές αποκλίσεις από την αναμενόμενη τιμή που οφείλονται σε διάσπαρτες πληροφορίες που επηρεάζουν άλλους επενδυτές λίγο και άλλους πολύ. Μερικά από αυτά ενσωματώνονται μέσα στη χρονοσειρά και χρησιμοποιούνται ξανά ως πληροφορία για τις επόμενες περιόδους. Οι επενδυτές που επηρεάζονται περισσότερο είναι αυτοί που έχουν βραχυχρόνιους ορίζοντες ρευστοποίησης.

Στον πραγματικό κόσμο, θα πρέπει να υπολογίσουμε και έναν τυχαίο (στοχαστικό) παράγοντα  $\eta(t)$  ο οποίος επηρεάζει τη χρονοσειρά. Έτσι, αντί να παρατηρήσουμε  $x(t)$  τελικά παρατηρούμε

$$y(t) = x(t) + \eta(t) \quad (33)$$

Εδώ έχουμε να κάνουμε με θόρυβο από παρατηρήσεις που δρα προσθετικά. Ο σκοπός των μεθόδων που θα περιγράψουμε είναι να διαχωρίσουμε το  $y(t)$  στο  $x(t)$  και στο  $\eta(t)$ . Αφού υπολογίσουμε το  $x(t)$ , τότες είναι πολύ πιο εύκολο να μάθουμε για τη δυναμική του συστήματος  $s(t)$ .

$$x(t) = y(t) - \eta(t) \quad (34)$$

Το  $\eta(t)$  υποθέτουμε ότι ακολουθεί κανονική, αν και στην πραγματικότητα η κατανομή του μας είναι άγνωστη.

$$\eta(t) \approx \eta(t) \sim N(0, \sigma_\eta^2) \quad (35)$$

Ορίζουμε ως λόγο θορύβου προς σήμα (noise-to-signal ratio) την ποσότητα

$$NSR = \frac{\sigma_\eta^2}{\sigma_x^2} \quad (36)$$

Όπου :  $\sigma_\eta$  είναι η τυπική απόκλιση του θορύβου

$\sigma_\chi$  είναι η τυπική απόκλιση της χωρίς θόρυβο χρονοσειράς

Εναλλακτικά χρησιμοποιούμε το λόγο σήματος προς θόρυβο (signal-to-noise ratio) την ποσότητα

$$SNR = \frac{\sigma_\chi^2}{\sigma_\eta^2} \quad (37)$$

Η ισχύρτητα του σήματος μετριέται και σε decibel(dB)

$$SNR(dB) = 10 \cdot \log SNR \text{ dB} \quad (38)$$

## Κεφάλαιο 3

### Προβλέψεις

Διακρίνουμε τις προβλέψεις σε βραχυχρόνιες και μακροχρόνιες. Στις μακροχρόνιες προβλέψεις μας ενδιαφέρει να μάθουμε πώς περίπου θα κυμανθεί η χρονοσειρά, ενώ στις βραχυχρόνιες προβλέψεις μας ενδιαφέρει η ακρίβεια των αποτελεσμάτων. Σχετικά με τις μακροχρόνιες προβλέψεις, μια στάσιμη χρονοσειρά γνωρίζουμε πολύ καλά πού περίπου θα κυμανθεί. Αναμένεται να ακολουθήσει το σταθερό μέσο όρο γύρω από τον οποίο κυμαίνεται. Το ενδιαφέρον μας θα επικεντρωθεί στις βραχυχρόνιες προβλέψεις και, περισσότερο στη πρόβλεψη της επόμενης τιμής της χρονοσειράς.

Ο εκτιμητής σφάλματος που θα χρησιμοποιήσουμε ορίζεται

$$E = \sqrt{\frac{\sum(\hat{x}-x)^2}{\sum(x-\bar{x})^2}} \quad (39)$$

όπου:  $\hat{x}$  είναι η τιμή που προβλέψαμε για την αντίστοιχη τιμή  $x$  της χρονοσειράς

$\bar{x}$  είναι ο μέσος όρος όλων των παρατηρήσεων μέχρι και την τιμή  $x$  της χρονοσειράς.

Ο εκτιμητής αυτός δημιουργήθηκε σύμφωνα με την εξής σκέψη: εάν η χρονοσειρά είναι τυχαία τότε η καλύτερη πρόβλεψη που μπορούμε να πάρουμε είναι ο μέσος όρος. Το  $E$  θα πρέπει να κυμαίνεται στο διάστημα  $(0,1)$ , ενώ θεωρητικά κυμαίνεται στο διάστημα  $[0, +\infty)$ .  $E=0$  σημαίνει ότι οι προβλέψεις μας είναι απόλυτα ακριβείς. Όταν  $E \approx 1$  τότε οι προβλέψεις μας δε διαφέρουν πολύ σε ακρίβεια από το να δεχτούμε το μέσο όρο σαν πρόβλεψη. Αυτό σημαίνει ότι πιθανότατα να έχουμε να κάνουμε με τυχαία χρονοσειρά. Το υπόδειγμα AR(1) θα είναι το μέτρο σύγκρισης.

Εμείς θα περιγράψουμε τις τοπικές γραμμικές (local linear) προβλέψεις, τις προβλέψεις με RBFs και τις προβλέψεις με τοπικό σταθμισμένο όρο (local weighted average). Όλες έχουν μια κοινή βάση: παρατηρούν τη νέα θέση των κοντινότερων γειτόνων του σημείου που θέλουμε να προβλέψουμε έτσι, ώστε να αποφασίσουν ποια θα είναι η προβλεπόμενη τιμή.

Για πρόβλεψη σε παραπάνω από 1 βήμα (ας πούμε  $p$  βήματα) στο μέλλον υπάρχουν δύο διαφορετικές αντιμετώπισεις:

1. Απευθείας (direct): προβλέπουμε έχοντας καταγράψει τη θέση των κοντινότερων γειτόνων μετά από  $p$  βήματα στο μέλλον.
2. Επαναληπτική (iterative): Προβλέπουμε την επόμενη θέση, στη συνέχεια βρίσκουμε τους κοντινότερους γείτονες της νέας θέσης, και επαναλαμβάνουμε τη διαδικασία αυτή για τις υπόλοιπες  $p-1$  φορές.

## Τοπικές γραμμικές

Ο πιο απλός τρόπος για να προβλέψουμε τη χρονοσειρά είναι ο εξής: από το τελευταίο σημείο θα βρούμε τον κοντινότερο γείτονα και η νέα του τιμή θα είναι και η πρόβλεψη που θα κάνουμε. Όταν το μήκος της χρονοσειράς τείνει στο άπειρο, τότε ο κοντινότερος γείτονας θα βρίσκεται απείρως κοντά στο τελευταίο σημείο και το σφάλμα θα τείνει στο 0. Όλα τα παραπάνω ισχύουν υπό την απουσία του θορύβου. Όμως, η πεπερασμένη (πολλές φορές μικρή) χρονοσειρά που διαθέτουμε και η παρουσία του θορύβου, έδωσαν ώθηση στην ανάπτυξη μεθόδων όπως αυτές παρακάτω.

Η μέθοδος των Farmer & Sidorowich, που χρονολογείται από το 1987, λαμβάνει υπόψη όχι μόνο την κίνηση του κοντινότερου γείτονα, αλλά και την κίνηση όλης της γειτονίας του σημείου που θέλουμε να προβλέψουμε. Με τον τρόπο αυτό, αν ο κοντινότερος γείτονας, λόγω θορύβου, μετακινηθεί προς διαφορετική κατεύθυνση από την κατεύθυνση της γειτονίας, τότε το σφάλμα που θα μπορούσαμε να κάνουμε θα είναι αρκετά μειωμένο. Η μέθοδος (OLS) συνοψίζεται στα παρακάτω βήματα:

1. Για το τελευταίο σημείο  $x_T$  διαλέγουμε τους  $k$  κοντινότερους γείτονες  $x_j$ ,  $j=1(1)k$ . Χρησιμοποιούμε την Ευκλείδεια νόρμα. (Η τελευταία παρατήρηση της χρονοσειράς είναι η  $x_{T+(m-1)T} \equiv x_{N-1}$ .)
2. Φτιάχνουμε ένα γραμμικό σύστημα  $k$  εξισώσεων με  $m$  αγνώστους. Οι γραμμές του πίνακα των συντελεστών  $A$  είναι συνιστώσες των  $x_j$ . Το διάνυσμα των αποτελεσμάτων  $b$  είναι η συνιστώσα  $m$  του διανύσματος  $x_{j+1}$ .
3. Λύνουμε το γραμμικό σύστημα  $Ac=b$  με SVD. Το σύστημα είναι συνήθως υπερορισμένο (overdetermined), και μέσω της SVD λύνεται με την μέθοδο των ελαχίστων τετραγώνων.
4. Η προβλεπόμενη τιμή είναι

$$x_{N+1} = \bar{b} + \sum_{i=1}^m c_i (x_i - \bar{x}_i) \quad (40)$$

Μερικές παρατηρήσεις:

- Παρουσιάζεται ξανά η ίδια μέθοδος και πολλές εφαρμογές της.
- Η διάσταση εμπύθισης ( $m$ ) που θα χρησιμοποιήσουμε για την ανακατασκευή θα πρέπει να είναι αρκετά μεγάλη .
- Το πλήθος των  $k$  κοντινότερων γειτόνων θα πρέπει να είναι  $k \geq m + 1$ .
- Αν πάρουμε παρά πολλούς κοντινότερους γείτονες, ξεφεύγουμε από τη μελέτη της γειτονίας και τα αποτελέσματα μας χάνουν την ακρίβειά τους.
- Εμπειρικά αποτελέσματα εφαρμογών με υψηλότερου βαθμού πολυώνυμα δεν είχαν κάποια σημαντική βελτίωση στα αποτελέσματα των προβλέψεων.

Προτάθηκαν παραλλαγές της μεθόδου αυτής που χρησιμεύουν όταν υπάρχει θόρυβος στη χρονοσειρά. Εμείς από όλες αυτές θα αναλύσουμε την Παλινδρόμηση Κυρίων συνιστωσών (Principal Components Regression, PCR). Το βασικό πρόβλημα με αυτή τη μέθοδο είναι η σωστή επιλογή γειτόνων. Θα πρέπει να διαλέξουμε γείτονες οι οποίοι θα έχουν μια σαφή δυναμική. Έχουμε ακόμη μια παραλλαγή που κατά κάποιο τρόπο λύνει το πρόβλημα της κακής επιλογής των γειτόνων. Εξομαλύνει τη χρονοσειρά με το να μηδενίζει τις υψηλές συχνότητες σε χρονικά παράθυρα 32 ή 64 παρατηρήσεων. Στη συνέχεια χρησιμοποιεί PCR.

Είναι προφανές ότι το  $b$  έχει θόρυβο. Όσο πιο καλά μπορέσουμε να ανακατασκευάσουμε το καθαρό  $b$  με το πολλαπλασιασμό  $A_c$  τόσο μικραίνει το σφάλμα μας. Για να αποφύγουμε τις μεγάλες διακυμάνσεις στο  $b$ , που προφανώς είναι ο θόρυβος, αρκεί να μη λάβουμε υπόψη μας τις τελευταίες ιδιάζουσες τιμές. Αυτές αντιστοιχούν στα τελευταία ιδιάζοντα διανύσματα τα οποία έχουν τη μεγαλύτερη διακύμανση.

Μια δεύτερη παραλλαγή, που δίνει παρόμοια αποτελέσματα με την PCR, είναι η Παλινδρόμηση Πλαγίας(Ridge regression,RR).

Με την παραλλαγή αυτή δεν αγνοούμε κάποιες ιδιάζουσες τιμές και τις υπόλοιπες τις κρατάμε στο ακέραιο, όπως κάναμε με την PCR. Αντίθετα, μειώνουμε περισσότερο τις λιγότερο σημαντικές (μικρές) ιδιάζουσες τιμές, ενώ μειώνουμε λιγότερο τις περισσότερες σημαντικές (μεγάλες).

Η RR δέχεται μια παράμετρο  $\mu^2$  στο διάστημα  $[0, +\infty)$ . Όταν  $\mu^2=0$  η RR εκφυλίζεται σε OLS ενώ όταν  $\mu^2 \rightarrow \infty$  το αποτέλεσμα που παίρνουμε είναι ο μέσος όρος του  $b$ . Για απλότητα, μπορούμε να δώσουμε στο  $\mu$  την τιμή ενός ποσοστού του αθροίσματος των τετραγώνων των ιδιάζουσών τιμών του  $A$ . Η κάθε ιδιάζουσα τιμή  $\sigma_i$  πολλαπλασιάζεται με ένα συντελεστή

$$f_i = \frac{\sigma_i^2 + \mu^2}{\sigma_i^2} \quad (41)$$

έτσι ώστε, όταν αντιστρέφεται για να συμβάλει στη λύση του συστήματος των γραμμικών εξισώσεων, η τιμή του συρρικνώνεται.

