



**ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ**  
**ΣΧΟΛΗ ΑΓΡΟΝΟΜΩΝ & ΤΟΠΟΓΡΑΦΩΝ**  
**ΜΗΧΑΝΙΚΩΝ**  
**ΤΟΜΕΑΣ ΣΥΓΚΟΙΝΩΝΙΑΚΩΝ ΕΡΓΩΝ**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΕΠΑΝΑΛΗΠΤΙΚΟΣ ΑΛΓΟΡΙΘΜΟΣ ΜΕΓΙΣΤΟΠΟΙΗΣΗΣ ΤΗΣ  
ΕΝΤΡΟΠΙΑΣ ΓΙΑ ΤΗΝ ΕΚΤΙΜΗΣΗ ΠΙΝΑΚΩΝ Π-Π

ΜΟΥΡΤΖΙΝΗΣ ΠΑΝΑΓΙΩΤΗΣ

ΕΠΙΒΛΕΠΩΝ: Επίκουρος Καθ. ΒΥΘΟΥΛΚΑΣ ΠΕΤΡΟΣ

ΑΘΗΝΑ, ΑΠΡΙΛΙΟΣ 2008

## Περίληψη

Οι πίνακες Προέλευσης-Προορισμού είναι απαραίτητοι στα μοντέλα προσομοίωσης για τον προσδιορισμό της συμπεριφοράς μέσα σε δίκτυο. Για ένα μεγάλο αριθμό εφαρμογών οι συμβατικές μέθοδοι για την εκτίμηση του πίνακα Π-Π είναι πολύ ακριβές για να χρησιμοποιηθούν. Για να αντιμετωπιστεί το πρόβλημα αυτό, αναπτύχθηκαν εναλλακτικές μέθοδοι, για την εκτίμηση του πίνακα Π-Π από μετρήσεις κυκλοφοριακών φόρτων σε οδικούς συνδέσμους. Σε αυτή τη διπλωματική εργασία οι εναλλακτικές μέθοδοι που εξετάζονται είναι το μοντέλο μεγιστοποίησης της εντροπίας και ο γενικευμένος αλγόριθμος εκτίμησης πινάκων Π-Π για γραμμικά δίκτυα. Στην συνέχεια αναπτύσσεται μια εναλλακτική μέθοδος εκτίμησης των πινάκων Π-Π. Όλα τα μοντέλα εφαρμόστηκαν σε πραγματικά δεδομένα. Έγινε αξιολόγηση, ανάλυση και σύγκριση των αποτελεσμάτων που προέκυψαν ώστε να εντοπιστούν τα πλεονεκτήματα και μειονεκτήματα των μεθόδων αυτών.

## **Abstract**

Origin-Destination matrices are required in order to model traffic routing behavior in networks. For a large number of applications conventional methods for estimating an origin destination matrix become too expensive to use. To overcome this problem, alternative methods for estimating an origin destination matrix from traffic counts, have been developed. In this thesis, two alternative methods are investigated; the maximum entropy model and the generalized algorithm for estimating transit ODs from traffic counts. A new algorithm for estimating an origin destination matrix is then described. All the models have been tested with real data. Then we evaluated and compared the results in order to find the advantages and disadvantages of the models.

Περίληψη .....	2
Abstract .....	3
1. Πίνακες Προέλευσης-Προορισμού .....	6
1.1. Εισαγωγικά .....	6
1.2. Στόχοι της διπλωματικής εργασίας .....	7
1.3. Δομή της διπλωματικής εργασίας .....	8
2. Το πρόβλημα της εκτίμησης πινάκων Π-Π .....	11
2.1. Εισαγωγή.....	11
2.2. Ανεξαρτησία και ασυνέπεια των μετρήσεων κυκλοφορίας.....	13
2.3. Προσέγγιση της ελαχιστοποίησης της πληροφορίας.....	15
2.4. Μια φόρμα μεγιστοποίησης της εντροπίας.....	18
2.4.1 Μερικές ιδιότητες του μοντέλου.....	24
2.5. Επιλύοντας το πολυαναλογικό πρόβλημα .....	26
3. Εναλλακτικές μορφές του βασικού μοντέλου μεγιστοποίησης της εντροπίας.....	28
3.1. Το βασικό μοντέλο εντροπίας.....	28
3.2. Μοντέλο 2 .....	29
3.3. Μοντέλο 3 .....	30
3.4. Μοντέλο 4 .....	31
3.5. Μοντέλο 5 .....	32
3.6. Μοντέλο 6 .....	33
3.7. Μοντέλο 7 .....	34
3.8. Μοντέλο 8 .....	36
3.9. Σύνοψη.....	36
4. Ένας γενικευμένος και αποτελεσματικός αλγόριθμος εκτίμησης πινάκων Π-Π από μετρήσεις επιβατών.....	39
4.1. Εισαγωγή.....	39
4.2. Βασικές παραδοχές .....	40
4.3. Τυποποίηση του αλγορίθμου .....	42
4.4. Συμπεράσματα .....	45
5. Εφαρμογές.....	46
5.1. Εισαγωγή.....	46
5.2. Στοιχεία της έρευνας για τα δεδομένα .....	49
5.2.1. Στοιχεία υποδομής .....	49
5.2.2. Μεθοδολογία έρευνας.....	49
5.2.3. Σκοπός της έρευνας.....	50
5.2.4. Χαρακτηριστικά λειτουργίας.....	51
5.3. Εφαρμογή του αλγορίθμου Li, Cassidy .....	54
5.4. Εφαρμογή των μοντέλων μεγιστοποίησης εντροπίας.....	67
5.4.1. Matlab .....	67
5.4.2. Αρχικός πίνακας.....	68
5.4.3. Αποτελέσματα.....	71
5.4.4. Ακρίβεια των μοντέλων σε σχέση με την κλίμακα του προβλήματος.....	81
5.4.5. Συμπεράσματα .....	86
5.5. Συμπεράσματα .....	88
6. Επαναληπτικός αλγόριθμος μεγιστοποίησης της εντροπίας.....	92
6.1. Εισαγωγή.....	92
6.2. Ο αλγόριθμος .....	92
6.3. Γενικευμένη μορφή του επαναληπτικού αλγορίθμου.....	99
6.4. Συμπεράσματα για τον επαναληπτικό αλγόριθμο .....	103
7. Εφαρμογή του επαναληπτικού αλγορίθμου μεγιστοποίησης της εντροπίας .....	104

7.1. Εφαρμογή.....	104
7.2. Συγκρίσεις.....	106
7.3. Εφαρμογή σε διαφορετικό πρόβλημα.....	108
7.4. Συμπεράσματα.....	111
8. Συμπεράσματα και προτάσεις.....	112
8.1. Συμπεράσματα.....	112
8.2. Προτάσεις.....	113
Βιβλιογραφικές αναφορές.....	114
Παράρτημα Α.....	115
Παράρτημα Β.....	117
Ο αλγόριθμος.....	117
Παρατηρήσεις.....	117
Ιδιότητες σύγκλισης.....	118
Προέλευση από την μέθοδο του Newton.....	119
Παράρτημα Γ.....	121
Το πρόβλημα.....	121
Η λύση.....	121
Επιλογή της παραμέτρου $\lambda$ .....	122

# 1. Πίνακες Προέλευσης-Προορισμού

## 1.1. Εισαγωγικά

Η ανάγκη για μετακίνηση είναι αποτέλεσμα της διασποράς των χρήσεων γης στο χώρο. Στην σημερινή εποχή οι μετακινήσεις έχουν αυξηθεί κατά πολύ, λόγω των ραγδαίων εξελίξεων και την ανάπτυξη σε όλους τους τομείς. Η αύξηση αυτή των μετακινήσεων όμως δημιουργεί πολλά κυκλοφοριακά προβλήματα ενώ η ανεκτικότητα της κυκλοφοριακής συμφόρησης και του χαμηλού επιπέδου εξυπηρέτησης ολοένα και μειώνεται. Η απάντηση σε αυτά τα προβλήματα δεν μπορεί να είναι η συνεχής κατασκευή καινούργιων υποδομών. Θα πρέπει κατ' αρχήν να γίνει πλήρης καταγραφή των προβλημάτων και των αιτιών που τα προκαλούν. Στη συνέχεια με κατάλληλη επεξεργασία των δεδομένων και την ανάπτυξη κατάλληλων μοντέλων θα πρέπει τα μεταφορικά συστήματα να βελτιστοποιηθούν. Αυτό φυσικά προϋποθέτει συστηματική θεώρηση και ανάλυση των προβλημάτων.

Η βασική πληροφορία για την ανάλυση αυτή είναι ο πίνακας Προέλευσης-Προορισμού. Τα μοντέλα προσομοίωσης, έχοντας την πληροφορία αυτή και τα χαρακτηριστικά του δικτύου προσπαθούν να προβλέψουν που θα επικρατήσουν συνθήκες κυκλοφοριακές συμφόρησης.

Οι πίνακες Π-Π είναι δύο διαστάσεων και περιέχουν πληροφορίες για τις μετακινήσεις που γίνονται μεταξύ των διαφόρων ζωνών της περιοχής μελέτης. Στις γραμμές του πίνακα δίνονται οι ζώνες προέλευσης και στις στήλες οι ζώνες προορισμού. Οι τιμές των κελιών έχουν κατευθυντικό νόημα δείχνοντας τον αριθμό των μετακινήσεων από την ζώνη προορισμού στη ζώνη προέλευσης. Τα κελία της διαγωνίου αναπαριστούν τις μετακινήσεις στο εσωτερικό κάθε ζώνης.

Η συγκέντρωση των απαραίτητων πληροφοριών ώστε να δημιουργηθούν αυτοί οι πίνακες μπορεί να γίνει με ποικίλους τρόπους. Η συμβατική μέθοδος για την εύρεση ενός πίνακα Προέλευσης-Προορισμού που αναπτύχθηκε για μεγάλης κλίμακας έρευνες χρησιμοποιεί ένα συνδυασμό από κατ' οίκον συνεντεύξεις και έρευνα πεδίου.

Εντούτοις, το υψηλό οικονομικό κόστος, οι μεγάλες απαιτήσεις σε ανθρώπινο δυναμικό, το μεγάλο χρονικό διάστημα που χρειάζεται για την επεξεργασία των δεδομένων και τα σφάλματα που προκύπτουν από την δειγματοληψία αποκλείει την χρήση αυτής της προσέγγισης για τις περισσότερες εφαρμογές. Εναλλακτικές μέθοδοι που προτείνονται για μικρότερης κλίμακας έρευνες είναι η περιορισμένη έρευνα πεδίου, έρευνες πινακίδων κυκλοφορίας, αεροφωτογραφίες, αλλά θεωρούνται επίσης υψηλού κόστους, όσο αναφορά τις απαιτήσεις σε ανθρώπινο δυναμικό και την επεξεργασία δεδομένων. Όλες αυτές οι μέθοδοι όμως, με πιθανή εξαίρεση τις αεροφωτογραφίες, βασίζονται σε δειγματοληψία και κατά συνέπεια για την εκτίμηση του πλήρους πίνακα Π-Π σημαίνει αναγωγή των στοιχείων στο σύνολο του πληθυσμού. Επίσης όλες οι παραπάνω μέθοδοι έχουν ένα ακόμα σημαντικό μειονέκτημα. Δεν έχουν την ικανότητα να λαμβάνουν υπόψη τις μεταβολές διαφόρων παραγόντων (αλλαγή αριθμού μετακινήσεων, αλλαγή στην γεωμετρία του δικτύου,...) που έχουν άμεση επίπτωση στην κατανομή των μετακινήσεων. Ως εκ τούτου προηγούμενα μετρημένοι πίνακες Π-Π θεωρούνται γρήγορα ξεπερασμένοι και προκύπτει η ανάγκη να επαναληφθεί εκ νέου μία από τις προαναφερθείσες μεθόδους που όπως προείπαμε έχουν απαγορευτικό κόστος.

Επειδή όμως σε κάθε έρευνα ή μελέτη υπάρχει περιορισμός σε χρήματα, χρόνο και εργατικό δυναμικό οι ερευνητές άρχισαν να ψάχνουν εναλλακτικές μεθόδους προσδιορισμού αυτών των πινάκων. Λαμβάνοντας υπόψη τα παραπάνω συμπεράνουμε την αναγκαιότητα χρήσης απλοποιημένων μεθόδων για τον εύκολο, γρήγορο και πιο οικονομικό προσδιορισμό πινάκων Π-Π. Αρκετές τέτοιες μέθοδοι βασίζονται σε μετρήσεις φόρτων κυκλοφορίας στους οδικούς συνδέσμους, ένα από τα πιο κοινά στοιχεία κυκλοφοριακής πληροφορίας. Τα στοιχεία αυτά είναι σχετικά ανέξοδο να συλλεχθούν και συνήθως χρησιμοποιούνται για διάφορους σκοπούς (μελέτες ατυχημάτων, σχεδιασμούς συντήρησης, βελτίωση κόμβων).

## **1.2. Στόχοι της διπλωματικής εργασίας**

Όπως περιγράψαμε και παραπάνω, η γρήγορη και ακριβής εκτίμηση των πινάκων Π-Π είναι αναγκαία για αποτελεσματικό σχεδιασμό και διαχείριση των συστημάτων μεταφορών. Σε αυτή την εργασία για την εκτίμηση των πινάκων Π-Π θα

χρησιμοποιήσουμε μοντέλα μεγιστοποίησης της εντροπίας και ένα γενικευμένο αλγόριθμο εκτίμησης πινάκων Π-Π για γραμμικά δίκτυα.

Οι στόχοι της διπλωματικής εργασίας είναι δύο:

- Να αξιολογήσει τις μεθόδους προσδιορισμού πινάκων Π-Π από μετρήσεις φόρτων σε οδικούς συνδέσμους και να εντοπίσει τα πλεονεκτήματα και μειονεκτήματα των μεθόδων αυτών.
- Με βάση τα συμπεράσματα αυτής της αξιολόγησης, να προτείνει εναλλακτικές μορφές μοντέλων που παρέχουν μεγαλύτερη ακρίβεια.

### **1.3. Δομή της διπλωματικής εργασίας**

Για να επιτύχουμε τους στόχους που περιγράψαμε, έγινε μια ανασκόπηση της βιβλιογραφίας ώστε να κατανοήσουμε την θεωρία που βασίζονται οι προσεγγίσεις για την εκτίμηση πινάκων Π-Π από μετρήσεις κυκλοφοριακών φόρτων σε οδικούς συνδέσμους. Στην συνέχεια έγινε αξιολόγηση αυτών των προσεγγίσεων και προτείνεται μια εναλλακτική μορφή που προσφέρει μεγαλύτερη ακρίβεια. Η δομή της εργασίας παρουσιάζεται στο παρακάτω διάγραμμα ροής.

Στο κεφάλαιο 2 αναπτύσσεται η θεωρία της ελαχιστοποίησης της πληροφορίας και μεγιστοποίησης της εντροπίας και καταλήγουμε στην αντικειμενική συνάρτηση του βασικού μοντέλου.

Στο κεφάλαιο 3 παρουσιάζονται κάποιες εναλλακτικές μορφές του βασικού μοντέλου μεγιστοποίησης της εντροπίας.

Στο κεφάλαιο 4 αναπτύσσεται ο γενικευμένος αλγόριθμος εκτίμησης πινάκων Π-Π για γραμμικά δίκτυα.

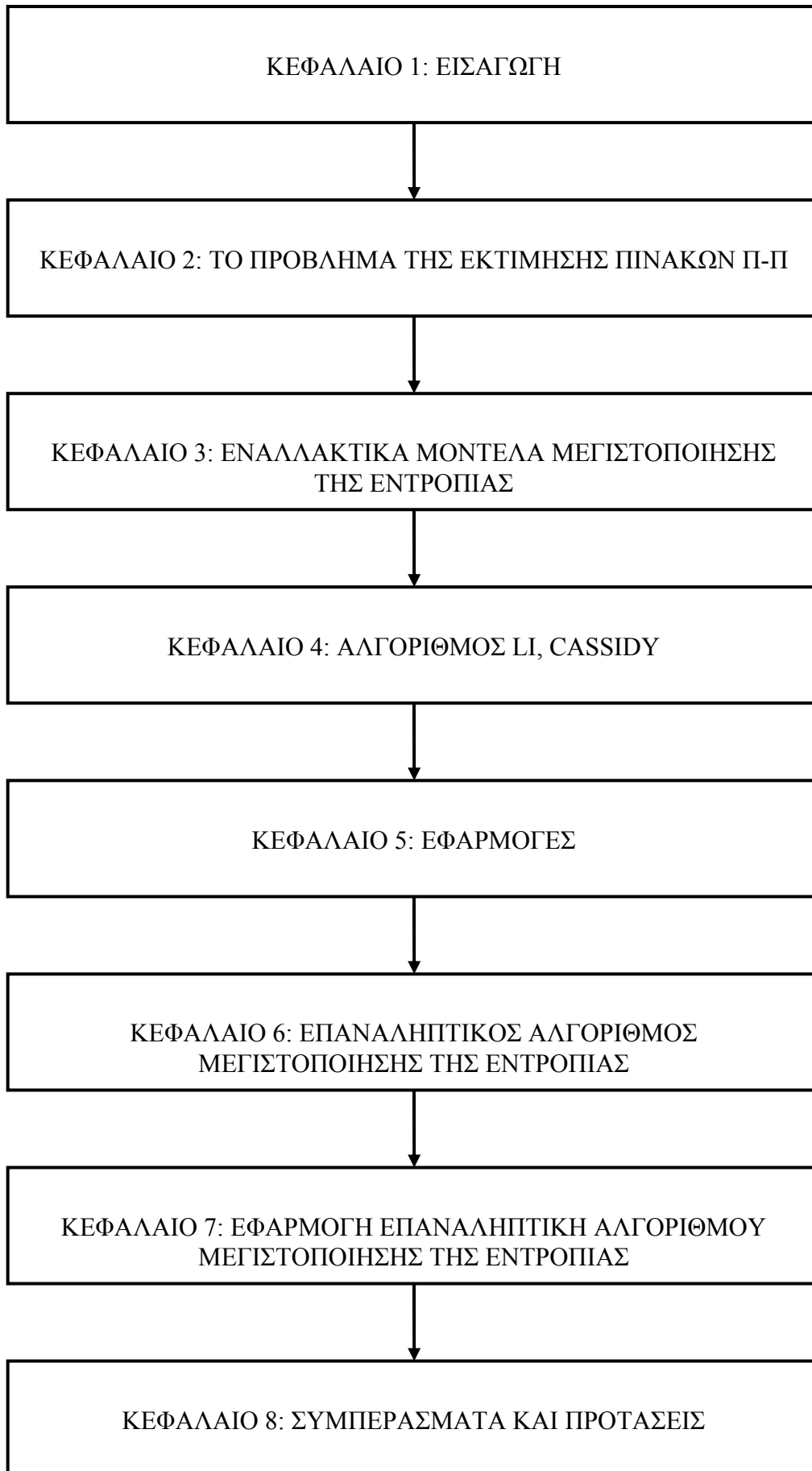


Στο κεφάλαιο 5 γίνεται εφαρμογή των μοντέλων μεγιστοποίησης της εντροπίας και του γενικευμένου αλγορίθμου των Li, Cassidy σε ένα πραγματικό πρόβλημα ώστε να αξιολογήσουμε την ακρίβεια των αποτελεσμάτων που δίνουν.

Στο κεφάλαιο 6 προτείνεται ένας εναλλακτικός τρόπο προσέγγισης του προβλήματος, αναπτύσσοντας έναν δικό μας επαναληπτικό αλγόριθμο μεγιστοποίησης της εντροπίας για την επίλυση του προβλήματος εκτίμησης πινάκων Π-Π.

Στο κεφάλαιο 7 γίνεται εφαρμογή του επαναληπτικού αλγορίθμου μεγιστοποίησης της εντροπίας.

Στο κεφάλαιο 8 συγκεντρώνουμε τα συμπεράσματα που έχουν προκύψει στα προηγούμενα κεφάλαια και κάνουμε προτάσεις για περαιτέρω έρευνα.



## 2. Το πρόβλημα της εκτίμησης πινάκων Π-Π

### 2.1. Εισαγωγή

Ας θεωρήσουμε μια περιοχή μελέτης η οποία έχει χωριστεί σε  $n$  ζώνες. Τις μετακινήσεις που γίνονται στο εσωτερικό των ζωνών αυτών δεν τις λαμβάνουμε υπόψη μας. Οι ζώνες αυτές μπορεί να αντιπροσωπεύουν χωριά, πόλεις ή μέρη πόλεων, συγκεκριμένες περιοχές μεγάλου ενδιαφέροντος ή ακόμα και στάσεις Μέσων Μαζικής Μεταφοράς.

Ένα πίνακας Π-Π είναι δύο διαστάσεων και ο αριθμός των μετακινήσεων από τη ζώνη  $i$ , που είναι η προέλευση (Origin) στη ζώνη  $j$ , που είναι ο προορισμός (Destination) συμβολίζεται με  $T_{ij}$ . Επίσης για κάθε πίνακα ισχύει ότι το άθροισμα όλων των μετακινήσεων είναι ίσο με το άθροισμα των παραγόμενων μετακινήσεων και με το άθροισμα των ελκυόμενων μετακινήσεων. Δηλαδή ισχύει:

$$\sum_i \sum_j T_{ij} = \sum_i O_i = \sum_j D_j \quad (2.1.1)$$

Μια πολύ σημαντική παράμετρος στην εκτίμηση πίνακα Π-Π από μετρήσεις κυκλοφορίας είναι προσδιορισμός των ζευγών προέλευσης-προορισμού των οποίων οι μετακινήσεις χρησιμοποιούν μία συγκεκριμένη διαδρομή. Έστω η μεταβλητή  $p_{ij}^a$  που αναπαριστά την πιθανότητα να γίνει μια μετακίνηση από την ζώνη  $i$  στην ζώνη  $j$  χρησιμοποιώντας τον οδικό σύνδεσμο  $a$ . Γενικά ισχύει:

$$0 \leq p_{ij}^a \leq 1 \quad (2.1.2)$$

$$\sum_k p_{ij}^k = 1 \quad \text{όπου } k \text{ είναι όλες οι διαδρομές των συνδυασμών του } i \text{ με το } j.$$

Αν θεωρήσουμε τον φόρτο στον οδικό σύνδεσμο  $a$  ως  $V_a$ , η θεμελιώδης εξίσωση για την εκτίμηση ενός πίνακα Προέλευσης-Προορισμού από μετρήσεις κυκλοφορίας είναι:

$$V_a = \sum_i \sum_j p_{ij}^a \cdot T_{ij} \quad (2.1.3)$$

Το πρότυπο καταμερισμού στο δίκτυο (traffic model) που θα χρησιμοποιήσουμε, ώστε η εκτίμηση που θα κάνουμε για τον πίνακα να είναι ρεαλιστική, είναι πολύ σημαντικός για τον προσδιορισμό του  $p_{ij}^a$ . Τα πρότυπα καταμερισμού στο δίκτυο χωρίζονται γενικά στις ακόλουθες κατηγορίες (Sheffi).

- Αναλογικός καταμερισμός (proportional assignment). Στη συγκεκριμένη περίπτωση η  $p_{ij}^a$  θεωρούμε ότι είναι ανεξάρτητη από τους κυκλοφοριακούς φόρτους (π.χ. περίπτωση κυκλοφοριακής συμφόρησης). Δηλαδή ο αριθμός των οδηγών που επιλέγουν κάθε διαδρομή εξαρτάται από τα χαρακτηριστικά της διαδρομής και τις υποθέσεις που κάνει ο κάθε οδηγός και όχι από τα επίπεδα του φόρτου κυκλοφορίας. Η τιμή της  $p_{ij}^a$  μπορεί να καθοριστεί ανεξάρτητα από τον πίνακα Π-Π πριν ακόμα γίνει οποιαδήποτε εκτίμηση του. Οι περισσότεροι αλγόριθμοι καταμερισμού ανήκουν σε αυτή την κατηγορία (Burnel, Dial). Ο αλγόριθμος όλα η τίποτα (all-or-nothing) είναι μια ιδιαίτερη περίπτωση αναλογικού καταμερισμού ο οποίος έχει την εξής μορφή:

$$p_{ij}^a = \begin{cases} 0 & \text{αν η διαδρομή } a \text{ δεν χρησιμοποιείται} \\ 1 & \text{αν η διαδρομή } a \text{ χρησιμοποιείται} \end{cases}$$

- Καταμερισμός περιορισμού χωρητικότητας (Capacity restrained assignment). Η συγκεκριμένη κατανομή θεωρείται πιο ρεαλιστική στην περίπτωση που η επίδραση της κυκλοφοριακής συμφόρησης θεωρείται η πιο σημαντική παράμετρος στην λήψη της απόφασης για το ποια διαδρομή θα επιλεγεί από τον κάθε οδηγό. Σε αυτές τις περιπτώσεις χρησιμοποιούνται συχνά μέθοδοι καταμερισμού ισορροπίας (traffic equilibrium assignment). Η τιμή της  $p_{ij}^a$  εξαρτάται από τα επίπεδα του φόρτου κάθε διαδρομής και δεν μπορεί να προσδιοριστεί ανεξάρτητα από την διαδικασία εκτίμησης του πίνακα Π-Π.

Αν υποθέσουμε ότι γνωρίζουμε τις τιμές της  $p_{ij}^a$  για κάθε διαδρομή και κάθε ζεύγος  $i, j$  του πίνακα, χρειαζόμαστε  $N^2$  ανεξάρτητες μετρήσεις για να καθορίσουμε μοναδικά ένα πίνακα Π-Π διαστάσεων  $N \times N$ . Πρακτικά όμως στις περισσότερες περιπτώσεις ο αριθμός των μετρήσεων που θα είναι διαθέσιμος θα είναι πολύ μικρότερος.

