

ΕΘΝΙΚΟ ΚΑΙ ΚΑΠΟΔΙΣΤΡΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

Τεχνικές Ανάλυσης και Πρόβλεψης Αγορών

Δημήτριος Α. Μπούρας
Γεώργιος Α. Κουκουβάος

Επιβλέπων Βαρουτάς Δημήτριος, Επίκουρος Καθηγητής

ΑΘΗΝΑ

ΝΟΕΜΒΡΙΟΣ 2015

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

Τεχνικές Ανάλυσης και Πρόβλεψης Αγορών

Δημήτριος Α. Μπούρας

ΑΜ: 1115200800067

Γεώργιος Α. Κουκουβάος

ΑΜ: 1115200800203

Επιβλέπων Βαρουτάς Δημήτριος, Επίκουρος Καθηγητής

ΠΕΡΙΛΗΨΗ

Στην παρούσα πτυχιακή εργασία παρουσιάζεται το θέμα των Τεχνικών αναλύσεων και προβλέψεων των αγορών. Η πρόκληση που ανακύπτει και ο στοχαστικός χαρακτήρας του προβλήματος καθιστούν τον τομέα των προβλέψεων ιδιαίτερα δύσκολο αλλά συγχρόνως και ενδιαφέρον. Ανεξάρτητα από τις συνθήκες ή τους χρονικούς ορίζοντες που εμπλέκονται , οι προβλέψεις αποτελούν σημαντικό βοήθημα για τον αποτελεσματικό και αποδοτικό σχεδιασμό. Οι προβλέψεις καταστάσεων διαφέρουν σε μεγάλο βαθμό στους χρονικούς ορίζοντες , στους παράγοντες που καθορίζουν τα πραγματικά αποτελέσματα , τα είδη των δεδομένων και σε πολλές άλλες πτυχές. Επιπλέον οι μέθοδοι πρόβλεψης μπορεί να είναι πολύ απλές όπως χρησιμοποιώντας την πιο πρόσφατη παρατήρηση ως πρόβλεψη , η οποία ονομάζεται ‘Απλοϊκή μέθοδος’ (Naive Method) ή εξαιρετικά πολύπλοκη όπως η ‘Πολυμετάβλητη παλινδρόμηση’ .Επιπρόσθετα παρουσιάζονται διάφορες μέθοδοι εξομάλυνσης , οι μέθοδοι πρόβλεψης συνεχούς ζήτησης, οι στατιστικές μέθοδοι πρόβλεψης , τα μοντέλα παλινδρόμησης και η ανάλυση ARIMA οι οποίες χρησιμοποιούνται για τη διαμόρφωση βραχυπρόθεσμων ή μακροπρόθεσμων προβλέψεων και μελλοντικών τιμών μιας χρονοσειράς

ΘΕΜΑΤΙΚΗ ΠΕΡΙΟΧΗ: Τεχνικές αναλύσεις και προβλέψεις αγορών
ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ: μοντέλα , πρόβλεψη , επιχείρηση , μέθοδοι , στατιστικές , αγορές , ανάλυση

ABSTRACT

This final paper develops the subject of Technical analysis and Forecasting of the markets. The challenge that appears and the contemplative character of the matter make the area of Forecasting specially difficult but at the same time very interesting. Regardless the circumstances and the time limits , forecasting is an important aid for the effective and efficient design. Conditional forecasting differs in the time limits , the factors that define the true results , the variety of data and many other aspects. Additionally the methods of forecasting can be very simple like using the most recent observation as forecast ,which is called Naïve Method or extremely complicated like 'Multivariate Regression' . In addition this paper develops different methods .Such as the method of forecasting constant demand , statistic methods of forecasting , regression models and ARIMA analysis , which are being used for conformation of short term or long term forecasting and future numbers of the timeline.

SUBJECT AREA: Technical analysis and Forecasting of the markets

KEYWORDS: forecasting, enterprise , methods , statistics , market , analysis

ΠΕΡΙΕΧΟΜΕΝΑ

1. ΕΙΣΑΓΩΓΗ 10

2. ΤΥΠΟΙ ΠΡΟΒΛΕΨΗΣ 13

3. ΜΕΘΟΔΟΣ ΝΑΪVE ΚΑΙ ΟΙ ΠΑΡΑΛΛΑΓΕΣ ΤΗΣ 14

3.1. Πρώτη παραλλαγή 14

3.2. Δεύτερη παραλλαγή 14

3.3. Τρίτη παραλλαγή 15

4. ΜΕΘΟΔΟΣ ΑΠΛΟΥ ΜΕΣΟΥ ΟΡΟΥ 16

5. ΜΕΘΟΔΟΣ ΑΠΛΟΥ ΚΙΝΗΤΟΥ ΜΕΣΟΥ ΟΡΟΥ 18

6. ΜΕΘΟΔΟΣ ΔΙΠΛΟΥ ΚΙΝΗΤΟΥ ΜΕΣΟΥ ΟΡΟΥ 21

7. ΜΕΘΟΔΟΣ ΑΠΛΗΣ ΕΚΘΕΤΙΚΗΣ ΕΞΟΜΑΛΥΝΣΗΣ 22

8. ΜΕΘΟΔΟΣ ΓΡΑΜΜΙΚΗΣ ΤΑΣΗΣ (HOLT EXPONENTIAL SMOOTHING) 23

9. ΜΕΘΟΔΟΣ ΜΗ ΓΡΑΜΜΙΚΗΣ ΤΑΣΗΣ 24

10. ΜΕΘΟΔΟΣ ΔΙΠΛΗΣ ΕΚΘΕΤΙΚΗΣ ΕΞΟΜΑΛΥΝΣΗΣ 26

11. ΜΕΘΟΔΟΣ ΤΡΙΠΛΗΣ ΕΚΘΕΤΙΚΗΣ ΕΞΟΜΑΛΥΝΣΗΣ 28

12. ΔΙΑΣΠΑΣΗ ΧΡΟΝΟΣΕΙΡΩΝ 29

12.1. Γραμμικό μοντέλο (Linear) 32

12.2. Παραβολικό μοντέλο (quadratic) 33

12.3. Εκθετικό μοντέλο (exponential) 33
12.4. Μοντέλο Pearl-Reed Logistic 34
12.5. Μοντέλο Gompertz 34
12.6. Διάσπαση χρονοσειράς – Πρόβλεψη 34
13. ΑΠΛΗ ΓΡΑΜΜΙΚΗ ΠΑΛΙΝΔΡΟΜΗΣΗ 35
13.1. Παραδοχές της γραμμικής παλινδρόμησης 36
13.2. Ασθενείς εξωγενείς 36
13.3. Γραμμικότητα 37
13.4. Συνεχής Μεταβλητότητα (αλλιώς ομοσκεδαστικότητα) 37
13.5. Ανεξαρτησία λαθών 38
13.6. Έλλειψη πολυσυγγραμμικότητας στους προγνωστικούς παράγοντες. 38
14. ΠΟΛΥΜΕΤΑΒΛΗΤΗ ΠΑΛΙΝΔΡΟΜΗΣΗ 40
14.1. Μερικές μεταβλητές πρόβλεψης 40
14.2. Πίνακας Συσχέτισης 41
14.3. Μοντέλο Πολυμετάβλητης Παλινδρόμησης 43
15. ARIMA 44
16. BASS 46
17. ΣΥΜΠΕΡΑΣΜΑΤΑ 47
ΑΝΑΦΟΡΕΣ 48

ΚΑΤΑΛΟΓΟΣ ΣΧΗΜΑΤΩΝ

Σχήμα 1: Προβλέψεις με τη Μέθοδο Κινητού Μέσου όρου.....19

Σχήμα 2:Γραφική Παράσταση Χρονοσειράς.....31

ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ

Πίνακας 1: Προβλέψεις με τη μέθοδο Κινητου Μέσου όρου 18

Πίνακας 2: Πίνακας Συσχέτισης..... 42

Πίνακας 3: Πίνακας Συντελεστής Συσχέτισης για τα δεδομένα..... 42

ΠΡΟΛΟΓΟΣ

Οι διαδικασίες ανάλυσης και πρόβλεψης των αγορών αποτελούν δομικά τμήματα μιας επιχειρησιακής μελέτης, ενώ έχουν ευρείες εφαρμογές στη βιομηχανία, στην οικονομία του περιβάλλοντος, στην αγροτική οικονομία κλπ. Αναφορικά με την αγορά των τηλεπικοινωνιών, αποτελούν αναπόσπαστο τμήμα της ανάπτυξης επιχειρηματικών σχεδίων (business plans), καθοδηγώντας τη λήψη κρίσιμων αποφάσεων που αφορούν στα επενδυτικά σχέδια και στη διαδικασία προώθησης τηλεπικοινωνιακών προϊόντων και υπηρεσιών, σε ένα ιδιαίτερα ανταγωνιστικό περιβάλλον. Στο πλαίσιο αυτό, τα στελέχη που έχουν την ικανότητα να κάνουν χρήση της κατάλληλης πληροφορίας και των υπαρχόντων μεθοδολογιών ώστε να αναλύσουν την αγορά και να προβλέψουν τις μελλοντικές καταστάσεις της, υποστηρίζοντας της διαδικασία λήψης κρίσιμων αποφάσεων, τυγχάνουν υψηλής ζήτησης από τους οργανισμούς

<Πρόβλεψη>, μια λέξη που ακούγεται όλο και πιο συχνά στη σημερινή εποχή. Κατά ορισμό, πρόβλεψη είναι η διαδικασία της εκτίμησης ενός μεγέθους σε άγνωστες καταστάσεις. Καθ΄ όλη τη διάρκεια του μαθήματος θα χρησιμοποιήσουμε μεθόδους (μαθηματικές) για την πρόβλεψη τάσεων στην αγορά αβέβαιης φύσης.

Επίσης πρέπει να έχουμε πάντα στο μυαλό μας, ότι η κρίση μας πρέπει να χρησιμοποιείται παράλληλα με τα αριθμητικά αποτελέσματα, αν θέλουμε στο τέλος να έχουμε καλές προβλέψεις, αλλά και να χρησιμοποιήσουμε τις προβλέψεις με τον πιο ωφέλιμο τρόπο. Όλοι οι άνθρωποι από διαφορετική σκοπιά και για διαφορετικούς λόγους ο καθένας αναζητούν μια πρόβλεψη. Μια εκτίμηση που θα τους βοηθήσει να λάβουν μια απόφαση με μικρότερο ρίσκο. Μικρότερο ρίσκο στην δουλειά τους, σε μία επικείμενη αγορά, στον προγραμματισμό των διακοπών τους, στην επιχείρησή τους, στη ζωή τους.

1. ΕΙΣΑΓΩΓΗ

Κατά τη διάρκεια του 17 αιώνα οι προβλέψεις αποτελούσαν τη καινοτομία της εποχής και τις χρησιμοποιούσαν αρκετοί, οι οποίοι ήθελαν να κάνουν κάποια εκτίμηση του αποτελέσματος του εγχειρήματός τους. Οι ραγδαίες αλλαγές σε όλους τους τομείς της καθημερινής ζωής και η αβεβαιότητα για την εξέλιξη των καταστάσεων στο μέλλον, έχουν κεντρίσει το ενδιαφέρον τόσο του επιστημονικού–ακαδημαϊκού κόσμου, όσο και του επιχειρηματικού κόσμου. Από την πλευρά των ακαδημαϊκών δημιουργήθηκαν πολλές διαφορετικές μέθοδοι πρόβλεψης, κάποιες εξ΄ ολοκλήρου βασιζόμενες σε θεωρητικό υπόβαθρο και κάποιες άλλες βασιζόμενες περισσότερο στα διαθέσιμα τεχνολογικά μέσα και στις δυνατότητες, που αυτά διαθέτουν. Οι διάφορες μέθοδοι επιλέγονται και εφαρμόζονται ανάλογα με συγκεκριμένες παραμέτρους, μεταβλητές, ή παράγοντες, όπως είναι:

- Το αντικείμενο της πρόβλεψης.
- Ο χρονικός ορίζοντας της πρόβλεψης.
- Τα ήδη υπάρχοντα δεδομένα και το πρότυπο συμπεριφοράς αυτών.
- Το κόστος για την παραγωγή των προβλέψεων.
- Η ευκολία χρήσης της επιλεγμένης μεθόδου πρόβλεψης.
- Η αποδεδειγμένη αξιοπιστία της μεθόδου.

Στην πραγματικότητα αυτό σημαίνει ότι η αληθινή αξία μιας μεθόδου αποτιμάται, όταν εφαρμόζεται σε πραγματικές καταστάσεις. Έτσι, η ολοκληρωμένη εικόνα για μια μέθοδο δημιουργείται μόνο, όταν εφαρμοστεί η θεωρία στην πράξη. Μια συνεργασία λοιπόν ακαδημαϊκού και επιχειρηματικού κόσμου μπορεί να επιφέρει καλά αποτελέσματα. Και αυτό μπορεί να το καταλάβει πολύ απλά ο οποιοσδήποτε, πόσο καλύτερα μοντέλα πρόβλεψης μπορούν να προκύψουν, αν αυτά σχεδιαστούν βάσει των ιδιαίτερων αναγκών της εκάστοτε επιχείρησης. Από την πλευρά του επιχειρησιακού περιβάλλοντος είναι ολοφάνερη η σημαντικότητα της πρόβλεψης, αφού μια επιχείρηση βρίσκεται καθημερινά στην ανάγκη για λήψη μιας απόφασης αντιμετωπίζοντας, όσο το δυνατόν περισσότερο την αβεβαιότητα του μέλλοντος. Η πρόβλεψη διαδραματίζει πρωτεύοντα ρόλο σε όποια δραστηριότητα λήψης αποφάσεων, καθορισμού, σχεδιασμού και υλοποίησης στρατηγικής, χρονικού προγραμματισμού, καθώς και σχεδιασμού πολιτικής λειτουργίας της επιχείρησης. Είναι σημαντικό να τονίσουμε, ότι έστω και μία πολύ μικρή βελτίωση στην ακρίβεια των προβλέψεων μπορεί να αποφέρει τεράστια ποσά κέρδους σε μία επιχείρηση. Για αυτό το λόγο οι διάφορες υπηρεσίες πρόβλεψης και τα διάφορα προϊόντα για την παραγωγή προβλέψεων, που γενικά κυκλοφορούν στην αγορά, γνωρίζουν μεγάλη ζήτηση και ταυτόχρονα αναπτύσσονται νέες διαδικασίες πρόβλεψης, λόγω της ανάγκης για ακριβείς προβλέψεις και ιδιαίτερα σε επιχειρήσεις.