## Ακτινικές Συναρτήσεις Βάσης

Εδώ το σκεπτικό αλλάζει. Χρειαζόμαστε μια πολιτική για την τοποθέτηση των κέντρων, μια ακτινική συνάρτηση βάσης (RBF) και να αποφασίσουμε αν θα έχουμε ολική ή τοπική προσέγγιση. Η μέθοδος προτάθηκε ως ολική με τοποθέτηση ενός κέντρου πάνω σε κάθε σημείο του χώρου των φάσεων αλλά η υπολογιστική ισχύς που χρειαζόταν ήταν τεράστια. Όταν εφαρμόστηκε για πρώτη φορά, η προσέγγιση ήταν τοπική με τοποθέτηση κέντρων πάνω στους κοντινότερους γείτονες. Μια από τις RBFs που προτάθηκαν είναι η

$$X_j(x) = (\|x - x_j\|^2 + c^2)^{-\beta} \quad (42)$$

Αυτή θα χρησιμοποιήσουμε και παρακάτω. Το  $c$  θα είναι ο μέσος όρος των αποστάσεων του  $x$  από τα  $x_j$ . Θα πρέπει  $\beta > 0$  για να έχει η μέθοδος την ιδιότητα να μπορεί να χρησιμοποιηθεί τοπικά. Η μέθοδος συνοψίζεται στα παρακάτω βήματα:

1. Για το τελευταίο σημείο  $x_T$  διαλέγουμε τους  $k$  κοντινότερους γείτονες  $x_j$ ,  $j=1(1)k$ . Χρησιμοποιούμε την Ευκλείδεια νόρμα. (Η τελευταία παρατήρηση της χρονοσειράς είναι η  $x_{T+(m-1)T} \equiv x_N$ .)
2. Τοποθετούμε ένα κέντρο πάνω στο κάθε κοντινό γείτονα.
3. Φτιάχνουμε ένα γραμμικό σύστημα  $k$  εξισώσεων με  $k$  αγνώστους. Η γραμμή  $i$  του πίνακα των συντελεστών  $A$  είναι  $X_j(x_i)$  από την σχέση (42). Το διάνυσμα των αποτελεσμάτων  $b$  είναι η συνιστώσα  $m$  του διανύσματος  $x_{j+1}$ .
4. Θα πρέπει να εξασφαλίσουμε ότι το  $b$  έχει μέσο όρο 0. Αφαιρούμε λοιπόν το μέσο όρο του  $b$  από κάθε συνιστώσα του.
5. Λύνουμε το γραμμικό σύστημα  $Ac=b$ . Προτείνεται η SVD.
6. Η προβλεπόμενη τιμή είναι το εσωτερικό γινόμενο του  $X_j(x_T)$  με το  $c$  συν ο μέσος όρος που αφαιρέσαμε στο 4ο βήμα, δηλαδή

$$x_{N+1} = \bar{b} + \sum_{i=1}^k c_k X_k(x_T) \quad (43)$$

## Εκτίμηση Πυκνότητας Πυρήνα

Η μέθοδος εκτίμησης πυκνότητας πυρήνα (kernel density estimation, KDE) είναι λίγο πιο απλή από τις τοπικές γραμμικές προβλέψεις. Είναι ουσιαστικά ένας σταθμισμένος μέσος όρος. Η μέθοδος συνοψίζεται στα παρακάτω βήματα:

1. Για το τελευταίο σημείο  $x_T$  διαλέγουμε τους  $k$  κοντινότερους γείτονες  $x_j$ ,  $j=1(1)k$ . Χρησιμοποιούμε την Ευκλείδεια νόρμα. (Η τελευταία παρατήρηση της χρονοσειράς είναι η  $x_{T+(m-1)T} \equiv x_N$ .)
2. Στις αποστάσεις  $r_j$  δίνεται μια βαρύτητα

$$\varphi(r_i) = \exp\left(-\frac{r_i^2}{c^2}\right) \quad (44)$$

όπου  $c$  είναι ο μέσος όρος των  $r_i$

ή οποιαδήποτε άλλη βαρύτητα που φθίνει γρήγορα στο  $0^+$  όσο μεγαλώνει η απόσταση  $r_i$ .

3. Σταθμίζουμε τις βαρύτητες έτσι ώστε  $\sum w_i = 1$

$$w_i = \frac{\varphi(r_i)}{\sum_{j=1}^k \varphi(r_j)} \quad (45)$$

4. Η πρόβλεψη δίνεται από τον τύπο

$$x_{N+1} = \sum_{j=1}^k w_j x_{j+p(m-1) \cdot T} \quad (46)$$

Η μέθοδος αυτή έχει διάφορες παραλλαγές. Και βλέπουμε που έχει τον περιορισμό  $k=m+1$  και τα  $x_j$  να περικλείουν το  $x_T$ . Ο περιορισμός να περικλείουν κάνει τη μέθοδο πολύ αργή καθώς αυξάνεται το  $m$ . Σε μια άλλη μέθοδο, που υπολογίζει όλα τα σημεία της χρονοσειράς ( $k=T-p$ ) με κύρια βαρύτητα στο κοντινότερα. Δίνοντας σημασία σε όλα τα σημεία του χώρου των φάσεων, έστω και μικρή, χάνει η μέθοδος την ακρίβεια της και εξαρτάται από υποκειμενικούς παράγοντες. Εμείς λίγο πιο πάνω δώσαμε τα κοινά σημεία των παραλλαγών που είναι και η καλύτερη αντιστάθμιση μεταξύ χρόνου επεξεργασίας και ακρίβειας. Βλέπουμε σύμφωνα με τη μέθοδο αυτή ότι: η πρόβλεψη δίνεται από το μέσο όρο των μετακινούμενων σημείων, σταθμισμένο ως προς την αρχική απόσταση.

## Οι προβλέψεις στις χρηματοοικονομικές χρονοσειρές

Σχετικά με τις χρηματοοικονομικές σειρές, που τέτοιες είναι και οι σειρές της ναυλαγοράς, το ενδιαφέρον είναι στραμμένο στις μεθόδους της μη γραμμικής δυναμικής και των τεχνητών νευρωνικών δικτύων. Μας απασχολεί κυρίως αν ισχύει η υπόθεση του τυχαίου περιπάτου. Η υπόθεση αυτή, στην ασθενή της μορφή, θέτει ότι

$$x_{t+1} = x_t + \eta_t \quad (47)$$

όπου :  $x_t$  είναι ο λογάριθμος της τιμής τη χρονική στιγμή  $t$

$\eta_t$  είναι μια διαδικασία λευκού θορύβου

Η μη γραμμική θεώρηση των πραγμάτων πηγάζει από το γεγονός ότι: όλοι οι επενδυτές δεν σκέπτονται με τον ίδιο τρόπο. Για παράδειγμα, άλλοι έχουν μακροπρόθεσμες βλέψεις, άλλοι μεσοπρόθεσμες και άλλοι βραχυπρόθεσμες. Επίσης, κάποια πληροφορία ένας επενδυτής μπορεί να την εκλάβει ως ευκαιρία ενώ άλλος να την εκλάβει ως απειλή. Άλλοι επενδυτές πάλι, φοβούνται περισσότερο την πτώση της αξίας του χαρτοφυλακίου τους ενώ άλλοι τη φοβούνται λιγότερο. Πόσο ικανές όμως είναι οι υπάρχουσες μη γραμμικές μέθοδοι να προβλέψουν τις



χρηματοοικονομικές χρονοσειρές; Οι αναφορές μας δίνουν μια πρώτη γεύση της ικανότητας αυτής. Το συμπέρασμα που βγαίνει από τις αναφορές αυτές είναι ότι περίπου μέχρι το 20 % της διακύμανσης εξηγείται (προβλέπεται) από τη μη γραμμική δυναμική. Το υπόλοιπο 80%, ίσως και παραπάνω που απομένει είναι θόρυβος.

Οι Bajo-Rubio et al (Bajo-Rubio et al., 1992) (επιχείρησαν πρόβλεψη εκτός δείγματος (forecast) με την απλή μέθοδο (Farmer & Sidorowich). Τα αποτελέσματα που δίνουν είναι σημαντικά καλύτερα από το RW. Στα αποτελέσματα επίσης παρατηρούμε μεγάλη ανομοιομορφία. Το καλύτερο E που πέτυχαν ήταν περίπου 50%.

Οι Diebold & Nason (Diebold & Nason, 1990) χρησιμοποίησαν τοπικούς σταθμισμένους μέσους όρους για πρόβλεψη εβδομαδιαίων τιμών συναλλάγματος εντός και εκτός δείγματος. Σε αρκετές περιπτώσεις ο τυχαίος περίπατος έδωσε καλύτερες προβλέψεις. Το κύριο χαρακτηριστικό της προσπάθειας αυτής είναι ότι έγιναν πολλές δοκιμές στις ελεύθερες παραμέτρους. Από το γεγονός αυτό, κατά συνέπεια, μας δίνεται η δυνατότητα να πάρουμε μια συνολική εικόνα της κατάστασης για τη χρονοσειρά που δοκίμασαν.

Στην περίπτωση των Jaditz & Sayers η βελτίωση είναι ένα μικρό ποσοστό του τυχαίου περιπάτου (περίπου 8%). Η συνεισφορά των τοπικών γραμμικών μεθόδων στην πρόβλεψη δεν είναι ευδιάκριτη.

Οι Lisi & Schiavo (Lisi & Schiavo, 1999) χρησιμοποίησαν την παραπάνω τροποποιημένη μέθοδο συγκρίνοντάς τη με τεχνητά νευρωνικά δίκτυα (ΤΝΔ). Το συμπέρασμα ήταν ότι έχουμε καλύτερα αποτελέσματα σε σχέση με τον τυχαίο περίπατο και με τη χασοτική μέθοδο και με τα ΤΝΔ. Επίσης τα ΤΝΔ έδωσαν ελαφρώς καλύτερα αποτελέσματα.

Αν υποθέσουμε ότι θέτουμε σε κάθε χρηματοοικονομική χρονοσειρά χρονική υστέρηση ίση με 1, το μόνο που NSR. Στο κεφάλαιο το σχετικό με τη μείωση θορύβου, είδαμε το πως μπορούμε να εκτιμήσουμε το NSR. Στο κεφάλαιο το σχετικό με τις διαστάσεις είδαμε το πως μπορούμε να αναγνωρίσουμε τη σωστή διάσταση συσχέτισης με γνωστή την τυπική απόκλιση του θορύβου.

## Κεφάλαιο 4

Στο κεφάλαιο αυτό παρατίθενται τα πειραματικά στοιχεία που εξάχθηκαν από την χρήση του προγράμματος NLTSA για την διερεύνηση του εάν παρουσιάζει η χρονοσειρά των spot τιμών των Sueszmax tanker, με δείκτη **UKC - ECC 130K DTY** μιας δεκαετίας περίπου, χαοτική συμπεριφορά.

Σε αυτό το σημείο πρέπει να αποσαφηνιστούν οι έννοιες των παραπάνω δεικτών. Ο δείκτης UKC-ECC φανερώνει την διαδρομή πάνω στην οποία κινούνται τα πλοία, τα οποία έχουν ως αφητηρία τις ακτές της Μεγάλης Βρετανίας και ως προορισμό τις ανατολικές ακτές των Ηνωμένων Πολιτειών. Ο δείκτης 130K φανερώνει το βάρος του ωφέλιμου φορτίου (deadweight) το οποίο ανέρχεται μέγιστα σε εκατό τριάντα χιλιάδες τόνους. Το DTY συμβολίζει ότι ο δείκτης αυτός εμπεριέχει τις ακάθαρτες διαδρομές (dirty routes). Για την δημοσίευση του παραπάνω δείκτη, τα χρηματιστήρια έχουν αναθέσει σε εταιρίες μεσιτών (Shipbrokers) την καθημερινή εκτίμηση των διαδρομών. Η βασική υποχρέωση των εταιριών αυτών, οι οποίες λειτουργούν ανεξάρτητα μεταξύ τους, είναι να εκτιμήσουν και να προσδιορίσουν το επίπεδο της αγοράς σε καθημερινή βάση για κάθε δείκτη και φυσικά για τις διαδρομές που ορίζουν τα Χρηματιστήρια. Οι εταιρίες αυτές αποτελούν για τα Χρηματιστήρια τους εκτιμητές (panelists), εκείνους τους συνεργάτες δηλαδή που δημοσιεύουν τους δείκτες. Η επιλογή των εταιριών αυτών γίνεται με βάση τα παρακάτω κριτήρια:

- Όπως αναφέρθηκε και παραπάνω πρόκειται κυρίως για εταιρίες μεσιτών (shipbrokers).
- Οι εκτιμητές θα πρέπει να είναι ανταγωνιστικές και επαγγελματικές εταιρίες, οι οποίες είναι ενεργά συμμετοχές στις αγορές για τις οποίες εκδίδουν αναφορές, με το κατάλληλο πάντα προσωπικό για το ρόλο αυτό.
- Τα Χρηματιστήρια προσπαθούν να διατηρήσουν ένα γεωγραφικό διασκορπισμό των εκτιμητών.
- Τα Χρηματιστήρια προσπαθούν να αποφύγουν τον διορισμό εκτιμητών, οι οποίοι είναι αποκλειστικοί αντιπρόσωποι ναυλωτών, οι οποίοι επηρεάζουν σε σημαντικό βαθμό το σχετικό εμπόριο (ουσιαστικά τις συγκεκριμένες διαδρομές).