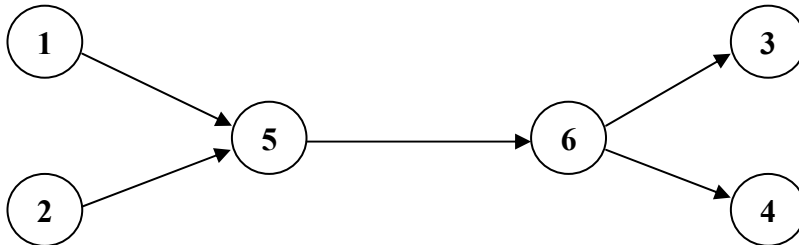
## 2.2. Ανεξαρτησία και ασυνέπεια των μετρήσεων κυκλοφορίας

Όπως διαπιστώσαμε για το πρόβλημα, προκύπτει ότι η ακρίβεια εξαρτάται από την ακρίβεια των μετρήσεων των κυκλοφοριακών φόρτων. Επισημαίνεται όμως ότι οι μετρήσεις κυκλοφορίας εμπεριέχουν σφάλματα. Κάποιες από τις σημαντικές πηγές σφαλμάτων είναι η απλοποιημένη περιγραφή του δικτύου, τα λάθη στις μετρήσεις και το γεγονός ότι δεν γίνονται όλες οι μετρήσεις την ίδια ώρα.

Επίσης υπάρχουν κάποια προβλήματα που σχετίζονται με την χρήση τους στην εκτίμηση των πινάκων Π-Π. Τα δύο πιο σημαντικά από αυτά είναι:

- **Ανεξαρτησία μετρήσεων.** Κάποιες μετρήσεις ίσως αποτύχουν να προσθέσουν κάποια πληροφορία. Δεν θα είναι όλες οι εξισώσεις (1) ανεξάρτητες αφού για κάθε κόμβο  $m$  ισχύει η αρχή της κυκλοφοριακής ροής που εκφράζεται από την σχέση  $\sum_l V_{lm} - \sum_k V_{mk} = 0$ , όπου  $V_{lm}$  είναι η ροή προς τον κόμβο και  $V_{mk}$  η ροή κυκλοφορίας από τον κόμβο. Κατά συνέπεια, εάν γνωρίζουμε όλες τις ροές από τον κόμβο  $m$  και τις ροές προς τον κόμβο εκτός ενός, τότε αυτόματα μπορούμε να προσδιορίσουμε αυτή την ροή. Για παράδειγμα, στο παρακάτω σχήμα 2.1 ο φόρτος  $V_{56}$  δεν περιέχει επιπλέον πληροφορία αν γνωρίζουμε ήδη τους φόρτους  $V_{15}$  και  $V_{25}$ .
- **Ασυνέπεια μετρήσεων.** Λάθη και ασυγχρόνιστες μετρήσεις είναι πιθανό να οδηγήσουν σε ασυνέπεια των ροών κυκλοφορίας. Τεχνικές για την αντιμετώπιση αυτού του προβλήματος υπάρχουν αλλά δεν το εξαλείφουν. Η πιο συνηθισμένη μέθοδος που χρησιμοποιείτε για την αντιμετώπιση του προβλήματος της ασυνέπειας της ροής είναι η δημιουργία μοντέλων

εκτίμησης του πίνακα Π-Π που να δέχονται ασυνεχείς ροές κυκλοφορίας. Παραδείγματος χάριν στο σχήμα μπορεί να ισχύει  $V_{56} \neq V_{15} + V_{25}$  αν οι μετρήσεις στους οδικούς συνδέσμους 56, 15, 25 δεν έχουν γίνει την ίδια χρονική περίοδο.



**Σχήμα 2.1**

Γενικά, πρέπει να αναφέρουμε ότι δεν είναι δυνατόν καθοριστεί μοναδικά ένας πίνακας Π-Π από μετρήσεις κυκλοφορίας και μόνο. Αυτό συμβαίνει γιατί συνήθως ο αριθμός  $(N^2 - N)$  των άγνωστων τιμών  $(T_{ij})$ , για  $\forall i, j$  και  $i \neq j$ , ενός πίνακα είναι κατά πολύ μεγαλύτερος απ' ότι οι μετρήσεις φόρτων και κατ' επέκταση οι σχέσεις που υπάρχουν.

Οι περισσότερες από τις εναλλακτικές προσεγγίσεις για το πρόβλημα της εκτίμησης του πίνακα Π-Π από μετρήσεις κυκλοφορίας υποθέτουν ότι η συμπεριφορά αυτών που μετακινούνται μπορεί να αντιπροσωπευτεί από τα μοντέλα βαρύτητας. Με αυτό τον τρόπο ο αριθμός των αγνώστων μειώνεται ραγδαία και συχνά το πρόβλημα γίνεται υπερορισμένο αντί για υποορισμένο. Ο Robillard (1975) πρότεινε ένα γενικευμένο μοντέλο βαρύτητας. Ο Low (1972), Hogberg (1976) και ο Holm et al. (1976) πρότειναν πιο συμβατικά μοντέλα βαρύτητας που απαιτούσαν επιπλέον δεδομένα, τα οποία όμως είναι εύκολο να συγκεντρωθούν. Ο Nguyen (1977) πρότεινε μια προσέγγιση καταμερισμού ισορροπίας, αλλά η μέθοδος αυτή απέτυχε να βρει ένα μοναδικό πίνακα εξαιτίας του υποορισμού του προβλήματος.

Η μεγαλύτερη κριτική που δέχονται τα παραπάνω μοντέλα είναι ότι αναγκάζοντας τον πίνακα των μετακινήσεων να ακολουθήσει τα πρότυπα των μοντέλων βαρύτητας, δεν χρησιμοποιείται όλη την πληροφορία που περιέχεται στις μετρήσεις κυκλοφορίας.

Μια καλύτερη προσέγγιση θα ήταν να λυθεί το υποορισμένο πρόβλημα με την ελαχιστοποίηση της εξωτερικής πληροφορίας. Αυτό μπορεί να γίνει με δύο τρόπους: πρώτον, με ελαχιστοποίηση της πληροφορίας και δεύτερον, με μεγιστοποίηση της εντροπίας (Henk J. Van Zuylen, Luis G. Willumsen, 1980).

### 2.3. Προσέγγιση της ελαχιστοποίησης της πληροφορίας

Ένα πιθανό αρχικό σημείο είναι να υπολογίσουμε την πληροφορία που περιέχεται σε ένα πίνακα μετακινήσεων  $T_{ij}$ . Από την στιγμή που η διαθέσιμη πληροφορία που βρίσκεται στις μετρήσεις κυκλοφορίας είναι ανεπαρκής για τον προσδιορισμό ενός πίνακα μετακινήσεων, είναι λογικό να επιλεγεί ένας πίνακας μετακινήσεων ο οποίος προσθέτει όσο το δυνατόν λιγότερη πληροφορία στην γνώση που περιέχεται στην εξίσωση (1). Η προσέγγιση αυτή έχει ακολουθηθεί από τον Zuylen (1978) χρησιμοποιώντας την μέτρηση πληροφορίας του Brillouin (1956).

Η πληροφορία η οποία υπάρχει σε ένα σετ  $N$  παρατηρήσεων όπου η κατάσταση  $k$  έχει παρατηρηθεί  $n_k$  φορές έχει οριστεί από τον Brillouin (1956) ως:

$$I = \log_e N! \pi_k \left[ \frac{q_k^{n_k}}{n_k!} \right] \quad (2.3.1)$$

όπου  $q_k$  είναι η *a priori* πιθανότητα να παρατηρηθεί η κατάσταση  $k$ . Αν οι παρατηρήσεις είναι μετρήσεις σε κάποια συγκεκριμένη διαδρομή είναι πιθανό να ορίσουμε την κατάσταση  $ij$  σαν την κατάσταση στην οποία τα οχήματα τα οποία έχουν παρατηρηθεί έχουν ταξιδέψει μεταξύ της προέλευσης  $i$  και του προορισμού  $j$ . Έτσι:

$$n_{ij}^a = T_{ij} p_{ij}^a \quad (2.3.2)$$

Μπορούμε επίσης να εκφράσουμε την *a priori* πιθανότητα να παρατηρηθεί η κατάσταση  $ij$  σε μία διαδρομή  $a$  ως την λειτουργία της *a priori* πληροφορίας του πίνακα Π-Π σαν

$$q_{ij}^a = \frac{t_{ij} p_{ij}^a}{\sum_{ij} t_{ij} p_{ij}^a} \quad (2.3.3)$$

όπου  $t_{ij}$  είναι ο *a priori* αριθμός των μετακινήσεων μεταξύ των  $i$  και  $j$  που δίνονται, για παράδειγμα, έναν παλαιό πίνακα Π-Π.

Η πληροφορία που περιέχεται στους φόρτους  $V_a$  ενός οδικού συνδέσμου είναι

$$I_a = -\log_e V_a! \prod_{ij} \left\{ \frac{t_{ij} p_{ij}^a}{S^a} \right\}^{T_{ij} p_{ij}^a} / \prod_{ij} (T_{ij} p_{ij}^a)! \quad (2.3.4)$$

όπου

$$S^a = \sum_{ij} t_{ij} p_{ij}^a \quad (2.3.5)$$

Χρησιμοποιώντας τον μετασχηματισμό του Stirling (Παράρτημα Α) προκύπτει

$$I_a \cong \sum_{ij} T_{ij} p_{ij}^a \log_e \frac{T_{ij} S^a}{V_a t_{ij}} \quad (2.3.6)$$

Αθροίζοντας όλες τις διαδρομές του δικτύου με μετρήσεις προκύπτει

$$I_a = \sum_a \sum_{ij} T_{ij} p_{ij}^a \log_e \frac{T_{ij} S^a}{V_a t_{ij}} \quad (2.3.7)$$

που είναι η συνολική πληροφορία που περιέχεται στις κυκλοφοριακούς φόρτους που έχουν μετρηθεί. Το πρόβλημα της εύρεσης του πίνακα Π-Π που να είναι συνεπής με



τις παρατηρήσεις και να προσθέτει μια ελάχιστη καινούργια πληροφορία, είναι το ίδιο με το να ελαχιστοποιήσουμε το  $I_a$  υπό τους περιορισμούς των φόρτων κυκλοφορίας (2.1.2).

Την λύση σε αυτό το πρόβλημα μπορούμε να την βρούμε διαφορίζοντας τον Lagrangean.

$$\mathfrak{L} = \sum_a \sum_{ij} T_{ij} p_{ij}^a \log_e \frac{T_{ij} S^a}{V_a t_{ij}} + \sum_a \lambda_a \left( \sum_{ij} T_{ij} p_{ij}^a - V_a \right) \quad (2.3.8)$$

όπου  $\lambda_a$  είναι ο πολλαπλασιαστής του Lagrangean για τον φόρτο στο  $a$ . Κατά συνέπεια:

$$\frac{\delta \mathfrak{L}}{\delta T_{ij}} = \sum_a p_{ij}^a \log_e \left[ \frac{T_{ij} S^a}{V_a t_{ij}} \right] + \sum_a T_{ij} p_{ij}^a \frac{1}{T_{ij}} + \sum_a \lambda_a p_{ij}^a \quad (2.3.9)$$

$$\frac{\delta \mathfrak{L}}{\delta T_{ij}} = \sum_a p_{ij}^a \log_e \left[ \frac{T_{ij} S^a}{V_a t_{ij}} \right] + \sum_a p_{ij}^a (1 + \lambda_a) = 0 \quad (2.3.10)$$

$$\log_e \left\{ \prod_a \left[ \frac{T_{ij} S^a}{V_a t_{ij}} \right]^{p_{ij}^a} \right\} = - \sum_a p_{ij}^a (1 + \lambda_a) \quad (2.3.11)$$

$$\left( \frac{T_{ij}}{t_{ij}} \right)^{\sum_a p_{ij}^a} = \prod_a \left( \frac{V_a}{S^a} \right)^{p_{ij}^a} \exp \left[ - \sum_a p_{ij}^a (1 + \lambda_a) \right] \quad (2.3.12)$$

Θέτοντας

$$g_{ij} = \sum_a p_{ij}^a \quad (2.3.13)$$

προκύπτει

$$T_{ij} = t_{ij} \prod_a \left\{ \frac{V_a}{S^a} e^{-(1+\lambda_a)} \right\}^{p_{ij}^a / g_{ij}} \quad (2.3.14)$$

Αντικαθιστώντας

$$X_a = \frac{V_a}{S^a} e^{-(1+\lambda_a)} \quad (2.3.15)$$

προκύπτει ότι

$$T_{ij} = t_{ij} \prod_a X_a^{p_{ij}^a / g_{ij}} \quad (2.3.16)$$

όπου το  $X_a$  μπορεί να υπολογιστεί από την εξίσωση (1). Τώρα το πρόβλημα έχει την μορφή ενός πολυαναλογικού μοντέλου (multi-proportional model).

## 2.4. Μια φόρμα μεγιστοποίησης της εντροπίας

Ο Willumsen (1978a, 1978b) αντιμετώπισε το ίδιο πρόβλημα ακολουθώντας μια προσέγγιση μεγιστοποίησης της εντροπίας και παρήγαγε ένα παρόμοιο μοντέλο.

Πριν προχωρήσουμε όμως στην παρουσίαση του μοντέλου θα θέλαμε να αποσαφηνίσουμε την φύση του προβλήματος. Εμείς θέλουμε να εξετάσουμε ένα σύστημα το οποίο αποτελείται από ένα δίκτυο και έναν αριθμό ανεξάρτητων ταξιδιωτών που μετακινούνται σε αυτό. Το ζητούμενο για εμάς είναι να προσδιορίσουμε τον πίνακα Π-Π για το σύστημα αυτό. Μοναδική λύση όμως δεν μπορεί να υπάρξει λόγω του υποορισμού του προβλήματος, και επομένως είμαστε αναγκασμένοι να δεχτούμε ως λύση τον πιο πιθανό πίνακα Π-Π. Οι μεταβλητές του συστήματος μας είναι ο αριθμός των μετακινήσεις  $T_{ij}$ , που γίνονται από το  $i$  στο  $j$ . Έτσι εμείς θέλουμε να προσδιορίσουμε τον πιο πιθανό πίνακα Π-Π, τον πίνακα δηλαδή ο οποίος προκύπτει από τους περισσότερους δυνατούς συνδυασμούς των

μετακινήσεων. Ο αριθμός των διαφορετικών συνδυασμών των στοιχείων ενός συνόλου μπορεί να υπολογιστεί, χρησιμοποιώντας την θεωρία της Συνδυαστικής Μαθηματικής.

Ένα χαρακτηριστικό της Συνδυαστικής Μαθηματικής είναι ο Δυωνυμικός Συνδυαστικός Συντελεστής (Binomial coefficient). Εάν έχουμε ένα μη αρνητικό ακέραιο  $n$  και έναν ακέραιο  $k$ , ο συντελεστής αυτός ορίζεται ως ο φυσικός αριθμός

$$\binom{n}{k} = \frac{n \cdot (n-1) \Lambda (n-k+1)}{k \cdot (k-1) \Lambda 1} = \frac{n!}{k!(n-k)!} \quad \text{εάν } n \geq k \geq 0$$

$$\binom{n}{k} = 0 \quad \text{εάν } k < 0 \text{ ή } k > n$$

Μια επέκταση του Δυωνυμικού Συνδυαστικού Συντελεστή είναι ο Πολυωνυμικός Συνδυαστικός Συντελεστής (Multinomial coefficient) για τον οποίο ισχύει

$$\binom{n}{k_1, k_2, \dots, k_m} = \frac{n!}{k_1! k_2! \Lambda k_m!} = \binom{k_1}{k_1} \binom{k_1 + k_2}{k_2} \Lambda \binom{k_1 + k_2 + \Lambda + k_m}{k_m} = \prod_{i=1}^m \binom{\sum_{j=1}^i k_j}{k_i}$$

και η τελική του μορφή είναι

$$\prod_{i=1}^m \binom{\sum_{j=1}^i k_j}{k_i} = \prod_{i=1}^m \binom{n}{k_i} = \frac{n!}{\prod_{i=1}^m k_i!}$$

$$\text{με } n = \sum_{i=1}^m k_i$$

Το νόημα του Πολυωνυμικού Συνδυαστικού Συντελεστή είναι ότι μας δείχνει τον αριθμό των διακριτών τρόπων που μπορούν να συνδυαστούν τα  $n$  στοιχεία ενός συνόλου.

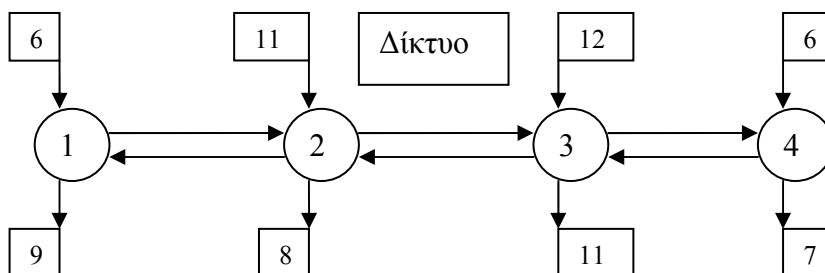
Σύμφωνα λοιπόν με τα παραπάνω ισχύει ότι ο Πολυωνυμικός Συνδυαστικός Συντελεστής για ένα σετ μετακινήσεων  $\{T_{ij}\}$  είναι:

$$W(\{T_{ij}\}) = \frac{T!}{\prod_{ij} T_{ij}!} \quad (2.4.1)$$

όπου  $T$  είναι το άθροισμα των μετακινήσεων, δηλαδή  $T = \sum_{ij} T_{ij}$ .

Σύμφωνα λοιπόν με το θεώρημα αυτό, εάν ο συντελεστής  $W(\{T_{ij}\})$  είναι μεγαλύτερος για ένα σετ μετακινήσεων  $\{T_{ij}\}$  απ' ότι για ένα άλλο, σημαίνει ότι το πρώτο σετ μπορεί να προκύψει από περισσότερους συνδυασμούς μετακινήσεων απ' ότι το δεύτερο και συνεπώς είναι πιο πιθανό να συμβεί.

Στη συνέχεια παρουσιάζουμε μια μικρή εφαρμογή (σχήμα 2.2) για να γίνει πιο κατανοητό. Έστω ότι έχουμε ένα γραμμικό δίκτυο διπλής κατεύθυνσης στο οποίο γνωρίζουμε τον συνολικό αριθμό των μετακινήσεων και τους φόρτους του δικτύου.



**Σχήμα 2.2**

Με βάση λοιπόν τους περιορισμούς που έχουμε (όπως φαίνονται στο σχήμα 2.2), παρουσιάζουμε δύο από τους πιθανούς πίνακες Π-Π που αντιστοιχούν στο δίκτυο μας.

**Πίνακας Α**

	1	2	3	4	sum
1	0	2	3	1	6

<b>2</b>	2	0	5	4	<b>11</b>
<b>3</b>	6	4	0	2	<b>12</b>
<b>4</b>	1	2	3	0	<b>6</b>
<b>sum</b>	<b>9</b>	<b>8</b>	<b>11</b>	<b>7</b>	<b>35</b>

**Πίνακας Β**

	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>sum</b>
<b>1</b>	0	4	1	1	<b>6</b>
<b>2</b>	2	0	6	3	<b>11</b>
<b>3</b>	5	4	0	3	<b>12</b>
<b>4</b>	2	0	4	0	<b>6</b>
<b>sum</b>	<b>9</b>	<b>8</b>	<b>11</b>	<b>7</b>	<b>35</b>

Για τους παραπάνω πίνακες ισχύει:

$$W(\{T_{ij}^A\}) = \frac{T^A!}{\prod_{ij} T_{ij}^A!} = 3.6 \cdot 10^{29}$$

$$W(\{T_{ij}^B\}) = \frac{T^B!}{\prod_{ij} T_{ij}^B!} = 6 \cdot 10^{28}$$

Άρα, σύμφωνα με τα αποτελέσματα, ο πιο πιθανός πίνακας να ισχύει είναι ο Α αφού μπορεί να προκύψει από περισσότερους συνδυασμούς.

Επομένως, για να προσδιορίσουμε τον πιθανότερο πίνακα των μετακινήσεων, ο οποίος είναι και ζητούμενος, θα πρέπει να μεγιστοποιήσουμε το  $W(\{T_{ij}\})$  ή μια μονοτονική συνάρτηση του. Έτσι λογαριθμίζοντας, από μια συνάρτηση γινομένων των αγνώστων καταλήγουμε σε μια συνάρτηση αθροίσματος των αγνώστων που είναι ευκολότερο να αναλυθεί.

$$S = \log_e W(\{T_{ij}\}) \tag{2.4.2}$$

Χρησιμοποιώντας τον μετασχηματισμό του Stirling (Παράρτημα Α), σύμφωνα με τον οποίο ισχύει  $\ln(X!) = X\ln(X) - X$ , το πρόβλημα γίνεται:

$$\text{Max } S = \log_e T! - \sum_{ij} (T_{ij} \log_e T_{ij} - T_{ij}) \quad (2.4.3)$$

Συνεπώς ο αριθμός των μετακινήσεων  $T$  είναι σταθερός και επομένως μπορεί να αποκλειστεί από το πρόβλημα μεγιστοποίησης. Συνεπώς το πρόβλημα γίνεται:

$$\text{maximise } S' = - \sum_{ij} (T_{ij} \log_e T_{ij} - T_{ij}) \quad (2.4.4)$$

με τον περιορισμό  $V_a - \sum_{ij} T_{ij} p_{ij}^a = 0$  (1) για όλες τις μετρήσεις φόρτων σε οδικούς συνδέσμους.

Το παραπάνω πρόβλημα μεγιστοποίησης με περιορισμό μπορεί να επιλυθεί υπολογίζοντας τα στάσιμα σημεία (stationary points) της Lagrangean του προβλήματος. Η Lagrangean εκφράζεται από την ακόλουθη σχέση

$$\mathfrak{L} = \sum_{ij} (T_{ij} \log_e T_{ij} - T_{ij}) + \sum_a \lambda_a \left( V_a - \sum_{ij} T_{ij} p_{ij}^a \right) \quad (2.4.5)$$

και παραγωγίζοντας την παραπάνω σχέση μπορούμε να υπολογίσουμε τα στάσιμα σημεία

$$\frac{\delta \mathfrak{L}}{\delta T_{ij}} = -\log_e T_{ij} - \sum_a \lambda_a p_{ij}^a = 0 \quad (2.4.6)$$

Κατά συνέπεια

$$T_{ij} = \exp \left[ - \sum_a \lambda_a p_{ij}^a \right] \quad (2.4.7)$$

και θέτοντας

$$e^{-\lambda_a} = X_a \quad (2.4.8)$$

έχουμε

$$T_{ij} = \prod_a X_a^{p_{ij}^a} \quad (2.4.9)$$

Αυτό είναι μία περίπτωση πολύαναλογικού προβλήματος (multi-proportional problem).

Μπορούμε επίσης να επεκτείνουμε αυτό το μοντέλο ώστε να κάνει χρήση παλαιότερης πληροφορίας ακολουθώντας την μεθοδολογία που πρότειναν οι Batty και March (1976) και Nathanson (1978). Αντικαθιστούμε την εξίσωση (2.4.1) με την:

$$W'(T_{ij}, t_{ij}) = \frac{T!}{\prod_{ij} T_{ij}!} \prod_{ij} q_{ij}^{T_{ij}} \quad (2.4.10)$$

όπου

$$q_{ij} = \frac{t_{ij}}{\sum_{ij} t_{ij}} \quad (2.4.11)$$

και όπου  $t_{ij}$ , η υπάρχουσα πληροφορία του παλαιού πίνακα Π-Π.

Στη συνέχεια, ακολουθώντας βήματα ανάλογα με αυτά που μας οδήγησαν από την εξίσωση (2.4.1) έως την εξίσωση (2.4.9) βρίσκουμε ότι ισχύει:

$$T_{ij} = t_{ij} \prod_a X_a^{p_{ij}^a} \quad (2.4.12)$$

και

$$X_a = \left( \sum_{ij} t_{ij} \right)^{1/L} e^{-\lambda_a} \quad (2.4.13)$$

όπου  $L$  ο αριθμός των μετρημένων οδικών συνδέσμων.

Φυσικά όπου δεν υπάρχει *a priori* πληροφορία που να αφορά τον πίνακα των μετακινήσεων, το μοντέλο (2.4.12) επανέρχεται στο μοντέλο (2.4.1).

### 2.4.1 Μερικές ιδιότητες του μοντέλου

Η κύρια διαφορά ανάμεσα στο μοντέλο της εξίσωσης (2.4.1) και της (2.4.10)

εντοπίζεται στον λόγο  $\frac{p_{ij}^a}{\sum_a p_{ij}^a}$  για το μοντέλο του Van Zuylen και στον παράγοντα

$p_{ij}^a$  στο μοντέλο του Willumsen. Αυτά έχουν την μορφή βαρών που σχετίζονται με τις παρατηρήσεις στον οδικό σύνδεσμο  $a$ . Η ομοιότητα δεν προκαλεί έκπληξη, καθώς η σύνδεση μεταξύ της μεγιστοποίησης της εντροπίας και της ελαχιστοποίησης της πληροφορίας είναι γνωστή. Στην πραγματικότητα, είναι δυνατόν να δείξουμε ότι η βασική διαφορά μεταξύ των μοντέλων βρίσκεται σε αυτό που θεωρείται να είναι η μονάδα της παρατήρησης: το μοντέλο του Van Zuylen χρησιμοποιεί ένα μετρημένο όχημα και το μοντέλο του Willumsen μία μετακίνηση.