Για να μπορέσει όμως κάποιος να χρησιμοποιήσει προς όφελός του τις προβλέψεις, πρέπει να μπορεί να κάνει την πιο κατάλληλη επιλογή. Και αυτό είναι απόρροια της πείρας. Τα διοικητικά στελέχη μεγάλων επιχειρήσεων πρέπει να διαθέτουν την κατάλληλη εμπειρία στις καταστάσεις, που έχουν να αντιμετωπίσουν και οι αποφάσεις τους να μην επηρεάζονται από προσωπικά ενδιαφέροντα, φιλοδοξίες, σκοπιμότητες και μονόπλευρη εξέταση του περιβάλλοντος. Για να επιτευχθεί ο εκάστοτε στόχος πρέπει οι αποφάσεις τους να είναι όσο το δυνατόν πιο αντικειμενικές και να μην βασίζονται μόνο στα “πειραματικά”, ή μόνο στα θεωρητικά αποτελέσματα, αλλά να έχουν την ικανότητα να συγκρίνουν και να συνδυάζουν τα αποτελέσματα και των δύο και να λαμβάνουν υπόψη τους τόσο τις οικονομικές καταστάσεις του κλάδου τους όσο και των χωρών που θα ενεργοποιηθούν.

Σε κάποιες περιπτώσεις υπάρχει μεγάλη δυσαρέσκεια σχετικά με τα πολύ μεγάλα σφάλματα των προβλέψεων, όπως και στην αδυναμία των διαφόρων μεθόδων να προβλέψουν και να προειδοποιήσουν έγκαιρα για επερχόμενες αλλαγές και ανωμαλίες. Στην πραγματικότητα όμως το περιβάλλον, στο οποίο ζούμε, χαρακτηρίζεται από συνεχείς μεταβολές, εξελίξεις, μεγάλες διακυμάνσεις και απρόσμενες καταστάσεις και συνθήκες, που αυξάνουν την ανάγκη για προβλέψεις. Ο κύριος λόγος που υπάρχει αυτή η αιωρούμενη δυσανασχέτηση και οι δυσμενείς κριτικές για τις προβλέψεις, είναι ότι πολλοί πιστεύουν, ότι μπορούν να προβλεφθούν τα πάντα. Αυτό όμως καταλήγει να είναι ένας μύθος. Η ύπαρξη σφαλμάτων είναι αναπόφευκτη. Για αυτό το λόγο υπάρχουν αποκλίσεις μεταξύ προβλέψεων και πραγματικών αποτελεσμάτων. Το μόνο που μπορεί να γίνει, για να είναι πιο ωφέλιμα τα αποτελέσματα, είναι να γίνεται, όσο είναι εφικτό φυσικά, ρεαλιστική εκτίμηση των παραμέτρων. Θα ασχοληθούμε, λοιπόν, με τις μεθόδους που χρησιμοποιούνται για την πρόβλεψη των αβέβαιων επιχειρηματικών τάσεων, έτσι ώστε να βοηθήσουν το management να λάβει καλύτερες αποφάσεις. Οι μέθοδοι αυτές αφορούν συνήθως στη μελέτη ιστορικών στοιχείων και στην επεξεργασία τους και γίνονται προβλέψεις μέσω των κατάλληλων υποδειγμάτων. Κατά τη διάρκεια του μαθήματος θα αναφερθούν λοιπόν διάφοροι τρόποι εξαγωγής προβλέψεων, που βασίζονται σε λογικές μεθόδους διαχείρισης δεδομένων, τα οποία προέρχονται από ιστορικά γεγονότα. Η ανάλυση παλινδρόμησης και η αυτοπαλινδρόμηση είναι παραδείγματα διαδικασιών πρόβλεψης, που βασίζονται σε δεδομένα και θα συζητηθούν παρακάτω. Αυτές οι διαδικασίες έχουν αποδειχθεί ότι είναι πολύ αποτελεσματικές και συχνά εμφανίζονται σε διάφορα λογισμικά πακέτα.

Μαζί με την ανάπτυξη των μεθόδων, που βασίζονται στα δεδομένα, ο ρόλος της ανθρώπινης κρίσης έχει αυξηθεί σημαντικά κατά τα τελευταία 25 χρόνια. Χωρίς κανένα ιστορικό δεδομένο, η ανθρώπινη κρίση μπορεί να είναι ο μόνος τρόπος, για να γίνουν προβλέψεις για το μέλλον. Σε περιπτώσεις όπου τα δεδομένα είναι διαθέσιμα, η κρίση πρέπει να χρησιμοποιείται για επανεξέταση, αλλά και για να τροποποιεί τις προβλέψεις, που εξάγονται από ποσοτικές διαδικασίες. Με τον πολλαπλασιασμό των υπολογιστών και τη διαθεσιμότητα των εξελιγμένων λογισμικών πακέτων οι προβλέψεις εξάγονται πολύ εύκολα. Ωστόσο, αυτή η ευκολία των υπολογισμών δεν αντικαθιστά την σκέψη και τη κρίση. Η έλλειψη της εποπτείας και η καταχρηστική χρήση των τεχνικών πρόβλεψης μπορεί να οδηγήσει μια επιχείρηση σε δαπανηρές αποφάσεις . Από την μια μεριά, βρίσκεται το στέλεχος που λόγω άγνοιας και φόβου των ποσοτικών μεθόδων και των

υπολογιστών στηρίζεται αποκλειστικά στη διαίσθηση του. Από την άλλη μεριά, βρίσκεται η πρόγνωση του αναλυτή που βασίζεται στις τελευταίες εξελιγμένες τεχνικές, αλλά είτε δεν έχει την ικανότητα, ή την εμπειρία, ή είναι απρόθυμος να συνδυάσει τη διαδικασία πρόβλεψης με τις ανάγκες της επιχείρησης, ή και του κράτους ανάλογα με το αντικείμενο, ή και να κρίνει αν αυτό, που πρόβλεψε, μπορεί να είναι αληθινό ή όχι. Συνοψίζοντας λοιπόν ο πιο αποτελεσματικός αναλυτής πρέπει να είναι σε θέση να διαμορφώσει ένα συνδυασμό ποσοτικών και ποιοτικών δεδομένων, προκειμένου να αποφεύγονται μεγάλα σφάλματα. Θεωρούμε ότι οι ποσοτικές μέθοδοι πρόβλεψης, που θα συζητηθούν στη συνέχεια, είναι μόνο η αρχή για την αποτελεσματική πρόβλεψη των αποτελεσμάτων μιας επιχείρησης. Η ανάλυση, η ανθρώπινη κρίση, η κοινή λογική και η εμπειρία πρέπει να κατευθύνουν τη διαδικασία της πρόβλεψης.

Αν λάβουμε υπόψη μας ότι μία τράπεζα για να κάνει ρεαλιστικά πλάνα πρέπει να προβλέψει μελλοντικές ισορροπίες μεταξύ δανείων και καταθέσεων, ότι οι κεντρικές τράπεζες, για να ορίσουν τα επιτόκια, πρέπει να προβλέψουν μελλοντικές αυξήσεις ΑΕΠ, ή και υφέσεις, όπως και τις πληθωριστικές πιέσεις, ένας Operational Manager για να σχεδιάσει τη παραγωγή πρέπει να διαθέτει προβλέψεις πωλήσεων, ή μία εταιρεία για να ορίσει το προσωπικό της πρέπει να προβλέψει μελλοντικές πωλήσεις των υπηρεσιών της γίνεται εμφανές ότι οι προβλέψεις είναι αναγκαίες.

2. ΤΥΠΟΙ ΠΡΟΒΛΕΨΗΣ

Οι διαδικασίες πρόβλεψης μπορεί αρχικά να διαχωριστούν σε μακροπρόθεσμες, ή βραχυπρόθεσμες. Οι μακροπρόθεσμες προβλέψεις, είναι αναγκαίες για να χαράξουν την γενική πορεία και στρατηγική της επιχείρησης ή του κράτους. Συνεπώς, σ΄ αυτές δίνουν ιδιαίτερη έμφαση τα ανώτατα διοικητικά στελέχη. Οι βραχυπρόθεσμες προβλέψεις χρειάζονται για τη χάραξη άμεσων στρατηγικών και χρησιμοποιούνται από τα μεσαία στελέχη και τους μάνατζερ, έτσι ώστε να ανταποκριθούν στις ανάγκες του άμεσου μέλλοντος . Οι προβλέψεις μπορούν επίσης να ταξινομηθούν σε μίκρο και μάκρο. Για παράδειγμα, ένας διευθυντής μπορεί να ενδιαφέρεται να προβλέψει τον αριθμό των εργαζομένων, που απαιτούνται για τους επόμενους μήνες στην επιχείρηση του (μίκρο πρόβλεψη), ενώ η κυβέρνηση θέλει να προβλέψει το συνολικό αριθμό των ατόμων, που θα απασχοληθούν σε ολόκληρη τη χώρα (μάκρο πρόβλεψη). Οι διαδικασίες πρόβλεψης μπορούν επίσης να ταξινομηθούν ανάλογα με το αν είναι περισσότερο ποσοτικές, ή ποιοτικές. Από την μια μεριά, μία καθαρά ποιοτική τεχνική είναι μία, που δεν απαιτεί καμία εμφανή διαχείριση των δεδομένων. Μόνο η κρίση του αναλυτή απαιτείται. Από την άλλη μεριά, οι καθαρά ποσοτικές τεχνικές δεν απαιτούν ανθρώπινη κρίση. Είναι μηχανικές μαθηματικές διαδικασίες που καταλήγουν σε ποσοτικά αποτελέσματα. Ορισμένες ποσοτικές διαδικασίες, φυσικά, απαιτούν μια πολύ πιο εκλεπτυσμένη διαχείριση των στοιχείων από ότι άλλες.

3. ΜΕΘΟΔΟΣ ΝΑΪVE ΚΑΙ ΟΙ ΠΑΡΑΛΛΑΓΕΣ ΤΗΣ

Είναι η πιο απλή μέθοδος και αποτελεί σημείο αναφοράς για τις άλλες μεθόδους πρόβλεψης. Εφαρμόζεται σε περιπτώσεις όπου τα δεδομένα είναι απαλλαγμένα από τάση. Χρησιμοποιείται και ως benchmark για τις άλλες μεθόδους, που πρέπει να δίνουν αποτελέσματα πιο ακριβή από τη Naïve. Η Naïve κάνει πρόβλεψη για μια επόμενη τιμή και δίνει την τελευταία γνωστή παρατήρηση, γι' αυτό λέγεται και «no change forecast». Έχει καλές επιδόσεις για πρόβλεψη μιας περιόδου μπροστά, καθώς η τιμή πρόβλεψης δε διαφέρει σημαντικά από την τελευταία παρατήρηση.

3.1. Πρώτη παραλλαγή

Η μέθοδος Naïve περιγράφεται από την ακόλουθη εξίσωση:

$$Y_{t+1} = Y_t$$

όπου,

- Y_{t+1}: η πρόβλεψη για την περίοδο t + 1.
- Y_t : η παρατήρηση την χρονική περίοδο t.

3.2. Δεύτερη παραλλαγή

Στην τεχνική αυτή υπολογίζονται οι διαφορές πάνω στις πραγματικές τιμές και προστίθενται σε αυτές.

$$Y_{t+1} = Y_t + (Y_t - Y_{t-1})$$

3.3. Τρίτη παραλλαγή

Στη τεχνική αυτή χρησιμοποιείται ποσοστιαία μεταβολή και όχι οι διαφορές, όπως έγινε στην προηγούμενη.

$$Y_{t+1} = Y_t \frac{Y_t}{Y_{t-1}}$$

4. ΜΕΘΟΔΟΣ ΑΠΛΟΥ ΜΕΣΟΥ ΟΡΟΥ

Η πρόβλεψη χρονοσειρών με τη μέθοδο των μέσων εμπεριέχει τον υπολογισμό του μέσου όρου του δείγματος παρατηρήσεων, καθώς και τη χρησιμοποίηση αυτού του μέσου σαν πρόβλεψη για την επόμενη περίοδο. Ο αριθμός των παρατηρήσεων του δείγματος που συμπεριλαμβάνονται στον υπολογισμό του μέσου προσδιορίζεται στην αρχή της διαδικασίας πρόβλεψης. Αλγεβρικά ο μέσος όρος εκφράζεται:

$$Y_{t+1=\frac{1}{t}}\sum_{i=1}^{t} Y_i$$

Ο μέσος όρος προκύπτει από την μέθοδο των ελαχίστων τετραγώνων εφαρμοσμένη στο άθροισμα τετραγωνικών σφαλμάτων για σταθερά SSE=∑(Y −Y)^2. Ελαχιστοποιώντας αυτή τη σχέση παίρνουμε ως σημείο ελαχιστοποίησης τον μέσο όρο.Η μέθοδος πρόβλεψης του απλού μέσου λειτουργεί με καλά αποτελέσματα, όταν οι συνθήκες που παράγουν τις χρονοσειρές έχουν σταθεροποιηθεί και το περιβάλλον παραμένει αμετάβλητο. Ο απλός μέσος χρησιμοποιεί όλα τα παρελθόντα δεδομένα, ώστε να γίνει πρόβλεψη για την επόμενη περίοδο.

5. ΜΕΘΟΔΟΣ ΑΠΛΟΥ ΚΙΝΗΤΟΥ ΜΕΣΟΥ ΟΡΟΥ

Χρησιμοποιούμε τον κινητό μέσο στην περίπτωση που θέλουμε να συγκεντρωθούμε περισσότερο στα πρόσφατα γεγονότα. Σε αυτή τη διαδικασία, καθώς μια νέα παρατήρηση γίνεται διαθέσιμη, ένας νέος μέσος όρος μπορεί να υπολογιστεί, στον οποίο παραλείπεται η πιο παλιά παρατήρηση, προκειμένου να συμπεριληφθεί η πιο πρόσφατη.