Η επιλογή των διαδρομών που καθορίζουν τους δείκτες, καθώς και το ειδικό βάρος της κάθε διαδρομής γίνεται από τα Χρηματιστήρια. Πως όμως κάνουν τις εκτιμήσεις τους οι εκτιμητές; Για να δώσουν τις εκτιμήσεις τους οι εκτιμητές, λαμβάνουν υπόψη τους όλες τις πληροφορίες της αγοράς που είναι διαθέσιμες σ' αυτούς σε καθημερινή βάση την ώρα της αναφοράς τους στα διάφορα χρηματιστήρια, οι οποίες βέβαια πρέπει να είναι προσαρμοσμένες στους κανόνες της διαδρομής. Οι εκτιμητές πρέπει να λαμβάνουν υπόψη τους τα παρακάτω:

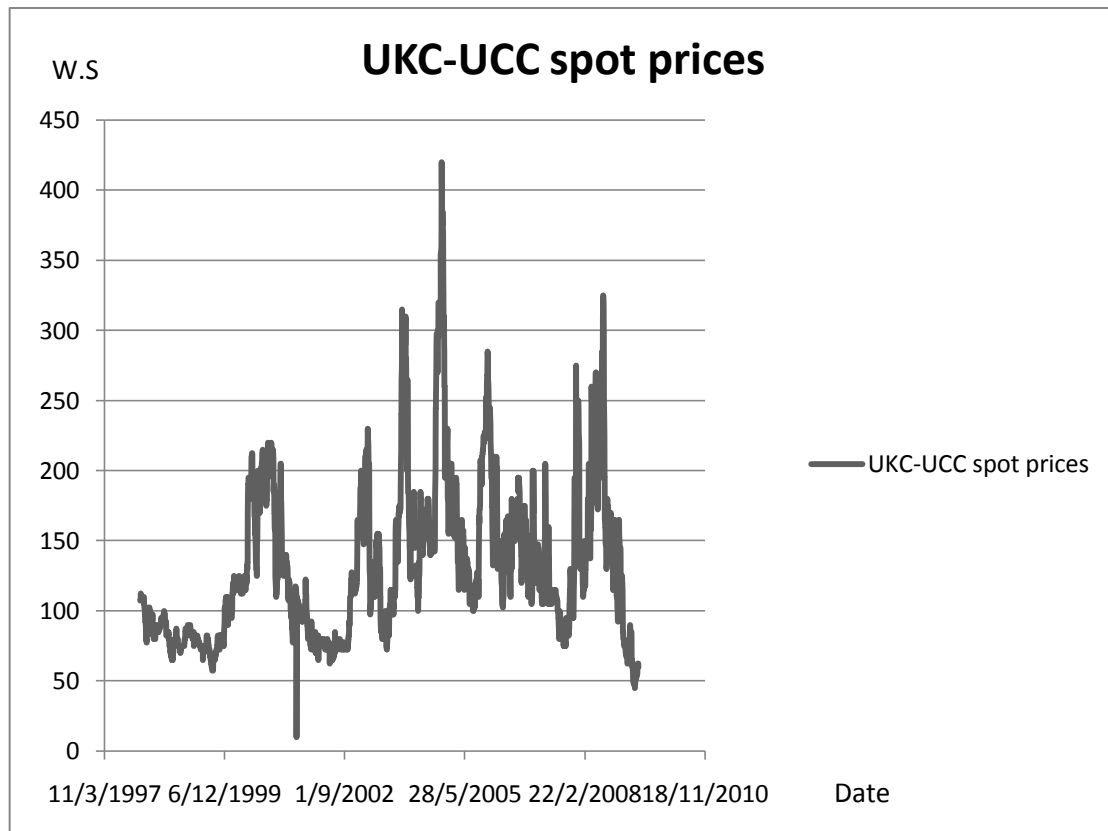
- Τα τελευταία συμβόλαια προς διαπραγμάτευση, κρίνοντας μόνοι τους τη συσχέτιση των πληροφοριών αυτών με τους δείκτες.
- Τις τρέχουσες διαπραγματεύσεις, μη ξεχνώντας την άμεση επιρροή που μπορούν να έχουν στην αγορά.
- Την προσφορά και τη ζήτηση, των πλοίων και του φορτίου αντίστοιχα.

Οι εκτιμητές θα πρέπει να μην λαμβάνουν υπόψη τους τα παρακάτω:

- Τις οποιεσδήποτε υποθέσεις για τα κίνητρα των αποφάσεων τόσο των ναυλωτών όσο και των πλοιοκτητών για κάποιο συγκεκριμένο συμβόλαιο.
- Μελλοντικές κινήσεις της αγοράς οι οποίες δεν έχουν σχέση με αυτά που ένας εκτιμητής εκτιμάει.

Αξίζει σε αυτό το σημείο να αναφερθούμε στις πρακτικές που χρησιμοποιούνται για την αποφυγή της κακής διαχείρισης τη αγοράς. Ούτε τα χρηματιστήρια αλλά ούτε και οι υπάλληλοί τους έχουν το δικαίωμα να διαπραγματευτούν σε κάποια αγορά μελλοντικών ναύλων (freight futures market). Επίσης δεν θα πρέπει να διαθέσουν οποιαδήποτε πληροφορία που να προέρχεται από τις αναφορές των εκτιμητών ή να επιτρέψουν την αποκάλυψη της ταυτότητας κάποιου εκτιμητή σε οποιοδήποτε εξωτερικό φορέα από τα χρηματιστήρια, ή σε ειδικές περιπτώσεις στους διευθυντές των Χρηματιστηρίων. Κανένας περιορισμός δεν τίθεται στις εταιρίες (Shipbrokers, κτλ) και στους υπαλλήλους τους, είτε αυτοί εργάζονται σαν μεσίτες (brokers), είτε σαν κυρίαρχοι της ναυλαγοράς. Παρολαυτά υπάρχει ο όρος για τον διορισμό των εταιριών αυτών και των υπαλλήλων τους που ασχολούνται με το θέμα αυτό, να αποκαλύπτουν στα Χρηματιστήρια την ανάμειξη τους με αυτές τις αγορές. Παρομοίως, οι πληροφορίες που είναι διαθέσιμες στο προσωπικό, στα πλαίσια των καθηκόντων τους δεν θα πρέπει να περάσουν ή να τεθούν προς διάθεση εκτός των Χρηματιστηρίων. Οι καθημερινές αναφορές των εκτιμητών είναι αυστηρά εμπιστευτικές μεταξύ των Χρηματιστηρίων και των νομίμων αντιπροσώπων των εταιριών. Οι εκτιμητές δεν θα πρέπει να θέτουν προς διάθεση αυτές τις αναφορές σε κανένα άλλο εκτός του ανταποκριτή, δηλαδή ούτε στο υπόλοιπο προσωπικό της εταιρίας, ούτε ιδιαίτερα στο προσωπικό και τα άλλα τμήματα που ασχολούνται με τις αγορές αυτές.

Οι τιμές που καταγράφηκαν τα χρόνια 1998-2009 παρουσιάζονται διαγραμματικά.

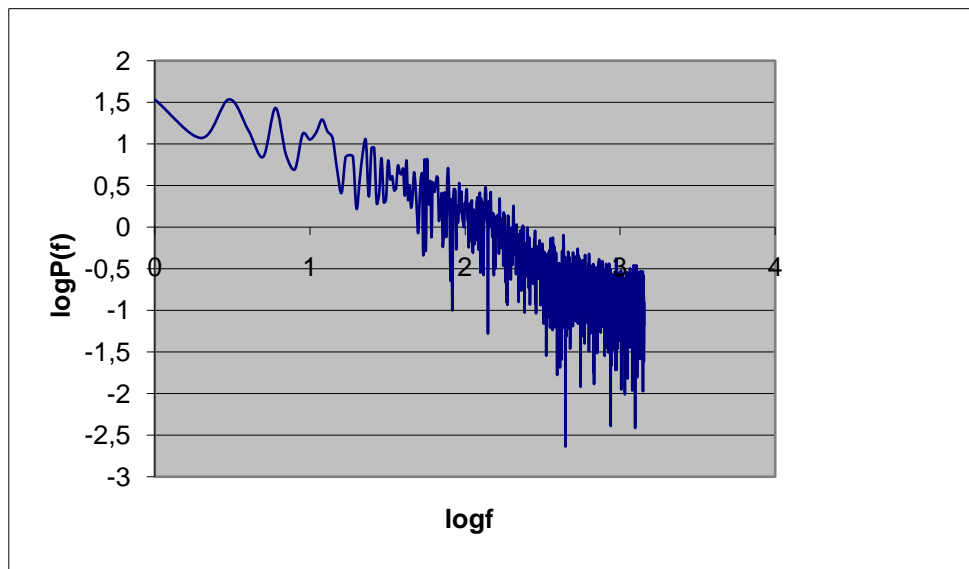


**Σχήμα 20:** Διαγραμματική παρουσίαση της υπό μελέτης χρονοσειράς

## Διερεύνηση Χαοτικής Συμπεριφοράς

Αρχικά θα παρατεθεί το φάσμα ισχύος (power spectrum). Μέσω του προγράμματος βρίσκουμε το  $P(f)$  και  $\log P(f)$  με βάση τους αλγόριθμους που έχουν παρατεθεί στο κεφάλαιο 2.

Με βάση τις τιμές αυτές εξάγεται το παρακάτω διάγραμμα .



**Σχήμα 21:** Παράσταση του φάσματος ισχύος

Από την αρνητική κλίση που παρουσιάζει το φάσμα ισχύος συμπεραίνουμε πως έχουμε πιθανά χαοτική χρονοσειρά, αν και υπάρχουν ορισμένα είδη θορύβου που έχουν αρνητική κλίση.

Με τη χρήση του προγράμματος μπορέσαμε να προσδιορίσουμε τη διάσταση συσχέτισης . Η διάσταση αυτή είναι που χρησιμοποιούμε κατά κόρον για να μας δείξει τη διάσταση του συστήματος. Από την διαδικασία εμβύθισης πήραμε τα εξής στοιχεία:

**Πίνακας 1:** Διάσταση εμφύθισης σε σχέση με την διάσταση συσχέτισης

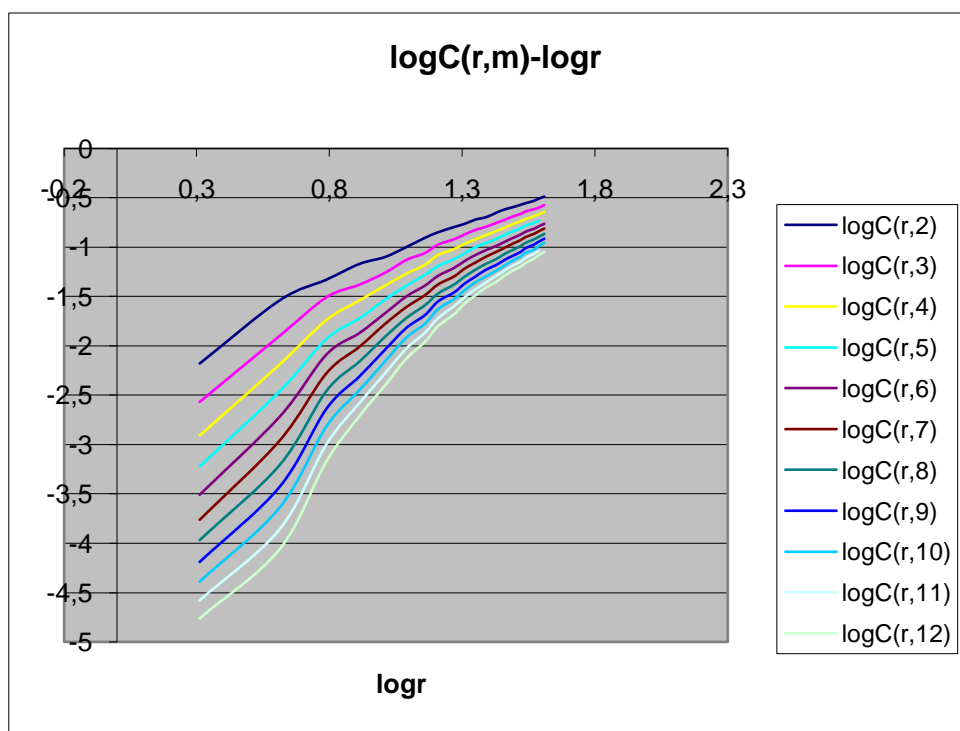
Embedding	Correlation
Dimension	Dimension
-----	-----
2	0,944
3	0,998
4	1,14
5	1,21
6	1,2
7	1,28
8	1,35
9	1,42
10	1,49
11	1,55
12	1,62