Ένας τρόπος να το δείξουμε αυτό είναι να παράγουμε το μοντέλο του Van Zuylen με την μέθοδο μεγιστοποίησης της εντροπίας. Κάθε μετακίνηση από το  $i$  στο  $j$  μετράται  $g_{ij} = \sum_a p_{ij}^a$  φορές κατά μέσο όρο.

Επομένως

$$\tilde{T} = \sum_{ij} g_{ij} T_{ij} \quad (2.4.1.1)$$



οι  $\tilde{T}$  μετρήσεις οχημάτων που καταμερίζονται σε ζευγάρια Προέλευσης-Προορισμού. Ακολουθώντας τον Wilson (1970) ξανά, ο αριθμός των πιθανών τρόπων να γίνει αυτό μπορεί να χρησιμοποιηθεί για να βρεθεί ο πιο πιθανός πίνακας των μετακινήσεων. Η *a priori* πιθανότητα, ένα καταμετρημένο όχημα να μετακινηθεί στην πραγματικότητα από το  $i$  στο  $j$  είναι

$$q_{ij} = \frac{g_{ij}t_{ij}}{\sum_{ij} g_{ij}t_{ij}} = \frac{g_{ij}t_{ij}}{\tilde{t}} \quad (2.4.1.2)$$

Τότε το πρόβλημα είναι να μεγιστοποιήσουμε

$$\tilde{S} = \log_e \tilde{T}! \prod_{ij} \frac{\left( \frac{g_{ij}t_{ij}}{\tilde{t}} \right)}{\left( g_{ij}T_{ij} \right)!} \quad (2.4.1.3)$$

$$\tilde{S} \cong \tilde{T} \left( \log_e \frac{\tilde{T}}{\tilde{t}} - 1 \right) - \sum_{ij} g_{ij}T_{ij} \left( \log_e \frac{T_{ij}}{t_{ij}} - 1 \right) \quad (2.4.1.4)$$

Με μια περεταίρω μεγιστοποίηση του  $\tilde{S}$  υπό τον περιορισμό  $V_a = \sum_i \sum_j p_{ij}^a T_{ij}$ , και χρησιμοποιώντας τους πολλαπλασιαστές του Lagrangean, θα προκύψει πάλι η εξίσωση (2.3.16).

Είναι θέμα προς συζήτηση το αν η κατανομή των μετρήσεων κυκλοφορίας σε ζεύγη προέλευσης-προορισμού ή ο προσδιορισμός του πιο πιθανού πίνακα Π-Π ο οποίος είναι συνεπής με τους παρατηρούμενους φόρτους, αντιπροσωπεύει καλύτερα το πραγματικό πρόβλημα. Και το  $S'$  και το  $\tilde{S}$  δίνουν ένα μέγεθος της διαφοράς μεταξύ των ενός πίνακα των μετακινήσεων  $\{T_{ij}\}$  και ενός *a priori* πίνακα  $\{t_{ij}\}$ . Όπως προτείνει ο Murchland (1978), δεν υπάρχει ισχυρό θεωρητικό επιχείρημα που να μας κάνει να προτιμήσουμε κάποια από τις παραπάνω μεθόδους. Πρακτικά θέματα όπως

υπολογιστικά προβλήματα και ευκολία στη χρήση είναι πιθανό να είναι πιο σημαντικά.

Και τα δύο μοντέλα που παρουσιάσαμε παραπάνω έχουν ως αποτέλεσμα ένα πολυαναλογικό πρόβλημα (multi-proportional problem). Επίσης, παρατηρούμε ότι δεν είναι αναγκαίο να έχουμε μετρήσεις για όλους τους οδικούς συνδέσμους του δικτύου. Φυσικά όσο περισσότερες μετρήσεις έχουμε τόσο καλύτερη θα είναι η εκτίμηση του πίνακα.

## 2.5. Επιλύοντας το πολυαναλογικό πρόβλημα

Ο Murchland (1977) μελέτησε το πολυαναλογικό πρόβλημα και πρότεινε έναν αλγόριθμο για την λύση του, ο οποίος μπορεί να εύκολα να προσαρμοστεί στο δικό μας πρόβλημα. Ο αλγόριθμος αυτός χρειάζεται την επαναλαμβανόμενη τροποποίηση ενός αρχικού πίνακα έτσι ώστε οι εκτιμώμενες ροές κυκλοφορίας να γίνουν ίσες με τις πραγματικές. Χρειάζεται να δημιουργήσουμε μια λίστα με οδικούς συνδέσμους  $\tilde{V}_a$  από  $a = 1$  έως  $L$ . Ο αλγόριθμος για το μοντέλο (2.4.12) είναι:

1. Βρίσκουμε, χρησιμοποιώντας μια κατάλληλη μέθοδο καταμερισμού, τις τιμές του  $p_{ij}^a$  και ορίζουμε τον αριθμό των επαναλήψεων  $n = 0$ .
2. Θέτουμε  $X_a^n = 1$  για όλους τους οδικούς συνδέσμους.
3. Θέτουμε  $a = 0$ .
4. Αυξάνουμε τον αριθμό του οδικού συνδέσμου  $a$  κατά ένα. Παίρνουμε τον οδικό σύνδεσμο  $a$  και υπολογίζουμε τις εκτιμώμενες ροές κυκλοφορίας.

$$V_a^n = \sum_{ij} t_{ij} \left( \prod_a X_a^{np_{ij}^a} \right) p_{ij}^a \quad (2.5.1)$$

και

$$X_a^{n+1} = X_a^n \cdot Y_a^{p_{ij}^a} \quad (2.5.2)$$

όπου το  $Y_a$  προκύπτει από την επίλυση της ακόλουθης σχέσης

$$\tilde{V}_a = \sum_{ij} p_{ij}^a t_{ij} \left( \prod_a X_a^{np_{ij}^a} \right) \cdot Y_a^{p_{ij}^a} \quad (2.5.3)$$

5. Αν η λίστα των οδικών συνδέσμων δεν έχει εξαντληθεί ( $a < L$ ) προχωρούμε βήμα (4) αλλιώς ελέγχουμε για σύγκλιση στο βήμα (6).
6. Αν όλες οι εκτιμούμενες κυκλοφοριακές ροές  $V_a^n$  είναι αρκετά κοντά στις πραγματικές ( $\pm 5\%$ ) σταματάμε, αλλιώς αυξάνουμε το  $n$  κατά μία μονάδα, προχωρούμε στο βήμα (3) και επεξεργαζόμαστε την λίστα των οδικών συνδέσμων ξανά.

Παρότι ο Murchland (1977) απέδειξε ότι το πολυαναλογικό πρόβλημα έχει μοναδική λύση, η απόλυτη σύγκλιση του αλγορίθμου δεν έχει ικανοποιητικά αποδειχθεί ακόμα.

### 3. Εναλλακτικές μορφές του βασικού μοντέλου μεγιστοποίησης της εντροπίας

Στο παραπάνω κεφάλαιο παρουσιάσαμε αναλυτικά τον τρόπο με τον οποίο ο Van Zuylen (1978), ακολουθώντας την μέθοδο της ελαχιστοποίησης της πληροφορίας και ο Willumsen (1978a, 1978b), ακολουθώντας την μέθοδο της μεγιστοποίησης της εντροπίας, κατέληξαν στο ίδιο μοντέλο για την εκτίμηση πινάκων Π-Π. Στο κεφάλαιο αυτό, κάνοντας διάφορες παραδοχές και απλοποιήσεις για το μοντέλο αυτό, θα παράγουμε κάποιες εναλλακτικές μορφές του.

#### 3.1. Το βασικό μοντέλο εντροπίας

Η αντικειμενική συνάρτηση (objective function) του μοντέλου αυτού για ένα σύνολο μετακινήσεων  $(T_{ij})$  είναι η:

$$\text{Maximize : } Z(T_{ij}) = \frac{T!}{\prod_{ij} (T_{ij}!)} \quad (3.1.1.a)$$

$$\text{όπου } T = \sum_{ij} (T_{ij})$$

Στην περίπτωση τώρα που θέλουμε να χρησιμοποιήσουμε και την ύπαρξη παλαιότερης πληροφορίας, η αντικειμενική συνάρτηση του μοντέλου γίνεται :

$$\text{Maximize : } Z(T_{ij}) = \frac{T!}{\prod_{ij} (T_{ij}!)} \prod_{ij} \left( \frac{t_{ij}}{\sum_{ij} t_{ij}} \right)^{T_{ij}} \quad (3.1.1.β)$$

όπου  $T = \sum_{ij} (T_{ij})$  και  $t_{ij}$  είναι τα στοιχεία του αρχικού πίνακα (από τον οποίο θα αντλήσουμε την πληροφορία).

### 3.2. Μοντέλο 2

Για να απλοποιήσουμε την μορφή του μοντέλου (3.1.1.α), θα λογαριθμίσουμε την αντικειμενική του συνάρτηση. Προκύπτει:

$$\text{Maximize : } Z(T_{ij}) = Ln\left(\frac{T!}{\prod_{ij} T_{ij}!}\right) = Ln(T!) - Ln\left(\prod_{ij} T_{ij}!\right) = Ln(T!) - \sum_{ij} Ln(T_{ij}!) \quad (3.2.1)$$

Εάν εφαρμόσουμε στην σχέση (3.2.1) τον μετασχηματισμό του Stirling (Παράρτημα Α) για τον οποίο ισχύει:

$$Ln(x!) = xLn(x) - x \quad \text{για μεγάλες τιμές του } x$$

προκύπτει η παρακάτω σχέση:

$$\text{Maximize : } Z(T_{ij}) = TLn(T) - T - \sum_{ij} (T_{ij}Ln(T_{ij}) - T_{ij}) \quad (3.2.2)$$

Στην περίπτωση που θέλουμε να χρησιμοποιήσουμε και την ύπαρξη παλαιότερης πληροφορίας θα χρησιμοποιήσουμε την σχέση (3.1.1.β). Εάν την λογαριθμίσουμε προκύπτει:

$$\text{Maximize : } Z(T_{ij}, t_{ij}) = Ln\left(\frac{T!}{\prod_{ij} (T_{ij}!)} \prod_{ij} \left(\frac{t_{ij}}{\sum_{ij} t_{ij}}\right)^{T_{ij}}\right) \quad (3.2.3)$$

Χρησιμοποιώντας τώρα τον μετασχηματισμό του Stirling για την (3.2.3) έχουμε:

$$\text{Maximize : } TLn\left(\frac{T}{t}\right) - T - \sum_{ij} \left(T_{ij}Ln\left(\frac{T_{ij}}{t_{ij}}\right) - T_{ij}\right) \quad (3.2.4)$$

όπου

$$T = \sum_{ij} T_{ij} \quad \text{και} \quad t = \sum_{ij} t_{ij}$$

### 3.3. Μοντέλο 3

Σκοπός μας εδώ είναι να απλοποιήσουμε την μορφή του βασικού μοντέλου εντροπίας. Έχοντας ως βάση την εξίσωση (3.1.1.β) μπορούμε να κάνουμε τις ακόλουθες πράξεις:

$$\text{Maximize : } Z(T_{ij}, t_{ij}) = \frac{T!}{\prod_{ij} (T_{ij}!)} \frac{\prod_{ij} t_{ij}^{T_{ij}}}{\left(\sum_{ij} t_{ij}\right)^{\sum T_{ij}}} \quad (3.3.1)$$

$$\text{Maximize : } Z(T_{ij}, t_{ij}) = \frac{T! \left(\sum_{ij} t_{ij}\right)^{-\sum T_{ij}}}{\prod_{ij} (T_{ij}!)} \prod_{ij} t_{ij}^{T_{ij}} \quad (3.3.2)$$

Γνωρίζοντας όμως ότι  $T$  είναι το άθροισμα των  $T_{ij}$  και  $t$  το άθροισμα των  $t_{ij}$  έχουμε

$$\text{Maximize : } Z(T_{ij}, t_{ij}) = \frac{T! \cdot t^{-T}}{\prod_{ij} (T_{ij}!)} \prod_{ij} t_{ij}^{T_{ij}} \quad (3.3.3)$$

Βασιζόμενοι στην υπόθεση του Willumsen (1984), ότι δηλαδή ο συνολικός αριθμός των μετακινήσεων  $(T, t)$  είναι σταθερός, μπορούμε να απλοποιήσουμε περαιτέρω την παραπάνω σχέση

$$\text{Maximize : } Z(T_{ij}, t_{ij}) = \frac{\prod_{ij} t_{ij}^{T_{ij}}}{\prod_{ij} (T_{ij}!)} \quad (3.3.4)$$

Ένα τώρα λογαριθμήσουμε την σχέση (3.3.4) και χρησιμοποιήσουμε και τον μετασχηματισμό του Stirling, η σχέση που προκύπτει είναι ή εξής:

$$\text{Maximize : } Z(T_{ij}, t_{ij}) = -\sum_{ij} \left( T_{ij} \text{Ln} \left( \frac{T_{ij}}{t_{ij}} \right) - T_{ij} \right) \quad (3.3.5)$$

Υπό τους περιορισμούς:

$$V_a = \sum_{ij} T_{ij} p_{ij}^a \quad \forall a$$

$$T_{ij} \geq 0 \quad \forall i, j$$

$$\text{Οπου } T = \sum_{ij} T_{ij} \quad \text{και} \quad t = \sum_{ij} t_{ij}$$

### 3.4. Μοντέλο 4

Ο Willumsen (1981, 1984) αφαίρεσε ένα σταθερό όρο από το μοντέλο 3 (άθροισμα των  $t_{ij}$ ) όπως φαίνεται παρακάτω. Αυτό δεν αλλάζει το αποτέλεσμα απλά απλοποίησε την μορφή της σχέσης. Ξεκινώντας από την σχέση (3.3.5) προκύπτει:

$$\text{Maximize : } Z(T_{ij}, t_{ij}) = -\sum_{ij} \left( T_{ij} \text{Ln} \left( \frac{T_{ij}}{t_{ij}} \right) - T_{ij} \right) - \sum_{ij} t_{ij} \quad (3.4.1)$$

$$\text{Maximize : } Z(T_{ij}, t_{ij}) = -\sum_{ij} \left( T_{ij} \text{Ln} \left( \frac{T_{ij}}{t_{ij}} \right) - T_{ij} + t_{ij} \right) \quad (3.4.2)$$

Υπό τους περιορισμούς:

$$V_a = \sum_{ij} T_{ij} p_{ij}^a \quad \forall a$$

$$T_{ij} \geq 0 \quad \forall i, j$$

$$\text{Όπου } T = \sum_{ij} T_{ij} \text{ και } t = \sum_{ij} t_{ij}$$

### 3.5. Μοντέλο 5

Με βάση την εξίσωση (3.4.2) και αντικαθιστώντας το ανάπτυγμα του Taylor για το  $\ln(x)$  όταν  $x \geq 0.5$ , για το οποίο ισχύει

$$\ln(x) = \left(\frac{x-1}{x}\right) + \frac{1}{2}\left(\frac{x-1}{x}\right)^2 + \frac{1}{3}\left(\frac{x-1}{x}\right)^3 + \dots \quad (3.5.1)$$

συμπεριλαμβάνοντας μόνο τους δύο πρώτους όρους, θεωρώντας ότι το  $T_{ij}$  είναι περίπου ίδιο με το  $t_{ij}$  (δηλ.  $T_{ij}/t_{ij} \approx 1$ ) προκύπτει ότι:

$$\text{Maximize: } Z(T_{ij}, t_{ij}) = - \sum_{ij} \left( T_{ij} \left( \frac{\frac{T_{ij}-1}{t_{ij}}}{\frac{T_{ij}}{t_{ij}}} \right) + \frac{T_{ij}}{2} \left( \frac{\frac{T_{ij}-1}{t_{ij}}}{\frac{T_{ij}}{t_{ij}}} \right)^2 - T_{ij} + t_{ij} \right) \quad (3.5.2)$$

$$\text{Maximize: } Z(T_{ij}, t_{ij}) = - \sum_{ij} \left( T_{ij} - t_{ij} + \frac{T_{ij}}{2} \left( \frac{\frac{T_{ij}-1}{t_{ij}}}{\frac{T_{ij}}{t_{ij}}} \right)^2 - T_{ij} + t_{ij} \right) \quad (3.5.3)$$

$$\text{Maximize: } Z(T_{ij}, t_{ij}) = - \sum_{ij} \left( T_{ij} - t_{ij} + \frac{1}{2T_{ij}} (T_{ij} - t_{ij})^2 - T_{ij} + t_{ij} \right) \quad (3.5.4)$$



$$\text{Minimize : } Z(T_{ij}, t_{ij}) = \sum_{ij} \left( \frac{1}{2T_{ij}} (T_{ij} - t_{ij})^2 \right) \quad (3.5.5)$$

Υπό τους περιορισμούς

$$V_a = \sum_{ij} T_{ij} p_{ij}^a \quad \forall a$$

$$T_{ij} \geq 0 \quad \forall a$$

$$\text{Όπου } \frac{T_{ij}}{t_{ij}} \approx 1$$

### 3.6. Μοντέλο 6

Σύμφωνα με τον Paramahamsan (1999), αφού το μοντέλο 5 (3.5.5) βασίζεται στην υπόθεση ότι  $T_{ij}/t_{ij} \approx 1$ , μπορούμε να αντικαταστήσουμε τον όρο  $T_{ij}$  στον παρονομαστή με τον όρο  $t_{ij}$  ώστε να δημιουργήσουμε ένα απλό πρόβλημα βελτιστοποίησης δευτέρου βαθμού όπως φαίνεται παρακάτω:

$$\text{Minimize : } Z(T_{ij}, t_{ij}) = \sum_{ij} \left( \frac{1}{2t_{ij}} (T_{ij} - t_{ij})^2 \right) \quad (3.6.1)$$

Υπό τους περιορισμούς

$$V_a = \sum_{ij} T_{ij} p_{ij}^a \quad \forall a$$

$$T_{ij} \geq 0 \quad \forall a$$

$$\text{Όπου } \frac{T_{ij}}{t_{ij}} \approx 1$$

Το πλεονέκτημα αυτού του μοντέλου είναι ότι στον ο όρος  $t_{ij}$  στον παρονομαστή είναι γνωστός *a priori*, σε αντίθεση με το μοντέλο 5 που στον παρονομαστή είναι ο όρος  $T_{ij}$ , που είναι η μεταβλητή που ζητάμε. Το μοντέλο 6 αντιπροσωπεύει μια σταθμισμένη παλινδρόμηση στην οποία το τετραγωνικό σφάλμα του αρχικού πίνακα (seed matrix) είναι ελάχιστο.

### 3.7. Μοντέλο 7

Μια εναλλακτική προσέγγιση στην λύση του προβλήματος του προσδιορισμού του πίνακα Π-Π είναι να το σχεδιάσουμε με βάση κυκλοφοριακούς φόρτους αντί με πίνακα μετακινήσεων. Αυτό προτάθηκε από τους Willumsen και Van Zuylen (1980) όπως θα δείξουμε παρακάτω. Η προσέγγιση που ακολουθήθηκε από τον Van Zuylen (1978) βασίζεται στην μέτρηση της πληροφορίας του Brillouin (1956).

Η πληροφορία που περιέχεται σε ένα σύνολο  $N$  παρατηρήσεων όπου η κατάσταση  $k$  έχει παρατηρηθεί  $n_k$  φορές ορίζεται από τον Brillouin (1956) ως:

$$I = -Ln \left( N! \prod_k \left( \frac{q_k^{n_k}}{n_k!} \right) \right) \quad (3.7.1)$$

όπου  $q_k$  είναι η *a priori* πιθανότητα της παρατηρούμενης κατάστασης  $k$ . Αν οι παρατηρούμενες μετρήσεις είναι σε ένα συγκεκριμένο οδικό σύνδεσμο τότε μπορεί κάποιος να ορίσει την κατάσταση  $ij$  ως την κατάσταση στην οποία το παρατηρούμενο όχημα μετακινήθηκε από την ζώνη  $i$  στην ζώνη  $j$ . Έτσι,

$$n_{ij}^a = T_{ij} p_{ij}^a \quad (3.7.2)$$

Μπορούμε επίσης να εκφράσουμε την *a priori* πιθανότητα της παρατηρούμενης κατάστασης  $ij$  ως λειτουργία της *a priori* πληροφορίας για τον πίνακα Π-Π ως:

$$q_{ij}^a = \frac{t_{ij} p_{ij}^a}{\sum_{ij} t_{ij} p_{ij}^a} \quad (3.7.3)$$

όπου  $t_{ij}$  είναι ο *a priori* αριθμός των μετακινήσεων μεταξύ  $i$  και  $j$ . Η πληροφορία που περιέχεται στον φόρτο  $V_a$  της διαδρομής είναι:

$$I_a = -Ln \left( \frac{V_a! \prod_{ij} \left( \frac{t_{ij} p_{ij}^a}{v_a} \right)^{T_{ij} p_{ij}^a}}{\prod_{ij} (T_{ij} p_{ij}^a)!} \right) \quad (3.7.4)$$

Η σχέση γίνεται ως εξής:

$$\text{Minimize : } Z(T_{ij}, t_{ij}) = -\sum_a Ln \left( \frac{V_a! \prod_{ij} \left( \frac{t_{ij} p_{ij}^a}{v_a} \right)^{T_{ij} p_{ij}^a}}{\prod_{ij} (T_{ij} p_{ij}^a)!} \right) \quad (3.7.5)$$

$$\text{Maximize : } Z(T_{ij}, t_{ij}) = \prod_a \left( \frac{V_a!}{\prod_{ij} (T_{ij} p_{ij}^a)!} \prod_{ij} \left( \frac{t_{ij} p_{ij}^a}{v_a} \right)^{T_{ij} p_{ij}^a} \right) \quad (3.7.6)$$

υπό τους περιορισμούς

$$V_a = \sum_{ij} T_{ij} p_{ij}^a \quad \forall a$$

$$v_a = \sum_{ij} t_{ij} p_{ij}^a \quad \forall a$$

$$T_{ij} \geq 0 \quad \forall a$$

### 3.8. Μοντέλο 8

Χρησιμοποιώντας τον μετασχηματισμό του Stirling (Παράρτημα Α) θα απλοποιήσουμε την μορφή του μοντέλου 7 (3.7.6).

$$\text{Maximize: } Z(T_{ij}, t_{ij}) = \prod_a \left( \frac{V_a!}{\prod_{ij} (T_{ij} p_{ij}^a)} \prod_{ij} \left( \frac{t_{ij} p_{ij}^a}{v_a} \right)^{T_{ij} p_{ij}^a} \right) \quad (3.8.1)$$

$$\text{Minimize: } Z(T_{ij}, t_{ij}) = - \sum_a \text{Ln} \left( \frac{V_a! \prod_{ij} \left( \frac{t_{ij} p_{ij}^a}{v_a} \right)^{T_{ij} p_{ij}^a}}{\prod_{ij} (T_{ij} p_{ij}^a)} \right) \quad (3.8.2)$$

$$\text{Minimize: } Z(T_{ij}, t_{ij}) = - \sum_a \left( V_a \text{Ln}(V_a) - V_a + \sum_{ij} T_{ij} p_{ij}^a \text{Ln} \left( \frac{t_{ij} p_{ij}^a}{v_a T_{ij} p_{ij}^a} \right) + \sum_{ij} T_{ij} p_{ij}^a \right) \quad (3.8.3)$$

$$\text{Minimize: } Z(T_{ij}, t_{ij}) = - \sum_a \left( \sum_{ij} T_{ij} p_{ij}^a \text{Ln}(V_a) - V_a + \sum_{ij} T_{ij} p_{ij}^a \text{Ln} \left( \frac{t_{ij} p_{ij}^a}{v_a T_{ij} p_{ij}^a} \right) + V_a \right) \quad (3.8.4)$$

$$\text{Minimize: } Z(T_{ij}, t_{ij}) = - \sum_a \sum_{ij} T_{ij} p_{ij}^a \text{Ln} \left( \frac{V_a t_{ij}}{T_{ij} v_a} \right) \quad (3.8.5)$$

$$\text{Minimize: } Z(T_{ij}, t_{ij}) = \sum_a \sum_{ij} T_{ij} p_{ij}^a \text{Ln} \left( \frac{T_{ij} v_a}{V_a t_{ij}} \right) \quad (3.8.6)$$

### 3.9. Σύνοψη

Κάνοντας μια ανασκόπηση του κεφαλαίου βλέπουμε ότι μπορούμε να ομαδοποιήσουμε τα μοντέλα μας σε δύο ομάδες. Στη πρώτη ομάδα ανήκουν τα

μοντέλα των οποίων η αντικειμενική συνάρτηση έχει ως βάση το βασικό μοντέλο μεγιστοποίησης της εντροπίας και είναι τα μοντέλα 1 έως 6. Το μοντέλο 1 (3.1.1.β) είναι η αντικειμενική συνάρτηση της θεωρίας της μεγιστοποίησης της εντροπίας. Το μοντέλο 2 (3.2.4) προέκυψε λογαριθμίζοντας την (3.1.1.β) και εφαρμόζοντας την προσέγγιση του Stirling. Το μοντέλο 3 (3.3.5) είναι μια μορφή που προτείνει ο Willumsen βασιζόμενος στην παραδοχή ότι ο συνολικός αριθμός των μετακινήσεων είναι σταθερός. Το μοντέλο 4 (3.4.2) είναι μια απλοποιημένη μορφή του μοντέλου 3. Το μοντέλο 5 (3.5.5) προκύπτει από τη εφαρμογή του αναπτύγματος του Taylor στο μοντέλο 4. Το μοντέλο 6 (3.6.1) είναι μια τροποποίηση του μοντέλου 5, το οποίο μπορεί θεωρηθεί ένα σταθμισμένο πρόβλημα γραμμικής παλινδρόμησης.