Η μέθοδος του κινητού μέσου όρου είναι η βασικότερη μέθοδος πρόβλεψης και χρησιμοποιείται σε περιπτώσεις που η ζήτηση δεν παρουσιάζει διακυμάνσεις και η εποχικότητα δε λαμβάνεται υπόψη. Αρχικά υπολογίζεται ο μέσος όρος της ζήτησης για ένα συγκεκριμένο αριθμό περιόδων *t*. Ο εν λόγω μέσος όρος χρησιμοποιείται για την πρόβλεψη της ζήτησης της αμέσως επομένης περιόδου *t*+1. Εν συνεχεία, για την πρόβλεψη της ζήτησης της επομένης περιόδου υπολογίζεται ο μέσος όρος της ζήτησης των προηγούμενων *t* περιόδων συνυπολογίζοντας σε αυτές την πρόβλεψη της προηγούμενης περιόδου. Με αυτόν τον τρόπο σε κάθε βήμα προστίθεται μια νέα πρόβλεψη και αφαιρείται η παλαιότερη. Για αυτό το λόγο η μέθοδος ονομάζεται «κινητός» μέσος όρος.

$$F_{t+1} = \frac{D_t + D_{t-1} + \ldots + D_{t-n+1}}{n}$$

Όπου F_{t+1} η πρόβλεψη της ζήτησης για την περίοδο *t*+1, D_t η πραγματική ζήτηση περιόδου *t* και *n* συνολικός αριθμός περιόδων που συμμετέχουν στο μέσο όρο.

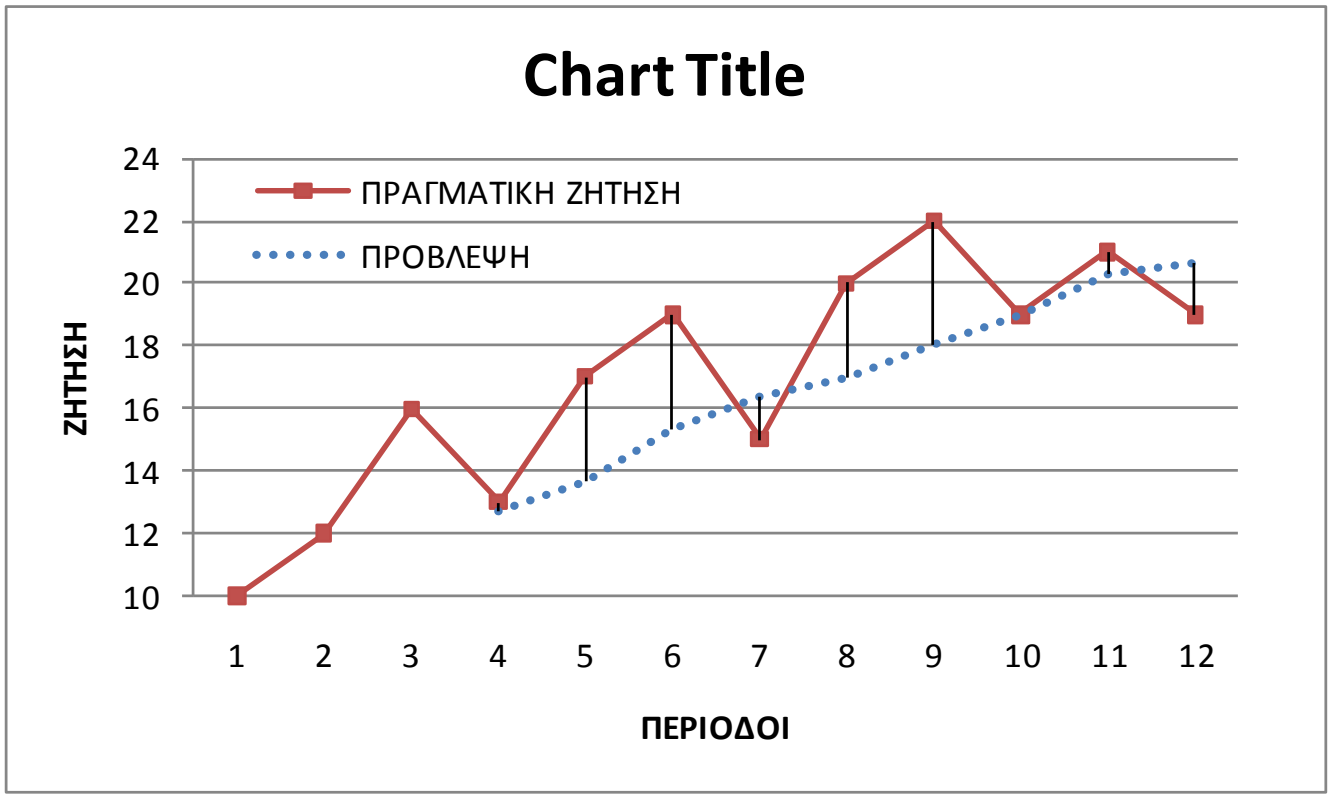
Πινάκας 1: Προβλέψεις με τη μέθοδο Κινητού Μέσου Όρου

Περίοδος	*Ζήτηση*	*Πρόβλεψη*
Ιανουάριος	10	
Φεβρουάριος	12	
Μάρτιος	16	
Απρίλιος	13	12,67
Μάιος	17	13,67
Ιούνιος	19	15,33
Ιούλιος	15	16,33
Αύγουστος	20	17,00
Σεπτέμβρι	22	18,00

ος Οκτώβριος	19	19,00
Νοέμβριος	21	20,33
Δεκέμβριος	19	20,67

Επίδοση Μοντέλου Κινητού Μέσου Όρου:

MAD=2,000
MSE=6,074
MAPE=10,61%



Σχημα 1: Προβλέψεις με τη μέθοδο Κινητού Μέσου Όρου

.

Τη μέθοδο αυτή τη χρησιμοποιούμε, όταν έχουμε τάση, δίνοντας καλύτερα αποτελέσματα από τον απλό μέσο όρο, χωρίς αυτό να σημαίνει ότι είναι και η πλέον

κατάλληλη μέθοδος για πρόβλεψη με τάση, γιατί δεν λαμβάνει υπόψη του την επίδραση του κάθε όρου ξεχωριστά αλλά θεωρεί την βαρύτητα του κάθε όρου στη σχέση σταθερή.

6. ΜΕΘΟΔΟΣ ΔΙΠΛΟΥ ΚΙΝΗΤΟΥ ΜΕΣΟΥ ΟΡΟΥ

Για να προβλέψουμε νέους όρους σε χρονοσειρές, που διαθέτουν γραμμική τάση, όπως η δική μας, χρησιμοποιούμε το διπλό κινητό μέσο. Επειδή με τη μέθοδο των ελαχίστων τετραγώνων κάθε φορά που εμφανίζονται νέες τιμές στις χρονοσειρές χρειάζεται να επανεκτιμήσουμε τις παραμέτρους του γραμμικού μοντέλου για να προχωρήσουμε σε νέες προβλέψεις, δεν είναι αναγκαίο να πάρουμε όλες τις προηγούμενες τιμές του μακρινού παρελθόντος, γιατί δεν έχουν κάποια σχέση με το παρόν. Οπότε προσπαθούμε να εκτιμήσουμε (προσαρμόσουμε) τη καμπύλη με ευθύγραμμα τμήματα.

Η μεθοδολογία που ακολουθούμε είναι η ακόλουθη. Στην αρχή παίρνουμε τον πρώτο κινητό μέσο και μετά το κινητό μέσο των πρώτων κινητών μέσων πάντα με τον ίδιο αριθμό όρων. Οι τιμές του πρώτου κινητού μέσου είναι μικρότερες από τις πραγματικές τιμές, ενώ οι τιμές του διπλού κινητού μέσου είναι μικρότερες και από

αυτές του πρώτου κινητού μέσου. Στο τέλος με τη βοήθεια μιας ευθείας κάνουμε την πρόβλεψη για p περιόδους μπροστά.

Η πρόβλεψη για την $t + p$ περίοδο, δίνεται από την παρακάτω σχέση.

$$Y_{t+p} = a_t + b_t p$$

με, $$a_t = M_t + (M_t + M'_t) = 2M_t + M'_t$$

και $$b_t = \frac{2}{k-1}(M_t - M'_t)$$

όπου, M_t = ο πρώτος κινητός μέσος.
M'_t= ο δεύτερος κινητός μέσος.
a_t = ο πρώτος κινητός μέσος συν την διαφορά του με το διπλό.
b_t = η κλίση της ευθείας πρόβλεψης από σημείο σε σημείο.

Οι τιμές του πρώτου και δεύτερου κινητού μέσου δίνονται από τις παρακάτω σχέσεις.

$$M_t = \frac{Y_t + Y_{t-1} + \cdots + Y_{t-k+1}}{k}$$

$$M'_t = \frac{M_t + M_{t-1} + \cdots + M_{t-k+1}}{k}$$

7. ΜΕΘΟΔΟΣ ΑΠΛΗΣ ΕΚΘΕΤΙΚΗΣ ΕΞΟΜΑΛΥΝΣΗΣ

Ενώ η μέθοδος του κινητού μέσου όρου λαμβάνει υπόψη της μόνο τις τελευταίες παρατηρήσεις, η μέθοδος της εκθετικής εξομάλυνσης (exponential smoothing) παίρνει υπόψη της όλες τις τιμές με διαφορετικό όμως βάρος. Η εξίσωση της απλής εκθετικής εξομάλυνσης (simple exponential smoothing) είναι:

$$Y_{t+1} = aY_t + (1-a)Y_t$$

Το α ονομάζεται σταθερά εξομάλυνσης και παίρνει τιμές από 0 έως 1. Η εκθετική εξομάλυνση μιας παραμέτρου είναι πολύ απλή μέθοδος, αφού μόνο μια τιμή, η πρόβλεψη της τελευταίας περιόδου, είναι αυτή που πρέπει να διασωθεί. Στην ουσία, ολόκληρη η χρονοσειρά εμπεριέχεται σ' αυτή την πρόβλεψη. Εάν εκφράσουμε το Y_t σε όρους της προηγηθείσας παρατήρησης Y_{t-1} και των τιμών της πρόβλεψης Y_{t-1}, τότε το ισοδύναμο για την πρόβλεψη της επόμενης περιόδου γίνεται:

$$Y_{t+1} = aY_t + (1-a)[aY_{t-1} + (1-a)Y_{t-1}]$$

ή πιο απλά,

$$Y_{t+1} = aY_t + a(1-a)Y_{t-1} + (1-a)^2 Y_{t-1}$$

Η νέα αυτή εξίσωση είναι μοντέλο δευτεροβάθμιας εκθετικής εξομάλυνσης μιας παραμέτρου. το όνομα αυτής της διαδικασίας προέρχεται από τις διαδοχικές σταθμίσεις $a, a(1-a), a(1-a)^2, a(1-a)^3$,..., οι οποίες μειώνονται εκθετικά. Η διαδικασία πρόβλεψης μπορεί να τροποποιηθεί οποιαδήποτε στιγμή με τη μεταβολή της τιμής της α. Μπορούμε να ξαναγράψουμε την εξίσωση ως εξής:

$$Y_{t+1} = Y_t + a(Y_t - Y_t) = Y_t + ae_t$$

Δηλαδή η πρόβλεψη στο t+1 είναι η πρόβλεψη στο t συν έναν παράγοντα επί της διαφοράς της πραγματικής τιμής μείον την τιμή της πρόβλεψης στο t, όπου το αet το σφάλμα πρόβλεψης για την περίοδο t. Επομένως, βλέπουμε ότι η πρόβλεψη που δίνεται από την εκθετική εξομάλυνση, είναι η παλαιά πρόβλεψη συν μια προσαρμογή για το σφάλμα, που έγινε στην τελευταία πρόβλεψη. Όταν το α βρίσκεται πλησίον του 1, η νέα πρόβλεψη περιέχει μια ουσιώδη προσαρμογή για το σφάλμα της προηγούμενης πρόβλεψης. Αντίθετα, εάν το α βρίσκεται πολύ κοντά στο 0, η νέα πρόβλεψη θα περιέχει μικρή μόνο προσαρμογή για το σφάλμα.

8. ΜΕΘΟΔΟΣ ΓΡΑΜΜΙΚΗΣ ΤΑΣΗΣ (Holt Exponential Smoothing)

Το μοντέλο εξομάλυνσης για γραμμική τάση (Holt 1957) είναι μια επέκταση της απλής εκθετικής που λαμβάνει υπόψη του την ύπαρξη τάσης στα δεδομένα και αποτελεί την πιο δημοφιλή παραδοχή. Οι προβλέψεις προκύπτουν απλά από την προέκταση μιας ευθείας γραμμής για οποιαδήποτε χρονική στιγμή στο μέλλον. Το μοντέλο περιγράφεται από τις εξισώσεις:

$e_t = Yt - Ft$

$S_t = S_t\text{-}1 + T_t\text{-}1 + \alpha \cdot e_t$

$T_t = T_t\text{-}1 + \beta \cdot e_t$

$F_t+m = S_t + m \cdot T_t$

Στις παραπάνω εξισώσεις et είναι το σφάλμα της περιόδου t, St το επίπεδο για την περίοδο t, Tt η τάση για την περίοδο t και Ft η πρόβλεψη για την περίοδο t. Η παράμετρος α είναι ο συντελεστής εξομάλυνσης του επιπέδου, ενώ η παράμετρος β ο συντελεστής εξομάλυνσης της τάσης και λαμβάνουν τιμές εντός του διαστήματος [0,1]. Με m συμβολίζεται ο χρονικός ορίζοντας της πρόβλεψης.

Όπως και στην περίπτωση της μεθόδου SES, έτσι και σε αυτή την περίπτωση η επιλογή του βέλτιστου συνδυασμού τιμών για τις παραμέτρους α και β βασίζεται στην ελαχιστοποίηση του μέσου τετραγωνικού σφάλματος (MSE). Συνήθως η βέλτιστη τιμή του συντελεστή β για την τάση είναι μικρότερη από την τιμή του συντελεστή α για το επίπεδο, κάτι που συμβαίνει διότι η τιμή της τάσης είναι μικρότερη από εκείνη του επιπέδου για κάθε περίοδο. Για την έναρξη της μεθόδου απαραίτητα είναι το αρχικό επίπεδο (S0) και η αρχική τάση (T0), η επιλογή των οποίων είναι ιδιαίτερα σημαντική για την ακρίβεια των παραγόμενων προβλέψεων. Το αρχικό επίπεδο υπολογίζεται όπως και στην απλή εκθετική εξομάλυνση.