Επίσης παρατίθενται οι πίνακες υπολογισμού του λογαρίθμου του ολοκληρώματος συσχέτισης έναντι λογαρίθμου της ακμής  $\log r - \log C(r,m)$

**Πίνακας 2:** Παρουσίαση των τιμών του λογαρίθμου της ακμής έναντι του λογαρίθμου του ολοκληρώματος συσχέτισης.

$\log r$	$\log C(r,2)$	$\log C(r,3)$	$\log C(r,4)$	$\log C(r,5)$	$\log C(r,6)$	$\log C(r,7)$	$\log C(r,8)$	$\log C(r,9)$	$\log C(r,10)$	$\log C(r,11)$	$\log C(r,12)$
0,312	-2,18	-2,57	-2,91	-3,22	-3,51	-3,76	-3,97	-4,2	-4,39	-4,58	-4,8
0,613	-1,54	-1,9	-2,19	-2,46	-2,72	-2,96	-3,21	-3,4	-3,64	-3,86	-4,1
0,789	-1,33	-1,51	-1,74	-1,93	-2,09	-2,28	-2,46	-2,6	-2,82	-3	-3,2
0,914	-1,17	-1,38	-1,54	-1,72	-1,87	-2,01	-2,16	-2,3	-2,45	-2,58	-2,7
1,01	-1,1	-1,26	-1,39	-1,53	-1,67	-1,78	-1,91	-2	-2,16	-2,28	-2,4
1,09	-1	-1,13	-1,27	-1,39	-1,5	-1,61	-1,72	-1,8	-1,92	-2,03	-2,1
1,16	-0,906	-1,06	-1,18	-1,28	-1,39	-1,49	-1,59	-1,7	-1,78	-1,87	-2
1,21	-0,849	-0,974	-1,08	-1,19	-1,29	-1,38	-1,47	-1,6	-1,63	-1,72	-1,8
1,27	-0,796	-0,92	-1,02	-1,12	-1,21	-1,3	-1,38	-1,5	-1,53	-1,6	-1,7
1,31	-0,764	-0,871	-0,966	-1,06	-1,14	-1,22	-1,3	-1,4	-1,44	-1,51	-1,6
1,35	-0,721	-0,825	-0,922	-1	-1,08	-1,16	-1,23	-1,3	-1,36	-1,43	-1,5
1,39	-0,697	-0,794	-0,878	-0,96	-1,03	-1,1	-1,17	-1,2	-1,3	-1,36	-1,4
1,43	-0,65	-0,753	-0,836	-0,92	-0,99	-1,05	-1,12	-1,2	-1,24	-1,29	-1,4
1,46	-0,618	-0,724	-0,804	-0,88	-0,95	-1,01	-1,07	-1,1	-1,19	-1,24	-1,3
1,49	-0,596	-0,69	-0,769	-0,84	-0,91	-0,97	-1,03	-1,1	-1,14	-1,19	-1,2
1,52	-0,571	-0,664	-0,738	-0,81	-0,87	-0,93	-0,99	-1,1	-1,1	-1,15	-1,2
1,54	-0,554	-0,637	-0,715	-0,78	-0,84	-0,9	-0,96	-1	-1,06	-1,11	-1,2
1,57	-0,531	-0,616	-0,688	-0,75	-0,81	-0,87	-0,92	-1	-1,03	-1,07	-1,1
1,59	-0,507	-0,598	-0,666	-0,73	-0,79	-0,84	-0,89	-0,9	-0,99	-1,04	-1,1
1,61	-0,487	-0,573	-0,644	-0,71	-0,76	-0,82	-0,87	-0,9	-0,96	-1,01	-1,1

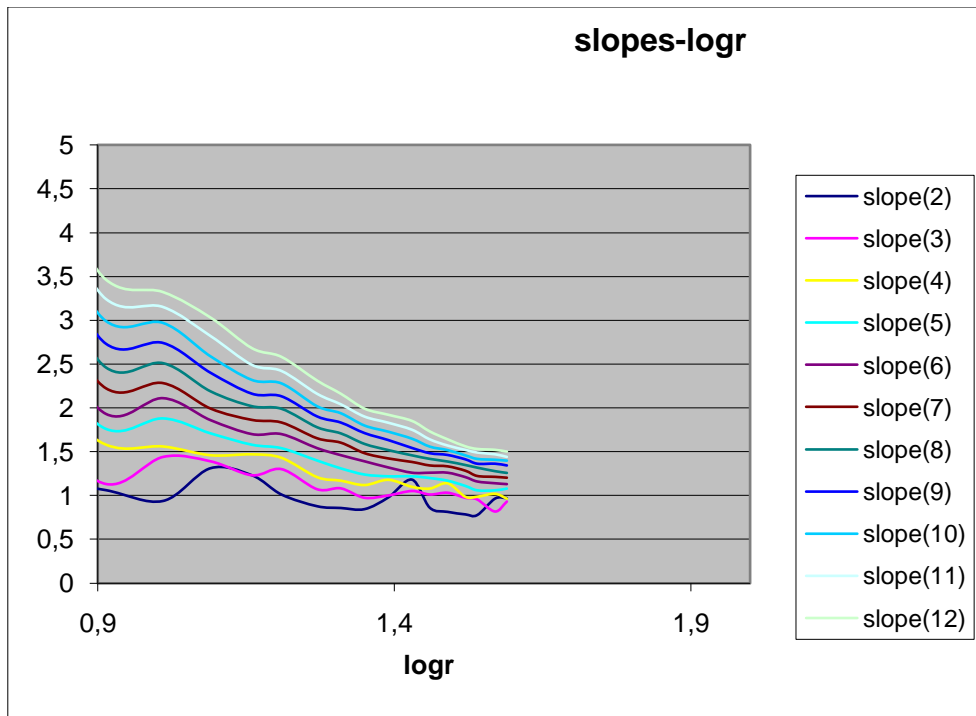
Και το σχήμα των αποτελεσμάτων



**Σχήμα 22:** Σχεδιάγραμμα του λογαρίθμου της ακμής έναντι του λογαρίθμου του ολοκληρώματος σχέτισης.

**Πίνακας 3:** Δίνονται οι κλίσεις έναντι του λογαρίθμου της ακμής

logr	slope(2)	slope(3)	slope(4)	slope(5)	slope(6)	slope(7)	slope(8)	slope(9)	slope(10)	slope(11)	slope(12)
0,613	1,81	2,21	2,45	2,7	2,92	3,05	3,09	3,17	3,21	3,21	3,23
0,789	1,22	1,74	2,16	2,49	2,85	3,18	3,53	3,79	4,02	4,3	4,51
0,914	1,06	1,13	1,58	1,76	1,93	2,22	2,46	2,73	2,99	3,24	3,46
1,01	0,937	1,44	1,56	1,88	2,11	2,28	2,51	2,74	2,97	3,15	3,32
1,09	1,31	1,39	1,46	1,71	1,86	1,99	2,19	2,4	2,59	2,82	3,03
1,16	1,23	1,23	1,47	1,58	1,7	1,86	2,02	2,16	2,32	2,49	2,68
1,21	1,01	1,3	1,43	1,54	1,7	1,83	1,99	2,13	2,28	2,42	2,58
1,27	0,877	1,07	1,21	1,4	1,54	1,65	1,78	1,9	2,02	2,16	2,31
1,31	0,857	1,08	1,17	1,31	1,46	1,6	1,71	1,83	1,94	2,04	2,16
1,35	0,844	0,972	1,12	1,24	1,39	1,48	1,59	1,71	1,8	1,9	1,99
1,39	0,977	0,999	1,18	1,22	1,32	1,42	1,52	1,63	1,73	1,83	1,92
1,43	1,18	1,05	1,1	1,22	1,26	1,38	1,46	1,54	1,65	1,75	1,85
1,46	0,86	1,01	1,08	1,2	1,26	1,34	1,42	1,48	1,56	1,64	1,73
1,49	0,813	1,03	1,14	1,17	1,26	1,33	1,39	1,46	1,52	1,57	1,64
1,52	0,784	0,972	0,986	1,11	1,21	1,28	1,35	1,41	1,46	1,52	1,56
1,54	0,776	0,953	0,987	1,06	1,16	1,22	1,32	1,36	1,42	1,49	1,53
1,57	0,964	0,815	1,02	1,06	1,14	1,21	1,28	1,36	1,41	1,47	1,52
1,59	0,97	0,929	0,965	1,08	1,13	1,2	1,26	1,34	1,4	1,44	1,5

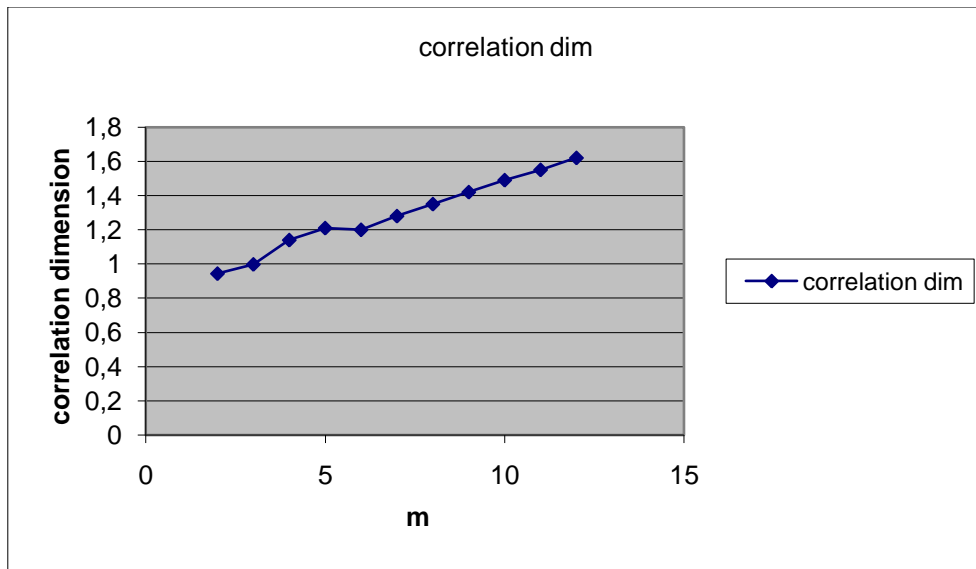


**Σχήμα 23:** Διάγραμμα των κλίσεων έναντι του λογαρίθμου της ακμής

**Πίνακας 4:** Δίνεται ο πίνακας διάστασης εμβύθισης σε σχέση με την διάσταση συσχέτισης

Embedding Dimension	Correlation Dimension
2	0,944
3	0,998
4	1,14
5	1,21
6	1,2
7	1,28
8	1,35
9	1,42
10	1,49
11	1,55
12	1,62





**Σχήμα 24:** Διάγραμμα της διάστασης εμφύθισης σε σχέση με την διάσταση συσχέτισης

Η προτεινόμενη τιμή της διάστασης συσχέτισης είναι 1.29. Με βάση το θεώρημα των Eckman & Ruelle η διάσταση συσχέτισης δεν πρέπει να ξεπερνά το  $2\log N$  όπου  $N$  είναι το μήκος της χρονοσειράς μας. Στην περίπτωση μας είναι  $2\log 2813=6.898$ , άρα καλυπτόμαστε.

Δεν πραγματοποιήθηκε παρουσίαση του ελκυστή της χρονοσειράς μας μια και η διάσταση της κινείται μεταξύ 1 και 2, οπότε δεν παρουσιάζει ενδιαφέρον και δεν θα προσδώσει χρήσιμα συμπεράσματα.