Στη δεύτερη ομάδα ανήκουν τα μοντέλα των οποίων η αντικειμενική συνάρτηση έχει ως βάση την σχέση της ελαχιστοποίησης της πληροφορίας και είναι τα μοντέλα 7 και 8. Το μοντέλο 7 (3.7.6) περιέχει στην αντικειμενική του συνάρτηση τους κυκλοφοριακούς φόρτους. Το μοντέλο 8 (3.8.6) προκύπτει λογαριθμίζοντας το μοντέλο 7.

Οι τελικές μορφές των μοντέλων είναι:

$$\text{Maximize : } Z(T_{ij}) = \frac{T!}{\prod_{ij} (T_{ij}!)} \prod_{ij} \left( \frac{t_{ij}}{\sum_{ij} t_{ij}} \right)^{T_{ij}} \quad (3.1.1.\beta)$$

$$\text{Maximize : } T \ln \left( \frac{T}{t} \right) - T - \sum_{ij} \left( T_{ij} \ln \left( \frac{T_{ij}}{t_{ij}} \right) - T_{ij} \right) \quad (3.2.4)$$

$$\text{Maximize : } Z(T_{ij}, t_{ij}) = - \sum_{ij} \left( T_{ij} \ln \left( \frac{T_{ij}}{t_{ij}} \right) - T_{ij} \right) \quad (3.3.5)$$

$$\text{Maximize : } Z(T_{ij}, t_{ij}) = - \sum_{ij} \left( T_{ij} \ln \left( \frac{T_{ij}}{t_{ij}} \right) - T_{ij} + t_{ij} \right) \quad (3.4.2)$$

$$\text{Minimize : } Z(T_{ij}, t_{ij}) = \sum_{ij} \left( \frac{1}{2T_{ij}} (T_{ij} - t_{ij})^2 \right) \quad (3.5.5)$$

$$\text{Minimize : } Z(T_{ij}, t_{ij}) = \sum_{ij} \left( \frac{1}{2t_{ij}} (T_{ij} - t_{ij})^2 \right) \quad (5.6.1)$$

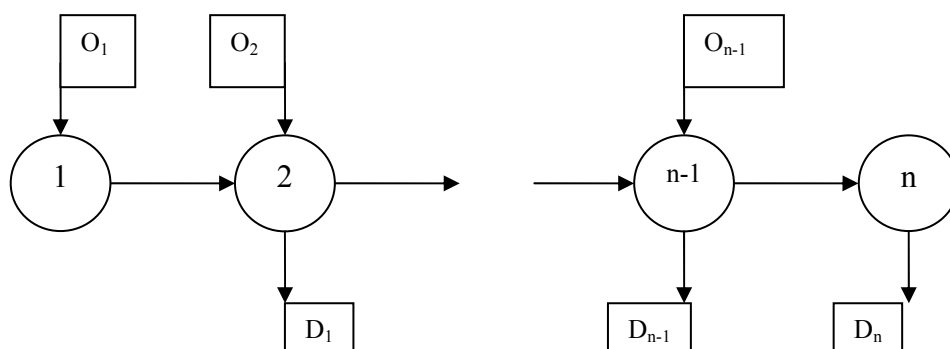
$$\text{Maximize : } Z(T_{ij}, t_{ij}) = \prod_a \left( \frac{V_a!}{\prod_{ij} (T_{ij} p_{ij}^a)} \prod_{ij} \left( \frac{t_{ij} p_{ij}^a}{v_a} \right)^{T_{ij} p_{ij}^a} \right) \quad (3.7.6)$$

$$\text{Minimize : } Z(T_{ij}, t_{ij}) = \sum_a \sum_{ij} T_{ij} p_{ij}^a \text{Ln} \left( \frac{T_{ij} v_a}{V_a t_{ij}} \right) \quad (3.8.6)$$

## 4. Ένας γενικευμένος και αποτελεσματικός αλγόριθμος εκτίμησης πινάκων Π-Π από μετρήσεις επιβατών

### 4.1. Εισαγωγή

Στο προηγούμενο κεφάλαιο παρουσιάσαμε και μελετήσαμε μια σειρά μαθηματικών μοντέλων που προσδιορίζουν τον πίνακα Π-Π εφαρμόζοντας τις αρχές της ελαχιστοποίησης της πληροφορίας και της μεγιστοποίησης της εντροπίας. Σε αυτό το κεφάλαιο θα παρουσιάσουμε μία εναλλακτική μέθοδο για τον προσδιορισμό του πίνακα Π-Π για γραμμές MMM που αναπτύχθηκε από τους Li, Cassidy (2006).



Σχήμα 4.1

Ο αλγόριθμος εκτιμά τον πίνακα Π-Π σε μια γραμμή MMM χρησιμοποιώντας μετρήσεις των επιβιβαζόμενων και αποβιβαζόμενων σε κάθε στάση κατά μήκος της γραμμής. Ο αλγόριθμος αυτός δεν προσδιορίζει μόνο τον πίνακα Π-Π για το συγκεκριμένο δρομολόγιο του οχήματος από το οποίο πάρθηκαν οι μετρήσεις επιβατών που επιβιβάστηκαν και αποβιβάστηκαν αλλά εκτιμά γενικότερα την πιθανότητα που έχουν οι επιβάτες να έχουν αποβιβαστούν σε κάθε στάση κατά μήκος κάθε μελλοντικής διαδρομής. Οι πιθανότητες αυτές είναι ικανές να παραμείνουν σταθερές κατά την διάρκεια μελλοντικών διαδρομών. Αυτές οι πιθανότητες, επιπλέον, εκτιμούνται με τέτοιο τρόπο ώστε να ανακλούν την λανθάνουσα τάση των επιβατών να ταξιδεύουν προς και από «σημαντικά κέντρα δραστηριότητας» (major

στάσεις). Επομένως ο αλγόριθμος αυτός είναι μια πιο ευρείας χρήσης μέθοδος από την μέθοδο εκτίμησης που είχε προταθεί από τον Tsygalnitsky (1977). Δεδομένου ότι ο αλγόριθμος αυτός δεν χρειάζεται αρχικό πίνακα Π-Π και ο αριθμός των επαναλήψεων που απαιτούνται για να παραχθούν οι εκτιμήσεις προσδιορίζεται a priori, είναι ευκολότερο να εφαρμοστεί και είναι πιο αποδοτικός υπολογιστικά από ότι τα μοντέλα μεγιστοποίησης της εντροπίας για την εκτίμηση πίνακα Προέλευσης - Προορισμού.

## 4.2. Βασικές παραδοχές

Μια βασική παραδοχή του αλγορίθμου είναι ότι οι στάσεις κατά μήκος μιας γραμμής ενός MMM χωρίζονται σε δύο κατηγορίες: σημαντικές και μη σημαντικές. Σημαντικές θεωρούνται οι στάσεις εκείνες όπου σημαντικός αριθμός επιβατών επιβιβάζονται και αποβιβάζονται στο MMM ενώ μη σημαντικές θεωρούνται εκείνες όπου μικρότερος αριθμός επιβατών επιβιβάζονται και αποβιβάζονται στο MMM.

Ας θεωρήσουμε ένα συγκεκριμένο δρομολόγιο σε μια γραμμή MMM που περιλαμβάνει σημαντικές και μη σημαντικές στάσεις. Την χρονική στιγμή ακριβώς πριν το όχημα φτάσει στην στάση  $s$ , έστω ότι  $a$  είναι ένας τυχαία επιλεγμένος επιβάτης ο οποίος είχε επιβιβαστεί σε σημαντική στάση. Η πιθανότητα αυτός ο επιβάτης να αποβιβαστεί στην στάση  $s$  είναι η  $p_{as}$ . Παράλληλα, έστω  $b$  να είναι ένας επιβάτης ο οποίος επιβιβάστηκε σε μη σημαντική στάση. Η πιθανότητα ο επιβάτης αυτός να αποβιβαστεί στην στάση  $s$  είναι  $p_{bs}$ . Στην στάση  $s$ , ισχύει

$$\frac{p_{as}}{p_{bs}} = \frac{1 - \omega}{\omega} \quad (4.2.1)$$

όπου  $\omega \in (0,1)$ . Εμείς θα υποθέσουμε ότι  $\omega = \omega_a$  για όλες τις σημαντικές στάσεις, διαφορετικά  $\omega = \omega_b$ . Η μέθοδος του Tsygalnitsky (1977) θεωρεί για όλες τις στάσεις  $\omega = 0.5$  που σημαίνει ότι δεν υπάρχει διαχωρισμός σε σημαντικές και μη σημαντικές στάσεις.



Οι επιβάτες που αποβιβάζονται στην στάση  $s$  «έλκονται» από τους επιβαίνοντες ακολουθώντας τον κανόνα που παρουσιάζεται στο παρακάτω θεώρημα.

**Θεώρημα:** Υποθέτουμε πώς ακριβώς πριν την στάση  $s$ , υπάρχουν  $N_a$  επιβάτες οι οποίοι έχουν επιβιβαστεί σε σημαντικές στάσεις και  $N_b$  επιβάτες που έχουν επιβιβαστεί σε μη σημαντικές στάσεις. Υποθέτουμε επίσης ότι ανάμεσα στους  $n$  επιβάτες που θα αποβιβαστούν στην στάση  $s$ , ένα υποσύνολο  $n_a$  επιβατών έχουν επιβιβαστεί σε σημαντικές στάσεις. Αν η (4.2.1) ισχύει, τότε η εκτίμηση του  $n_a$  δίνεται από τον τύπο.

$$E[n_a] = \frac{(1-\omega)N_a}{(1-\omega)N_a + \omega N_b} \quad (4.2.2)$$

**Απόδειξη:** Επιλέγουμε τυχαία έναν επιβάτη ακριβώς πριν από την στάση  $s$  και ορίζουμε τα ακόλουθα γεγονότα:

- $off$ : ο επιβάτης αποβιβάζεται στην στάση  $s$ ,
- $a$  ή  $b$ : ο επιβάτης επιβιβάστηκε σε σημαντική ή μη σημαντική στάση, αντίστοιχα.

Η εξ. (4.2.1) σημαίνει ότι

$$\frac{p(off|a)}{p(off|b)} = \frac{1-\omega}{\omega} \quad (4.2.3)$$

Αφού

$$p(a|off) = \frac{p(off|a) \cdot p(a)}{p(off)} \quad (4.2.4)$$

και

$$1 - p(a|off) = p(b|off) = \frac{p(off|b) \cdot p(b)}{p(off)} \quad (4.2.5)$$

έχουμε

$$\frac{p(a|off)}{1 - p(a|off)} = \frac{1 - \omega}{\omega} \frac{p(a)}{p(b)} \quad (4.2.6)$$

κατά συνέπεια,

$$p(a|off) = \frac{(1 - \omega)p(a)}{\omega \cdot p(b) + (1 - \omega)p(a)} \quad (7)$$

Αφού  $p(a) = \frac{N_a}{N_a + N_b}$ ,  $p(b) = \frac{N_b}{N_a + N_b}$ , και  $E[n_a] = n \cdot p(a|off)$  έχουμε την (4.2.2).

Αν  $n - N_b \leq E[n_a] \leq N_a$ , η εκτίμηση του  $n_a$  δίνεται από  $n_a = E[n_a]$ . Εντούτοις, η εξ.

(2) δίνει  $E[n_a] > N_a$  όταν  $\frac{\omega}{1 - \omega} < \frac{n - N_a}{N_b}$ . Σε αυτή την περίπτωση, θέτουμε

$n_a = N_a$ . Παρομοίως, όταν  $\frac{\omega}{1 - \omega} < \frac{n - N_b}{N_a}$ , η (2) δίνει  $n - E[n_a] > N_b$ . Σε αυτή την

περίπτωση, θέτουμε  $n_a = n - N_b$ .

### 4.3. Τυποποίηση του αλγορίθμου

Έχοντας μόλις περιγράψει πως ο αλγόριθμος «έλκει» αποβιβάζοντες επιβάτες για κάθε  $\omega$ , τώρα θα παρουσιάσουμε την διαδικασία επιλογής των  $\omega_a, \omega_b$  για να εκτιμήσουμε τον πίνακα Π-Π. Η μετρήσεις επιβιβάσεων και αποβιβάσεων επιβατών που χρειάζονται για την διαδικασία αυτή μπορεί να συλλεχθούν από πολλαπλές διαδρομές που γίνονται σε κάποια συγκεκριμένη χρονική στιγμή κατά την διάρκεια μιας ή περισσότερων ημερών.

Ας υποθέσουμε ότι χρησιμοποιούμε δεδομένα από ένα σετ  $M$  δρομολογίων για να εκτιμήσουμε τους πίνακες Π-Π για συγκεκριμένο εύρος τιμών των  $\omega_a, \omega_b$  και δεδομένα από ένα σετ  $J$  δρομολογίων για να επιλέξουμε τα  $\omega_a, \omega_b$  που δίνουν καλύτερη εκτίμηση του πίνακα Π-Π (τα δύο σετ διαδρομών μπορεί να είναι διαφορετικά ή και τα ίδια). Τέλος, υποθέτουμε ότι η διαδρομή έχει  $I$  στάσεις. Η επαναληπτική διαδικασία έχει ως εξής:

1. *Εκτίμηση πίνακα Π-Π για αρχική επιλογή των  $\omega_a, \omega_b$ .* Για κάθε  $M$  διαδρομή, λαμβάνουμε τον πίνακα Π-Π σύροντας αποβιβάζοντας επιβάτες σε κάθε στάση από το σύνολο των επιβατών χρησιμοποιώντας την λογική που περιγράφηκε παραπάνω. Σε κάθε σημαντική στάση, το  $n_a$  (και κατά συνέπεια το  $n_b = n - n_a$ ) υπολογίζονται για κάποια αρχική επιλογή του  $\omega_a$ . Η επιλεγμένη τιμή του  $\omega_b$  χρησιμοποιείται για τον υπολογισμό των  $n_a, n_b$  σε κάθε μη σημαντική στάση. Μετά υπολογίζεται ο μέσος όρος των πινάκων Π-Π για όλες τις  $M$  διαδρομές για να βρούμε τον συνολικό πίνακα Π-Π. Η καταλληλότητα του συνολικού αυτού πίνακα αξιολογείται *ex post* χρησιμοποιώντας τα ακόλουθα βήματα:
2. *Μετατροπή του συνολικού πίνακα Π-Π σε πιθανότητες αποβιβάσεων επιβατών.* Ένας πίνακας πιθανότητας αποβιβάσεων,  $P$ , παράγεται διαιρώντας κάθε στοιχείο του συνολικού πίνακα Π-Π με το άθροισμα της σειράς του πίνακα που ανήκει το στοιχείο. Για κάθε άθροισμα σειράς του πίνακα που είναι μηδέν, μπορεί κάποιος να συμπληρώσει την αντίστοιχη σειρά στον πίνακα  $P$  με τις πιθανότητες αποβιβάσεων από κάποια γειτονική στάση. Αν το σετ των  $J$  διαδρομών είναι υποσύνολο των  $M$  διαδρομών, τότε κάποιος μπορεί να συμπληρώσει την αντίστοιχη σειρά του  $P$  με μηδέν, αφού δεν γίνεται καμία επιβίβαση κατά την διάρκεια των  $J$  διαδρομών.
3. *Πρόβλεψη του αριθμού των αποβιβάσεων σε κάθε διαδρομή.* Για κάθε  $J$  διαδρομή, ο αριθμός αποβιβάσεων υπολογίζεται χρησιμοποιώντας τον αριθμό επιβιβάσεων της διαδρομής και τον πίνακα  $P$ . Ο αριθμός αποβιβάσεων,  $A_s$ ,

για κάθε στάση  $s$ , υπολογίζεται από τον τύπο  $A_s = \sum_{i=1}^I B_i P_{is}$ , όπου  $B_i$  είναι ο πραγματικός αριθμός επιβιβάσεων στην στάση  $i$  και  $P_{is}$  είναι η πιθανότητα οι επιβάτες που επιβιβάστηκαν στην στάση  $i$  να αποβιβαστούν στην  $s$  (όπως υπολογίστηκε στο βήμα 2).

4. *Εκτίμηση καταλληλότητας των πιθανοτήτων αποβιβάσεων.* Για να εκτιμηθεί η καταλληλότητα του πίνακα πιθανοτήτων αποβιβάσεων, κάποιος θα μπορούσε να υπολογίσει την (Ευκλείδεια) απόσταση μεταξύ του αριθμού των εκτιμώμενων αποβιβάσεων και του αριθμού μετρημένων αποβιβάσεων για κάθε διαδρομή. Αυτή η μέτρηση όμως έχει μικρή φυσική σημασία. Αντ' αυτού ο αλγόριθμος εκτιμά την καταλληλότητα βασιζόμενος στον μέσο αριθμό επιβατών μιας διαδρομής. Για κάθε  $J$  διαδρομή, η καταλληλότητα του πίνακα πιθανοτήτων αποβιβάσεων καθορίζεται συγκρίνοντας τον πραγματικό αριθμό επιβατών (λαμβάνετε από τον πραγματικό αριθμό επιβιβάσεων και αποβιβάσεων) με τον εκτιμώμενο μέσο αριθμό επιβατών (όπως υπολογίστηκε στο βήμα 3). Για να μετρήσουμε την συνολική καταλληλότητα του πίνακα πιθανοτήτων αποβιβάσεων για όλες τις διαδρομές, θέτουμε το πραγματικό μέσο αριθμό επιβατών για την  $J$  διαδρομή ως  $x^J$  και τον εκτιμώμενο μέσο αριθμό επιβατών για αυτή την διαδρομή ως  $x_{\omega_a, \omega_b}^J$ . Τότε χρησιμοποιούμε τον

τύπο  $D = \sqrt{\frac{1}{J} \sum_{j=1}^J (x_{\omega_a, \omega_b}^j - x^j)^2}$  σαν μια συνολική μέτρηση καταλληλότητας του πίνακα Π-Π σε συνάρτηση με δεδομένα  $\omega_a, \omega_b$ .

5. *Επανάληψη.* Τα  $\omega_a, \omega_b$  επιλέγονται από κοινού μέσο επανάληψης για να παράγουν τον πίνακα πιθανοτήτων αποβιβάσεων που αντιστοιχεί στην μικρότερη τιμή του  $D$ .

Οι Li, Cassidy εισάγουν και μια ακόμα παράμετρο στον αλγόριθμο τους. Η παράμετρος αυτή είναι ο καθορισμός της ελάχιστης απόστασης για την οποία κάποιος επιβάτης θα αποφασίσει να πάει με κάποιο άλλο τρόπο (π.χ. περπάτημα) στον προορισμό του αντί να χρησιμοποιήσει το μέσο μεταφοράς το οποίο εξετάζουμε. Εάν

θεωρήσουμε ότι έχουμε προσδιορίσει την ελάχιστη αυτή απόσταση, αυτό σημαίνει ότι εάν δύο στάσεις απέχουν μεταξύ τους ίση ή μικρότερη απόσταση από αυτή, οι επιβάτες δεν θα χρησιμοποιήσουν το επιβατικό μέσο για την μετακίνησή τους.

#### **4.4. Συμπεράσματα**

Ο αλγόριθμος των Li, Cassidy (2006) επιλύει σταδιακά το πρόβλημα υπολογίζοντας τον αριθμό των επιβιβαζόμενων και αποβιβαζόμενων σε κάθε στάση ακολουθώντας την κίνηση του οχήματος. Επίσης κάνει δύο βασικές παραδοχές: η πρώτη είναι ότι κατηγοριοποιεί τις στάσεις σε σημαντικές και μη σημαντικές ανάλογα με την «δραστηριότητα» που έχουν και η δεύτερη είναι η ελάχιστη απόσταση διαδρομής για την οποία κάποιος θα επιλέξει να μετακινηθεί με MMM.

Αντίθετα, τα μοντέλα μεγιστοποίησης της εντροπίας τυποποιούν το πρόβλημα με την μορφή μιας αντικειμενικής συνάρτησης και υπολογίζονται συγχρόνως οι τιμές όλων των αγνώστων μεταβλητών. Τα αποτελέσματα που προκύπτουν από την επίλυση των μοντέλων αυτών θα πρέπει να είναι σύμφωνα με τους περιορισμούς που έχουμε για το σύστημα που εξετάζουμε.

## 5. Εφαρμογές

### 5.1. Εισαγωγή

Στα κεφαλαία 3 και 4 παρουσιάσαμε δύο διαφορετικές μεθοδολογίες για την εκτίμηση πινάκων Π-Π από μετρήσεις φόρτων σε οδικούς συνδέσμους. Σε αυτό το κεφάλαιο θα εφαρμόσουμε τις μεθοδολογίες αυτές για ένα πρόβλημα πραγματικών διαστάσεων, του οποίου γνωρίζουμε την λύση, ώστε να προσδιορίσουμε την ακρίβεια του δίνουν αλλά και το πώς συμπεριφέρονται για διάφορες διαστάσεις του προβλήματος.

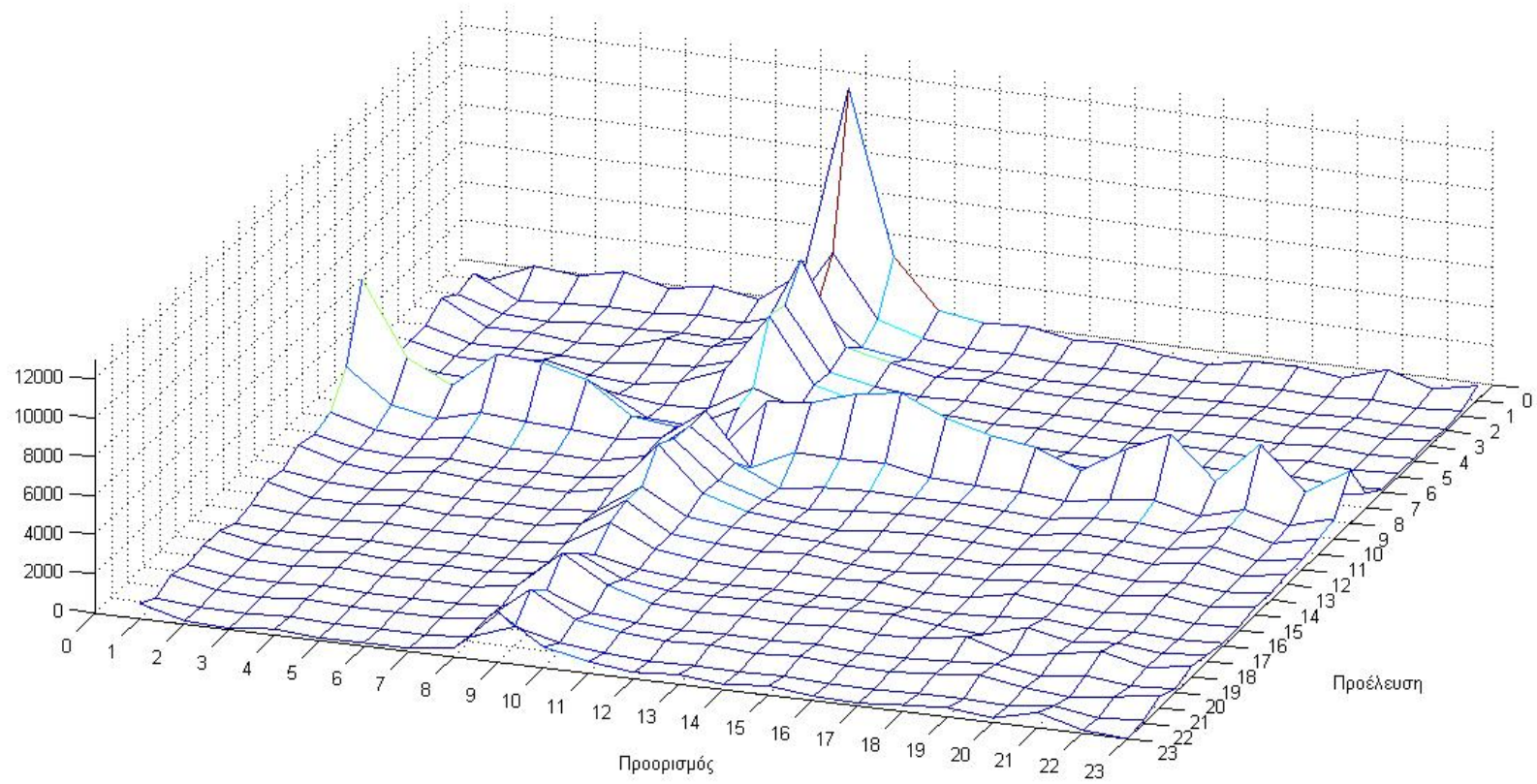
Το πρόβλημα αυτό θα είναι ο προσδιορισμός του πίνακα Π-Π για την γραμμή του ΗΣΑΠ. Ως δεδομένα εισόδου για τις εφαρμογές θα είναι ο αριθμός των εισερχομένων ( $O_i$ ) και ο αριθμός των εξερχομένων ( $D_i$ ) σε κάθε στάση.