Ως αρχική τάση μπορούν να χρησιμοποιηθούν:

- Η διαφορά της δεύτερης από την πρώτη παρατήρηση (Y2-Y1)
- Η διαφορά της v-οστής παρατήρησης από την πρώτη διαιρεμένη με v-1
- Η σταθερά της κλίσης από το μοντέλο απλής γραμμικής παλινδρόμησης

9. ΜΕΘΟΔΟΣ ΜΗ ΓΡΑΜΜΙΚΗΣ ΤΑΣΗΣ

Το μοντέλο μη γραμμικής τάσης (Gardner και McKenzie 1985) αποτελεί μια προσαρμογή του μοντέλου γραμμικής τάσης που χρησιμοποιείται σε περιπτώσεις μη γραμμικών τάσεων. Αυτό επιτυγχάνεται με την προσθήκη της παραμέτρου διόρθωσης της τάσης φ, η οποία ελέγχει το ρυθμό αύξησης των τιμών της τάσης σε μια χρονοσειρά. Οι εξισώσεις που περιγράφουν το μοντέλο μη γραμμικής τάσης είναι οι εξής:

$e_\tau = Y_t - F_t$

$S_t = S_t\text{-}1 + T_t\text{-}1 + \alpha \cdot e_t$

$T_t = T_t\text{-}1 + \beta \cdot e$

$F_{t+m} = S_t + \Sigma_i{}^1{}_{=1}\, \varphi^i \, {}^*T_t$

Όπως γίνεται φανερό οι εξισώσεις είναι ίδιες με εκείνες της μεθόδου Holt, πλην της τελευταίας, όπου αντί για τον υπολογισμό μιας γραμμικής αύξησης της τάσης μέσω του συντελεστή m, γίνεται ένας μη γραμμικός υπολογισμός της, με τη χρήση της παραμέτρου εξομάλυνσης φ. Η παράμετρος φ (σε αντίθεση με τις παραμέτρους α και β) μπορεί να λάβει και τιμές μεγαλύτερες της μονάδας. Για τις διάφορες τιμές της παραμέτρου φ το μοντέλο μη γραμμικής τάσης μπορεί να πάρει τις ακόλουθες μορφές:

- Για φ=0 προκύπτει το μοντέλο της απλής εκθετικής εξομάλυνσης, εφόσον η τάση δεν επηρεάζει τον καθορισμό των στατιστικών σημειακών προβλέψεων.
- Για 0<φ<1 προκύπτει το μοντέλο της φθίνουσας τάσης (damped exponential smoothing). Το μοντέλο αυτό χρησιμοποιείται κυρίως στην παραγωγή μεσοπρόθεσμων προβλέψεων, καθώς χαρακτηρίζεται από έλλειψη τάσης για υπεραισιοδοξία.
- Για φ=1 προκύπτει το μοντέλο γραμμικής τάσης, καθώς τη θέση του αθροίσματος παίρνει το γινόμενο m · Tt

- Για φ>1 προκύπτει το μοντέλο της εκθετικής τάσης, το οποίο χαρακτηρίζεται από μεγάλη θετική προκατάληψη και χρησιμοποιείται σε περιπτώσεις που ζητούμενο είναι η πρόβλεψη ζήτησης στην αρχή του κύκλου ζωής ενός προϊόντος ή μιας υπηρεσίας.

Η επιλογή του συντελεστή εξομάλυνσης φ είναι ιδιαίτερα σημαντική για την ακρίβεια των προβλέψεων. Συνήθως περιορίζεται στο διάστημα [0,1], κάτι που αποτρέπει την εσφαλμένη επιλογή του, που οδηγεί σε υπεραισιόδοξες προβλέψεις. Σε αυτή την

περίπτωση προκύπτει η μέθοδος DES (Damped Exponential Smoothing). Έπειτα, μπορεί να προσδιορισθεί ο βέλτιστος συνδυασμός των α, β και φ με βάση την ελαχιστοποίηση του μέσου τετραγωνικού σφάλματος.

Εμπειρικές μελέτες έχουν δείξει ότι το μοντέλο μη γραμμικής τάσης υπερτερεί του μοντέλου γραμμικής τάσης, καθώς παράγει προβλέψεις καλύτερης ακρίβειας. Το μοντέλο μη γραμμικής τάσης πλεονεκτεί έναντι των υπολοίπων μοντέλων εκθετικής εξομάλυνσης κυρίως στην εξαγωγή προβλέψεων μεγάλου χρονικού ορίζοντα, εφόσον η επιλογή κατάλληλης παραμέτρου φ του δίνει τη δυνατότητα να προσαρμόζεται καλύτερα στη φύση των εκάστοτε δεδομένων.

10. ΜΕΘΟΔΟΣ ΔΙΠΛΗΣ ΕΚΘΕΤΙΚΗΣ ΕΞΟΜΑΛΥΝΣΗΣ

Η μέθοδος αυτή ονομάζεται εκθετική εξομάλυνση με προσαρμογή στην τάση και χρησιμοποιείται όταν υπάρχει τάση στις παρατηρήσεις της χρονοσειράς, κάτι που συμβαίνει και με την δική μας χρονοσειρά. Έχει δύο παραμέτρους εξομάλυνσης, την παράμετρο α για την εξομάλυνση του επιπέδου των τιμών της χρονοσειράς και την παράμετρο β για την εξομάλυνση της τάσης, σε αντίθεση με τη μέθοδο της απλής εκθετικής εξομάλυνσης που έχει μόνο μία.

Η εξομάλυνση του επιπέδου των τιμών της χρονοσειράς γίνεται με την ακόλουθη σχέση:

$$L_t = aY_t + (1-a)L_{t-1} + T_{t-1}$$

όπου α είναι η σταθερά για την εξομάλυνση για 0≤α≤1, Lt οι εξομαλυνθείσες τιμές της χρονοσειράς, που προκύπτουν από την εξομάλυνση, για t=2, 3, ..., n, ενώ για t=1 ορίζεται ως αρχική συνθήκη L1=Y1. Η σχέση αυτή είναι όμοια με τη σχέση της απλής εκθετικής εξομάλυνσης εκτός από τον παράγοντα T t−1 , ο οποίος εισήχθη προκειμένου να ενσωματώσουμε τη τάση στη περίπτωση που υπάρχει.

Χωρίς τον παράγοντα αυτό η σχέση είναι ίδια με τη σχέση της απλής εκθετικής εξομάλυνσης. Το Level είναι η εξομαλυνθείσα τιμή των δεδομένων στο τέλος μιας περιόδου.

Η εξομάλυνση της τάσης γίνεται ως εξής:

$$T_t = \beta(L_t - L_{t-1}) + (1-\beta)T_{t-1}$$

όπου β, για 0≤β≤1, είναι η σταθερά για την εξομάλυνση της τάσης, Tt οι εξομαλυνθείσες τιμές της τάσης, για t=2, 3, ..., n, ενώ για t=1 ορίζεται ως αρχική συνθήκη η T1=0. Η παρούσα τάση σταθμίζεται με δύο παράγοντες: Ο πρώτος προέρχεται από τη διαδοχική διαφορά των εξομαλυνθέντων τιμών της χρονοσειράς ενώ ο δεύτερος από την εξομαλυνθείσα τιμή τάσης της προηγούμενης περιόδου. Το Trend λοιπόν είναι η εξομαλυνθείσα τιμή της αύξησης, ή της μείωσης των δεδομένων στο τέλος μιας περιόδου.

Η πρόβλεψη για την p μελλοντική περίοδο υπολογίζεται ως:

$$Y_{t+p} = L_t + pT_t$$

όπου p = 1, 2, 3, ….

Οι τιμές των παραμέτρων α και β για μια συγκεκριμένη χρονοσειρά είναι αυτές, που ελαχιστοποιούν την τιμή του κριτηρίου MSE, ή κάποιου άλλου κριτηρίου στα δεδομένα της χρονοσειράς. Οι σχέσεις αυτές προκύπτουν αν πάρουμε την αναμενόμενη τιμή της εξομαλυνθείσας τιμής στην απλή εκθετική εξομάλυνση και μετά ορίσουμε δεύτερη εξομάλυνση στην αναμενόμενη τιμή της πρώτης

εξομάλυνσης.

11. ΜΕΘΟΔΟΣ ΤΡΙΠΛΗΣ ΕΚΘΕΤΙΚΗΣ ΕΞΟΜΑΛΥΝΣΗΣ

Ένα βήμα παρακάτω πηγαίνει η μέθοδος του Winters, ο οποίος ανέπτυξε μια μέθοδο για την προσαρμογή της εποχικής, ή περιοδικής κίνησης μέσα στο πλαίσιο

της γραμμικής εκθετικής εξομάλυνσης με ή χωρίς τάση.

Για την ανάπτυξη της μεθόδου αυτής θεωρούμε το Multiplicative Model για την αναπαράσταση των δεδομένων μέσω των συνιστωσών.

$$Y = L\ x\ S$$

Μέσα στο L “κρύβεται” η τάση, η κυκλικότητα και η τυχαιότητα.

Προκειμένου να κάνουμε πρόβλεψη με αυτό το μοντέλο χρειαζόμαστε τέσσερις εξισώσεις:

Η εκθετικά εξομαλυνθείσα σειρά είναι η εξής:

$$L_t = a\frac{Y_t}{S_{t-s}} + (1-a)(L_{t-1} + T_{t-1})$$

Η εκτίμηση της εποχικότητας είναι η εξής:

$$S_t = \gamma\frac{Y_t}{L_t} + (1-\gamma)S_{t-s}$$

όπου S είναι ο παράγοντας προσαρμογής της εποχικότητας και το s είναι το μήκος της εποχικότητας.

Ο εκτίμηση της τάσης παραμένει η ίδια:

$$T_t = \beta(L_t - L_{t-1}) + (1-\beta)T_{t-1}$$

Προβλέποντας p περιόδους στο μέλλον, έχουμε:

$$Y_{t+p} = (L_t + pT_t)S_{t-s+p}$$

Με την πρώτη εξίσωση γίνονται επίκαιρες οι εξομαλυνθείσες τιμές της σειράς. Το L δεν εμπεριέχει την εποχικότητα. Στην εξίσωση αυτή το Yt διαιρείται δια του St-s που προσαρμόζει τις αρχικές παρατηρήσεις, Yt για εποχικότητα και αναιρεί τις επιδράσεις της εποχικότητας, όσο καλύτερα αυτές μπορεί να μετρηθούν από τη χρονολογική σειρά. Η δεύτερη εξίσωση δίνει την εκτίμηση της εποχικής συνιστώσας, Yt/Lt πολλαπλασιασμένη επί τη σταθερά γ συν την παλαιά εποχική

εκτίμηση, St-s πολλαπλασιασμένη επί (1–γ). Επομένως, η επικαιροποίηση των εποχικών εκτιμήσεων είναι από μόνη της μια διαδικασία εκθετικής εξομάλυνσης. Επίσης, ο Yt διαιρείται με Lt, προκειμένου να εκφραστεί η τιμή ως δείκτης παρά ως απόλυτο μέγεθος. Αυτό επιτρέπει την εύρεση του μέσου όρου των νέων εποχικών εκτιμήσεων με βάση τον εποχικό δείκτη της προηγούμενης περιόδου.

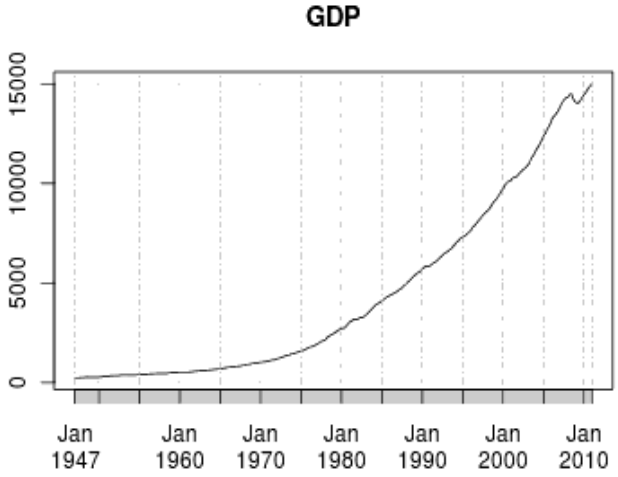
Η τρίτη εξίσωση εκφράζει τη σύγχρονη τιμή της συνιστώσας της τάσης, που επιτυγχάνεται με τη συνηθισμένη διαδικασία εκθετικής εξομάλυνσης. Τέλος, μετά από αυτή την εξίσωση λαμβάνουμε την εξίσωση για τις μελλοντικές περιόδους. Η διαφορά είναι ότι αυτή η εκτίμηση για τη μελλοντική περίοδο, t+p, πολλαπλασιάζεται επί St-s+p. Αυτός είναι ο τελικός διαθέσιμος εποχικός δείκτης και αποτελεί την προσαρμογή της πρόβλεψης για εποχικότητα. Οι αρχικές τιμές είναι S1 = 1.0, T1 = 0, L1 = Y1. Οι παράμετροι α,β,γ ορίζονται με τέτοιο τρόπο, ώστε το MSE να γίνει ελάχιστο κάτι που σ' αυτή την περίπτωση είναι πολύ

χρονοβόρο.