Μέσω του προγράμματος NLTSA προσδιορίσαμε τον μέγιστο εκθέτη Lyapunov εφαρμόζοντας μια εύρωστη μέθοδο. Με αυτή τη μέθοδο υπολογίζουμε τον μέσο όρο των τοπικών LE που είναι  $S(a) = \frac{1}{T-a} \sum_{i=1}^{T-a} \log_2 \left( \frac{1}{|U_{i,r}|} \sum_{j \in U_{i,r}} \text{dist}(x_i, x_j, a) \right)$

και στην συνέχεια τον μέγιστο LE που δίνεται από τον τύπο

$$\lambda_1(a) = S(a + 1) - S(a) \quad (51)$$

Και παίρνουμε το μέσο όρο των  $\lambda_1(a)$ . Εναλλακτικά υπολογίζουμε την κλίση της παλινδρόμησης

$$S(a) = \lambda_1 \cdot a + c \quad (52)$$

Όπου  $\lambda_1$  είναι η κλίση

$C$  είναι μια σταθερά που δεν μας ενδιαφέρει η τιμή της

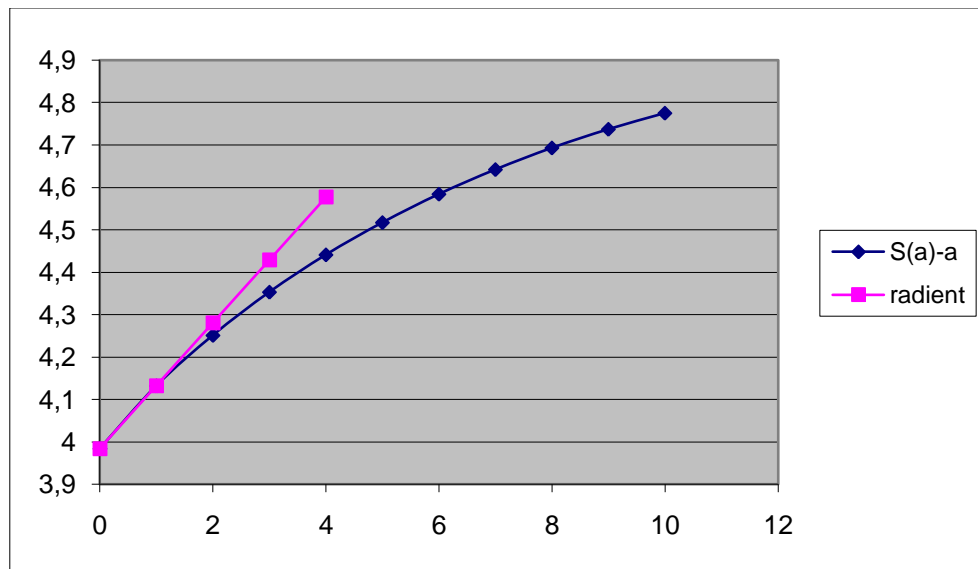
Σε μια περιοχή που το λάθος να είναι σχετικά μικρό. Ας πούμε  $R^2 \geq 98\%$ .

Στην συνέχεια παραθέτουμε τον πίνακα με τις τιμές της διάστασης εμφύθισης και τις τιμές των LLE και  $S(a)$ .

**Πίνακας 5:** Τιμές εκτίμησης του μέγιστου εκθέτη Lyapunov.

evol.	S(evol.)	LLE est.
0	3,984	
1	4,132	0,1483
2	4,251	0,1188
3	4,353	0,1016
4	4,441	0,08878
5	4,517	0,07582
6	4,584	0,06659
7	4,642	0,05811
8	4,693	0,05121
9	4,737	0,04409
10	4,775	0,03809

Αφού υπολογίσαμε το  $S(a)$  σχεδιάζουμε το διάγραμμα  $S(a) - a$  και την συνάρτηση  $y=0,1483x+3,3926$  που είναι η κλίση(radiant). Η τιμή 0,1453 είναι πολύ κοντά στην τιμή που χρησιμοποιήσαμε.



**Σχήμα 25:** Διάγραμμα  $S(a)-a$  και κλίσης

Από το διάγραμμα βλέπουμε πως οι γραφικές παραστάσεις ξεκινάνε μαζί και μεταξύ της διάστασης εμφύθισης 1 και 2 αλλάζουν κλίση. Άρα έχουμε διάσταση εμφύθισης  $m=2, t=1$ . Για  $m \geq 2$  σταματάμε να λαμβάνουμε υπόψη αυτές τις τιμές.

## Προβλέψεις

Οι λόγοι που επιχειρούμε να προβλέψουμε τις χρηματοοικονομικές χρονοσειρές είναι οι ανταμοιβές που μπορούμε να κερδίσουμε, το ενδιαφέρον που αυτές παράγουν, και η πρόκληση της προσπάθειας πρόβλεψης των πιο δύσκολων από τις χρονοσειρές. Οι χρηματοοικονομικές χρονοσειρές, μια από τις οποίες είναι και η υπό εξέταση χρονοσειρά μας, παράγονται από την αλληλεπίδραση εκατοντάδων χιλιάδων έξυπνων και με υψηλά κίνητρα ανθρώπων, ο καθένας από τους οποίους έχει πολλούς μήνες εάν όχι χρόνια εμπειρίας στην παρακολούθηση των αγορών και στην προσπάθεια να μαντέψει την επόμενη κίνηση, την επομένη τάση (trend), το επόμενο κύμα. Το ότι υπάρχει κάποια προβλεψιμότητα στις αγορές είναι ένα θαύμα. Η υπόθεση της ικανοποιητικής αγοράς (Efficient Market Hypothesis, Fama), που επικρατούσε μέχρι τις αρχές τις δεκαετίας του 90 ισχυριζόταν ότι η χρονική εξέλιξη των αγορών ήταν σαν ένας τυχαίος περίπατος. Η εξήγηση για αυτήν την υπόθεση δίδεται συνήθως μέσω του εξής παραδείγματος: έστω ότι ένα χαρτονόμισμα των δέκα ευρώ βρίσκεται στο πεζοδρόμιο, αυτό αμέσως θα βρισκόταν κάποιος να το αρπάξει. Επομένως με βάση αυτήν την λογική δεν υπάρχει ιστορία στην συμπεριφορά των αγορών. Ότι πληροφορίες θα μπορούσαν να χρησιμοποιηθούν από κάποιον χρηματιστή είναι παρούσες στις στιγμιαίες καταστάσεις της αγοράς και στην κοινή λογική. Στοιχεία χαμηλοδιάστατου χάους σε πολλές αγορές, και επομένως περιορισμένη προβλεψιμότητα (Medio), υπονόμεισε την υπόθεση της ικανοποιητικής αγοράς. Σε αυτήν την εργασία επιδιώκουμε να αποδείξουμε ότι μπορούμε να πραγματοποιήσουμε προβλέψεις στην αγορά βασιζόμενη στην ιστορία της.

Όλες οι προβλέψεις που έγιναν με τη χρήση του προγράμματος πραγματοποιήθηκαν εντός δείγματος έτσι ώστε να γνωρίζουμε τα αποτελέσματα και να μπορούμε να μετράμε το σφάλμα μας. Πραγματοποιήθηκαν 7 προβλέψεις μέσα σε 7 διαφορετικά σημεία της χρονοσειράς εφαρμόζοντας 5 διαφορετικές μεθόδους. Οι μέθοδοι αυτές είναι :

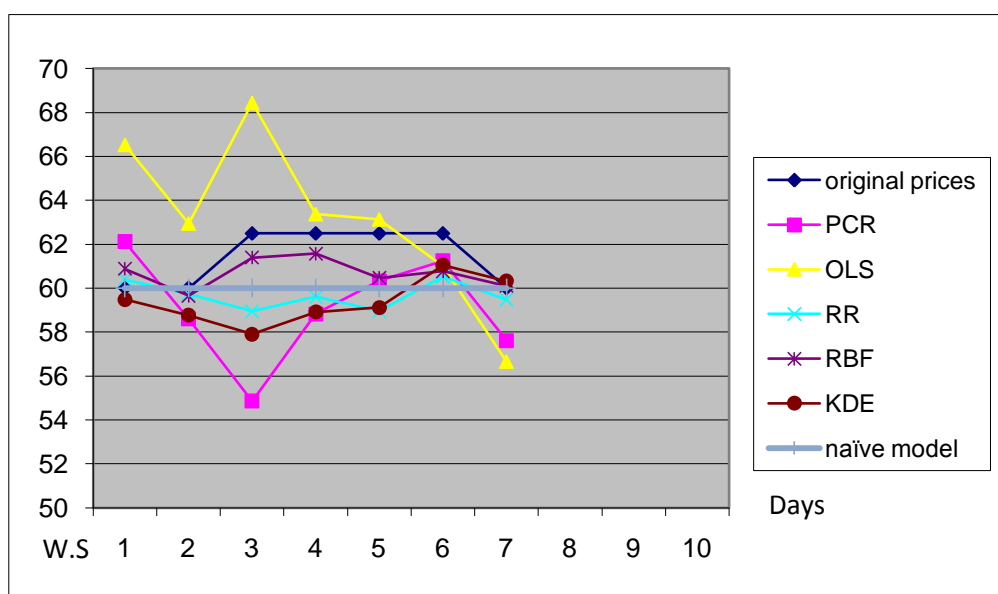
1. OLS-μέθοδος ελαχίστων τετραγώνων
2. RR-μέθοδος παλινδρόμησης πλαγίας
3. RBF- μέθοδος Ακτινικών συναρτήσεων Βάσης
4. KDE-Μέθοδος εκτίμησης πυρήνα
5. PCR-μέθοδος παλινδρόμησης κυρίων συνιστωσών(Principal components regression)
6. Naïve model-μέθοδος αθών μοντέλων

Η πρώτη πρόβλεψη πραγματοποιήθηκε σε 864 σημεία και έγινε πρόβλεψη 7 βημάτων.

**Πίνακας 6:** Παρουσιάζονται τα αποτελέσματα των προβλέψεων και με τις 5 μεθόδους .

PCR	OLS	RR	RBF	KDE	naïve model	original prices
62,1239	66,5338	60,3947	60,8774	59,4872	60	60
58,593	62,9477	59,7368	59,649	58,7683	60	60
54,8523	68,4497	58,9474	61,3915	57,9028	60	62,5
58,8252	63,3698	59,6053	61,5758	58,91	60	62,5
60,3346	63,1324	58,9474	60,463	59,1244	60	62,5
61,2449	60,9949	60,5263	60,7723	61,046	60	62,5
57,6096	56,6467	59,4737	60,0984	60,3275	60	60
E: 0.0506295	E: 0.053615	E: 0.0328227	E: 0.0169342	E: 0.0374213	E:	
					0,026847	

Όπου E είναι εκτιμητής σφάλματος των προβλέψεων της κάθε μεθόδου όπως αυτός έχει οριστεί στην θεωρεία.



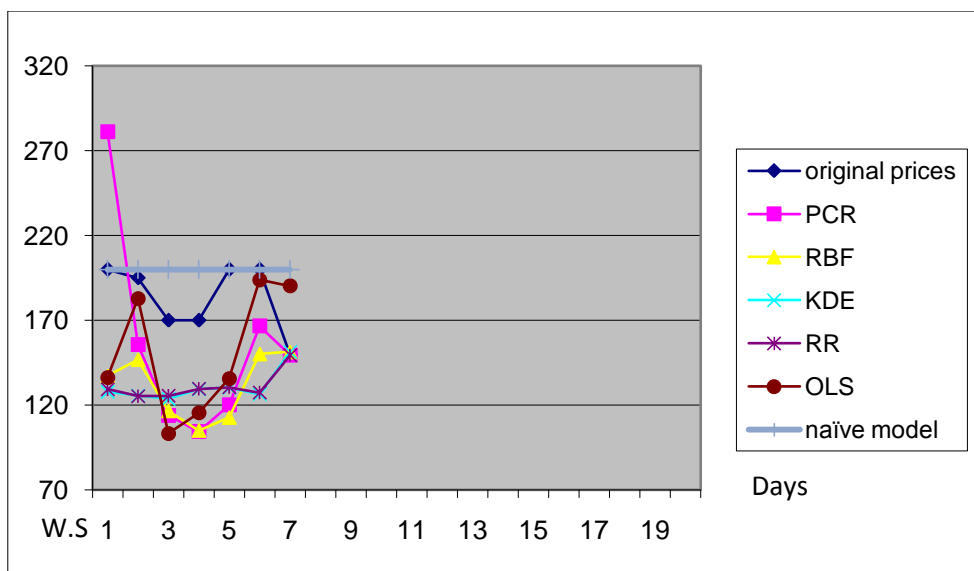
**Σχήμα 26:** Παρουσίαση των προβλέψεων για την πρώτη πρόβλεψη εντός δείγματος.