Παρακάτω παρουσιάζουμε τον γνωστό πίνακα, τον οποίο και θέλουμε να προσδιορίσουμε και ένα γράφημα που δείχνει την διακύμανση των τιμών του.

## Π-Π μετακινήσεων μεταξύ σταθμών ΗΣΑΠ – σύνολο ημέρας

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
		Πειρ	Ν.Φάλ	Μοσχ	Καλλ	Ταύρ	Πετρ	Θης	Μοναστ	Ομόν	Βικτ	Απτ	Αγ.Νικ	Κ.Πατ	Αγ.Ελευθ	Α.Πατ	Περισ	Πευκ	Νέα Ιων	Ν.Ηράκλ	Ειρ	Μαρ	Κ.Α.Τ	Κηφ
1	Πειραιάς	0	990	777	1257	685	1079	670	2370	12056	3733	1225	876	973	562	639	323	123	576	592	283	988	281	725
2	Ν.Φάληρο	1124	0	146	252	195	403	267	934	4424	1278	573	336	343	186	162	128	65	156	187	94	235	79	237
3	Μοσχάτο	842	130	0	208	180	257	211	644	2474	616	306	153	189	84	117	63	39	89	86	17	108	35	83
4	Καλλιθέα	1449	234	241	0	101	760	607	1209	5647	1424	907	457	561	328	262	248	99	215	279	145	304	112	280
5	Ταύρος	814	181	218	139	0	260	381	1021	4115	950	477	254	314	190	148	117	59	154	162	69	232	78	153
6	Πετράλωνα	837	313	248	709	189	0	390	1279	4342	1083	496	302	346	195	240	114	68	260	195	85	250	78	190
7	Θησείο	659	249	152	562	380	289	0	109	1439	653	392	245	320	175	157	80	51	150	190	83	262	72	157
8	Μοναστηράκι	2032	813	533	1225	1213	1345	125	0	1169	1446	1257	904	1158	693	779	490	224	571	685	304	768	308	741
9	Ομόνοια	6340	2600	1458	3369	3209	2604	1018	736	0	2603	2828	3548	3876	2864	2253	1923	1025	2377	3450	1301	3495	1311	2694
10	Βικτώρια	2647	985	537	1147	956	839	639	959	2652	0	1061	1003	1075	603	479	430	248	577	834	301	918	395	1012
11	Αττική	1106	493	275	685	370	369	304	822	2640	1038	0	327	424	362	370	249	116	432	474	217	822	266	782
12	Αγ.Νικόλαος	767	302	160	363	179	235	212	592	2780	925	201	0	203	338	304	196	88	437	338	204	601	155	528
13	Κ.Πατήσια	1012	345	140	528	322	281	265	749	3306	1062	387	155	0	238	335	230	128	375	430	215	564	180	516
14	Αγ.Ελευθέριος	509	182	93	281	193	194	161	394	1905	644	393	270	189	0	150	170	104	447	287	146	336	138	354
15	Α.Πατήσια	726	169	137	338	156	233	172	508	1984	550	340	221	353	117	0	177	160	368	435	170	460	172	426
16	Περισσός	334	143	76	213	129	151	95	297	1645	560	267	154	216	171	258	0	42	489	253	108	302	102	197
17	Πευκάκια	144	52	29	107	40	47	55	128	797	317	94	78	117	91	136	40	0	106	296	89	173	67	166
18	Νέα Ιωνία	560	143	86	270	124	189	131	358	1647	593	293	298	368	332	402	460	192	0	629	169	401	143	357
19	Ν.Ηράκλειο	577	155	92	326	162	216	151	518	2454	1092	421	268	335	232	377	237	297	681	0	355	765	220	554
20	Ειρήνη	390	98	53	181	80	73	95	223	1333	441	223	118	194	149	242	89	83	202	335	0	313	56	190
21	Μαρούσι	677	189	76	266	152	154	164	386	2169	899	542	353	440	273	415	227	178	350	559	206	0	306	602
22	Κ.Α.Τ	267	84	42	162	64	87	63	230	1045	501	258	109	164	124	185	80	59	135	147	68	337	0	162
23	Κηφισιά	790	181	75	320	139	171	170	456	2552	1071	604	339	495	279	482	233	152	333	475	234	788	166	0

**Πίνακας 5.1 (Πηγή: )**



**Γράφημα 5.1**



## **5.2. Στοιχεία της έρευνας για τα δεδομένα**

Οι Ηλεκτρικοί Σιδηρόδρομοι Αθηνών-Πειραιώς (ΗΣΑΠ) Α.Ε. ιδρύθηκαν με το Ν. 352/1976, αναλαμβάνοντας τη λειτουργία της γραμμής αστικού σιδηροδρόμου και των γραμμών αστικών λεωφορειών της εταιρείας Ελληνικοί Ηλεκτρικοί Σιδηρόδρομοι (ΕΗΣ) Α.Ε. Η εταιρεία ανήκει εξ ολοκλήρου στο Δημόσιο και λειτουργεί ως νομικό πρόσωπο ιδιωτικού δικαίου. Ανήκει στις κοινωφελείς επιχειρήσεις και στο Δημόσιο Τομέα (Ν. 1256/82), και εποπτεύεται και ελέγχεται από το Υπουργείο και Επικοινωνιών.

### **5.2.1. Στοιχεία υποδομής**

Σύμφωνα με στοιχεία που δόθηκαν από τους ΗΣΑΠ, ο στόλος που είναι σε χρήση το 1996 αποτελείται από 222 οχήματα. Τα αμαξοστάσια των σιδηροδρόμων των ΗΣΑΠ διακρίνονται σε αμαξοστάσια επισκευής, που είναι δύο και σε αμαξοστάσια στάθμευσης, που είναι 11.

Η γραμμή σιδηροδρόμου των ΗΣΑΠ διασχίζει το λεκανοπέδιο με κατεύθυνση Βορρά-Νότο, έχει μήκος 26 χιλιομέτρων και 23 σταθμούς. Το εύρος της γραμμής είναι 1,434 μέτρα. Στο μεγαλύτερο τμήμα της είναι υπέργεια, εκτός από το υπόγειο τμήμα του κέντρου της Αθήνας (Μοναστηράκι-Ομόνοια-Βικτώρια-Αττική). Τα δρομολόγια οργανώνονται σε διάφορους τύπους ανάλογα με την περίοδο (Χειμερινή-Θερινή), την ημέρα (Δευτέρα-Παρασκευή ή Σάββατο-Κυριακή) και τις αργίες ή ιδιαίτερα περιστατικά (όπως αγώνες ποδοσφαίρου-μπάσκετ).

### **5.2.2. Μεθοδολογία έρευνας**

Η αυτοψία σε όλους τους σταθμούς έγινε κατά τον μήνα Ιούνιο του 1996, με σκοπό να καταγραφούν τα βασικά στοιχεία υποδομής και λειτουργίας, οι ελλείψεις και τα προβλήματα τους. Ιδιαίτερη έμφαση δόθηκε σε προβλήματα όπως:

- Πρόσβαση στο σταθμό από το οδικό δίκτυο.

- Κυκλοφορία και στάθμευση στην περιοχή του σταθμού, ικανοποίηση αναγκών στάθμευσης για μετεπιβίβαση στη γραμμή.
- Κίνηση των πεζών, τόσο εντός του σταθμού, όσο και στην γύρω περιοχή.
- Προσπελασιμότητα και εξυπηρέτηση του σταθμού από άλλα δημόσια μέσα μεταφοράς και εξυπηρέτηση των μετεπιβιβάσεων.
- Σταθμοί ταξί.

Χρησιμοποιήθηκαν χάρτες για την σήμανση των θέσεων μετεπιβίβασης και συντάχθηκαν σκαριφήματα για την κατάσταση γύρω από τους σταθμούς (πεζοδρόμια, χώροι στάθμευσης, κατευθύνσεις, κυκλοφορία, διαβάσεις υπόγειες και ισόπεδες, φωτεινοί σηματοδότες κλπ.)

### **5.2.3. Σκοπός της έρευνας**

Σκοπός της έρευνας ήταν να καθορίσει για την γραμμή των ΗΣΑΠ:

- Τις προελεύσεις και τους προορισμούς των επιβατών (σε επίπεδα σταθμών)
- Τον αριθμό των διακινούμενων επιβατών και την πληρότητα των συρμών μεταξύ των σταθμών.

Η έρευνα πραγματοποιήθηκε στο διάστημα από 28 έως 31 Μαΐου 1996, καλύπτοντας την συνολική ημερησία λειτουργία και των 23 σταθμών.

Από παρατηρητές που υπήρχαν στις εξόδους του κάθε σταθμού:

- Καταγράφηκαν σε ειδικό έντυπο οι αποβιβαζόμενοι επιβάτες ανά συρμό.
- Συλλέχθηκαν σε κάδους τα ακυρωμένα εισιτήρια.
- Καταγράφηκαν οι ενδείξεις των ακυρωτικών μηχανημάτων στην έναρξη και το πέρας της λειτουργίας του σταθμού, καθώς και 8 περιόδους μέτρησης.

### **Παρατηρήσεις-σχόλια**

Σχετικά με την διεξαγωγή της έρευνας μετακινήσεων επιβατών στους σταθμούς ΗΣΑΠ, θα μπορούσαν να επισημανθούν τα εξής:

- Έκτακτο περιστατικό που σημειώθηκε κατά την πρώτη ημέρα της έρευνας (28-5-1996) στο σταθμό Ν. Φαλήρου και είχε ως αποτέλεσμα την προσωρινή διακοπή της κίνησης των συρμών.
- Στο Ομονοίας, τοποθετήθηκαν ερευνητές μόνο στις κανονικές εξόδους του σταθμού. Κατά την διάρκεια της έρευνας, διαπιστώθηκε ότι ένα μεγάλο ποσοστό επιβατών που έβγαινε από τον σταθμό από τις σκάλες που χρησιμοποιούνται για την είσοδο και δεν καταμετρήθηκε. Για τον λόγο αυτό, στις 19-6-1996, έγινε πρόσθετη καταμέτρηση επιβατών στην Ομόνοια, στις εισόδους και κατά τις ώρες 07.00 έως 19.00.
- Στους σταθμούς Φαλήρου, Ταύρου, Θησείου και Ειρήνης λόγω της διαρρύθμισης με κοινές εξόδους, ανεξάρτητα από την κατεύθυνση του συρμού, η καταμέτρηση των επιβατών και συλλογή των εισιτηρίων έγινε για το σύνολο των διελεύσεων, χωρίς διαχωρισμό κατεύθυνσης.

#### **5.2.4. Χαρακτηριστικά λειτουργίας**

##### **Χρονοαπόσταση**

Με βάση τα αναλυτικά στοιχεία του ΗΣΑΠ που δόθηκαν από το Γραφείο Κίνησης της Διεύθυνσης Εκμετάλλευσης προέκυψε η χρονοαπόσταση (συχνότητα διελεύσεων) σε κάθε τμήμα γραμμής για τις περιόδους πρωινής και απογευματινής αιχμής 07.00-09.00 και 14.00-16.00. Η χρονοαπόσταση για το υπόλοιπο της ημέρας προέκυψε από το πηλίκο των ωρών διάρκειας της περιόδου (15,5 ώρες) προς το σύνολο των συρμών που διέρχονται από κάθε τμήμα της γραμμής σε μια συγκεκριμένη περίοδο. Ο αριθμός των συρμών υπολογίστηκε από τον αριθμό των διελεύσεων (ανά ώρα), για κάθε υποπερίοδο του Υπολοίπου Ημέρας (στοιχεία ΗΣΑΠ).

Η χρονοαπόσταση για το σύνολο της γραμμής υπολογίστηκε με βάση τον μέσο όρο των χρονοαποστάσεων των επιμέρους τμημάτων, αφού σταθμίστηκε με το χιλιομετρικό μήκος του αντίστοιχου τμήματος.

### **Χωρητικότητα-Ποσοστό όρθιων-Μέγιστη πληρότητα**

Με βάση τον αριθμό των συρμών και τις διαφορετικές συνθέσεις (στοιχεία ΗΣΑΠ) για κάθε τμήμα της γραμμής, υπολογίστηκε η συνολική χωρητικότητα συρμών και από εκεί η μέγιστη ωριαία μεταφορική ικανότητα για κάθε τμήμα και κάθε περίοδο. Το τελευταίο μέγεθος διπλασιάστηκε για να προκύψουν στοιχεία συνολικά για την γραμμή (για τις δύο κατευθύνσεις) και διαιρέθηκε με τον αριθμό των ωρών της περιόδου, για να προκύψει η μέση χωρητικότητα, για κάθε τμήμα πλήρους διαδρομής ανά ώρα.

Με βάση τις μέσες ωριαίες χωρητικότητες καθημένων ανά χρονική περίοδο, υπολογίστηκε η συνολική χωρητικότητα των τμημάτων της γραμμής, για το σύνολο της ημέρας.

Οι μέγιστες πληρότητες υπολογίστηκαν για κάθε τμήμα και για κάθε κατεύθυνση, με βάση τα αποτελέσματα που έδωσε ο πίνακας της επιβατικής κίνησης. Συγκεκριμένα διαιρέθηκε ο μέγιστος αριθμός μεταφερόμενων σε κάθε τμήμα με την αντίστοιχη χωρητικότητα.

Ο αριθμός των όρθιων υπολογίστηκε από την διαφορά του συνόλου των επιβιβαζόμενων και της χωρητικότητας καθημένων για κάθε τμήμα πλήρους διαδρομής και για κάθε κατεύθυνση.

### **Προέλευση-Προορισμός (μεταξύ σταθμών)**

Η μηχανογράφηση των πρωτογενών στοιχείων της έρευνας μετακινήσεων επιβατών αστικού σιδηροδρόμου έγινε με πρόγραμμα, το οποίο θέτει τους ακόλουθους περιορισμούς, ώστε να προλαμβάνονται σφάλματα πληκτρολόγησης ή εισαγωγή λανθασμένων στοιχείων:

- Κατά την εισαγωγή εισιτηρίων επιλέγεται ένας από τους προδηλωμένους σταθμούς.
- Επιτρέπονται τιμές ενδείξεων ακυρωτικών, μεγαλύτερες από τις ενδείξεις της προηγούμενης καταγραφής.
- Επιτρέπονται διαφορές ώρας ακύρωσης του εισιτηρίου και της αρχής ημώρου συλλογής, μικρότερες από 1.30 ώρα.
- Δεν επιτρέπεται ο αριθμός των εξερχόμενων επιβατών ανά συρμό να είναι μεγαλύτερος της χωρητικότητας των συρμών.

Σημειώνεται ότι στο σταθμό της Ομόνοιας, όπου διενεργήθηκαν συμπληρωματικές μετρήσεις στις εξόδους οι οποίες δεν είχαν ερευνηθεί κατά την ημέρα της κανονικής έρευνας, οι ελλείψεις των συμπληρωματικών μετρήσεων σε ορισμένα ημίωρα, συμπληρώθηκαν με βάση τη διακύμανση που προέκυψε από τις λοιπές θέσεις του σταθμού.

Τα μηχανογραφημένα πρωτογενή στοιχεία ελέγχθηκαν για τον εντοπισμό σφαλμάτων. Οι έλεγχοι που έγιναν, είναι οι ακόλουθοι:

- Ελέγχθηκε ο αριθμός των συλλεχθέντων εισιτηρίων ανά περίοδο, ο οποίος πρέπει να είναι μικρότερος ή ίσος από το σύνολο των εξερχόμενων επιβατών, όπως καταμετρήθηκαν από τους παρατηρητές.
- Ελέγχθηκε ο αριθμός εισερχόμενων, όπως προκύπτει από τις ενδείξεις των ακυρωτικών, ώστε να είναι μικρότερος ή ίσος από τον αριθμό των εξερχόμενων από τις μετρήσεις, για το άθροισμα κάθε ημέρας.

Συνολικά, έγινε εισαγωγή στοιχείων από 160.443 εισιτήρια, από τα οποία χρησιμοποιήθηκαν ως αποδεκτά 160.260. Από τα πρωτογενή στοιχεία του αρχείου της έρευνας και ειδικότερα από τα στοιχεία των συλλεχθέντων εισιτηρίων, προκύπτουν κατ' αρχάς οι πίνακες προέλευσης-προορισμού μεταξύ των σταθμών, για κάθε ημίωρο της έρευνας (39 ημίωρα), για το δείγμα (όσους απογράφηκαν πλήρως, δηλαδή παρέδωσαν τα εισιτήρια). Στη συνέχεια, υπολογίζονται οι συντελεστές αναγωγής των πινάκων προέλευσης-προορισμού, για το σύνολο των εξερχόμενων ανά ημίωρο (άθροισμα και των δύο κατευθύνσεων).

Τα παραπάνω στοιχεία ομαδοποιήθηκαν σε ομάδες περιόδων, ως εξής:

- Εκτός αιχμής (αρχή-7.00, 9.00-14.00, 16.00-τέλος)
- Πρωινής αιχμής(7.00-9.00)
- Απογευματινής αιχμής (14.00-16.00) και
- Σύνολο ημέρας

Για την ομαδοποίηση, προστέθηκαν τα στοιχεία των επί μέρους πινάκων και προέκυψαν οι πίνακες Π-Π μεταξύ σταθμών, ανά ομάδες περιόδων και για το σύνολο της ημέρας.

### **Επιβατική κίνηση**

Οι πίνακες της ωριαίας επιβατικής κίνησης ανά ομάδα περιόδων προκύπτουν από τους αντίστοιχους πίνακες Π-Π. Ο αριθμός των επιβιβαζόμενων είναι το άθροισμα της αντίστοιχης γραμμής του πίνακα, για τις εγγραφές της κάθε κατεύθυνσης και των αποβιβαζόμενων, το άθροισμα της αντίστοιχης.

### **5.3. Εφαρμογή του αλγορίθμου Li, Cassidy**

Αυτό που χρειαζόμαστε για να λειτουργήσει αποτελεσματικά ο αλγόριθμος, είναι να προσδιορίσουμε ποιες από τις στάσεις είναι σημαντικές. Η απόφαση αυτή είναι καθαρά υποκειμενική αφού οι Li, Cassidy (2006) δεν προσδιορίζουν σαφώς τα χαρακτηριστικά που θα πρέπει να έχουν οι στάσεις αυτές. Παρόλα αυτά ακολουθώντας το παράδειγμα των συγγραφέων, οι οποίοι θεώρησαν ως σημαντικές στάσεις αυτές με την μεγαλύτερη κυκλοφορία, εξετάσαμε τρεις περιπτώσεις: στην πρώτη θεωρήσαμε ως στάση σημαντική την Ομόνοια (στάση 9), στην δεύτερη περίπτωση θεωρήσαμε ως σημαντικές στάσεις τον Πειραιά και την Ομόνοια (στάση 1 και 9 αντίστοιχα) και στην τρίτη περίπτωση θεωρήσαμε ότι όλες οι στάσεις έχουν την ίδια βαρύτητα. Η επιλογή για τις δύο πρώτες περιπτώσεις είναι προφανής όπως φαίνεται και στο γράφημα 5.1 όπου είναι ξεκάθαρη η έντονη διακύμανση που

υπάρχει στην στάση 9, που είναι η Ομόνοια, και δευτερεύοντος στην στάση 1, που είναι ο Πειραιάς. Η διακύμανση στις υπόλοιπες στάσεις είναι αμελητέα και το γράφημα έχει την μορφή σχεδόν επίπεδης επιφάνειας. Με βάση λοιπόν αυτό το γράφημα θεωρούμε ότι δεν υπάρχουν άλλες στάσεις που να μπορούν να θεωρηθούν ως σημαντικές.

Ο αλγόριθμος, λόγω της απλότητάς του επιλύθηκε στο Microsoft Excel 2003 με την χρήση του εργαλείου Solver το οποίο βελτιστοποιεί μη γραμμικά προβλήματα με χρήση του αλγόριθμου Generalized Reduced Gradient (GRG2). Το Microsoft Excel Solver εφαρμόζει μια επαναληπτική αριθμητική μέθοδο όπου χρησιμοποιεί δοκιμαστικές τιμές για τα κελιά, των οποίων η τιμή μεταβάλλεται και παρατηρεί τα αποτελέσματα τα οποία υπολογίστηκαν από τα κελιά περιορισμού και το κελί βελτιστοποίησης. Κάθε δοκιμή τιμών ονομάζεται επανάληψη. Επειδή μια κατεξοχήν προσέγγιση δοκιμής και σφάλματος είναι χρονοβόρα, το Microsoft Excel Solver κάνει μια εκτενής ανάλυση των αποτελεσμάτων και του ρυθμού μεταβολής τους καθώς μεταβάλλονται τα δεδομένα εισόδου και καθοδηγεί ανάλογα τις τιμές των μεταβλητών για την επόμενη επανάληψη.

Στην δική μας περίπτωση τα μεταβαλλόμενα κελιά είναι αυτά που έχουν τις τιμές των συντελεστών  $\omega_a$ ,  $\omega_b$  και ως κελί βελτιστοποίησης είναι αυτό που έχει την τιμή του συντελεστή  $D$ . Οι τιμές που παίρνουν οι συντελεστές είναι από 0.1 έως 0.9 και με βήμα 0.1.

Στους παρακάτω πίνακες παρουσιάζονται τα δεδομένα εισόδου για τις τρεις περιπτώσεις που εξετάσαμε:

#### Πρώτη περίπτωση

		Επιβίβαση		Αποβίβαση	επιβάτες
στάσεις	σημαντική(1)/μη σημαντική(0)	σημαντική	μη σημαντική		
1	0	0	31783	0	
2	0	0	10680	990	31783
3	0	0	5959	923	41473
4	0	0	13945	1717	46509
5	0	0	9134	1161	58737
6	0	0	9913	2759	66710

7	0	0	4535	2526	73864
8	0	0	11497	7566	75873
9	1	35548	0	35666	79804
10	0	0	8936	13786	79686
11	0	0	4841	9522	74836
12	0	0	3392	8405	70155
13	0	0	3211	9782	65142
14	0	0	2132	6818	58571
15	0	0	2368	6395	53885
16	0	0	1493	4938	49858
17	0	0	897	2639	46413
18	0	0	1699	7779	44671
19	0	0	1894	9802	38591
20	0	0	559	4355	30683
21	0	0	908	12297	26887
22	0	0	162	4554	15498
23	0	0	0	11106	11106

**Πίνακας 5.2**

**Δεύτερη περίπτωση**

στάσεις	σημαντική(1)/μη σημαντική(0)	Επιβίβαση		Αποβίβαση	επιβάτες
		σημαντική	μη σημαντική		
1	1	31783	0	0	
2	0	0	10680	990	31783
3	0	0	5959	923	41473
4	0	0	13945	1717	46509
5	0	0	9134	1161	58737
6	0	0	9913	2759	66710
7	0	0	4535	2526	73864
8	0	0	11497	7566	75873
9	1	35548	0	35666	79804
10	0	0	8936	13786	79686
11	0	0	4841	9522	74836
12	0	0	3392	8405	70155
13	0	0	3211	9782	65142
14	0	0	2132	6818	58571
15	0	0	2368	6395	53885
16	0	0	1493	4938	49858
17	0	0	897	2639	46413
18	0	0	1699	7779	44671
19	0	0	1894	9802	38591
20	0	0	559	4355	30683
21	0	0	908	12297	26887
22	0	0	162	4554	15498
23	0	0	0	11106	11106



### Πίνακας 5.3

#### Τρίτη περίπτωση

στάσεις	σημαντική(1)/μη σημαντική(0)	Επιβίβαση σημαντική	μη σημαντική	Αποβίβαση	επιβάτες
1	0	0	31783	0	
2	0	0	10680	990	31783
3	0	0	5959	923	41473
4	0	0	13945	1717	46509
5	0	0	9134	1161	58737
6	0	0	9913	2759	66710
7	0	0	4535	2526	73864
8	0	0	11497	7566	75873
9	0	0	35548	35666	79804
10	0	0	8936	13786	79686
11	0	0	4841	9522	74836
12	0	0	3392	8405	70155
13	0	0	3211	9782	65142
14	0	0	2132	6818	58571
15	0	0	2368	6395	53885
16	0	0	1493	4938	49858
17	0	0	897	2639	46413
18	0	0	1699	7779	44671
19	0	0	1894	9802	38591
20	0	0	559	4355	30683
21	0	0	908	12297	26887
22	0	0	162	4554	15498
23	0	0	0	11106	11106

Πίνακας 5.4

Πρώτη περίπτωση: Σε αυτή την περίπτωση εξετάζουμε το πρόβλημα με «στάση αυξημένης ζήτησης» την Ομόνοια. Με την βελτιστοποίηση με το Microsoft Excel Solver προέκυψε ότι η βέλτιστη λύση είναι για τις τιμές των συντελεστών  $\omega_a = 0.5$  και  $\omega_b = 0.9$  για τις οποίες η τιμή του συντελεστή  $D$  πήρε την ελάχιστη τιμή της,  $D = 7.93902 \cdot 10^{-7}$ .

Δεύτερη περίπτωση: Σε αυτή την περίπτωση εξετάζουμε το πρόβλημα με «στάσεις αυξημένης ζήτησης» τον Πειραιά και την Ομόνοια. Με την βελτιστοποίηση με το Microsoft Excel Solver προέκυψε ότι η βέλτιστη λύση είναι για τις τιμές των

συντελεστών  $\omega_a = 0.5$  και  $\omega_b = 0.8$  για τις οποίες η τιμή του συντελεστή  $D$  πήρε την ελάχιστη τιμή της,  $D = 1.93832 \cdot 10^{-8}$ .