12. ΔΙΑΣΠΑΣΗ ΧΡΟΝΟΣΕΙΡΩΝ

Οι χαρακτηριστικές κινήσεις μίας χρονοσειράς μπορούν να διακριθούν σε τέσσερα κύρια είδη, τα οποία συχνά ονομάζονται συνιστώσες (components) της χρονοσειράς. Οι κινήσεις αυτές είναι οι μακροπρόθεσμες ή κύριες κινήσεις, οι κυκλικές κινήσεις ή μεταβολές, οι εποχικές κινήσεις ή μεταβολές και οι ακανόνιστες ή τυχαίες κινήσεις. Η διάσπαση χρονοσειρών (time series decomposition) στηρίζεται στην υπόθεση ότι οι τιμές μιας χρονοσειράς σχηματίζονται από τις παραπάνω συνιστώσες που τη συνθέτουν. Για τη δημιουργία των προβλέψεων με τη μέθοδο αυτή, η χρονοσειρά διασπάται στις ανωτέρω τέσσερις συνιστώσες και προσδιορίζεται η επιρροή που έχει καθένα από αυτά στη διαμόρφωση των τιμών της μεταβλητής. Οι μακροπρόθεσμες ή κύριες κινήσεις ή τάση αναφέρονται στη γενική κατεύθυνση που φαίνεται ότι ακολουθεί η γραφική παράσταση μίας χρονοσειράς κατά μία μεγάλη διάρκεια χρόνου. Σε πολλές περιπτώσεις οι τιμές των παρατηρήσεων ορισμένων χρονοσειρών τείνουν να αυξάνονται ή να μειώνονται με αρκετά σταθερό ρυθμό για μεγάλα χρονικά διαστήματα. Η συμπεριφορά αυτή εκφράζεται από την τάση που φανερώνει τη μακροχρόνια εξέλιξη της χρονοσειράς, η οποία μπορεί να είναι ανοδική ή καθοδική. Η τάση οφείλεται συνήθως σε πληθυσμιακές αλλαγές, σε τεχνολογικές αλλαγές, σε οικονομικούς παράγοντες, όπως π.χ. στον πληθωρισμό, στην αύξηση της παραγωγικότητας κ.α. Οι κυκλικές κινήσεις ή κυκλικότητα αναφέρονται με μακροπρόθεσμες ταλαντώσεις γύρω από τη γραμμή ή καμπύλη τάσης. Η κυκλικότητα εμφανίζεται ακανόνιστα με κυματοειδή μορφή και διαρκεί για χρονικό διάστημα πολύ μεγαλύτερο του έτους. Η συμπεριφορά αυτή των τιμών των χρονοσειρών αποδίδεται κυρίως στους οικονομικούς κύκλους, οι οποίοι οφείλονται σε μεταβαλλόμενες οικονομικές, τεχνολογικές και άλλες συνθήκες. Επειδή όμως οι οικονομικοί κύκλοι δεν εμφανίζονται με την ίδια περιοδικότητα ή και την ίδια μορφή, για το λόγο αυτό το στοιχείο της κυκλικότητας, σε αντίθεση με την τάση και την εποχικότητα, δεν θεωρείται ότι συμβάλλει άμεσα στη δημιουργία προβλέψεων. Ωστόσο, η κυκλικότητα μπορεί να χρησιμοποιηθεί για να προσδιοριστεί η μέχρι τώρα εξέλιξη των τιμών της χρονοσειράς. Οι εποχικές κινήσεις ή εποχικότητα αναφέρονται στην ταυτόσημη ή σχεδόν ταυτόσημη εξέλιξη που έχει μία χρονοσειρά κατά τη διάρκεια κάποιων συγκεκριμένων μηνών ή τριμήνων διαδοχικών ετών. Η εποχικότητα οφείλεται σε επαναλαμβανόμενα γεγονότα. Τα δεδομένα ορισμένων χρονοσειρών αναφέρονται σε χρονικές περιόδους μικρότερες του έτους, όπως μήνες ή τρίμηνα, με αποτέλεσμα να παρατηρούνται εποχικές διακυμάνσεις, οι οποίες εμφανίζονται κατά τη διάρκεια του έτους και επαναλαμβάνονται με την ίδια ή περίπου την ίδια μορφή από έτος σε έτος.Για παράδειγμα η μηνιαία κατανάλωση παγωτού είναι μεγαλύτερη κατά την καλοκαιρινή περίοδο και μικρότερη κατά την χειμερινή περίοδο. Γενικά, το φαινόμενο της εποχικότητας οφείλεται κυρίως σε μεταβολές του καιρού, σε πολιτικές της διοίκησης αναφορικά με περιόδους εκπτώσεων, καθώς και σε άλλους παράγοντες όπως θρησκευτικούς, κοινωνικούς κ.α. Οι εποχικές διακυμάνσεις, επειδή παρουσιάζονται με συστηματικό τρόπο συνήθως, μπορούν εύκολα να αναλυθούν και να προσδιοριστούν και κατά συνέπεια να χρησιμοποιηθούν για την πρόβλεψη των μελλοντικών τιμών της χρονοσειράς, κάτι που συμβαίνει άλλωστε και με την τάση.Οι ακανόνιστες ή τυχαίες κινήσεις αναφέρονται στις σποραδικές, ακανόνιστες (irregular) κινήσεις μιας χρονοσειράς λόγω τυχαίων παραγόντων και γεγονότων. Οι τυχαίες κινήσεις επηρεάζουν τις τιμές των χρονοσειρών κατά ένα τυχαίο και μη συστηματικό τρόπο, ο οποίος δεν μπορεί να προσδιοριστεί. Η συνιστώσα αυτή λοιπόν δεν είναι δυνατόν να χρησιμοποιηθεί στη διαμόρφωση των μελλοντικών τιμών των χρονοσειρών. Οι τυχαίες

κινήσεις οφείλονται σε όλους εκείνους τους τυχαίους και απρόσμενους παράγοντες που επηρεάζουν τις τιμές των χρονοσειρών και οι οποίοι δεν προσδιορίζονται από την τάση, την εποχικότητα και την κυκλικότητα. Οι παράγοντες αυτοί μπορεί να είναι πόλεμοι, σεισμοί, απρόσμενες καιρικές μεταβολές, απεργίες, διαδόσεις για συγκεκριμένο προϊόν, αιφνίδιες μεταβολές στις προτιμήσεις των καταναλωτών, απρόσμενες αλλαγές στη νομοθεσία κ.α. Σε μία συγκεκριμένη χρονοσειρά είναι δυνατόν να μην συνυπάρχουν και οι τέσσερις συνιστώσες αλλά μόνο κάποιες από αυτές. Η ανάλυση χρονοσειρών συνίσταται στην περιγραφή (εν γένει με μαθηματικό τρόπο) των συνιστωσών κινήσεων που υπάρχουν. Η γραφική παράσταση μιας χρονοσειράς απεικονίζεται στο παρακάτω διάγραμμα. Τα δεδομένα προέρχονται από τις μηνιαίες πωλήσεις μιας Χ εταιρείας για το χρονικό διάστημα 2000-2005. Οι πωλήσεις είναι σε χιλιάδες ευρώ.Στην παρακάτω χρονοσειρά συνυπάρχουν και οι τέσσερις συνιστώσες.



Σχημα 2: Γραφική παράσταση χρονοσειράς

Για την ανάλυση των χρονοσειρών χρησιμοποιούμε τους ακόλουθους συμβολισμούς:

Yt = Πραγματική τιμή της χρονοσειράς

Tt = Τάση

St = Εποχικότητα

Ct = Κυκλικότητα

It = Τυχαίες κινήσεις

όπου t = 1,2,3...,n.

Η εξέταση των στοιχείων αυτών γίνεται σύμφωνα με κάποιο μαθηματικό υπόδειγμα που

φανερώνει τον τρόπο με τον οποίο οι παρατηρήσεις της χρονοσειράς προσδιορίζονται από τις συνιστώσες της χρονοσειράς. Τα χρησιμοποιούμεναυποδείγματα είναι το προσθετικό μοντέλο (additive model) και το πολλαπλασιαστικό μοντέλο (multiplicative model).Στο προσθετικό μοντέλο οι πραγματικές τιμές της χρονοσειράς για κάθε περίοδο θεωρούνται ως το άθροισμα των τεσσάρων συνιστωσών και δημιουργούνται με τον ακόλουθο τρόπο:

Yt = Tt + St + Ct + It

Από την παραπάνω σχέση είναι φανερό ότι όλες οι συνιστώσες είναι εκφρασμένα στην ίδια μονάδα μέτρησης με εκείνη των παρατηρήσεων της χρονοσειράς. Αντίθετα στο πολλαπλασιαστικό μοντέλο οι πραγματικές τιμές της χρονοσειράς προσδιορίζονται από το γινόμενο των τεσσάρων συνιστωσών, δηλαδή ως ακολούθως:

Yt = Tt St Ct It

Στο μοντέλο αυτό μόνο η τάση είναι εκφρασμένη στην ίδια μονάδα μέτρησης με εκείνη της χρονοσειράς Yt ενώ τα στοιχεία Ct , St και It είναι δείκτες ανεξάρτητοι από μονάδες μέτρησης.Από τα δύο παραπάνω μοντέλα το προσθετικό μοντέλο χρησιμοποιείται λιγότερο συχνά στην πράξη, επειδή είναι δύσκολο στην ανάλυση του για υπολογιστικούς κυρίως λόγους. Επίσης βασίζεται στην υπόθεση ότι οι συνιστώσες της χρονοσειράς είναι ανεξάρτητες μεταξύ τους, που σημαίνει για παράδειγμα, ότι η τάση δεν επηρεάζει την εποχικότητα στον υπολογισμό των τιμών της χρονοσειράς. Η παραδοχή αυτή μπορεί να είναι σωστή κυρίως για φυσικά φαινόμενα, αλλά σπάνια ισχύει σε επιχειρησιακές και οικονομικές εφαρμογές, στις οποίες συνήθως η τάση επηρεάζει μεταξύ των άλλων και τις εποχικές διακυμάνσεις. Στη συνέχεια της διπλωματικής θα χρησιμοποιήσουμε το πολλαπλασιαστικό μοντέλο, δεδομένου ότι για τους παραπάνω πρακτικούς και θεωρητικούς λόγους το μοντέλο αυτό πλεονεκτει' του προσθετικού για την ανάλυση των οικονομικών χρονοσειρών.

Το πρώτο βήμα για την ανάλυση της τάσης είναι η γραφική παράσταση της χρονοσειράς, σε αριθμητική, σε ημι-λογαριθμική κλίμακα, ή άλλους μετασχηματισμούς. Εμείς θα εφαρμόσουμε το γραμμικό μοντέλο τάσης, το παραβολικό μοντέλο, to Pearl Reed Logistic και το μοντέλο Gombertz θα συγκρίνουμε τα αποτελέσματά τους και θα επιλέξουμε το κατάλληλο για να κάνουμε την προσαρμογή μας.

12.1. Γραμμικό μοντέλο (Linear)

Η κρίση (judgement) χρησιμοποιείται πολύ συχνά ως κριτήριο για τον προσδιορισμό του μοντέλου με την καλύτερη προσαρμογή. Μερικές φορές ο αναλυτής χαράζει απλώς μια γραμμή διαμέσου της σειράς των παρατηρήσεων, προκειμένου να αποκαλύψει το σχήμα και την κατεύθυνση της τάσης της χρονοσειράς. Ο αναλυτής, χαράζοντας τη γραμμή της τάσης, πρέπει να έχει την ικανότητα να αναγνωρίζει την επαναλαμβανόμενη εποχική διακύμανση και τους κύκλους, μέσω των οποίων διέρχεται η τάση. Με άλλα λόγια, θα πρέπει ο αναλυτής να έχει εξοικειωθεί σημαντικά με αυτές τις συγκεκριμένες σειρές, διαφορετικά η διαδικασία αυτή μπορεί να αποδειχθεί εξαιρετικά δύσκολη. Γι' αυτό οι διενεργούντες προβλέψεις συνήθως εμπιστεύονται την αντικειμενική μέθοδο των ελαχίστων τετραγώνων για να αποφύγουν την υποκειμενικότητα, που υπεισέρχεται στη

μέθοδο που βασίζεται στην κρίση τους. Για να αρχίσουμε, θα ορίσουμε την εκτιμηθείσα γραμμική εξίσωση ως εξής:

$$Y_t = a + bt$$

$\hat{Y}_t$ είναι η προβλεφθείσα τιμή τάσης για τη μεταβλητή Y για μια επιλεγείσα χρονική περίοδο t. α είναι η τιμή της τάσης, όταν t=0 (η διαφορά ύψους). b είναι η κλίση της ευθείας (η μεταβολή στο Y που 190 σχετίζεται με μια μοναδιαία μεταβολή στη t). Το t παριστάνει την παρατήρηση για την χρονική περίοδο.

12.2. Παραβολικό μοντέλο (quadratic)

Μία επιπρόσθετη καμπύλη τάσης είναι η παραβολή ελαχίστων τετραγώνων, που ορίζεται από τις σχέσεις:

$$y = a + bx + cx^2,$$

$$\sum y = na + b\sum x + c\sum x^2$$

$$\sum xy = a\sum x + b\sum x^2 + c\sum x^3$$

$$\sum x^2y = a\sum x^2 + b\sum x^3 + c\sum x^4$$

Οπότε, οποιαδήποτε τιμή πρόβλεψης, με ανεξάρτητη μεταβλητή το χρόνο, δίνεται από τη σχέση:

$$Y_t = a + bt + ct^2$$

12.3. Εκθετικό μοντέλο (exponential)

Η εκθετική καμπύλη τάσης, ανήκει στις καμπύλες που αναπαριστούν τάσεις, σε πολλές βιομηχανίες και γραμμές παραγωγής προϊόντων, που αυξάνονται με μικρότερη ταχύτητα στη φάση ωρίμανσης.

Η καμπύλη είναι της μορφής:

$$Y_t = b_0 {b_1}^t$$

12.4. Μοντέλο Pearl-Reed Logistic

Μία άλλη καμπύλη, την οποία θα εξετάσουμε, είναι η Pearl-Reed Logistic. Η καμπύλη αυτή εκφράζεται από την ακόλουθη σχέση.

$$Y_t = \frac{1000}{b_0 b_1 {b_2}^t}$$

12.5. Μοντέλο Gompertz

Η καμπύλη Gompertz είναι μια καμπύλη τύπου S. Εκφράζεται από την ακόλουθη σχέση.

$$Y_t = \theta_1 \exp(-\exp(\theta_2 - \theta_3 t))$$

12.6. Διάσπαση χρονοσειράς – Πρόβλεψη

Από τα παραπάνω μοντέλα, το μοντέλο Gompertz μας δίνει τα καλύτερα αποτελέσματα, όσον αφορά τους δείκτες αξιοπιστίας (MSE=0.324068). Άρα, αυτό είναι που θα χρησιμοποιήσουμε για να κάνουμε την αποσύνθεση της χρονοσειράς.

13. ΑΠΛΗ ΓΡΑΜΜΙΚΗ ΠΑΛΙΝΔΡΟΜΗΣΗ

Σε αυτή τη μέθοδο, θα πρέπει να σχεδιάσουμε-προσαρμόσουμε μία ευθεία γραμμή ανάμεσα στα σημεία δεδομένων που αναπαριστώνται μέσω διαγραμμάτων διασποράς.

Η ευθεία που προσαρμόζεται καλύτερα σε μία συλλογή Χ-Υ σημείων είναι η γραμμή αυτή, που ελαχιστοποιεί το άθροισμα των τετραγωνικών κάθετων αποστάσεων της γραμμής από τα προς μελέτη σημεία. Αυτή η γραμμή είναι γνωστή ως γραμμή ελαχίστων τετραγώνων ή προσαρμοσμένη γραμμή παλινδρόμησης (fitted regression line) και η εξίσωσή της ονομάζεται προσαρμοσμένη εξίσωση παλινδρόμησης (fitted regression equation).