Στο παραπάνω σχεδιάγραμμα παρουσιάζεται διαγραμματικά η απόκλιση της κάθε μεθόδου πρόβλεψης από τα πραγματικά δεδομένα.

Όπως και παραπάνω στη δεύτερη πρόβλεψη χρησιμοποιήθηκαν 2263 σημεία.

**Πίνακας 7:** Παρατίθεται πάλι πίνακας και οι προβλέψεις

original prices	PCR	RBF	KDE	RR	OLS	naïve model
200	281,151	137,232	127,824	129,404	136,312	200
195	155,673	146,841	125,64	125,208	182,823	200
170	114,006	116,608	123,691	125,417	103,27	200
170	104,407	105,063	129,297	129,583	115,353	200
200	120,082	112,673	130,241	130,417	135,61	200
200	166,613	150,189	126,352	127,292	193,829	200
150	149,496	151,364	151,161	149,375	190,189	200
	E: 1.0868	E: 0.9938	E: 1.0214	E: 1.0141	E: 0.8023	E:0,4517

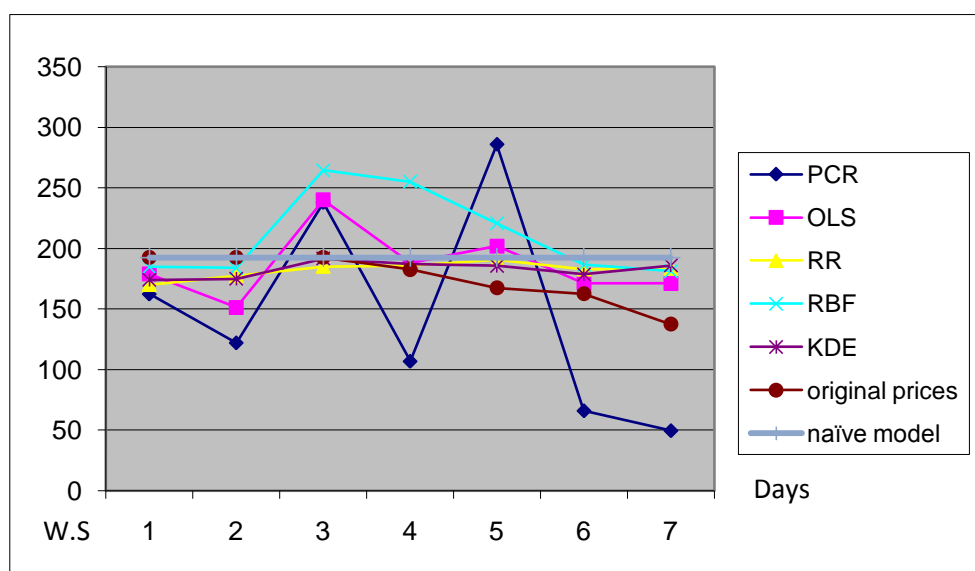


**Σχήμα 27:** Παρουσίαση των προβλέψεων για την δεύτερη πρόβλεψη εντός δείγματος

Στην Τρίτη μέτρηση 2585 σημείων εξήχθησαν τα παρακάτω αποτελέσματα:

**Πίνακας 8:** Παρουσιάζονται τα αποτελέσματα για την Τρίτη πρόβλεψη εντός δείγματος.

PCR	OLS	RR	RBF	KDE	naïve model	original prices
162,390	178,788	170	184,802	174,006	192,5	192,5
122,140	151,29	177,273	183,885	174,968	192,5	192,5
237,911	240,069	185	264,467	191,527	192,5	192,5
106,922	187,985	185,909	255,053	187,324	192,5	182,5
285,880	201,943	190,682	220,764	185,763	192,5	167,5
66,106	171,076	182,727	186,606	178,595	192,5	162,5
49,724	171,218	183,409	181,349	185,821	192,5	137,5
					E:	
E: 1.6449	E: 0.63131	E: 0.483693	E: 0.98221	E: 0.467179	0,534578	

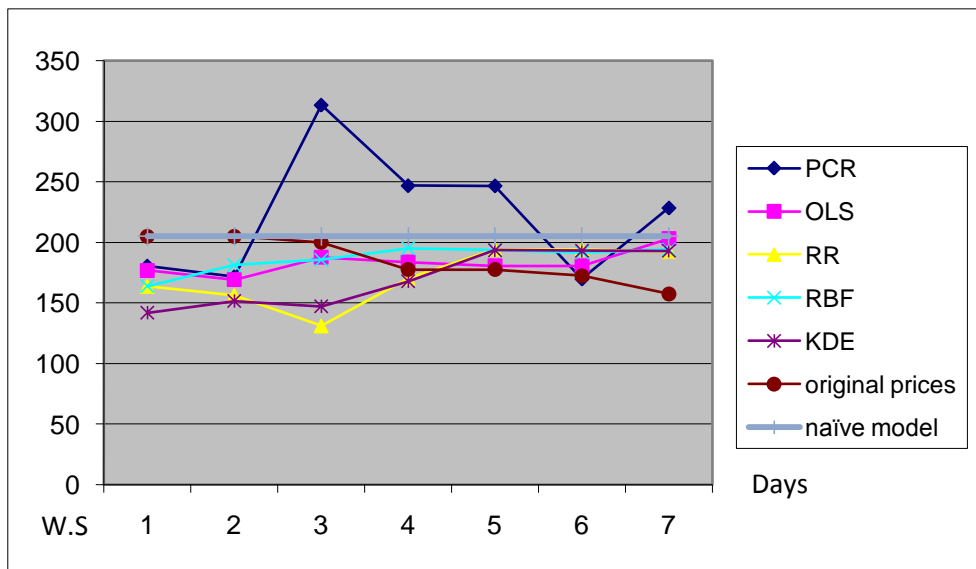


**Σχήμα 28:** Παρουσίαση των προβλέψεων για την τρίτη πρόβλεψη εντός δείγματος.

**Πίνακας 9:** Στην τέταρτη κατά σειρά μέτρηση 802 σημείων έχουμε τις ακόλουθες μετρήσεις

OLS	PCR	RR	RBF	KDE	naïve method	original prices
177	180	164	164	142	205	205
169	171,57	156	181	152	205	205
188	313	131	186	147	205	200
184	247	171	195	168	205	177,5
181	246,72	194	194	194	205	177,5
181	170	194	192	193	205	172,5
203	228	192,5	194	193	205	157,5
E: 0,08454	E: 0,816392	E: 0,311088	E: 0,105693	E: 0,302533	E: 0,351962	

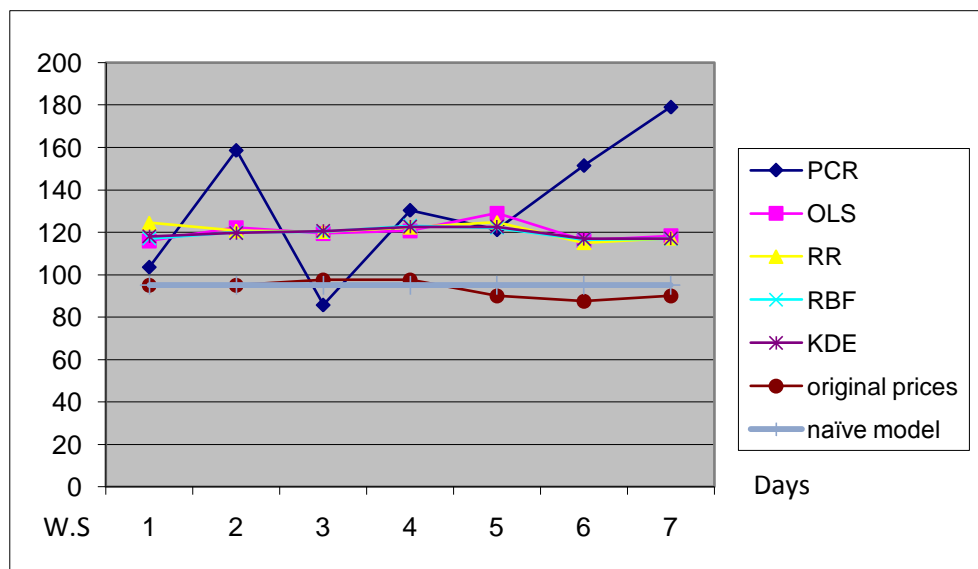
και παρουσιάζεται και το ακόλουθο σχεδιάγραμμα:



**Σχήμα 29:** Παρουσιάζονται τα αποτελέσματα για την Τέταρτη πρόβλεψη εντός δείγματος.

**Πίνακας 10:** Στην επόμενη μέτρηση 860 σημείων, που είναι η 5η κατά σειρά έχουμε τα ακόλουθα αποτελέσματα

OLS	PCR	RR	RBF	KDE	naïve method	original prices
116	104	124	117	118	95	95
122	159	120,698	120	120	95	95
119	86	120	120	121	95	97,5
121	130	122	123	123	95	97,5
129	121	125	122	123	95	90
116	151	115	116,562	117	95	87,5
118,33	179	117	117	117	95	90
E: 2,00002	E: 5,6716	E: 1,90593	E: 1,80988	E: 1,8294 5	E:  0,21312	

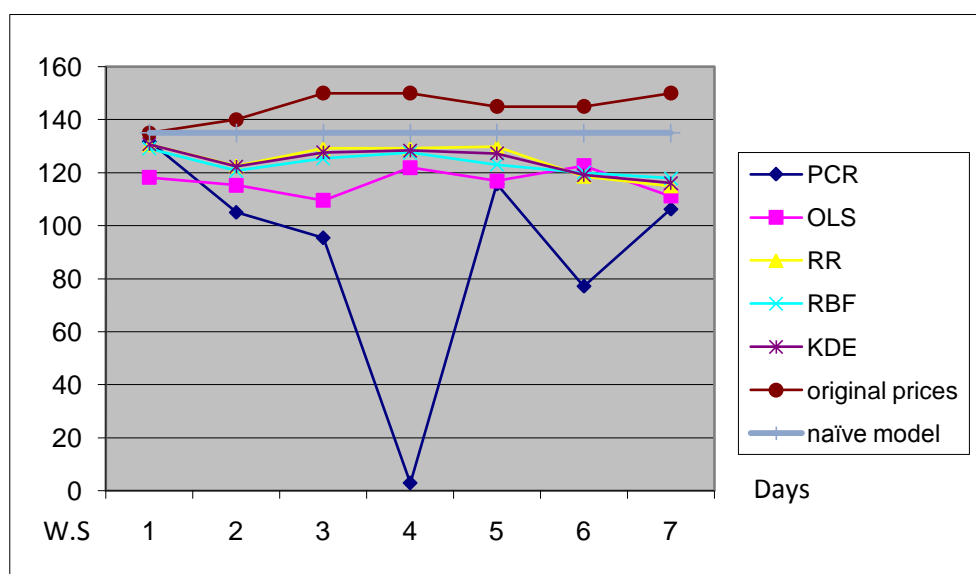


**Σχήμα 30:** Παρουσιάζονται τα αποτελέσματα για την πέμπτη πρόβλεψη εντός δείγματος.



**Πίνακας 11:** Στην 6 μέτρηση 1352 σημείων, έχουμε τα ακόλουθα αποτελέσματα

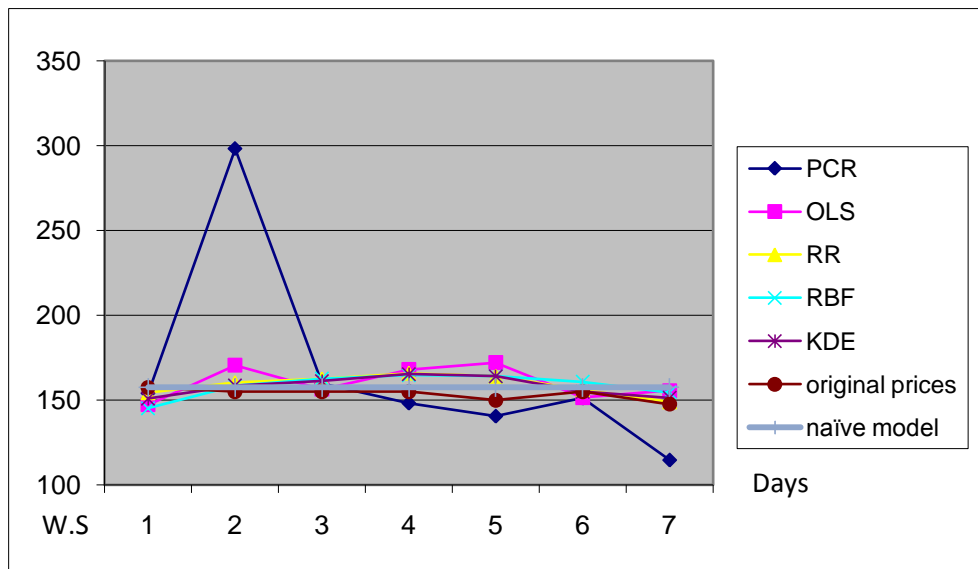
OLS	PCR	RR	RBF	KDE	naïve method	original prices
118	132	131	129	131	135	135
115	105	122,5	121	122	135	140
110	95	129	125	128	135	150
121,9	3	129	128	128	135	150
117	116	130	123	127	135	145
122,63	77	119	120	119	135	145
111	106	115	118	116	135	150
E: 0,658758	E: 3,54138	E: 0,636739	E: 0,550815	E: 0,606995	E: 0,33191306	