Τρίτη περίπτωση: Σε αυτή την περίπτωση εξετάζουμε το πρόβλημα θεωρώντας ότι όλες οι στάσεις έχουν το ίδιο «βάρος» και κατά συνέπεια ισχύει ότι  $\omega_a = 0.5$  και  $\omega_b = 0.5$ . Για αυτές τις τιμές των συντελεστών η τιμή του  $D$  πήρε την τιμή,  $D = 3.11113 \cdot 10^{-6}$ .

Τα αποτελέσματα που προέκυψαν παρουσιάζονται στους παρακάτω πίνακες

### Αποτελέσματα της πρώτης περίπτωσης – σύνολο ημέρας

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
		Πειρ	Ν.Φάλ	Μοσχ	Καλλ	Ταύρ	Πετρ	Θησ	Μοναστ	Ομόν	Βικτ	Αττ	Αγ.Νικ	Κ.Πατ	Αγ.Ελευθ	Α.Πατ	Περισ	Πευκ	Νέα Ιων	Ν.Ηράκλ	Ειρ	Μαρ	Κ.Α.Τ	Κηφ
1	Πειραιάς	0	990	685	1112	573	1176	932	2624	10588	3757	2012	1530	1579	952	796	532	254	698	698	162	134	0	0
2	Ν.Φάληρο	0	0	238	386	199	408	323	910	3672	1303	698	531	547	330	276	184	88	242	242	56	47	0	0
3	Μοσχάτο	0	0	0	220	113	233	184	519	2096	743	398	303	312	188	158	105	50	138	138	32	27	0	0
4	Καλλιθέα	0	0	0	0	276	565	448	1262	5092	1807	967	736	759	458	383	256	122	336	336	78	65	0	0
5	Ταύρος	0	0	0	0	0	378	299	843	3403	1207	646	492	507	306	256	171	82	224	224	52	43	0	0
6	Πετράλωνα	0	0	0	0	0	0	339	955	3852	1367	732	557	574	346	290	193	92	254	254	59	49	0	0
7	Θησείο	0	0	0	0	0	0	0	452	1825	647	347	264	272	164	137	92	44	120	120	28	23	0	0
8	Μοναστηράκι	0	0	0	0	0	0	0	0	5138	1823	976	742	766	462	386	258	123	339	339	79	65	0	0
9	Ομόνοια	0	0	0	0	0	0	0	0	0	1132	823	779	991	796	838	719	427	1331	2162	1551	9408	3646	10944
10	Βικτώρια	0	0	0	0	0	0	0	0	0	0	1923	1463	1509	910	761	508	243	668	668	155	129	0	0
11	Αττική	0	0	0	0	0	0	0	0	0	0	0	1010	1042	628	525	351	168	461	461	107	89	0	0
12	Αγ.Νικόλαος	0	0	0	0	0	0	0	0	0	0	0	0	922	556	465	311	149	408	408	95	79	0	0
13	Κ.Πατήσια	0	0	0	0	0	0	0	0	0	0	0	0	0	723	605	404	193	531	531	123	102	0	0
14	Αγ.Ελευθέριος	0	0	0	0	0	0	0	0	0	0	0	0	0	0	518	346	165	455	455	105	88	0	0
15	Α.Πατήσια	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	508	243	667	667	155	128	0	0
16	Περισσός	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	195	536	535	124	103	0	0
17	Πευκάκια	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	370	370	86	71	0	0
18	Νέα Ιωνία	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1193	276	230	0	0
19	Ν.Ηράκλειο	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1035	859	0	0
20	Ειρήνη	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	559	0	0
21	Μαρούσι	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	908	0
22	Κ.Α.Τ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	162
23	Κηφισιά	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

**Πίνακας 5.5**

### Αποτελέσματα της δεύτερης περίπτωσης – σύνολο ημέρας

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
		Πειρ	Ν.Φάλ	Μοσχ	Καλλ	Ταύρ	Πετρ	Θης	Μοναστ	Ομόν	Βικτ	Αττ	Αγ.Νικ	Κ.Πατ	Αγ.Ελευθ	Α.Πατ	Περισ	Πευκ	Νέα Ιων	Ν.Ηράκλ	Ειρ	Μαρ	Κ.Α.Τ	Κηφ
1	Πειραιάς	0	990	387	551	238	459	354	1004	12424	1277	843	748	898	666	655	523	291	872	1224	650	2324	1101	3304
2	Ν.Φάληρο	0	0	536	735	301	565	415	1133	3126	1285	618	444	437	249	200	129	60	164	161	43	79	0	0
3	Μοσχάτο	0	0	0	432	177	332	244	666	1837	755	363	261	257	146	118	76	35	97	94	25	46	0	0
4	Καλλιθέα	0	0	0	0	445	837	615	1680	4634	1905	916	658	648	369	297	192	89	243	238	63	117	0	0
5	Ταύρος	0	0	0	0	0	566	416	1136	3135	1289	620	445	438	249	201	130	61	165	161	43	79	0	0
6	Πετράλωνα	0	0	0	0	0	0	482	1315	3627	1491	717	515	507	289	232	150	70	191	186	50	91	0	0
7	Θησείο	0	0	0	0	0	0	0	632	1744	717	345	248	244	139	112	72	34	92	90	24	44	0	0
8	Μοναστηράκι	0	0	0	0	0	0	0	0	5138	2113	1015	729	718	409	329	212	99	270	264	70	129	0	0
9	Ομόνοια	0	0	0	0	0	0	0	0	2953	1949	1730	2075	1539	1515	1208	672	2016	2831	1504	5372	2545	7639	
10	Βικτώρια	0	0	0	0	0	0	0	0	0	2137	1535	1511	861	693	447	209	568	555	148	272	0	0	
11	Αττική	0	0	0	0	0	0	0	0	0	0	0	1093	1076	613	493	318	149	405	395	105	194	0	0
12	Αγ.Νικόλαος	0	0	0	0	0	0	0	0	0	0	0	0	974	555	446	288	135	366	358	95	175	0	0
13	Κ.Πατήσια	0	0	0	0	0	0	0	0	0	0	0	0	0	736	593	382	179	486	475	127	233	0	0
14	Αγ.Ελευθέριος	0	0	0	0	0	0	0	0	0	0	0	0	0	0	511	329	154	419	409	109	201	0	0
15	Α.Πατήσια	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	481	225	612	598	160	293	0	0
16	Περισσός	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	178	484	473	126	232	0	0
17	Πευκάκια	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	330	323	86	158	0	0
18	Νέα Ιωνία	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	967	258	474	0	0
19	Ν.Ηράκλειο	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	668	1226	0	0
20	Ειρήνη	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	559	0	0
21	Μαρούσι	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	908	0
22	Κ.Α.Τ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	162
23	Κηφισιά	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

**Πίνακας 5.6**

### Αποτελέσματα της τρίτης περίπτωσης – σύνολο ημέρας

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
		Πειρ	Ν.Φάλ	Μοσχ	Καλλ	Ταύρ	Πετρ	Θησ	Μοναστ	Ομόν	Βικτ	Αττ	Αγ.Νικ	Κ.Πατ	Αγ.Ελευθ	Α.Πατ	Περισ	Πευκ	Νέα Ιων	Ν.Ηράκλ	Ειρ	Μαρ	Κ.Α.Τ	Κηφ
1	Πειραιάς	0	990	685	1112	573	1176	932	2624	10588	2267	1379	1133	1250	824	742	546	282	815	982	409	1132	395	948
2	Ν.Φάληρο	0	0	238	386	199	408	323	910	3672	786	478	393	434	286	257	189	98	283	341	142	393	137	329
3	Μοσχάτο	0	0	0	220	113	233	184	519	2096	449	273	224	247	163	147	108	56	161	194	81	224	78	188
4	Καλλιθέα	0	0	0	0	276	565	448	1262	5092	1090	663	545	601	396	357	262	136	392	472	197	544	190	456
5	Ταύρος	0	0	0	0	0	378	299	843	3403	728	443	364	402	265	238	175	91	262	316	132	364	127	305
6	Πετράλωνα	0	0	0	0	0	0	339	955	3852	825	502	412	455	300	270	199	103	297	357	149	412	144	345
7	Θησείο	0	0	0	0	0	0	0	452	1825	391	238	195	215	142	128	94	49	140	169	71	195	68	163
8	Μοναστηράκι	0	0	0	0	0	0	0	0	5138	1100	669	550	607	400	360	265	137	396	477	199	549	191	460
9	Ομόνοια	0	0	0	0	0	0	0	0	0	6150	3741	3074	3391	2234	2013	1480	766	2212	2664	1111	3071	1071	2573
10	Βικτώρια	0	0	0	0	0	0	0	0	0	0	1137	934	1031	679	612	450	233	672	810	338	933	325	782
11	Αττική	0	0	0	0	0	0	0	0	0	0	0	580	640	422	380	279	144	417	503	210	579	202	485
12	Αγ.Νικόλαος	0	0	0	0	0	0	0	0	0	0	0	0	509	336	302	222	115	332	400	167	461	161	386
13	Κ.Πατήσια	0	0	0	0	0	0	0	0	0	0	0	0	0	374	337	248	128	370	446	186	514	179	430
14	Αγ.Ελευθέριος	0	0	0	0	0	0	0	0	0	0	0	0	0	0	253	186	96	278	335	140	386	135	323
15	Α.Πατήσια	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	235	121	350	422	176	487	170	408
16	Περισσός	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	85	245	295	123	340	119	285
17	Πευκάκια	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	156	188	78	217	76	182
18	Νέα Ιωνία	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	432	180	497	173	417
19	Ν.Ηράκλειο	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	269	743	259	623
20	Ειρήνη	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	256	89	214
21	Μαρούσι	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	267	641
22	Κ.Α.Τ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	162
23	Κηφισιά	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

**Πίνακας 5.7**

Τα αποτελέσματα αυτά δεν μπορούμε να πούμε ότι κρίνονται ικανοποιητικά. Παρατηρούμε ότι ο αλγόριθμος δεν κάνει καλή κατανομή του αριθμού των μετακινήσεων και παρουσιάζει κάποια «συστηματικά σφάλματα». Πιο συγκεκριμένα, βλέπουμε ότι μηδενίζει όλα τα κελιά των τελευταίων δύο στάσεων εκτός από αυτά που αντιστοιχούν στις σημαντικές στάσεις που έχουμε θέσει κάθε φορά. Με βάση αυτό μπορούμε να πούμε ότι ο αλγόριθμος δίνει μεγαλύτερη βαρύτητα σε αυτές τις στάσεις απ' ό,τι θα έπρεπε με αποτέλεσμα να αντιστοιχεί σε αυτές περισσότερες μετακινήσεις και να «υποτροφοδοτεί» τις υπόλοιπες στάσεις. Επίσης βλέπουμε ότι πάντα στο πρώτο κελί κάθε γραμμής δίνει τιμή πολύ μεγαλύτερη από την πραγματική. Οι δύο παραπάνω επισημάνσεις παρουσιάζονται στην πρώτη και δεύτερη περίπτωση που εμείς εξετάσαμε και οδηγούμαστε στο συμπέρασμα ότι είναι κάποια δυσλειτουργία του αλγορίθμου, πιθανότατα στο σημείο που δίνει βαρύτητα σε προκαθορισμένες στάσεις μέσω των συντελεστών  $\omega_a, \omega_b$  και δεν οφείλεται σε κάποιο ιδιαίτερο χαρακτηριστικό των δεδομένων μας. Το συμπέρασμα αυτό ενισχύεται και από τα αρκετά καλύτερα αποτελέσματα που έδωσε ο αλγόριθμος στην τρίτη περίπτωση, όπου δεν θεωρήσαμε σημαντική κάποια στάση. Σε αυτή την περίπτωση δεν μηδενίστηκε κανένα κελί του πίνακα και έκανε μια πολύ καλύτερη κατανομή στις μετακινήσεις.

Παρακάτω παρουσιάζονται κάποια στατιστικά μεγέθη και γραφήματα των αποτελεσμάτων

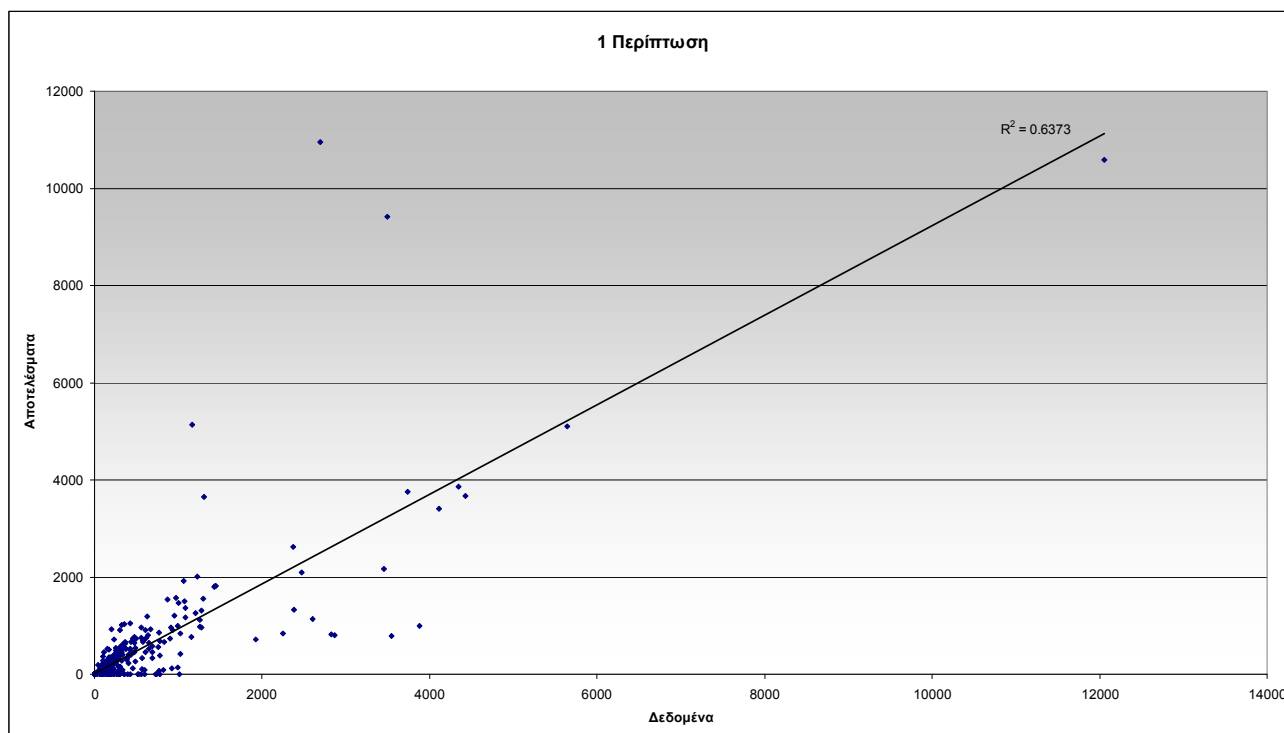
#### Πρώτη περίπτωση:

<u>Στατιστικά παλινδρόμησης</u>	
Πολλαπλό R	0.798
R Τετράγωνο	0.637
Προσαρμοσμένο R Τετράγωνο	0.637
Τυπικό σφάλμα	579.957
Μέγεθος δείγματος	529.000

<u>ΑΝΑΛΥΣΗ ΔΙΑΚΥΜΑΝΣΗΣ</u>					
	<i>βαθμοί ελευθερίας</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Σημαντικότητα F</i>
Παλινδρόμηση	1	311520256	311520256	926	0
Υπόλοιπο	527	177256710	336350		

	Συντελεστές	Τυπικό σφάλμα	t	τιμή-P	Κατώτερο 95%	Υψηλότερο 95%	Κατώτερο 95.0%	Υψηλότερο 95.0%
Τεταγμένη επί την αρχή	24.714	26.934	0.918	0.359	-28.198	77.625	-28.198	77.625
Μεταβλητή X 1	0.921	0.030	30.433	0.000	0.862	0.980	0.862	0.980

**Πίνακας 5.8**



**Γράφημα 5.2**

Δεύτερη περίπτωση:

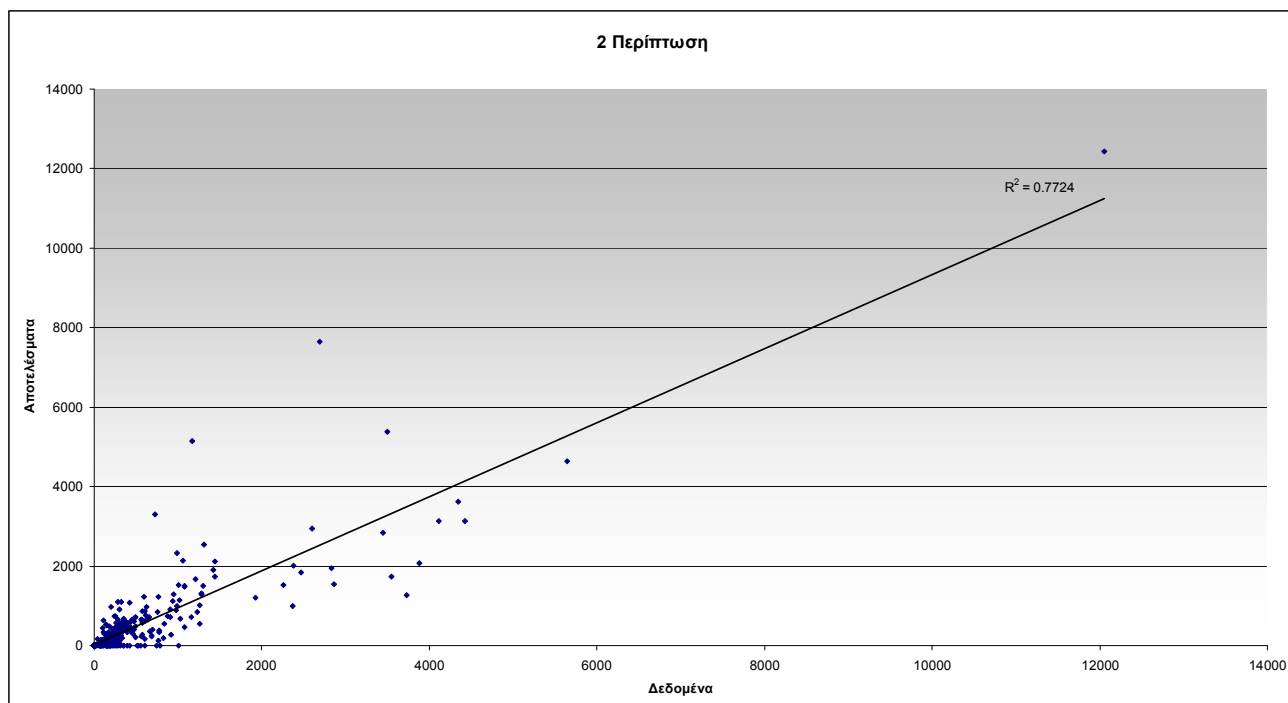
Στατιστικά παλινδρόμησης	
Πολλαπλό R	0.879
R Τετράγωνο	0.772
Προσαρμοσμένο R Τετράγωνο	0.772
Τυπικό σφάλμα	422.042
Μέγεθος δείγματος	529.000

ΑΝΑΛΥΣΗ  
ΔΙΑΚΥΜΑΝΣΗΣ

	βαθμοί ελευθερίας	SS	MS	F	Σημαντικότητα F
Παλινδρόμηση	1	318515461	318515461	1788	0
Υπόλοιπο	527	93869076	178120		
Σύνολο	528	412384537			

	Συντελεστές	Τυπικό σφάλμα	t	τιμή- P	Κατώτερο 95%	Υψηλότερο 95%	Κατώτερο 95.0%	Υψηλότερο 95.0%
Τεταγμένη επί την αρχή	21.503	19.600	1.097	0.273	-17.002	60.007	-17.002	60.007
Μεταβλητή X 1	0.931	0.022	42.287	0.000	0.888	0.975	0.888	0.975

**Πίνακας 5.9**



**Γράφημα 5.3**

Τρίτη περίπτωση:

Στατιστικά παλινδρόμησης	
Πολλαπλό R	0.941
R Τετράγωνο	0.885
Προσαρμοσμένο R Τετράγωνο	0.885
Τυπικό σφάλμα	271.643
Μέγεθος δείγματος	529.000

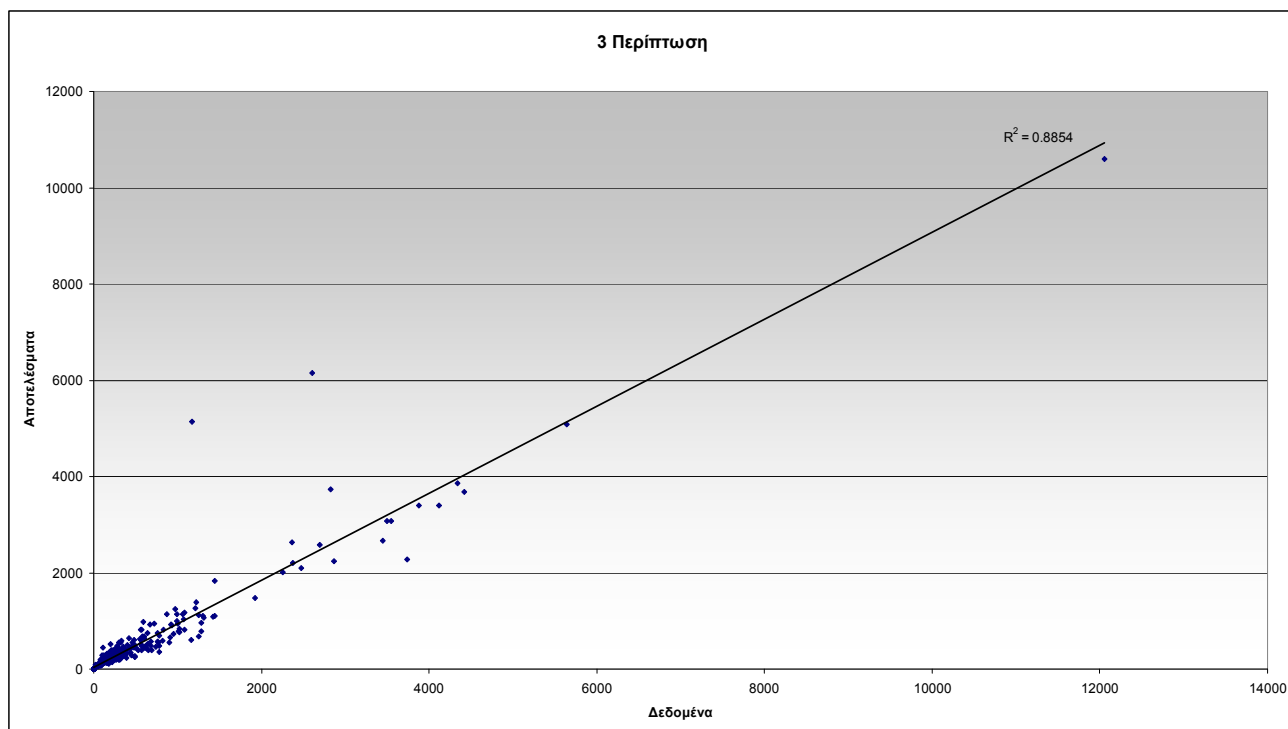


ΑΝΑΛΥΣΗ  
ΔΙΑΚΥΜΑΝΣΗΣ

	βαθμοί ελευθερίας	SS	MS	F	Σημαντικότητα F
Παλινδρόμηση	1	300347409	300347409	4070	0
Υπόλοιπο	527	38887145	73790		
Σύνολο	528	339234553			

	Συντελεστές	Τυπικό σφάλμα	t	τιμή- P	Κατώτερο 95%	Υψηλότερο 95%	Κατώτερο 95.0%	Υψηλότερο 95.0%
Τεταγμένη επί την αρχή	29.943	12.616	2.373	0.018	5.160	54.726	5.160	54.726
Μεταβλητή X 1	0.904	0.014	63.799	0.000	0.876	0.932	0.876	0.932

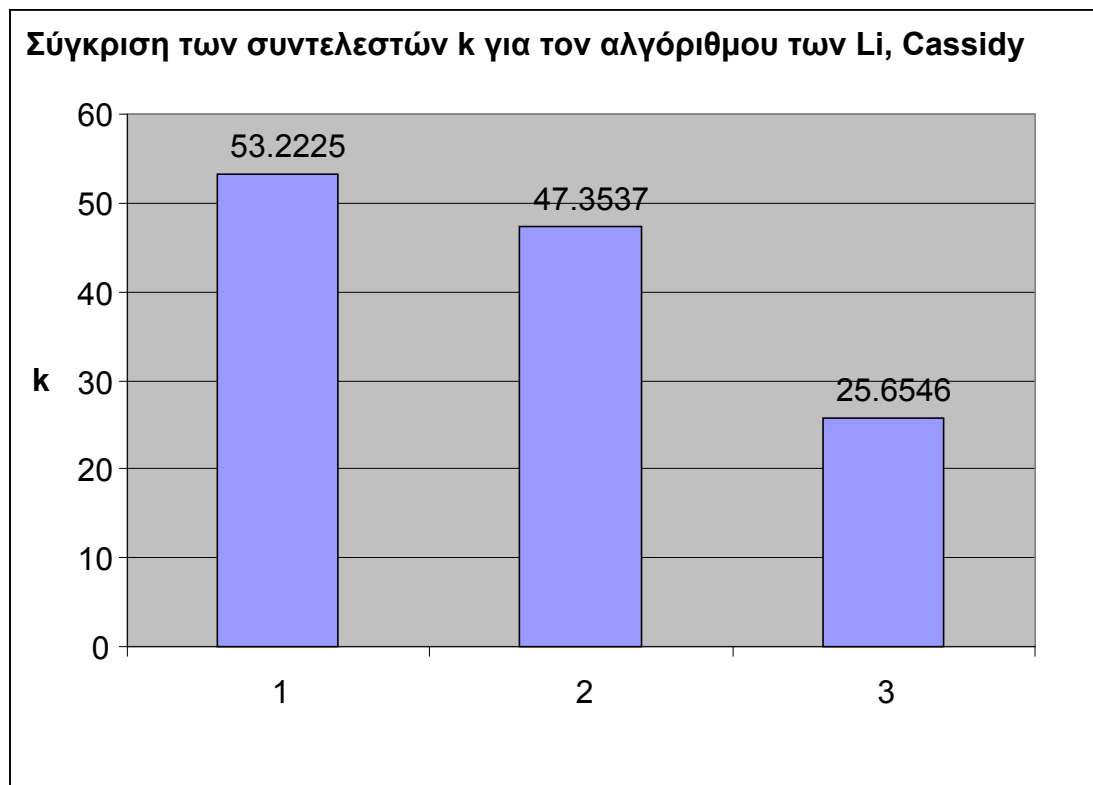
Πίνακας 5.10



Γράφημα 5.4

Στην συνέχεια παρουσιάζεται ένα συγκριτικό διάγραμμα του συντελεστή  $k$  (σταθμισμένη ποσοστιαία μεταβολή αποτελεσμάτων από τα δεδομένα) ανάμεσα στις τρεις περιπτώσεις που εξετάσαμε αλλά και με τις εξισώσεις μεγιστοποίησης της εντροπίας (για τους αντίστοιχους πάντα τριγωνικούς πίνακες). Έστω ο γνωστός μας πίνακας  $TT_{ij}$  και ο πίνακας που προσδιορίσαμε  $T_{ij}$ . Ο συντελεστής  $k$  δίνεται από τον ακόλουθο τύπο:

$$k = \sum_{ij} \left[ \text{abs} \left( \frac{T_{ij} - TT_{ij}}{t_{ij}} \cdot 100 \right) \cdot \frac{TT_{ij}}{\sum_{ij} TT_{ij}} \right]$$



**Γράφημα 5.5**

## 5.4. Εφαρμογή των μοντέλων μεγιστοποίησης εντροπίας

Αυτό που θέλουμε να προσδιορίσουμε σε αυτό το κεφάλαιο είναι η ακρίβεια των μοντέλων και ο τρόπος που συμπεριφέρονται ανάλογα με την κλίμακα του προβλήματος.