Η προσαρμοσμένη ευθεία γραμμή είναι της μορφής:

$$Y = b_0 + b_1 X$$

Η γραμμική παλινδρόμηση είναι μια προσέγγιση για τη μοντελοποίηση της σχέσης μεταξύ μιας βαθμωτής εξαρτημένης μεταβλητής Υ και μία ή περισσότερες επεξηγηματικές μεταβλητές (ή ανεξάρτητη μεταβλητή) συμβολίζεται Χ. Περίπτωση μιας επεξηγηματικής μεταβλητής ονομάζεται απλή γραμμική παλινδρόμηση. Για περισσότερες από μία επεξηγηματικές μεταβλητές, η διαδικασία ονομάζεται πολλαπλή γραμμική παλινδρόμηση.(Ο όρος αυτός θα πρέπει να διακρίνεται από πολυμεταβλητή γραμμική παλινδρόμηση, όπου πολλαπλά προβλέπουν συσχέτιση με εξαρτημένες μεταβλητές , αντί για μία ενιαία βαθμωτή μεταβλητή.)

Στην γραμμική παλινδρόμηση, τα δεδομένα μοντελοποιούνται χρησιμοποιώντας γραμμικές λειτουργίες προγνωστικά, και οι άγνωστες παράμετροι μοντέλου υπολογίζονται από τα δεδομένα. Τέτοια μοντέλα καλούνται γραμμικά μοντέλα. Συνηθέστερα, η γραμμική παλινδρόμηση αναφέρεται σε ένα μοντέλο στο οποίο ο υποθετικός μέσος όρος του Υ δεδομένης της αξίας του Χ είναι μια συνάρτηση αφινικών Χ λιγότερο συχνά, όπου η γραμμική παλινδρόμηση μπορεί να αναφέρεται σε ένα μοντέλο στο οποίο η διάμεσος, ή κάποιο άλλο ποσοστημόριο της υποθετικής διανομής y που δίνεται Χ εκφράζεται ως γραμμική συνάρτηση του Χ Όπως όλες τις μορφές ανάλυσης παλινδρόμησης, η γραμμική παλινδρόμηση επικεντρώνεται στους όρους κατανομής πιθανότητας του y που δίνονται Χ αντί για την από κοινού πιθανότητα διανομής του Υ και Χ, η οποία είναι η και η περιοχή της πολυμεταβλητής ανάλυσης.

Η γραμμική παλινδρόμηση ήταν ο πρώτος τύπος της ανάλυσης παλινδρόμησης που μελετήθηκε αυστηρά, και προορίζεται να χρησιμοποιηθεί εκτενώς σε πρακτικές εφαρμογές. Αυτό συμβαίνει επειδή τα μοντέλα που εξαρτώνται γραμμικά από άγνωστες παραμέτρους τους είναι πιο εύκολο να χωρέσουν από τα μοντέλα τα οποία είναι μη-γραμμικά με παραμέτρους τους και επειδή οι στατιστικές ιδιότητες των προκυπτόντων

εκτιμήσεων είναι εύκολο να προσδιοριστεί.

Αν ο στόχος είναι η πρόβλεψη, ή η μείωση, η γραμμική παλινδρόμηση μπορεί να χρησιμοποιηθεί για να χωρέσει ένα προγνωστικό μοντέλο σε ένα παρατηρούμενο δεδομένο με Χ και Υ τιμές. Μετά από την ανάπτυξη ενός τέτοιου μοντέλου, μια πρόσθετη τιμή του Χ είναι τότε χωρίς την συνοδευτική αξία του y, όπου το μοντέλο μπορεί να χρησιμοποιηθεί για να κάνει μια πρόβλεψη της τιμής του y.

Δεδομένης μια μεταβλητή y και ενός αριθμού μεταβλητών X1, ..., Xp που μπορεί να σχετίζονται με το y,η ανάλυση γραμμικής παλινδρόμησης μπορεί να εφαρμοστεί στην ποσοτικοποίηση της αντοχής της σχέσης μεταξύ Υ και του χj , προκειμένου να αξιολογηθεί η οποία σχέση χj με y καθόλου, και να προσδιορίσει ποιες υποκατηγορίες του χj περιέχουν περιττές πληροφορίες.

Τα μοντέλα γραμμικής παλινδρόμησης συχνά χρησιμοποιούνται κατά την προσέγγιση λιγότερων τετραγώνων, αλλά μπορεί επίσης να χρησιμοποιηθεί με άλλους τρόπους, όπως με ελαχιστοποίηση της "έλλειψη προσαρμογής" σε κάποιο άλλο πρότυπο (όπως με τουλάχιστον παλινδρόμηση της απόλυτης αποκλίσεις). Αντιστρόφως, τουλάχιστον η προσέγγιση με τα τετράγωνα μπορεί να χρησιμοποιηθεί για να χωρέσει τα μοντέλα που δεν είναι γραμμικά μοντέλα. Έτσι, αν και οι όροι "ελαχίστων τετραγώνων" και "γραμμικό μοντέλο" συνδέονται στενά, δεν είναι συνώνυμοι.

13.1. Παραδοχές της γραμμικής παλινδρόμησης

Πρότυπα μοντέλα γραμμικής παλινδρόμησης με τυποποιημένες τεχνικές εκτίμησης κάνουν μια σειρά από υποθέσεις σχετικά με τις μεταβλητές πρόβλεψης, τις μεταβλητές απόκρισης και τη σχέση τους. Πολλές επεκτάσεις έχουν αναπτυχθεί που επιτρέπουν κάθε μία από αυτές τις υποθέσεις να χαλαρώσουν (δηλαδή μειώνεται σε μια ηπιότερη μορφή), και σε ορισμένες περιπτώσεις θα εξαλειφθούν εντελώς. Μερικές μέθοδοι είναι αρκετά γενική ώστε να μπορούν να χαλαρώσουν πολλαπλές υποθέσεις ταυτόχρονα, και σε άλλες περιπτώσεις, αυτό μπορεί να επιτευχθεί με συνδυασμό διαφορετικών επεκτάσεων. Γενικά αυτές οι επεκτάσεις καταστούν την διαδικασία εκτίμησης πιο σύνθετη και χρονοβόρα, και μπορεί επίσης να απαιτεί περισσότερα δεδομένα για να παράγει ένα εξίσου ακριβό μοντέλο.

Τα ακόλουθα είναι οι κύριες παραδοχές που έγιναν από τα συνήθη μοντέλα γραμμικής παλινδρόμησης με τυποποιημένες τεχνικές εκτίμησης (π.χ. η διαδικασία ελαχίστων τετραγώνων):

13.2. Ασθενείς εξωγενείς

Αυτό ουσιαστικά σημαίνει ότι μπορούν να αντιμετωπιστούν οι μεταβλητές πρόβλεψης Χ ως σταθερές τιμές, παρά τυχαία μεταβλητή s. Αυτό σημαίνει, για παράδειγμα, ότι οι μεταβλητές πρόβλεψης θεωρείται ότι είναι χωρίς λάθη-που είναι, μη μολυσμένα με

σφάλματα μέτρησης. Παρά το γεγονός ότι η υπόθεση αυτή δεν είναι ρεαλιστική σε πολλές ρυθμίσεις, οδηγεί σε πολύ πιο δύσκολο μοντέλο για τα σφάλματα των μεταβλητών s.

13.3. Γραμμικότητα

Αυτό σημαίνει ότι η μέση τιμή της μεταβλητής απόκρισης είναι ένας γραμμικός συνδυασμός των παραμέτρων (συντελεστές παλινδρόμησης) και οι μεταβλητές πρόβλεψης. Σημειώστε ότι αυτή η υπόθεση είναι πολύ λιγότερο περιοριστική από ό, τι μπορεί αρχικά να φαίνεται. Επειδή οι μεταβλητές πρόβλεψης αντιμετωπίζονται ως σταθερές τιμές (βλέπε παραπάνω), γραμμικότητα είναι πραγματικά μόνο ένας περιορισμός σχετικά με τις παραμέτρους. Οι ίδιες οι μεταβλητές πρόβλεψης μπορεί να μετασχηματιστούν αυθαίρετα, και στην πραγματικότητα μπορούν να προστεθούν πολλαπλά αντίγραφα της ίδιας υποκείμενης μεταβλητής πρόβλεψης, το καθένα μετασχηματισμένο με διαφορετικό τρόπο. Αυτό το τέχνασμα έχει χρησιμοποιηθεί, για παράδειγμα, στη πολυωνυμική παλινδρόμηση, το οποίο χρησιμοποιεί γραμμική παλινδρόμηση για να χωρέσει τη μεταβλητή απόκριση ως μια αυθαίρετη λειτουργία μιας μεταβλητής πρόβλεψης. Αυτό κάνει την γραμμική παλινδρόμηση μία εξαιρετικά ισχυρή μέθοδο εξαγωγής συμπερασμάτων. Στην πραγματικότητα, τα μοντέλα όπως η πολυωνυμική παλινδρόμηση είναι συχνά «πολύ ισχυρά". Ως αποτέλεσμα, κάποια είδη νομιμοποίησης (μαθηματικά) | νομιμοποίησης θα πρέπει να χρησιμοποιούνται τυπικά για να αποφεύγονται οι υπερβολικές λύσεις που προέρχονται από τη διαδικασία εκτίμησης. Κοινά παραδείγματα είναι παλινδρόμηση κορυφογραμμής και λάσο παλινδρόμησης. Bayesianγραμμικής παλινδρόμησης μπορεί επίσης να χρησιμοποιηθεί, η οποία από τη φύση της είναι περισσότερο ή λιγότερο ανοσοποιητική στο πρόβλημα της υπερπροσαρμογής. (Στην πραγματικότητα, κορυφογραμμή παλινδρόμησης και λάσο παλινδρόμησης μπορεί τόσο να θεωρηθούν ως ειδικές περιπτώσεις Bayesian γραμμικής παλινδρόμησης, με συγκεκριμένους τύπους πριν από τη διανομή s που τοποθετούνται με τους συντελεστές παλινδρόμησης.)

13.4. Συνεχής Μεταβλητότητα (αλλιώς ομοσκεδαστικότητα)

Αυτό σημαίνει πως οι διαφορετικές μεταβλητές απόκρισης έχουν την ίδια μεταβλητότητα στα λάθη τους,ανεξάρτητα από τις τιμές των προβλεπόμενων μεταβλητών.Πρακτικά,αυτή η υπόθεση δεν ισχύει(δηλ. τα λάθη είναι ετεροσκεδαστικά)αν οι μεταβλητές απόκρισης μπορούν να ποικίλλουν σε μια ευρεία κλίμακα.Με σκοπό να αποφασίσουμε για τα ετερογενή λάθη μεταβλητότητας,ή πότε μια διάταξη υπολοίπων παραβιάζει τη δομή των υποθέσεων της ομοσκεδαστικότητας(το λάθος είναι εξ' ίσου μεταβλητό γύρω από την "καλύτερη-ταιριαστή γραμμή" για όλα τα σημεία του x),είναι συνετό να ψάξουμε για ένα αποτέλεσμα μεταξύ του υπόλοιπου λάθους και των προβλεπόμενων τιμών.Πρέπει να πούμε ότι θα υπάρχει μια συστηματική αλλαγή στο απόλυτο ή στο τετράγωνο των υπολοίπων όταν "συνωμοτούν" σε βάρος του προβλεπόμενου αποτελέσματος. Το λάθος δε θα είναι ομοιόμορφα κατανεμημένο στη γραμμή παλινδρόμησης.Η ετεροσκεδαστικότητα θα συνεισφέρει στο να ανεβάσει το μέσο όρο των διακριτών μεταβλητών γύρω από τα σημεία για να μια μοναδική μεταβλητή που ανακριβώς αναπαριστά όλες τις μεταβλητές της γραμμής. Συνεπώς,τα υπόλοιπα εμφανίζονται συσσωρευμένα και απλωμένα στις προβλεπόμενες

παραστάσεις τους για μεγαλύτερες ή μικρότερες τιμές για σημεία πάνω στη γραμμή παλινδρόμησης και το νόημα της παράστασης θα είναι λάθος.Τυπικά ,για παράδειγμα, μια μεταβλητή απόκρισης που η σύγκλισή της είναι μεγάλη θα έχει μεγαλύτερη διασπορά από μια που η σύγκλισή της είναι μικρή.Για παράδειγμα,ένας συγκεκριμένος άνθρωπος που το εισόδημά του προβλέπεται στα 100.000$ μπορεί εύκολα να έχει ένα πραγματικό εισόδημα 80.000$ ή 120.000$(μια τυπική απόκλιση γύρω στις 20.000$),όταν ένας άλλος άνθρωπος με προβλεπόμενο εισόδημα 10.000$ είναι απίθανο να έχει την ίδια τυπική απόκλιση των 20.000$,η οποία σημαίνει ότι το πραγματικό τους εισόδημα θα ποίκιλλε οπουδήποτε μεταξύ των -10.000$ και των 30.000$.(Στην πραγματικότητα, αυτό δείχνει, σε πολλές περιπτώσεις-συχνά στις ίδιες περιπτώσεις όπου η υπόθεση των κανονικά κατανεμημένων σφαλμάτων-η διακύμανση ή η τυπική απόκλιση θα έπρεπε να προβλέπεται να είναι ανάλογη με τη τιμή, παρά με τη συνέχεια.)Οι απλές μέθοδοι γραμμικής παλινδρόμησης και εστίασης δίνουν λιγότερη ακρίβεια στην εκτίμηση των παραμέτρων και χάνουν σημαντικές ποσότητες όπως σίγουρα λάθη όταν ουσιαστικά η ετεροσκεδαστικότητα υπάρχει. Ωστόσο, πολλές τεχνικές εστίασης (π.χ.η [proportional στάθμιση ελαχίστων τετραγώνων] και η ετεροσκεδαστικότητα-με συνεχή σφάλματα)μπορούν να χειριστούν την ετεροσκεδαστικότητα με έναν απλό γενικό τρόπο. Οι τεχνικές της Μπευζιανής γραμμικής παλινδρόμησης μπορούν ακόμα να χρησιμοποιηθούν όταν η διακύμανση φέρεται να είναι συνάρτηση της τιμής. Είναι ακόμα πιθανό σε μερικές περιπτώσεις να διορθώσεις ένα πρόβλημα με το να εφαρμόζεις μια αλλαγή στην διακριτή μεταβλητή(π.χ. να ταιριάξεις ένα λογάριθμο της διακριτής μεταβλητής χρησιμοποιώντας μια μέθοδο γραμμικής παλινδρόμησης, η οποία σημαίνει ότι η διακριτή μεταβλητή έχει μια ευρεία κανονική κατανομή από ότι μια απλή κανονική κατανομή.