**Σχήμα 31:** Παρουσιάζονται τα αποτελέσματα για την έκτη πρόβλεψη εντός δείγματος.

**Πίνακας 12:** Στην εβδομή και τελευταία μέτρηση 1675 σημείων, έχουμε

OLS	PCR	RR	RBF	KDE	naïve method	original prices
147	154	154	145	151	158	157,5
170	298	160	158	158,5	158	155
156	160	162,5	163	161,26	158	155
168	148	166	165	165	158	155
172	141	164	164	164	158	150
151	151	156	161	154,94	158	155
156	115	149	154	151	158	147,5
E: 0,211311	E: 1,36796	E: 0,122778	E: 0,0969239	E: 0,100243	E: 0,138113	



**Σχήμα 32:** Παρουσιάζονται τα αποτελέσματα για την εβδομή πρόβλεψη εντός δείγματος .

Τέλος πραγματοποιήθηκαν οι εξής μετρήσεις για την εύρεση της τάσης σε κάθε μέτρηση. Παρακολουθώντας τις ανόδους και τις καθόδους των τιμών των προβλέψεων τις συγκρίνουμε με τις ανόδους και τις καθόδους των πραγματικών τιμών, και έτσι βρίσκουμε ποσοστιαία κατά πόσο η πρόβλεψη ακολουθεί την τάση (trend) των πραγματικών τιμών. Επομένως έχουμε:

**Πίνακας 13:** Στην πρώτη μέτρηση έχουμε τα εξής αποτελέσματα

PCR				OLS			
output \actual	up	down	result	output \actual	up	down	result
up	0	0	0	up	1	0	1
down	1	1	2	down	0	1	1
		50%			100%	100%	
total		14%		total	14%	14%	
RBF				KDE			
output \actual	up	down	result	output \actual	up	down	result
up	1	0	1	up	0	0	0
down	0	1	1	down	1	1	2
	100%	100%				50%	
total	14%	14%		total		14%	

RR			
output \actual	up	down	result
up	0	0	0
down	1	1	2
		50%	
total		14%	

**Πίνακας 14:** Στην δεύτερη μέτρηση έχουμε τις ακόλουθες παρατηρήσεις

PCR				OLS			
output \actual	up	down	result	output \actual	up	down	result
up	1	1	2	up	1	1	2
down	1	3	4	down	0	2	2
	50%	75%			50%	100%	
total	43%	43%		total	14%	29%	
RBF				KDE			
output \actual	up	down	result	output \actual	up	down	result
up	1	2	3	up	1	1	2
down	0	1	1	down	0	1	1
	33%	100%			50%	100%	
total	14%	14%		total	14%	14%	

RR			
output \actual	up	down	result
up	1	1	2
down	0	1	1
	50%	100%	
total	14%	14%	

Πίνακας 15: Στην Τρίτη μέτρηση βρίσκουμε τα ακόλουθα

PCR				OLS			
output \actual	up	down	result	output \actual	up	down	result
up	0	1	1	up	0	1	1
down	0	3	3	down	0	2	2
		100%				100%	
total		43%		total	0	29%	
RBF				KDE			
output \actual	up	down	result	output \actual	up	down	result
up	0	0	0	up	0	1	1
down	0	3	3	down	0	3	3
		100%				100%	
total		43%		total		43%	
RR							
output \actual	up	down	result				
up	0	2	2				
down	0	2	2				
		100%					
total		29%					

**Πίνακας 16:** Στην τέταρτη μέτρηση έχουμε τα εξής

PCR				OLS			
output \actual	up	down	result	output \actual	up	down	result
up	0	2	2	up	0	2	2
down	0	2	2	down	0	1	1
		1				100%	
total		29%		total	0	14%	
RBF				KDE			
output \actual	up	down	result	output \actual	up	down	result
up	0	3	3	up	0	1	1
down	0	1	1	down	0	1	1
	0	100%				100%	
total	0	14%		total		14%	
RR							
output \actual	up	down	result				
up	0	0	0				
down	0	1	1				
		100%					
total		14%					

**Πίνακας 17:** Στην Πέμπτη μέτρηση έχουμε τα εξής αποτελέσματα

PCR				OLS			
output \actual	up	down	result	output \actual	up	down	result
up	1	1	2	up	1	1	2
down	1	1	2	down	1	1	2
	50%	50%			50%	50%	
total	14%	14%		total	14%	14%	
RBF				KDE			
output \actual	up	down	result	output \actual	up	down	result
up	1	0	1	up	2	0	2
down	0	2	2	down	0	1	1
	100%	100%			100%	100%	
total	14%	29%		total	29%	14%	

RR			
output \actual	up	down	result
up	1	1	2
down	1	1	2
	50%	50%	
total	14%	14%	

**Πίνακας 18:** Στην έκτη μέτρηση παρουσιάζονται τα ακόλουθα

PCR				OLS			
output \actual	up	down	result	output \actual	up	down	result
up	1	1	2	up	0	0	0
down	2	0	2	down	3	1	4
	50%					25%	
total	14%			total		14%	
RBF				KDE			
output \actual	up	down	result	output \actual	up	down	result
up	1	0	1	up	1	0	1
down	2	1	3	down	2	1	3
	100%	33%			100%	33%	
total	14%	14%		total	14%	14%	
RR							
output \actual	up	down	result				
up	1	0	1				
down	2	0	2				
	100%						
total	14%						



**Πίνακας 19:** Στην έβδομη και τελευταία μέτρηση έχουμε τα εξής αποτελέσματα.

PCR				OLS			
output \actual	up	down	result	output \actual	up	down	result
up	1	0	1	up	0	3	3
down	0	2	2	down	1	0	1
	100%	100%					
total	14%	29%		total			
RBF				KDE			
output \actual	up	down	result	output \actual	up	down	result
up	0	1	1	up	0	1	1
down	1	1	2	down	1	1	2
		50%				50%	
total		14%		total		14%	

RR			
output \actual	up	down	result
up	0	1	1
down	1	1	2
		50%	
total		14%	

## Συμπεράσματα

Στην πρώτη μέτρηση παρατηρούμε από το σχεδιάγραμμα ότι όλες οι μετρήσεις έχουν αποκλίσεις από τις πραγματικές τιμές είναι μικρές. Από το σφάλμα E βλέπουμε ότι η μέθοδος με το μεγαλύτερο σφάλμα είναι η OLS. Η μέθοδος με την μικρότερη απόκλιση είναι η RBF με  $E:0.01693$  δηλαδή με 17% περίπου απόκλιση από τις πραγματικές τιμές.

Στην δεύτερη μέτρηση παρατηρούμε ότι έχουμε μεγαλύτερες αποκλίσεις. Μια ικασία που μπορεί να γίνει είναι ότι οι μέθοδοι πρόβλεψης που χρησιμοποιούμε δεν μπορούν να δώσουν καλά αποτελέσματα όταν οι πραγματικές τιμές έχουν βραχυπρόθεσμα απότομες αυξομειώσεις. Η μέθοδος με την μικρότερη απόκλιση ήταν η naïve method με  $E:0.4517$ . Η μέθοδος με την μεγαλύτερη απόκλιση ήταν η PCR με σφάλμα  $E:1.08688$ .

Στην τρίτη μέτρηση παρατηρούμε ότι έχουμε ακόμα μεγαλύτερες αυξομειώσεις. Η μέθοδος με το μικρότερο σφάλμα είναι και πάλι η KDE με σφάλμα  $E: 0.467179$ . Η μέθοδος με το μεγαλύτερο σφάλμα είναι η PCR με  $E:1.6445$ .

Στην τέταρτη μέτρηση παρατηρούνται ομαλές αυξομειώσεις. Η μέθοδος με το μεγαλύτερο σφάλμα είναι η PCR με  $E:0.81639$ . Η μέθοδος με το μικρότερο σφάλμα είναι η OLS με  $E: 0.08454$ .

Στην πέμπτη κατά σειρά μέτρηση παρουσιάζονται μεγάλα σφάλματα. Η μέθοδος με το μικρότερο σφάλμα είναι η naïve method με  $E:0.2131$ . Η μέθοδος με το μεγαλύτερο λάθος είναι η PCR με  $E:5.6716$ .

Στην έκτη μέτρηση έχουμε τα εξής αποτελέσματα. Η μέθοδος με το μικρότερο σφάλμα είναι η naïve method με  $E:0.3319$ . Η μέθοδος από την άλλη με το μεγαλύτερο σφάλμα είναι η PCR με  $E:3.5414$ .

Στην έβδομη και τελευταία μέτρηση έχουμε τις ακόλουθες παρατηρήσεις. Η μέθοδος με το μικρότερο σφάλμα είναι η naïve method με  $E:0.1381$ . Η μέθοδος με το μεγαλύτερο σφάλμα είναι η PCR με  $E:1.3680$ .

Στην πρώτη μέτρηση οι μέθοδοι πρόβλεψης έδωσαν ικανοποιητικά αποτελέσματα ενώ στην δεύτερη και στην τρίτη είχαμε μεγάλες αυξομειώσεις μεταξύ των μεθόδων. Στην τέταρτη και έβδομη μέτρηση οι προβλέψεις με τις περισσότερες μεθόδους παρουσιάζονται να ακολουθούν το trend των original τιμών. Και στην πέμπτη και στην 6 γίνεται το ίδιο πράγμα αλλά με μεγαλύτερα σφάλματα. Η naïve method που είναι η πιο συντηρητική και η πιο απλή έχει τις καλύτερες προσεγγίσεις από όλες τις άλλες και σε 4 από τις 7 μετρήσεις.

Επίσης έγιναν οι ακόλουθες παρατηρήσεις :

1. Όταν οι πραγματικές τιμές διατηρούν μια σταθερή κλίση, τότε οι μέθοδοι πρόβλεψης παρουσιάζουν επιτυχία.
2. Όταν τα δεδομένα που χρησιμοποιούμε είναι λίγα για την πρόβλεψη τότε το σφάλμα των μεθόδων είναι μεγάλο. Χαρακτηριστικό παράδειγμα είναι η 5<sup>η</sup> μέτρηση.
3. Αν και υπάρχει μεγάλο σφάλμα από τα πραγματικά δεδομένα οι περισσότερες από τις μεθόδους ακολουθούν το σωστό trend(4 στις 5 περιπτώσεις).
4. Η μέθοδος πρόβλεψης που απέχει περισσότερο στην σωστή πρόβλεψη του trend αλλά και παρουσιάζει συνήθως το μεγαλύτερο σφάλμα είναι η PCR.

Με την αναλυτικότερη μέθοδο εύρεσης του trend που παρουσιάζεται στους πίνακες 16 έως 22 έχουμε τα ακόλουθα συμπεράσματα:

- Στην πρώτη πρόβλεψη οι μέθοδοι που ακολουθούν καλύτερα το trend με 14% είναι η OLS και η PCR
- Στην 2<sup>η</sup> πρόβλεψη η PCR ακολουθεί καλύτερα το trend με 43% και στα up και στα down.
- Στην 3<sup>η</sup> πρόβλεψη οι PCR,RBF και KDE ακολουθούν καλύτερα το trend με 43% στα down.Στην συγκεκριμένη πρόβλεψη οι πραγματικές τιμές δεν παρουσίαζαν καθόλου up.
- Στην 4<sup>η</sup> πρόβλεψη έχουμε πάλι 0 στα up στις πραγματικές τιμές η PCR είναι καλύτερη στα down.
- Στην 5<sup>η</sup> πρόβλεψη η RBF έχει 14% ταύτιση στα up με τις πραγματικές τιμές και 29% στα down.
- Στην 6<sup>η</sup> πρόβλεψη οι RBF και KDE είναι οι μέθοδοι με την καλύτερη προσέγγιση του trend των πραγματικών τιμών με 14% up και κάτω.
- Στην 7<sup>η</sup> και τελευταία πρόβλεψη η PCR είναι η καλύτερη στο trend με 14% στο up και 29% down.