Τα μοντέλα που συγκεντρώσαμε από την εργασία του Paramahamsan (1999) προσπαθούν να προσδιορίσουν τον πίνακα Π-Π έχοντας ως αφετηρία έναν αρχικό πίνακα Π-Π της περιοχής μελέτης και κάποιες μετρήσεις φόρτων κυκλοφορίας των οδικών συνδέσμων των ζωνών. Ουσιαστικά αυτό που γίνεται είναι μία επικαιροποίηση του αρχικού πίνακα Π-Π.

Πρώτα ας δούμε την διάσταση του προβλήματος δηλαδή τους αγνώστους που πρέπει να προσδιοριστούν και τα δεδομένα που υπάρχουν. Γενικά ισχύει ότι οι άγνωστοι που πρέπει να προσδιοριστούν σε ένα πρόβλημα εκτίμησης ενός πίνακα Π-Π είναι  $N^2 - N$ , δηλαδή όσα τα κελία του τετραγωνικού πίνακα μείον τα κελία της διαγωνίου (οι ενδοζωνικές μετακινήσεις δεν λαμβάνονται υπόψη). Στην δική μας περίπτωση ισχύει ότι οι άγνωστοι είναι  $23^2 - 23$  δηλαδή 506 και οι γνωστές μας σχέσεις είναι  $23 \cdot 2 = 46$ , όπου είναι και η ιδανική περίπτωση αφού γνωρίζουμε όλους τους φόρτους (σε κάθε σταθμό γνωρίζουμε πόσοι επιβάτες επιβιβάζονται και πόσοι αποβιβάζονται). Βλέπουμε λοιπόν ότι έχουμε να αντιμετωπίσουμε ένα υποορισμένο πρόβλημα.

### 5.4.1. Matlab

Η επίλυση των εξισώσεων έγινε με το υπολογιστικό πακέτο Matlab (Matrix Laboratory), το οποίο είναι ένα πρόγραμμα υψηλών επιδόσεων και κατασκευάζεται από την The Math Works Inc. Περιλαμβάνει ένα πλήθος από εργαλειοθήκες, η κάθε μια από τις οποίες είναι κατασκευασμένη για την αντιμετώπιση συγκεκριμένων κατηγοριών προβλημάτων καθώς και ένα πακέτο προσομοίωσης δυναμικών συστημάτων (Simulink). Εμείς χρησιμοποιήσαμε την εργαλειοθήκη βελτιστοποίησης, η οποία περιέχει ρουτίνες βελτιστοποίησης για διάφορα συστήματα. Πιο

συγκεκριμένα χρησιμοποιήθηκε η ρουτίνα βελτιστοποίησης του προγράμματος για μη γραμμικά συστήματα υπό περιορισμούς (fmincon).

Το Matlab χρησιμοποιεί τον αλγόριθμο Gauss-Newton (Παράρτημα Β) σαν προεπιλογή για να λύνει μη γραμμικές εξισώσεις. Αν ο αλγόριθμος Gauss-Newton δεν δώσει καλή σύγκλιση τότε το πρόγραμμα αυτόματα αλλάζει και χρησιμοποιεί τον αλγόριθμο Levenberg-Marquardt (LM) (Παράρτημα Γ). Ο αλγόριθμος LM είναι μια τροποποίηση του αλγόριθμου Gauss-Newton, ο οποίος είναι πιο δυνατός και χρειάζεται λιγότερο υπολογιστικό χρόνο. Αυτό που θα πρέπει να τονίσουμε εδώ όμως είναι το πρόβλημα των αρχικών τιμών που έχουν οι αλγόριθμοι επίλυσης. Το πρόγραμμα με τους αλγόριθμους που χρησιμοποιεί προσπαθεί να αντιμετωπίσει το πρόβλημα αυτό αλλά πάντα υπάρχει ο κίνδυνος η λύση που θα δοθεί να είναι κάποιο τοπικό μέγιστο ή ελάχιστο αντί για το ζητούμενο ολικό. Ο μοναδικός τρόπος να αντιμετωπιστεί αυτό το πρόβλημα είναι η δοκιμή διαφορετικών αρχικών τιμών. Σημαντικό είναι οι αρχικές τιμές να είναι από όλο το εύρος των δυνατών τιμών που μπορούν να πάρουν οι μεταβλητές. Αν το αποτέλεσμα είναι το ίδιο για τις διαφορετικές αυτές αρχικές τιμές, τότε σημαίνει ότι έχουμε φτάσει στην επιθυμητή λύση.

#### **5.4.2. Αρχικός πίνακας**

Όπως έχουμε πει, τα μοντέλα της εντροπίας κάνουν επικαιροποίηση σε έναν αρχικό πίνακα Π-Π με βάση κυκλοφοριακούς φόρτους. Εμείς γνωρίζαμε εξ αρχής τον «ζητούμενο» πίνακα Π-Π και τους κυκλοφοριακούς φόρτους. Αυτό που δεν είχαμε ήταν ένας αρχικός πίνακας ώστε να τον χρησιμοποιήσουμε ως βάση. Για να αντιμετωπίσουμε αυτό το πρόβλημα κάναμε το εξής: διαιρέσαμε τον γνωστό πίνακα που είχαμε με έναν τυχαίο αριθμό που είχε τιμή από 0.7 έως 1.3. Το εύρος αυτό δεν επιλέχθηκε τυχαία.

Η επικαιροποίηση θα πρέπει να γίνεται ανά τακτά χρονικά διαστήματα. Αυτή είναι μια πολύ σημαντική προϋπόθεση που εξασφαλίζει σε μεγάλο βαθμό την χρονική συνοχή των δεδομένων που χρησιμοποιούμε, δηλαδή τον αρχικό πίνακα Π-Π και τους καινούργιους φόρτους κυκλοφορίας που έχουν μετρηθεί. Επειδή ο αρχικός πίνακας

και οι φόρτοι κυκλοφορίας είναι δεδομένα τα οποία έχουν μετρηθεί σε διαφορετικές χρονικές περιόδους, υπάρχει ο κίνδυνος να μην αναφέρονται στο ίδιο σύστημα, λόγω διαφόρων μεταβολών που μπορεί να έχουν γίνει σε αυτό (δίκτυο, μέγεθος πληθυσμού, ζήτηση των μετακινήσεων). Παραδείγματος χάριν δεν γίνεται να επικαιροποιήσουμε ένα πίνακα Π-Π μετά από δέκα χρόνια γιατί το δίκτυο από το οποίο θα έχουμε μετρήσει τους κυκλοφοριακούς φόρτους δεν θα έχει καμία σχέση με το δίκτυο στο οποίο αντιστοιχεί ο αρχικός πίνακας λόγω των αλλαγών που θα έχουν μεσολαβήσει. Έτσι θέτοντας ένα μικρό χρονικό όριο στο οποίο μπορούν να γίνονται οι ενημερώσεις, προσπαθούμε να εξασφαλίσουμε την σταθερότητα του συστήματος, έστω και με κάποιες παραδοχές.

Οπότε διαιρώντας κάθε κελί του γνωστού μας πίνακα με τυχαίο αριθμό που έχει τιμή από 0.7 έως 1.3 δημιουργούμε ένα τυχαίο, παλαιό πίνακα που η τιμή του κάθε κελιού του έχει μια μέγιστη απόκλιση της τάξεως του  $\pm 30\%$ .

Ο αρχικός πίνακας που δημιουργήθηκε από τον πίνακα 5.1 είναι ο παρακάτω.

### Αρχικός πίνακας - σύνολο ημέρα

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
		Πειρ	Ν.Φάλ	Μοσχ	Καλλ	Ταύρ	Πετρ	Θησ	Μοναστ	Ομόν	Βικτ	Αττ	Αγ.Νικ	Κ.Πατ	Αγ.Ελευθ	Α.Πατ	Περισ	Πευκ	Νέα Ιωνία	Ν.Ηράκ	Ειρ	Μαρ	Κ.Α.Τ	Κηφ
1	Πειραιάς	0	1238	777	967	979	1199	744	2633	12056	3111	1021	730	973	468	581	461	95	480	455	314	988	255	906
2	Ν.Φάληρο	865	0	146	360	217	403	297	1168	4022	1826	637	336	381	186	231	160	65	156	156	104	235	72	215
3	Μοσχάτο	648	163	0	231	200	257	176	644	2249	513	383	128	172	70	167	90	35	111	72	13	98	27	64
4	Καλλιθέα	1449	334	185	0	92	585	552	1099	5134	1295	907	352	623	273	218	248	76	269	254	132	338	93	233
5	Ταύρος	678	151	242	199	0	371	293	1276	5879	1357	596	231	314	238	148	98	59	171	125	53	232	60	118
6	Πετράλωνα	1046	241	207	886	145	0	300	1599	3618	1083	382	275	266	279	240	163	68	236	279	71	227	71	158
7	Θησείο	732	311	190	511	543	289	0	156	1799	594	392	350	320	175	174	114	51	167	173	119	374	90	174
8	Μοναστηράκι	1693	625	761	1114	1103	1921	125	0	1461	2066	1257	822	965	990	599	700	280	634	761	338	768	280	674
9	Ομόνοια	7044	2167	1325	2592	2468	3720	925	613	0	2603	2175	4435	3230	3182	1878	1923	854	2161	3450	1183	4993	1008	3368
10	Βικτώρια	3309	1094	671	1147	1195	763	581	1199	2652	0	1061	1003	1075	670	368	478	207	824	642	274	765	395	1446
11	Αττική	1229	704	344	527	463	527	380	632	2200	1038	0	252	424	302	285	226	116	393	395	217	913	242	978
12	Αγ.Νικόλαος	852	252	123	363	163	214	236	493	2317	841	168	0	226	307	380	163	110	437	260	185	859	129	587
13	Κ.Πατήσια	1446	383	175	480	460	351	265	832	4723	817	553	141	0	238	305	230	183	313	614	215	434	257	469
14	Αγ.Ελευθέριος	566	228	78	281	175	162	134	563	1732	585	491	270	236	0	167	189	104	373	261	162	336	106	443
15	Α.Πατήσια	558	211	114	423	142	194	132	462	1984	458	378	221	441	106	0	221	123	307	395	131	418	172	473
16	Περισσός	418	159	84	266	108	126	79	228	1265	800	243	140	196	132	323	0	60	489	195	98	275	93	281
17	Πευκάκια	160	47	32	153	44	52	79	107	797	352	134	111	90	91	105	31	0	82	329	89	157	96	208
18	Νέα Ιωνία	509	159	123	386	113	145	131	511	1267	741	266	426	283	415	335	418	175	0	524	130	401	159	510
19	Ν.Ηράκλειο	721	172	131	296	125	216	137	398	3068	840	601	383	258	290	377	263	228	757	0	394	638	169	693
20	Ειρήνη	488	75	66	181	73	104	79	319	1212	401	203	98	277	135	303	68	83	168	372	0	285	43	173
21	Μαρούσι	615	189	95	266	169	118	126	386	2410	817	452	294	440	303	461	227	254	438	430	229	0	340	547
22	Κ.Α.Τ	243	93	32	203	53	73	90	256	1161	501	215	121	137	95	154	62	74	169	163	76	259	0	203
23	Κηφισιά	718	165	83	356	174	171	142	380	2320	824	671	261	450	279	438	333	217	303	396	334	716	138	0

Πίνακας 5.11

### 5.4.3. Αποτελέσματα

Όπως έχουμε πει, για να επιλύσουμε τα μοντέλα χρησιμοποιήσαμε το υπολογιστικό πακέτο Matlab. Ο προγραμματισμός έγινε με γενικευμένη μορφή ώστε να μπορεί να επιλυθεί οποιασδήποτε κλίμακας πρόβλημα με ελάχιστες τροποποιήσεις του προγράμματος. Ως βάση χρησιμοποιήσαμε τον αρχικό πίνακα που προσδιορίσαμε παραπάνω. Οι περιορισμοί μας ήταν τα αθροίσματα των γραμμών και των στηλών του πίνακα που θέλουμε να προσδιορίσουμε.

Τα αποτελέσματα που προέκυψαν παρουσιάζονται στους πίνακες 5.12, 5.13, 5.14. Αυτό που παρατηρούμε είναι ότι για τα μοντέλα 1 έως 6 ο εκτιμώμενος πίνακας που προκύπτει είναι κοινός. Αυτό οφείλεται στο γεγονός ότι τα μοντέλα αυτά είναι στην ουσία διαφορετικές μορφές του βασικού μοντέλου μεγιστοποίησης της εντροπίας. Επίσης παρατηρούμε ότι το ίδιο ισχύει και για τα μοντέλα 7 και 8 αφού και αυτά αποτελούν διαφορετικές μορφές ενός μοντέλου.

### Αποτελέσματα για από τα μοντέλα 1 έως 6 – σύνολο ημέρας

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
		Πειρ	Ν.Φάλ	Μοσχ	Καλλ	Ταύρ	Πετρ	Θησ	Μοναστ	Ομόν	Βικτ	Αττ	Αγ.Νικ	Κ.Πατ	Αγ.Ελευθ	Α.Πατ	Περισ	Πευκ	Νέα Ιωνία	Ν.Ηράκ	Ειρ	Μαρ	Κ.Α.Τ	Κηφ
1	Πειραιάς	0	1414	706	1571	979	1541	838	1975	13396	4666	1531	674	1081	803	710	359	112	443	455	236	898	234	659
2	Ν.Φάληρο	1249	0	133	252	177	448	223	718	6320	983	441	305	286	155	203	98	65	223	208	85	336	99	215
3	Μοσχάτο	1203	144	0	208	150	257	301	537	2749	513	437	128	158	76	98	53	35	74	123	24	108	32	64
4	Καλλιθέα	1317	213	301	0	112	760	607	1343	8067	2034	1296	381	701	298	218	191	124	307	279	112	276	102	255
5	Ταύρος	904	226	218	107	0	325	346	785	5144	950	681	254	242	158	148	146	59	140	147	63	290	87	219
6	Πετράλωνα	1046	447	276	1013	236	0	355	1279	4342	1083	451	275	384	177	200	163	68	260	163	71	357	71	238
7	Θησείο	732	311	190	803	317	361	0	84	1107	502	302	272	457	135	121	62	64	115	173	104	374	80	121
8	Μοναστηράκι	2258	739	533	1361	1516	1223	139	0	899	1446	1143	822	1654	693	779	445	249	571	685	304	698	342	674
9	Ομόνοια	5764	3714	1122	2808	4011	2003	783	736	0	2366	2175	3942	4307	3182	2253	2747	932	1828	2654	1001	4993	1873	2072
10	Βικτώρια	2941	985	537	1639	869	1049	913	1066	2040	0	884	912	896	861	479	478	248	444	642	274	1148	304	843
11	Αττική	851	493	393	527	308	410	234	1028	2200	1153	0	273	326	402	370	226	129	617	395	197	1174	242	652
12	Αγ.Νικόλαος	1096	378	145	363	224	294	212	740	2527	712	223	0	185	282	234	178	110	336	260	227	601	119	440
13	Κ.Πατήσια	1124	383	108	440	460	351	294	624	4723	1517	323	194	0	198	305	230	128	417	538	195	470	150	645
14	Αγ.Ελευθέριος	636	152	85	234	193	243	230	303	2721	495	302	245	270	0	150	142	116	447	221	122	305	138	443
15	Α.Πατήσια	1037	188	171	307	156	194	156	635	1526	688	486	170	272	90	0	177	160	409	395	131	511	132	387
16	Περισσός	278	159	63	194	129	168	73	248	1265	622	297	128	166	143	287	0	35	445	316	154	275	93	152
17	Πευκάκια	206	58	29	119	36	36	46	128	797	396	72	71	106	76	124	50	0	118	269	89	247	74	208
18	Νέα Ιωνία	431	130	86	338	95	236	101	298	1497	847	244	426	368	474	335	460	160	0	699	154	446	119	510
19	Ν.Ηράκλειο	444	172	92	326	135	216	151	576	3506	840	421	206	419	193	539	198	330	973	0	444	588	183	554
20	Ειρήνη	433	98	44	151	67	81	95	279	1904	441	186	91	216	166	220	74	64	202	479	0	348	56	190
21	Μαρούσι	752	210	63	242	169	220	164	386	1668	692	417	321	400	341	346	175	137	389	508	172	0	278	860
22	Κ.Α.Τ	205	84	60	180	91	97	53	288	950	455	235	136	234	95	154	62	49	135	163	62	421	0	231
23	Κηφισιά	790	201	107	457	126	214	243	570	3646	1530	671	308	381	279	371	179	190	256	396	180	788	166	0

**Πίνακας 5.12**



**Αποτελέσματα για τα μοντέλα 7 και 8 με κατεύθυνση Πειραιάς-Κηφισιά – σύνολο ημέρας**

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
		Πειρ	Ν.Φάλ	Μοσχ	Καλλ	Ταύρ	Πετρ	Θησ	Μοναστ	Ομόν	Βικτ	Αττ	Αγ.Νικ	Κ.Πατ	Αγ.Ελευθ	Α.Πατ	Περισ	Πευκ	Νέα Ιωνία	Ν.Ηράκ	Ειρ	Μαρ	Κ.Α.Τ	Κηφ
1	Πειραιάς	0	990	785	1241	600	1197	561	1863	14068	2795	883	814	972	422	657	281	190	666	562	183	862	340	853
2	Ν.Φάληρο	0	0	138	248	216	410	315	807	4512	1331	644	381	247	209	66	210	72	90	134	61	264	47	277
3	Μοσχάτο	0	0	0	227	205	52	282	792	2516	594	309	274	147	96	60	61	29	40	12	52	118	22	70
4	Καλλιθέα	0	0	0	0	140	918	588	1599	4628	1945	751	473	784	254	193	298	90	158	153	120	408	113	331
5	Ταύρος	0	0	0	0	0	183	340	871	4838	1075	434	295	274	91	93	121	22	35	109	0	199	15	138
6	Πετράλωνα	0	0	0	0	0	0	439	1539	3392	1062	615	417	339	323	261	186	144	223	131	77	347	139	280
7	Θησείο	0	0	0	0	0	0	0	96	1096	623	309	370	304	230	215	148	114	94	197	86	350	135	168
8	Μοναστηράκι	0	0	0	0	0	0	0	0	617	1942	1033	837	1602	932	761	482	136	340	550	213	1091	323	637
9	Ομόνοια	0	0	0	0	0	0	0	0	2419	3368	2987	3118	2746	2653	1535	760	2806	4734	1712	2623	885	3204	
10	Βικτώρια	0	0	0	0	0	0	0	0	0	0	1177	1160	1442	570	384	305	145	674	465	168	870	458	1119
11	Αττική	0	0	0	0	0	0	0	0	0	0	0	397	355	328	417	257	143	385	602	170	966	262	560
12	Αγ.Νικόλαος	0	0	0	0	0	0	0	0	0	0	0	0	197	310	290	237	119	530	238	202	548	147	573
13	Κ.Πατήσια	0	0	0	0	0	0	0	0	0	0	0	0	0	308	244	238	141	301	472	193	692	262	360
14	Αγ.Ελευθέριος	0	0	0	0	0	0	0	0	0	0	0	0	0	0	101	316	190	351	175	152	413	160	274
15	Α.Πατήσια	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	264	263	493	303	146	421	185	293
16	Περισσός	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	81	476	155	120	388	116	158
17	Πευκάκια	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	117	251	71	234	122	101
18	Νέα Ιωνία	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	559	185	395	201	360
19	Ν.Ηράκλειο	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	444	784	250	416
20	Ειρήνη	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	325	102	132
21	Μαρούσι	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	269	639
22	Κ.Α.Τ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	162
23	Κηφισιά	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

**Πίνακας 5.13**

**Αποτελέσματα για τα μοντέλα 7 και 8 με κατεύθυνση Κηφισιά-Πειραιάς – σύνολο ημέρας**

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
		Κηφ	Κ.Α.Τ	Μαρ	Ειρ	Ν.Ηράκλ	Νέα Ιωνία	Πευκ	Περισ	Α.Πατ	Αγ.Ελευθ	Κ.Πατ	Αγ.Νικ	Αττ	Βικτ	Ομόν	Μοναστ	Θησ	Πετρ	Ταύρ	Καλλ	Μοσχ	Ν.Φάλ	Πειρα
1	Κηφισιά	0	166	722	186	526	322	182	196	462	274	604	247	555	812	2912	403	221	116	118	221	71	133	1056
2	Κ.Α.Τ	0	0	403	70	177	204	97	84	150	160	141	113	298	521	800	276	85	51	76	127	55	95	229
3	Μαρούσι	0	0	0	252	559	374	198	229	502	250	429	474	552	1276	1394	340	149	94	184	305	87	251	779
4	Ειρήνη	0	0	0	0	255	178	69	109	221	122	263	63	194	476	1625	192	63	42	108	227	67	75	253
5	Ν.Ηράκλειο	0	0	0	0	0	623	306	278	340	157	222	244	411	1062	3210	446	81	139	233	209	59	81	490
6	Νέα Ιωνία	0	0	0	0	0	0	110	403	497	307	455	331	317	672	2019	235	57	84	110	301	65	52	433
7	Πευκάκια	0	0	0	0	0	0	0	68	117	120	126	67	121	365	543	117	84	48	83	148	45	87	133
8	Περισσός	0	0	0	0	0	0	0	0	208	235	177	180	205	605	1773	217	115	62	187	204	45	152	342
9	Α.Πατήσια	0	0	0	0	0	0	0	0	0	143	310	217	348	705	1957	412	204	215	179	375	175	201	565
10	Αγ.Ελευθέριος	0	0	0	0	0	0	0	0	0	0	144	179	320	641	2453	303	153	180	215	235	73	153	358
11	Κ.Πατήσια	0	0	0	0	0	0	0	0	0	0	0	249	419	861	3453	796	239	228	337	554	189	322	906
12	Αγ.Νικόλαος	0	0	0	0	0	0	0	0	0	0	0	0	284	935	1971	840	249	322	300	361	241	323	890
13	Αττική	0	0	0	0	0	0	0	0	0	0	0	0	0	761	2338	713	338	457	450	706	193	660	1485
14	Βικτώρια	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2460	1142	778	738	833	1098	491	1159	2662
15	Ομόνοια	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	923	962	3042	3126	3535	1516	2341	5889
16	Μοναστηράκι	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	43	1383	881	1019	374	818	2768
17	Θησείο	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	276	439	676	119	189	592
18	Πετράλωνα	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	199	723	227	414	732
19	Ταύρος	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	167	247	254	683
20	Καλλιθέα	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	382	281	1260
21	Μοσχάτο	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	972
22	Ν.Φάληρο	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1124
23	Πειραιάς	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