13.5. Ανεξαρτησία λαθών

Αυτό προϋποθέτει ότι τα σφάλματα των μεταβλητών απόκρισης είναι ασύνδετα μεταξύ τους. (Πραγματική στατιστική ανεξαρτησία. Είναι μια ισχυρότερη κατάσταση από την απλή έλλειψη συσχέτισης και συχνά δεν είναι απαραίτητη, αν και μπορεί να αξιοποιηθεί, εάν είναι γνωστό για να κρατήσει) Ορισμένες μέθοδοι είναι σε θέση να χειρίζονται σφάλματα, αν και συνήθως απαιτούν πολύ περισσότερα δεδομένα, εκτός αν κάποιο είδος νομιμοποίησης χρησιμοποιείται για την πόλωση του μοντέλου για την παραδοχή ασυσχέτιστες λάθη. Μπεϋζιανή γραμμική παλινδρόμηση είναι ένας γενικός τρόπος χειρισμού αυτού του θέματος.

13.6. Έλλειψη πολυσυγγραμμικότητας στους προγνωστικούς παράγοντες.

Για πρότυπο ελαχίστων τετραγώνων μέθοδοι εκτίμησης, η μήτρα σχεδιασμού X πρέπει να έχει πλήρη βαθμίδα στήλης p ! Αλλιώς, έχουμε μια κατάσταση που είναι γνωστή ως πολυσυγγραμμικότητα στις μεταβλητές πρόβλεψης. Αυτό μπορεί να προκληθεί από την κατοχή δύο ή περισσότερων τέλεια συσχετισμένων μεταβλητών πρόβλεψης (π.χ. αν η ίδια μεταβλητή πρόβλεψης λανθασμένα δίνεται δύο φορές, είτε χωρίς μετατροπή ενός από τα αντίγραφα ή με τη μετατροπή ενός από τα αντίγραφα γραμμικά). Μπορεί επίσης

να συμβεί εάν υπάρχουν πάρα πολύ λίγα διαθέσιμα δεδομένα σε σχέση με τον αριθμό των παραμέτρων που πρέπει να υπολογίζεται (π.χ. λιγότερα σημεία δεδομένων από συντελεστές παλινδρόμησης). Στην περίπτωση πολυσυγγραμμικότητας, το διάνυσμα παραμέτρων β θα είναι μη αναγνωρίσιμο - δεν έχει μοναδική λύση. Στην καλύτερη περίπτωση θα είναι σε θέση να εντοπίσει κάποιες από τις παραμέτρους, δηλαδή περιορίζετε η αξία του σε κάποιο γραμμικό υπόχωρο του 'R' p . Μέθοδοι για την τοποθέτηση γραμμικών μοντέλων με πολυσυγγραμμικότητα έχουν αναπτυχθεί? μερικοί απαιτούν επιπλέον παραδοχές, όπως η «επίδραση αραιότητας» -δηλαδή ένα μεγάλο μέρος από τα αποτελέσματα είναι ακριβώς μηδέν. Σημειώστε ότι τα περισσότερα υπολογιστικά ακριβά επαναληπτικών αλγορίθμων για την εκτίμηση των παραμέτρων, όπως αυτές που χρησιμοποιούνται σε γενικευμένο γραμμικό μοντέλο s, δεν υποφέρουν από αυτό το πρόβλημα, και στην πραγματικότητα αυτό είναι απολύτως φυσιολογικό κατά τον χειρισμό [[κατηγορικά δεδομένα | κατηγορηματικά αποτιμώνται]] προγνωστικοί παράγοντες σε μια ξεχωριστή μεταβλητή δείκτη προγνωστικός δείκτης για κάθε πιθανή κατηγορία, η οποία εισάγει αναπόφευκτα πολυσυγγραμμικότητα.

Πέρα από αυτές τις υποθέσεις, αρκετές άλλες στατιστικές ιδιότητες των δεδομένων επηρεάζει σημαντικά την απόδοση των διαφορετικών μεθόδων εκτίμησης:

- Η στατιστική σχέση μεταξύ των όρων σφάλματος και τις ερμηνευτικές μεταβλητές παίζει σημαντικό ρόλο στον καθορισμό του κατά πόσον η διαδικασία εκτίμησης έχει επιθυμητές ιδιότητες δειγματοληψίας, όπως είναι αμερόληπτοι και συνεπείς.
- Η ρύθμιση ή κατανομής πιθανότητας του «x» »έχει μια σημαντική επιρροή των μεταβλητών πρόβλεψης σχετικά με την ακρίβεια των εκτιμήσεων των β. Δειγματοληψία και σχεδιασμός των πειραμάτων είναι τα ανεπτυγμένα υποπεδία των στατιστικών στοιχείων που παρέχουν καθοδήγηση για τη συλλογή δεδομένων με τέτοιο τρόπο για να επιτευχθεί μια ακριβής εκτίμηση της β .

14. ΠΟΛΥΜΕΤΑΒΛΗΤΗ ΠΑΛΙΝΔΡΟΜΗΣΗ

Στην απλή γραμμική παλινδρόμηση, ερευνήθηκε η σχέση μεταξύ μίας μόνο ανεξάρτητης μεταβλητής και μίας εξαρτημένης μεταβλητής. Η σχέση μεταξύ δύο μεταβλητών συχνά επιτρέπει σε κάποιον να προβλέψει με ακρίβεια την εξαρτημένη μεταβλητή, όταν έχει γνώση της ανεξάρτητης. Δυστυχώς, πολλές πραγματικές καταστάσεις πρόβλεψης δεν είναι τόσο απλές. Περισσότερες από μία ανεξάρτητες μεταβλητές είναι συχνά απαραίτητες, έτσι ώστε να προβλέψει κανείς την εξαρτημένη μεταβλητή με ακρίβεια. Τα μοντέλα παλινδρόμησης με περισσότερες από μία ανεξάρτητες μεταβλητές ονομάζονται μοντέλα πολυμεταβλητής παλινδρόμησης. Οι περισσότερες από τις έννοιες που παρουσιάστηκαν στην απλή γραμμική παλινδρόμηση μεταφέρονται και στην πολυμεταβλητή παλινδρόμηση. Ωστόσο, μερικές νέες ιδέες εισάγονται επειδή χρησιμοποιούνται περισσότερες από μία ανεξάρτητες μεταβλητές για την πρόβλεψη της εξαρτημένης. Η πολυμεταβλητή παλινδρόμηση περιλαμβάνει τη χρήση περισσότερων από μία ανεξάρτητων μεταβλητών για την πρόβλεψη μίας εξαρτημένης μεταβλητής.

14.1. Μερικές μεταβλητές πρόβλεψης

Ας θεωρήσουμε ένα παράδειγμα, κατά το οποίο μία εταιρία παραγωγής γάλακτος στην Αμερική θέλει να προβλέψει την ποσότητα γαλονιών γάλακτος, που πουλάει, ανάλογα με την τιμή, που χρεώνει το γαλόνι. Εφαρμόζοντας απλή γραμμική παλινδρόμηση στα δεδομένα που είχε συγκεντρώσει κατέληξε στο συμπέρασμα, ότι το πρόβλημα της πρόβλεψης δεν είναι ολοκληρωτικά ακριβές. Μπορεί να ερμηνευθεί μόνο το 75% (r^2) των διαφορών στα γαλόνια του γάλακτος, που πωλούνται, χρησιμοποιώντας μόνο μία ανεξάρτητη μεταβλητή. Το 25% ($1-r^2$) της

συνολικής μεταβλητότητας δεν μπορεί να ερμηνευθεί. Με άλλα λόγια, γνωρίζοντας μόνο την τιμή του προϊόντος, γνωρίζει κανείς το 75%, αυτών που χρειάζεται, προκειμένου να προβλέψει σωστά την ποσότητα των πωλήσεων. Για να γίνουν πιο ακριβείς οι προβλέψεις, πρέπει να βρεθεί και άλλη μεταβλητή πρόβλεψης, που θα επιτρέψει να ερμηνευθεί περισσότερη από τη συνολική μεταβλητότητα. Αν επιτευχθεί η μείωση της μη ερμηνευμένης μεταβλητότητας, οι προβλέψεις θα περιέχουν λιγότερη αβεβαιότητα και θα είναι πιο ακριβείς. Πρέπει να γίνει έρευνα, για να ανακαλυφθεί μία ακόμα ανεξάρτητη μεταβλητή, που να σχετίζεται με την ποσότητα των πωλήσεων γαλονιών γάλακτος. Ωστόσο, αυτή η νέα ανεξάρτητη μεταβλητή (predictor) δεν μπορεί να σχετίζεται τόσο πολύ με την ανεξάρτητη μεταβλητή (τιμή ανά γαλόνι), που χρησιμοποιείται ήδη. Αν οι δυο ανεξάρτητες μεταβλητές σχετίζονται πολύ μεταξύ τους, θα ερμηνεύουν την ίδια μεταβολή και η προσθήκη της δεύτερης μεταβλητής δεν θα βελτιώσει την πρόβλεψη. Η απλή λύση στο πρόβλημα των δύο υψηλά συσχετισμένων ανεξάρτητων μεταβλητών είναι μερικώς το να μη χρησιμοποιούνται και οι δύο μαζί. Το πρόβλημα αυτό (multicollinearity) θα συζητηθεί αργότερα σε αυτό το κεφάλαιο.

14.2. Πίνακας Συσχέτισης

Η εταιρία αποφασίζει, ότι και τα διαφημιστικά έξοδα μπορεί να βοηθήσουν στο να βελτιωθεί η πρόβλεψη των εβδομαδιαίων πωλήσεων. Η διερεύνηση της σχέσης μεταξύ των διαφημιστικών εξόδων, του όγκου των πωλήσεων και της τιμής ανά γαλόνι, γίνεται μέσω ενός πίνακα συσχέτισης (correlation matrix). Ο πίνακας συσχέτισης κατασκευάζεται υπολογίζοντας τους συντελεστές απλής συσχέτισης για κάθε συνδυασμό ζευγαριών των μεταβλητών . Ένα παράδειγμα του πίνακα συσχέτισης παρουσιάζεται στον ακόλουθο Πίνακα. Ο συντελεστής συσχέτισης που υποδεικνύει τη σχέση μεταξύ των μεταβλητών 1 και 2 συμβολίζεται με r12. Να σημειωθεί ότι το πρώτο ψηφίο 1, επίσης αναφέρεται και στη σειρά ενώ το δεύτερο 2, αναφέρεται στη στήλη του πίνακα. Αυτή η προσέγγιση επιτρέπει σε

κάποιον να καθορίσει με μία ματιά, τη σχέση μεταξύ δύο μεταβλητών. Φυσικά, η συσχέτιση μεταξύ, για παράδειγμα, των μεταβλητών 1 και 2 είναι ακριβώς η ίδια με τη συσχέτιση των μεταβλητών 2 και 1. Γι' αυτό ισχύει: r12=r21. Έτσι, το μισό του πίνακα είναι απαραίτητο. Επιπλέον, η συσχέτιση μιας μεταβλητής με τον εαυτό της είναι πάντα 1, έτσι για παράδειγμα, r11=r22=r33=1. Η εταιρία τρέχει τα δεδομένα της στον υπολογιστή, και ο πίνακας συσχέτισης φαίνεται στον παρακάτω Πίνακα. Μία διερεύνηση της σχέσης μεταξύ των διαφημιστικών εξόδων, του όγκου των πωλήσεων και της τιμής ανά γαλόνι, δείχνει ότι η νέα ανεξάρτητη μεταβλητή πρέπει να συνεισφέρει σε βελτιωμένη πρόβλεψη. Ο πίνακας συσχέτισης δείχνει ότι τα διαφημιστικά έξοδα έχουν υψηλή θετική σχέση (r13=0,89) με την εξαρτημένη μεταβλητή, τον όγκο των πωλήσεων, και μία μέτρια

αρνητική σχέση (r23=-0,65) με την ανεξάρτητη μεταβλητή, τη τιμή ανά γαλόνι. Αυτός ο συνδυασμός σχέσεων πρέπει να επιτρέψει στα διαφημιστικά έξοδα να ερμηνεύουν κάποια από τη συνολική μεταβολή η οποία δεν ερμηνεύεται ήδη από την τιμή ανά γαλόνι. Όπως θα φανεί και στη συνέχεια, όταν χρησιμοποιηθούν ταυτόχρονα και η τιμή ανά γαλόνι και τα διαφημιστικά έξοδα για την εκτίμηση του όγκου των πωλήσεων, το R^2 αυξάνει στο 93,2%. Η ανάλυση του πίνακα συσχέτισης είναι ένα σημαντικό αρχικό βήμα στη λύση κάθε προβλήματος που περιλαμβάνει πολλαπλές ανεξάρτητες μεταβλητές.