Σε αυτήν την εργασία λοιπόν αποδείξαμε πως η χρονοσειρά μας, που αποτελείται από τις τιμές της ναυλαγοράς πλοίων Τανκερ, παρουσιάζει χαοτική συμπεριφορά. Το αποτέλεσμα αυτό συμφωνεί με την διδακτορική διατριβή της κ. Μαρινίκης Ψιφιά. Στην διδακτορική διατριβή αποδείχθηκε με την χρήση του εργαλείου BDS πως η ναυλαγορά και στην χρονοαύλωση και στην ναύλωση ανά ταξίδι παρουσιάζει χαοτική συμπεριφορά, και προτρέπουν τους μελετητές να επιχειρήσουν να εφαρμόσουν μεθόδους πρόβλεψης σε αυτές. Στην διπλωματική εργασία αυτή έκτος από την απόδειξη της χαοτικής συμπεριφοράς μέσω του φάσματος ισχύος και της εύρεσης των μεγίστων εκθετών Lyapunov, εφαρμόσαμε 5 διαφορετικές μεθόδους πρόβλεψης. Οι προβλέψεις μας έγιναν όλες εντός δείγματος (predictions) και όχι

εκτός δείγματος (forecastle), προκειμένου να έχουμε την δυνατότητα να διασταυρώσουμε κατά πόσο αυτές οι προβλέψεις προσεγγίζουν την πραγματικότητα.

Με βάση τα αποτελέσματα των μεθόδων πρόβλεψης που παρατίθενται παραπάνω η πιο αποδοτική μέθοδος πρόβλεψης είναι η αθώα μέθοδος (naïve method), μια μέθοδος η οποία με απλά λόγια θεωρεί ότι η τιμή της χρονοσειράς θα παραμείνει ως έχει. Επομένως καμιά χασοτική μέθοδος πρόβλεψης δεν μας έδωσε τα επιθυμητά αποτελέσματα. Μπορούμε να ισχυριστούμε ότι ως ένα βαθμό αυτές οι μέθοδοι πρόβλεψης κατάφεραν να προσεγγίσουν την τάση της αγοράς (trend) αν και σε βραχυπρόθεσμες προβλέψεις. Δεν είναι επομένως εφικτή η πρόβλεψη της χρονοσειράς μας, και κατ' επέκταση οι τιμές της ναυλαγοράς που μελετήσαμε με τις προαναφερθείσες μεθόδους.

Σε αυτό στο σημείο πρέπει να αναφερθούμε στα πειραματικά αποτελέσματα της διδακτορικής διατριβής του Andrew Nicola Edmonds (Edmonds). Η μέθοδος πρόβλεψης, που ο ίδιος ανέπτυξε με την χρήση νευρωνικών δικτύων, εφαρμόστηκε σε οικονομικά δεδομένα. Συγκεκριμένα εφάρμοσε την μέθοδο πρόβλεψης σε 5 χρονοσειρές παρμένες από την αγορά συναλλάγματος. Η δειγματοληψία γινόταν ανά λεπτό κατά την διάρκεια μίας μέρας. Οι τιμές που χρησιμοποίησε ήταν οι spot τιμές του Γερμανικού μάρκου έναντι του δολαρίου DM-\$, της λύρας Αγγλίας έναντι του δολαρίου £-\$, και τα τριμηνιαία επιτόκια της αγγλικής λύρας, του δολαρίου και του μάρκου αντίστοιχα, τα οποία παρουσιάζουν χασοτική συμπεριφορά (Medio). Τα αποτελέσματα από τις προβλέψεις στις χρονοσειρές αυτές ήταν ικανοποιητικά, αλλά σε καμία περίπτωση δεν ήταν αρκετά προκειμένου να χρησιμοποιηθούν στην αγορά ούτως ώστε κάποιος να βγάλει χρήματα από αυτά. Αναφέρεται χαρακτηριστικά στα συμπεράσματα ότι ένα σύστημα πρόβλεψης το οποίο παρουσιάζει 65% επιτυχία πρόβλεψης μπορεί να έχει σαν αποτέλεσμα να οδηγήσει σε απώλεια χρημάτων και ένα άλλο με 40% επιτυχία πρόβλεψης να είναι επικερδές.

## **Προτάσεις για περαιτέρω έρευνα**

Στην αρχή αυτής της εργασίας αναγνωρίστηκε η ανάγκη για πρόβλεψη και ανάλυση χρονοσειρών παρμένες από τον πραγματικό κόσμο που παρουσιάζουν χάος και άλλες που έχουν μη γραμμική συμπεριφορά. Ενώ υπήρχαν εργαλεία να πραγματοποιήσουν αυτό το έργο, έπασχαν από 2 οπισθοδρομήσεις: την ανάγκη για την επιλογή εμπειρικά ενός set από παραμέτρους που ήταν ζωτικές για την απόδοση της πρόβλεψης, και το ποσό της υπολογιστικής ισχύος που απαιτείτο. Οι παράμετροι έχουν επιλεγθεί χωρίς να καταφύγουν στην θεωρία και τον αιρετισμό πρέπει να υπάρχει η αμφιβολία για την ποιότητα των αποτελεσμάτων.

Για περαιτέρω έρευνα προτείνουμε την χρήση Νευρωνικών δικτύων. η βασική ιδέα σε αυτά είναι η χρήση εξομοιωμένων νευρώνων σε αυτά, δηλαδή η ύπαρξη απλών μονάδων επεξεργασίας οι οποίες μπορούν να χρησιμοποιηθούν μαζί με έναν τρόπο πανομοιότυπο με τον ανθρώπινο εγκέφαλο προκειμένου να έχουμε ένα είδος τεχνητής

νοημοσύνης. Βασισμένοι στην βιολογία και στο πως λειτουργεί ο ανθρώπινος εγκέφαλος βγάλαμε το συμπέρασμα ότι η μάθηση πρέπει να διαμοιράζεται σε ένα αριθμό από απλές υπολογιστικές μονάδες και να διαμορφώνει τη συμπεριφορά των συνδέσεων των μονάδων ή και τις μονάδες τις ίδιες. Νευρωνικά δίκτυα έχουν χρησιμοποιηθεί για την πρόβλεψη της αγοράς συναλλάγματος με αρκετά μεγάλο ποσοστό επιτυχίας πρόβλεψης, με βάση στην εργασία που έχει εκδώσει ο Andrew Nicola Edmonds.

Μια ουσιώδης λύση με προοπτική μπορεί να είναι η προσθήκη στον αλγόριθμο να ξεχνάει. Κάθε σημείο εμπύθισης παλαιότερο από μια συγκεκριμένη ημερομηνία θα μπορούσε να αγνοηθεί εντελώς στην παραγωγή προβλέψεων. Θα ήταν ενδιαφέρον να χρησιμοποιηθούν νευρωνικά δίκτυα για την έρευνα παρουσίας χάους στις χρονοσειρές που παράγονται στην ναυλαγορά.

## Βιβλιογραφία

1. *TIME SERIES PREDICTION USING SUPERVISED LEARNING AND TOOLS FROM CHAOS THEORY*. A thesis submitted to the faculty of science and computing, University of Luton, in partial fulfillment of the requirements for the degree of doctor of philosophy Author : Andrew Nicola Edmonds, 1996.
2. *Χάος. Ανάλυση και πρόβλεψη χρονοσειρών . Εφαρμογές στο χρηματιστήριο Αθηνών με απλά παραδείγματα*. Σύγγραμμα των Κώστα Συριόπουλου και Αλέξανδρου Λεοντίση, 1996.
3. j. Doyne Farmer and John j. Sidorowitch, *Predicting Chaotic Time Series*, Physical Review, 1987.
4. Hurst, H.E, Black, R.P and Simaika, Y.M., *Long-Term storage: An Experimental Study*, Constable-London, 1965.
5. Kaplan D., Glass L., *Coarse-grained embeddings of time series: random walks, Gaussian random processes, and deterministic chaos*, 1991.
6. Kennel, M. Brown, R. Abarbanel H.D.I, *Determining embedding dimension for phase space reconstruction using a geometrical construction*, 1992.
7. Lorenz E., *Deterministic non-periodic flow*, *journal of atmospheric science*, 20, pp.130-146, 1963.
8. M.C. Mackey & L. Glass, *Oscillation and chaos in Physiological Control Systems*, Science, 197, 287, 1977.
9. Medio, Alfredo. *Chaotic Dynamics: Theory and application to Economics*. Chapter 14 1992 Cambridge University press, 1992.
10. Ruelle D., *Small random perturbations of dynamical systems and the definition of attractors*, Commun Math. Phys. 137, pp 82, 1981.
11. Shannon C.E. & Weaver W. *The mathematical theory of communication*. Urbana II. University of Illinois press 1949.
12. Takens F., *Detecting Strange attractors in turbulence*. Lecture Notes in Mathematics, Vol 898 (Warwick 1980) eds D.A. Rand, I.S. Young pp. 368-381 Berlin: Springer Verlag, 1981.
13. Wolf, A., Swift, J.B., Swinney, H.L., Vastano, J.A., *Determining Lyapunov exponents from a time series*. Physica 16D 285-317 1985
14. Turchin P., *Chaos & Stability in rodent population dynamics: evidence from a non-linear time-series analysis*. Oikos 24178-6168, 1993.
15. Eckman J.-P., Ruelle D., *Fundamental limitations for estimating dimensions and Lyapunov exponents in dynamical systems*, Physica D56, 1992.
16. Kantz H., *A robust method to estimate the maximal Lyapunov exponent of a time series*, Physics Letters A 185, 1994.
17. Kugiumtzis F., Medio A., *Is a random walk the best exchange rate predictor?*, International Journal of Forecasting 13, σελ 255-267, 1997.
18. Sauer T., *Time series prediction by using delay coordinate embedding*, in *Time Series Prediction: Forecasting the Future and Understanding the Past*, Santa

- Fe Institute of Science of Complexity XV, eds. Weigend A. S. and Gershenfeld N. A., Addison-Wesley, Reading, MA, σελ 175-194, 1993.
19. Abarbanel H., Brown R., Sidorowitch J., Tsimring Lev Sh., *The analysis of observed chaotic data in physical systems*, Reviews of Modern Physics 65 (4), σελ 1331-1392, 1993.
  20. Casdagli M., *Nonlinear prediction of chaotic time series*, Physica D35, σελ 335-356, 1989.
  21. Smith L. A., *Identification and prediction of low dimensional dynamics*, Physica D 58, σελ 50-76, 1992.
  22. Sugihara G., May R., *Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series*, Nature 344, σελ 734-740, 1990.
  23. Fama, E.F., *The behavior of Stock Market Prices*, journal of business, σελ 34-105, January 1965.
  24. Alexandros M. Goulielmos and Mariniki Psifia, *A study of trip and time charter freight rate indices: 1968-2003*, Maritime Policy & Management, Vol 34, No. 1, pages 55-67, February 2007.
  25. Mandelbrot, B., *The Fractal Geometry of Nature* : σελ 197, 1983.
  26. Grassberger P., Procaccia I., *Characterization of strange attractors*, Physical Review Letters A128(6-7), σελ 369-373, 1983
  27. Eckman J.-P., Ruelle D., *Fundamental limitations for estimating dimensions and Lyapunov exponents in dynamical systems*, Physica D56, σελ 185-187, 1992.
  28. Bajo-Rubio O., Fernandez-Rodriguez F., Sosvilla-Rivero S., *Chaotic behavior in exchange –rate series: first results for the Peseta-United States Dollar case*, Economics Letters 39, σελ 207-211, 1992.
  29. Diebold F. X., Nason J. A., *Nonparametric exchange rate prediction?*, Journal of International Economics 28, σελ 315-332, 1990.
  30. Lisi F., Schiavo R. A., *A comparison between neural networks and chaotic models for exchange rate prediction*, Computational Statistics & Data Analysis 30, σελ 87-102, 1999.
  31. Feigenbaum M. J., *Quantitative universality for a class of nonlinear transformations*, Journal of Statistical Physics 19, σελ 25-52 1978.
  32. Feigenbaum M. J., *The universal metric properties of nonlinear transformations*, Journal of Statistical Physics 21, σελ 669-706 1979.
  33. Feigenbaum M. J., *Universal behavior in nonlinear systems*, Los Alamos science 1, σελ 4-27 1980.
  34. Feigenbaum M. J., *The transition to chaos*, Chaos: The new science, J. Holte (ed.), σελ 45-53 , 1993.
  35. Brock W., Dechert D., Scheinkman J., *A test for independence based on correlation dimension*, University of Wisconsin-Madison, Social Science Research WP No. 8762, Econometric Reviews, 15, σελ 197-235, 1987.
  36. Peitgen H-O., H Jurgens, D. Saupe, *Chaos and fractals: new frontiers of science*, Berlin: Springer, 1992.