Πίνακας 2 : Πίνακας Συσχέτισης

Μεταβλητές	1	2	3
1	R_{11}	R_{12}	R_{13}
2	R_{21}	R_{22}	R_{23}
3	R_{31}	R_{32}	R_{33}

Πίνακας 3: Συντελεστής Συσχέτισης για τα Δεδομένα

Μεταβλητές	Πωλήσεις(1)	Αξία(2)	Διαφήμιση(3)
Πωλήσεις(1)	1,000	-0,863	0,891
Αξία(2)		1,000	-0,654
Διαφήμιση(3)			1,000

14.3. Μοντέλο Πολυμετάβλητης Παλινδρόμησης

Στην απλή παλινδρόμηση, η εξαρτημένη μεταβλητή συμβολίζεται με Υ και η ανεξάρτητη μεταβλητή με Χ. Στην ανάλυση πολυμεταβλητής παλινδρόμησης, τα Χ με κάτω δείκτες χρησιμοποιούνται για να συμβολίσουν τις ανεξάρτητες μεταβλητές. Η εξαρτημένη μεταβλητή ακόμα συμβολίζεται με Υ, και οι ανεξάρτητες μεταβλητές με Χ1, Χ2, ..., Χk. Όταν καθοριστεί τα αρχικό σύνολο ανεξάρτητων μεταβλητών, η σχέση μεταξύ του Υ και των Χ μπορεί να εκφραστεί σαν μοντέλο πολυμεταβλητής παλινδρόμησης. Στο μοντέλο πολυμεταβλητής παλινδρόμησης, η μέση παρατηρούμενη τιμή είναι η ακόλουθη γραμμική συνάρτηση των ανεξάρτητω ν μεταβλητών:

$$\mu_Y=\beta_0+ \beta_1 X_1+ \beta_2 X_2+ ...+ \beta_K X_K$$

Αυτή η έκφραση είναι η συνάρτηση πολυμεταβλητής παλινδρόμησης του πληθυσμού. Όπως και με την απλή γραμμική παλινδρόμηση, δεν μπορούμε απ' ευθείας να παρατηρήσουμε τη συνάρτηση παλινδρόμησης πληθυσμού, επειδή οι παρατηρούμενες τιμές του Υ διαφέρουν από τη μέση τιμή. Υποθέτουμε ότι τα Υ είναι κανονικά κατανεμημένα γύρω από τους μέσους τους με την ίδια τυπική απόκλιση, σ. Τα δεδομένα για την απλή γραμμική παλινδρόμηση αποτελούνται από τις παρατηρήσεις (Xi, Yi) των δύο μεταβλητών. Στην πολυμεταβλητή παλινδρόμηση, τα δεδομένα για κάθε περίπτωση αποτελούνται από μία παρατήρηση της εξαρτημένης μεταβλητής και μία παρατήρηση για κάθε μία από τις ανεξάρτητες μεταβλητές. Η i παρατήρηση της j μεταβλητής πρόβλεψης συμβολίζεται με Χij.

15. ARIMA

Τα μοντέλα πολλαπλής γραμμικής παλινδρόμησης περιγράφονται από την εξίσωση:

$$Y = bo + b1\ X\ 1 + b2\ X\ 2 + ... + b\ p\ X\ p + e$$

όπου Y η εξαρτημένη μεταβλητή και X1, X2, ... , Xp οι ανεξάρτητες μεταβλητές.

Η εξίσωση $Yt = bo + b1Yt\ -1 + b2Yt\ -2 + ... + b\ pYt\ -\ p + et$ είναι μία εξίσωση πολλαπλής παλινδρόμησης με τη διαφορά ότι οι ανεξάρτητες μεταβλητές είναι τιμές της χρονοσειράς σε προηγούμενες χρονικές περιόδους. Το παραπάνω υπόδειγμα ονομάζεται αυτοπαλίνδρομο υπόδειγμα (autoregression model) και συμβολίζεται AR(p). Συνεπώς τα υποδείγματα AR εκφράζουν το Yt σαν γραμμική συνάρτηση των p παρελθουσών πραγματικών τιμών του Yt.

Επίσης η εξίσωση $Yt = bo + et\ -\ b1et\ -1\ -\ b2\ et\ -2\ -\ ...\ -\ bq\ et\ -q$ είναι μία εξίσωση πολλαπλής παλινδρόμησης με τη διαφορά ότι οι ανεξάρτητες μεταβλητές είναι τιμές των σφαλμάτων προηγούμενων χρονικών περιόδων (διαφορά προβλεφθείσας τιμής από την πραγματική). Το παραπάνω υπόδειγμα ονομάζεται υπόδειγμα κινητού μέσου όρου (moving average) και συμβολίζεται MA(q). Συνεπώς τα υποδείγματα MA εκφράζουν το Yt σαν γραμμικό συνδυασμό παρελθόντων σφαλμάτων πρόβλεψης.Τα αυτοπαλίνδρομα υποδείγματα AR(p) μπορούν να συνδυαστούν αποτελεσματικά με τα υποδείγματα κινητού μέσου όρου MA(q) και να σχηματίσουν μια χρήσιμη ομάδα υποδειγμάτων χρονοσειρών τα οποία ονομάζονται αυτοπαλίνδρομα υποδείγματα κινητού μέσου όρου (autoregressive moving average models) και συμβολίζονται ARMA(p,q). Τα υποδείγματα αυτά μπορούν να χρησιμοποιηθούν μόνο για σταθερές (στάσιμες) χρονοσειρές. Ωστόσο μπορούν να επεκταθούν και σε μη σταθερές (μη στάσιμες) χρονοσειρές με τη χρήση της μεθόδου της διαφόρισης. Σε αυτήν την περίπτωση ονομάζονται ολοκληρωμένα (integrated I) αυτοπαλίνδρομα μοντέλα κινητού μέσου όρου (autoregressive integrated moving average models) και συμβολίζονται ως ARIMA(p,d,q) όπου p είναι η τάξη του αυτοπαλίνδρομου υποδείγματος, d η τάξη της διαφόρισης για την επίτευξη σταθερότητας και q η τάξη του κινητού μέσου όρου υποδείγματος.Το μοντέλο λευκού θορύβου (white noise) εκφράζεται ως ARIMA(0,0,0) ενώ το μοντέλο τυχαίου περιπάτου ως ARIMA(0,1,0).

Τα ολοκληρωμένα αυτοπαλίνδρομα υποδείγματα κινητού μέσου όρου (Autoregressive Integrated Moving Average) είναι στοχαστικά μαθηματικά μοντέλα τα οποία χρησιμοποιούνται για την περιγραφή της διαχρονικής εξέλιξης κάποιου φυσικού μεγέθους. Τα στοχαστικά μοντέλα περιλαμβάνουν τον τυχαίο παράγοντα (τυχαίο σφάλμα ή σφάλμα πρόβλεψης), τις τιμές του μεγέθους οι οποίες εμφανίζονται σε προηγούμενες χρονικές στιγμές και ίσως κάποιους άλλους στοχαστικούς παράγοντες. Το υπόδειγμα που προκύπτει είναι ένας γραμμικός συνδυασμός των παραπάνω ποσοτήτων. Τα υποδείγματα ARIMA έχουν μελετηθεί εκτεταμένα από τους Box και Jenkins σε βαθμό που τα ονόματα των παραπάνω να είναι σχεδόν συνώνυμα με τις ARIMA διαδικασίες και τις εφαρμογές τους στην ανάλυση και την πρόβλεψη

χρονοσειρών. Οι Box-Jenkins πρότειναν μια οικογένεια αλγεβρικών υποδειγμάτων πρόβλεψης από τα οποία επιλέγεται το πιο κατάλληλο για την πρόβλεψη μιας χρονοσειράς. Σε αυτά τα υποδείγματα οι προβλέψεις στηρίζονται στις προηγούμενες χρονικά τιμές της χρονοσειράς και στα πρότυπα συμπεριφοράς που εμφανίζει η χρονοσειρά.

Τα ARIMA (autoregressive integrated moving average) μοντέλα είναι μία κατηγορία γραμμικών μοντέλων, που είναι ικανά να αναπαραστήσουν στάσιμες και μη στάσιμες χρονοσειρές. Τα ARIMA μοντέλα δεν περιλαμβάνουν ανεξάρτητες μεταβλητές στην κατασκευή τους. Αντί για αυτό κάνουν χρήση της διαθέσιμης πληροφορίας, που κρύβουν οι χρονοσειρές, για τη δημιουργία προβλέψεων.

Τα ARIMA μοντέλα συνδυάζουν τις ιδιότητες τριών διαφορετικών υπο- μοντέλων (υπο-συστημάτων):

- Αυτοπαλινδρόμησης (autoregression).
- Ολοκλήρωσης (integration).
- Εξομάλυνσης με κινητό μέσο (moving average).

16. BASS

Το τελευταίο μοντέλο της μελέτης μας είναι αυτό του Bass. Η θεωρία αυτή εφαρμόστηκε για την περιγραφή της αύξησης των πωλήσεων κατά τη διάρκεια του χρόνου, μετά την εισαγωγή των προϊόντων στην αγορά και μελετά το κατά πόσο η υιοθέτηση ενός αγαθού από την αγορά εξαρτάται από κοινωνικές συμπεριφορές.

Σε ένα κοινωνικό σύστημα κάποια άτομα αποφασίζουν να υιοθετήσουν μία καινοτομία ανεξάρτητα από κάποια άλλα άτομα. Αυτοί ονομάζονται “Innovators”. Εκτός από τους “Innovators” οι υπόλοιποι επηρεάζονται ως προς το χρόνο υιοθέτησης μιας καινοτομίας από την πίεση του κοινωνικού συστήματος και η πίεση αυτή αυξάνει ολοένα και περισσότερο για τους καθυστερημένους.

Για να αναπτύξουμε τη θεωρία BASS δεχόμαστε μία βασική υπόθεση: “Η πιθανότητα μία αρχική αγορά να γίνει στο χρόνο T, δεδομένου ότι δεν έχει προηγηθεί κάποια άλλη αγορά, είναι μία γραμμική συνάρτηση των προηγούμενων αγοραστών”. Ή με λίγα λόγια ο Bass ορίζει την αναλογία, κατά την οποία οι υποψήφιοι αγοραστές, οι οποίοι δεν έχουν υιοθετήσει ακόμη, κατά τη χρονική περίοδο μεταξύ T και T+ΔT με ΔT → 0, αυξάνονται αναλογικά με τον πληθυσμό, ο οποίος έχει ήδη υιοθετήσει. Η αναλογία είναι ο συντελεστής διάχυσης.

Το βασικό μοντέλο, όσον αφορά τη διείσδυση (σχετικές πωλήσεις), είναι:

$$S(T) = pm + (q - p)Y(T) - \frac{q}{m}[Y(T)]^2$$

Για να υπολογίσουμε τα p, q και m γράφουμε το διακριτό ανάλογο ως:

$$S_T = a + bY_{T-1} + cY_{T-1}{}^2$$

17. ΣΥΜΠΕΡΑΣΜΑΤΑ

Αρχικά στην παρούσα πτυχιακή μελετήθηκαν διάφορες μέθοδοι εξομάλυνσης , οι μέθοδοι πρόβλεψης συνεχούς ζήτησης , οι στατιστικές μέθοδοι πρόβλεψης, τα μοντέλα παλινδρόμησης και η ανάλυση ARIMA οι οποίες χρησιμοποιούνται για τη διαμόρφωση βραχυπρόθεσμων ή μακροπρόθεσμων προβλέψεων και μελλοντικών τιμών μιας χρονοσειράς. Οι μέθοδοι εξομάλυνσης γίνονται σε σύντομο χρονικό διάστημα και με χαμηλό κόστος , ενώ βασίζονται σε απλές υπολογιστικές πράξεις και δεν απαιτούν μεγάλο αριθμό παρατηρήσεων της χρονοσειράς. Παρουσιάστηκε η ανάλυση ARIMA η οποία αποτελεί μια στατιστικά εξειδικευμένη προσέγγιση στην ανάλυση. Ο σκοπός της πρόβλεψης είναι να βοηθήσει τη διεύθυνση να λάβει αποφάσεις , να προγραμματίσει τις απαιτήσεις για πρώτες ύλες και γενικά να εξελιχθεί η επιχείρηση μέσα από προβλεπόμενες και στοχευμένες κινήσεις. Τέλος για να επιτευχθεί η καλύτερη δυνατή πρόβλεψη θα πρέπει να επιλεγεί η καταλληλότερη μέθοδος προβλέψεων. Αντιλαμβανόμαστε λοιπόν την αναγκαιότητα του ερευνητή να διαθέτει εμπειρία , οξυδέρκεια , διορατικότητα για να καταλήξει στην κατάλληλη μέθοδο.

ΑΝΑΦΟΡΕΣ

[1] Armano G, Marchesi M, Murru A, (2005) A hybrid genetic-neural architecture for stock indexes forecasting, Information Sciences (170)1: 3–33

[2] Atsalakis George, Valavanis Kimon, (2009). Surveying stock market forecasting techniques– Part II: Soft computing methods. Expert Systems with Applications (36) 5932–5941

[3] Atsalakis G. & Ucenic C. (2007). Forecasting the tourism demand using ANFIS for assuring successful strategies in the view of sustainable development in the tourism sector. Proceedings of the 2nd IASME/WSEAS International Conference on Energy & Environment, (pp. 213-219). Portorose, Slovenia.

[4] Atsalakis G. and Valavanis K. (2010). Forecasting stock trends using a combined technical analysis and neuro-fuzzy based approach. Journal of Financial Decision Making, 6(1):79-94

[5] Coates Vary, Farooque Mahmud, Klavans Richard, Lapid Koty, Linstone Harold, Pistorius Carl, Porter Alan (2001), On the Future of Technological Forecasting , Technological Forecasting and Social Change (67), 1–17 7. Committee on Forecasting Future Disruptive Technologies; National Research Council (2009)

[6] Persistent Forecasting of Disruptive Technologies, ISBN: 0-309-11661-9, 136 pages

[7] De Gooijer Jan G, Hyndman J. Rob (2006), 25 years of time series forecasting, Journal of Forecasting 10. Firat Ayse Kaya, Lee Wei, Woon Stuart Madnick (2008), Technological Forecasting – A Review. Working Paper CISL# 2008-15, Composite Information Systems Laboratory (CISL) Sloan School of Management, Room E53-320 Massachusetts Institute of Technology Cambridge

[8] Green Kesten, J. Scott Armstrong (2007), Structured analogies for forecasting Business and Economic Forecasting, International Journal of Forecasting (23) 365–376

[9] Hasson F, Keeney S, Mckenna H (2000). Research guidelines for the Delphi survey technique. J Adv Nurs, (32):1008–1015 144

[10] Higgins L. Matthew (2011). Advances in Economic Forecasting. Western Michigan University, Upjohn Institute for Employment Research.

[11] Hyndman R, Athanasopoulos G, (2013), Forecasting: Principles and Practice, https://www.otexts.org/book/fpp.

[12] Kourentzes Nikolaos, Petropoulos Fotios, Trapero Juan R (2013, June). Improving forecasting by estimating time series structural components across multiple frequencies. Paper presented at ISF, Seoul Korea.

[13] Kucharavy Dmitry, Roland De Guio, IAMOT (2008), Technological Forecasting and Assessment of barriers for emerging technologies, International Association for Management of Technology, Proceedings , LGECO - Design Engineering Laboratory, INSA Strasbourg

[14] Lawrence Michael, Goodwin Paul, O'Connor Marcus, O¨nkal Dilek (2006). Judgmental forecasting: A review of progress over the last 25years. International Journal of Forecasting, (22) 493 – 518