**Πίνακας 5.14**

Με βάση τους πίνακες των αποτελεσμάτων έχουμε να κάνουμε κάποιες παρατηρήσεις. Γενικά τα αποτελέσματα χαρακτηρίζονται αρκετά καλά ειδικά αν συνυπολογίσει κανείς το γεγονός ότι προσδιορίζουμε 506 αγνώστους με 46 σχέσεις. Αυτό που πρέπει να πούμε είναι ότι παρατηρείται μια απόκλιση στους περιορισμούς. Αυτό το πρόβλημα παρουσιάζεται στα μοντέλα 1 έως 6 ενώ στα μοντέλα 7 και 8 οι περιορισμοί τηρούνται απόλυτα. Παρακάτω φαίνεται η απόκλιση των περιορισμών:

### Αθροίσματα γραμμών (Προέλευση)

Στάσεις	Δεδομένα	Αποτελέσματα	% μεταβολή	απόλυτη % μεταβολή	Σταθμισμένη % μεταβολή
1	31783	35282	11.01	11.01	1.165
2	11804	13222	12.01	12.01	0.472
3	6931	7472	7.80	7.80	0.180
4	15869	19294	21.59	21.59	1.141
5	10486	11640	11.00	11.00	0.384
6	12209	12954	6.10	6.10	0.248
7	6826	6786	-0.59	0.59	0.013
8	18783	19173	2.08	2.08	0.130
9	56882	57266	0.68	0.68	0.128
10	20297	20450	0.75	0.75	0.051
11	12943	12600	-2.65	2.65	0.114
12	10108	9885	-2.21	2.21	0.074
13	11763	13817	17.46	17.46	0.684
14	7540	8192	8.65	8.65	0.217
15	8372	8379	0.08	0.08	0.002
16	6202	5688	-8.28	8.28	0.171
17	3169	3355	5.87	5.87	0.062
18	8145	8454	3.79	3.79	0.103
19	10485	11505	9.73	9.73	0.340
20	5161	5883	13.99	13.99	0.241
21	9583	8910	-7.03	7.03	0.224
22	4373	4441	1.56	1.56	0.023
23	10505	12049	14.70	14.70	0.514
<b>άθροισμα</b>	300219		<b>Μέσος Όρος</b>	7.37	6.683

Πίνακας 5.15

### Αθροίσματα σειρών (Προορισμός)

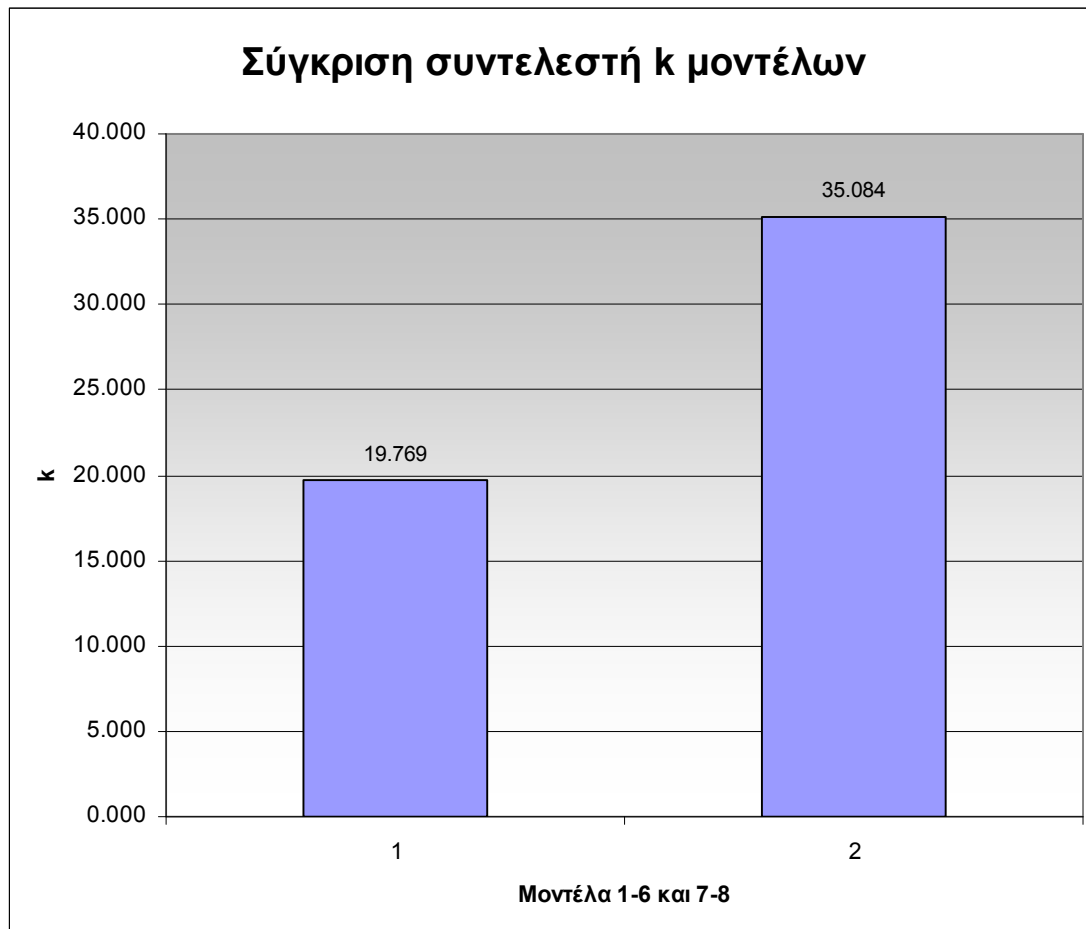
Στάσεις	Δεδομένα	Αποτελέσματα	% μεταβολή	απόλυτη % μεταβολή	Σταθμισμένη % μεταβολή
1	24603	25698	4.45	4.45	0.365
2	9031	10900	20.69	20.69	0.622
3	5644	5462	-3.22	3.22	0.061
4	12908	13638	5.66	5.66	0.243
5	9218	10558	14.53	14.53	0.446
6	10236	10726	4.79	4.79	0.163
7	6346	6556	3.31	3.31	0.070
8	14922	14625	-1.99	1.99	0.099
9	64575	72995	13.04	13.04	2.805
10	23479	24933	6.19	6.19	0.484
11	13545	13217	-2.42	2.42	0.109
12	10768	10533	-2.18	2.18	0.078
13	12653	13508	6.76	6.76	0.285
14	8586	9279	8.07	8.07	0.231
15	8892	8641	-2.82	2.82	0.084
16	6304	6892	9.33	9.33	0.196
17	3600	3563	-1.03	1.03	0.012
18	9480	9149	-3.49	3.49	0.110
19	11318	10167	-10.17	10.17	0.383
20	4863	4399	-9.55	9.55	0.155
21	13422	15653	16.62	16.62	0.743
22	4720	4974	5.39	5.39	0.085
23	11106	10630	-4.29	4.29	0.159
<b>άθροισμα</b>	300219		<b>Μέσος Όρος</b>	6.96	7.988

Πίνακας 5.15

Οι αποκλίσεις αυτές βλέπουμε ότι είναι της τάξης του 7% και προφανώς είναι μια επιπλέον πηγή σφάλματος που επηρεάζει την ακρίβεια των αποτελεσμάτων. Πρέπει να πούμε ότι αυτή η απόκλιση οφείλεται στον αλγόριθμο επίλυσης που χρησιμοποιεί το Matlab, αφού αυτό το πρόβλημα παρουσιαζόταν και όταν δοκιμάσαμε και άλλους τυχαίους αρχικούς πίνακες, αλλά και στο μέγεθος του προβλήματος αφού για τα μοντέλα 7 και 8 που χρησιμοποιούν τριγωνικό πίνακα (το πρόβλημα γίνεται πολύ μικρότερο) οι περιορισμοί τηρούνται.

Στην συνέχεια παρουσιάζεται ένα συγκριτικό διάγραμμα του συντελεστή  $k$  (σταθμισμένη ποσοστιαία μεταβολή αποτελεσμάτων από τα δεδομένα) ανάμεσα στις εξισώσεις που εξετάσαμε.

Το διάγραμμα που προκύπτει είναι το ακόλουθο (για τα μοντέλα 7 και 8 είναι το συνολικό των δύο τριγωνικών πινάκων που δίνουν):



**Γράφημα 5.6**

Στην συνέχεια θα δείξουμε κάποια στατιστικά μεγέθη και γραφήματα για τα αποτελέσματα.

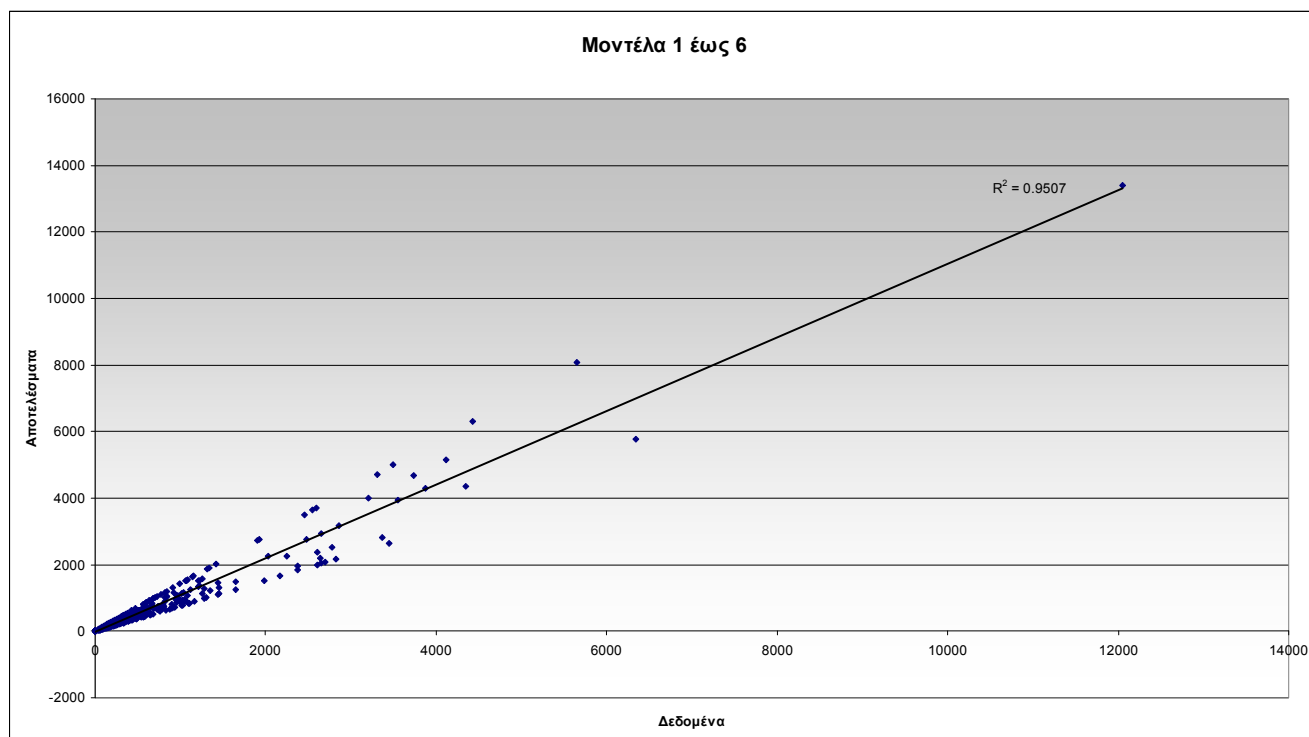
### Μοντέλα 1 έως 6:

Στατιστικά παλινδρόμησης	
Πολυπλό R	0.975
R Τετράγωνο	0.951
Προσαρμοσμένο R Τετράγωνο	0.951
Τυπικό σφάλμα	233.564
Μέγεθος δείγματος	529.000

ΑΝΑΛΥΣΗ ΔΙΑΚΥΜΑΝΣΗΣ					
	βαθμοί ελευθερίας	SS	MS	F	Σημαντικότητα F
Παλινδρόμηση	1	554398383	554398383	10163	0
Υπόλοιπο	527	28748935	54552		
Σύνολο	528	583147318			

	Συντελεστές	Τυπικό σφάλμα	t	τιμή-P	Κατώτερο 95%	Υψηλότερο 95%	Κατώτερο 95.0%	Υψηλότερο 95.0%
Τεταγμένη επί την αρχή	-29.969	11.917	-2.515	0.012	-53.380	-6.559	-53.380	-6.559
Μεταβλητή X 1	1.108	0.011	100.810	0.000	1.086	1.129	1.086	1.129

**Πίνακας 5.16**



**Γράφημα 5.7**

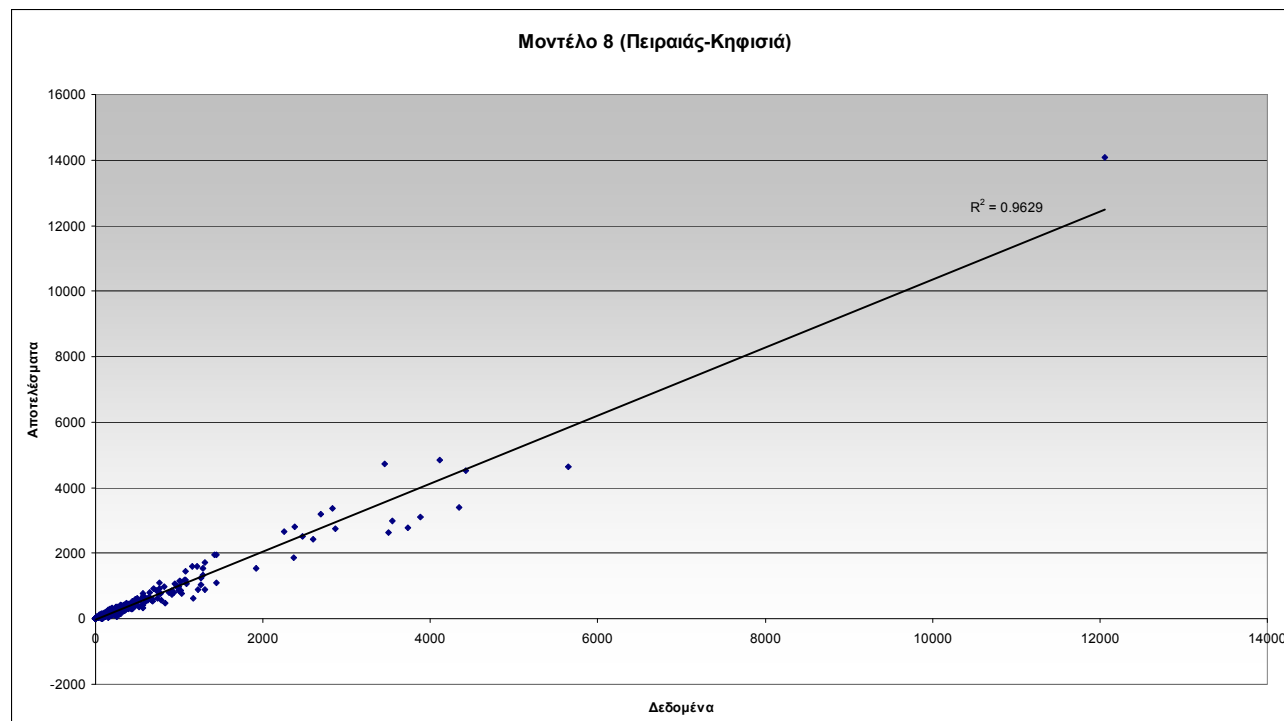
### Μοντέλα 7 και 8 (Πειραιάς-Κηφισιά):

Στατιστικά παλινδρόμησης	
Πολλαπλό R	0.981
R Τετράγωνο	0.963
Προσαρμοσμένο R Τετράγωνο	0.963
Τυπικό σφάλμα	169.995
Μέγεθος δείγματος	529.000

ΑΝΑΛΥΣΗ ΔΙΑΚΥΜΑΝΣΗΣ					
	βαθμοί ελευθερίας	SS	MS	F	Σημαντικότητα F
Παλινδρόμηση	1	394834771	394834771	13663	0
Υπόλοιπο	527	15229405	28898		
Σύνολο	528	410064176			

	Συντελεστές	Τυπικό σφάλμα	t	τιμή-P	Κατώτερο 95%	Υψηλότερο 95%	Κατώτερο 95.0%	Υψηλότερο 95.0%
Τεταγμένη επί την αρχή	-11.532	7.895	-1.461	0.145	-27.041	3.977	-27.041	3.977
Μεταβλητή X 1	1.037	0.009	116.888	0.000	1.019	1.054	1.019	1.054

**Πίνακας 5.16**



**Γράφημα 5.8**

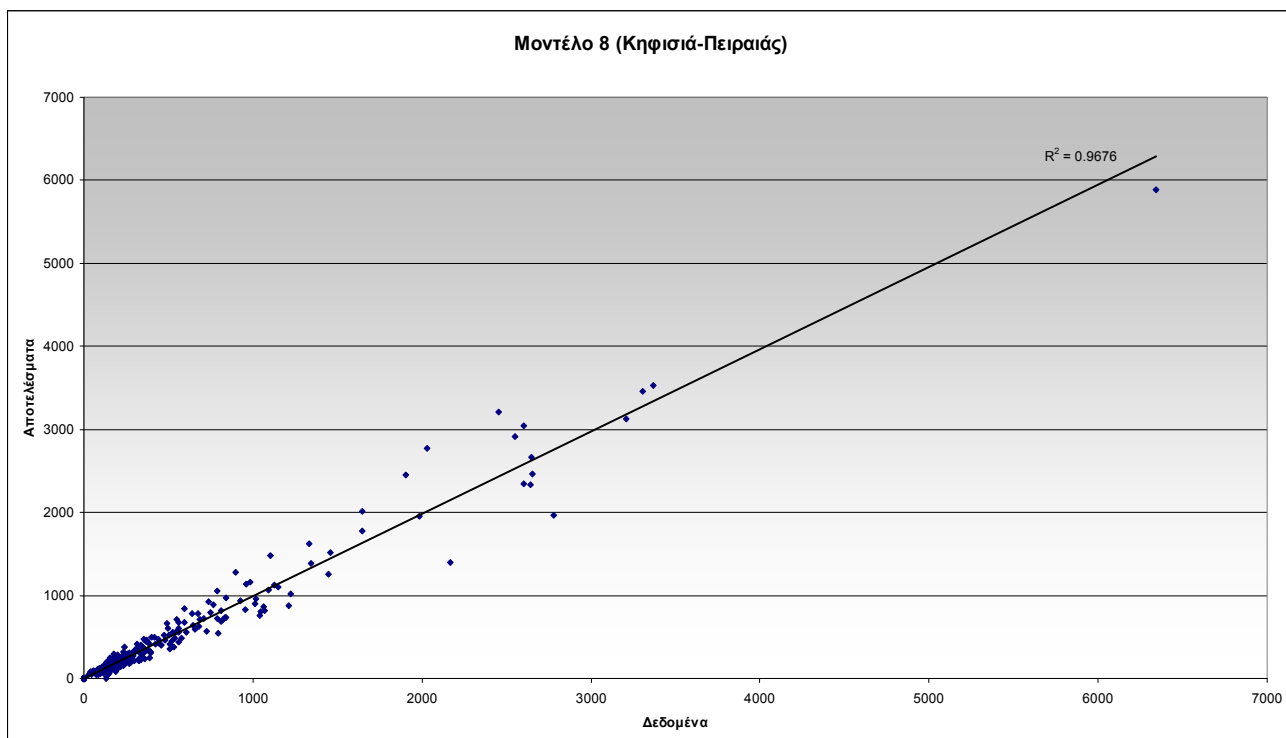
Μοντέλα 7 και 8 (Κηφισιά-Πειραιάς):

Στατιστικά παλινδρόμησης	
Πολλαπλό R	0.984
R Τετράγωνο	0.968
Προσαρμοσμένο R Τετράγωνο	0.968
Τυπικό σφάλμα	102.643
Μέγεθος δείγματος	529.000

ΑΝΑΛΥΣΗ ΔΙΑΚΥΜΑΝΣΗΣ					
	βαθμοί ελευθερίας	SS	MS	F	Σημαντικότητα F
Παλινδρόμηση	1	165843096	165843096	15741	0
Υπόλοιπο	527	5552275	10536		
Σύνολο	528	171395372			

	Συντελεστές	Τυπικό σφάλμα	t	τιμή-P	Κατώτερο 95%	Υψηλότερο 95%	Κατώτερο 95.0%	Υψηλότερο 95.0%
Τεταγμένη επί την αρχή	2.293	4.895	0.468	0.640	-7.324	11.909	-7.324	11.909
Μεταβλητή X 1	0.991	0.008	125.464	0.000	0.975	1.007	0.975	1.007

**Πίνακας 5.17**



**Γράφημα 5.9**



#### 5.4.4. Ακρίβεια των μοντέλων σε σχέση με την κλίμακα του προβλήματος

Στο προηγούμενο κεφάλαιο εφαρμόσαμε τα μοντέλα μεγιστοποίησης της εντροπίας σε ένα πραγματικό πρόβλημα με σκοπό να προσδιορίσουμε την ακρίβεια που προσφέρουν. Σε αυτό το κεφάλαιο θέλουμε να διερευνήσουμε εάν η ακρίβεια που προσφέρουν τα μοντέλα αυτά έχει κάποια συσχέτιση με το μέγεθος του προβλήματος που επιλύουν την κάθε φορά. Δηλαδή θα εξετάσουμε την ακρίβεια που προσφέρουν για διάφορα μεγέθη πινάκων Π-Π.

Από τα αποτελέσματα του προηγούμενου κεφαλαίου είδαμε ότι την καλύτερη ακρίβεια την δίνουν τα μοντέλα 1 έως 6. Εμείς επιλέξαμε το μοντέλο 6, που έχει την πιο απλή μορφή, και προσδιορίσαμε διαδοχικά πίνακες Π-Π μεγέθους  $6 \times 6$ ,  $10 \times 10$ ,  $14 \times 14$ ,  $19 \times 19$  και  $23 \times 23$ . Ο αρχικός πίνακας που χρησιμοποιήσαμε είναι κοινός για όλες τις περιπτώσεις, παίρνοντας τυχαία στοιχεία για κάθε περίπτωση και τους αντίστοιχους περιορισμούς. Παρακάτω παρουσιάζονται μερικά στατιστικά μεγέθη και γραφήματα για τις περιπτώσεις αυτές.

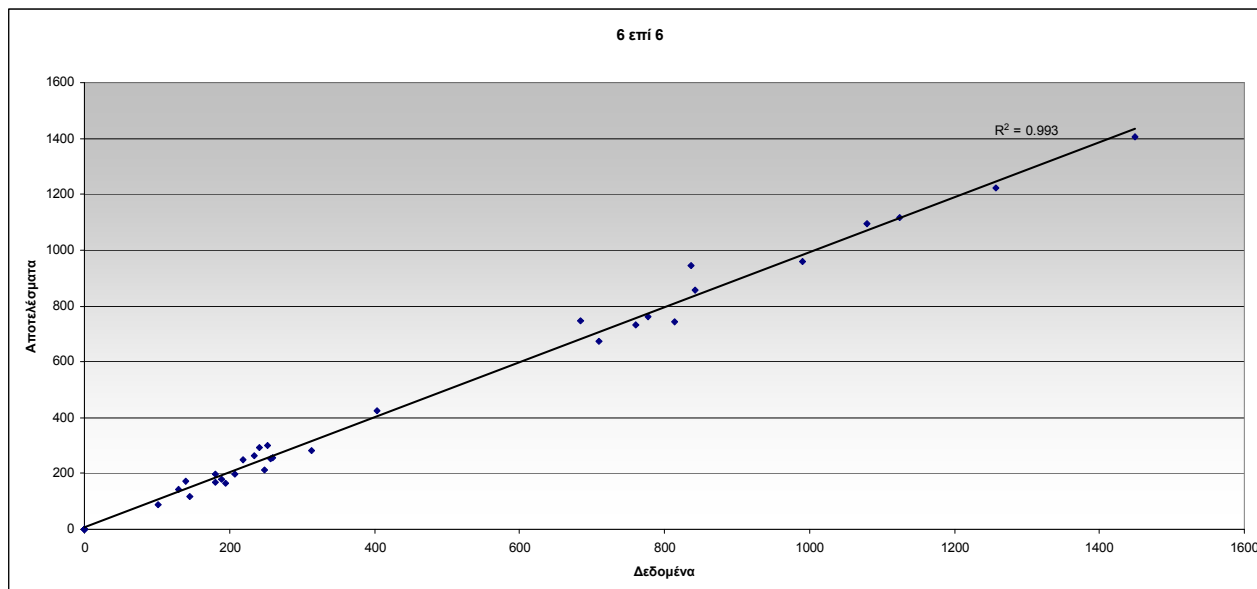
6 επί 6:

Στατιστικά παλινδρόμησης	
Πολλαπλό R	0.996
R Τετράγωνο	0.993
Προσαρμοσμένο R Τετράγωνο	0.993
Τυπικό σφάλμα	34.236
Μέγεθος δείγματος	36.000

ΑΝΑΛΥΣΗ ΔΙΑΚΥΜΑΝΣΗΣ					
	βαθμοί ελευθερίας	SS	MS	F	Σημαντικότητα F
Παλινδρόμηση	1	5651743	5651743	4822	0
Υπόλοιπο	34	39851	1172		
Σύνολο	35	5691595			

	Συντελεστές	Τυπικό σφάλμα	t	πιμή- P	Κατώτερο 95%	Υψηλότερο 95%	Κατώτερο 95.0%	Υψηλότερο 95.0%
Τεταγμένη επί την αρχή	5.763	8.283	0.696	0.491	-11.071	22.597	-11.071	22.597
Μεταβλητή X 1	0.986	0.014	69.440	0.000	0.958	1.015	0.958	1.015

**Πίνακας 5.18**



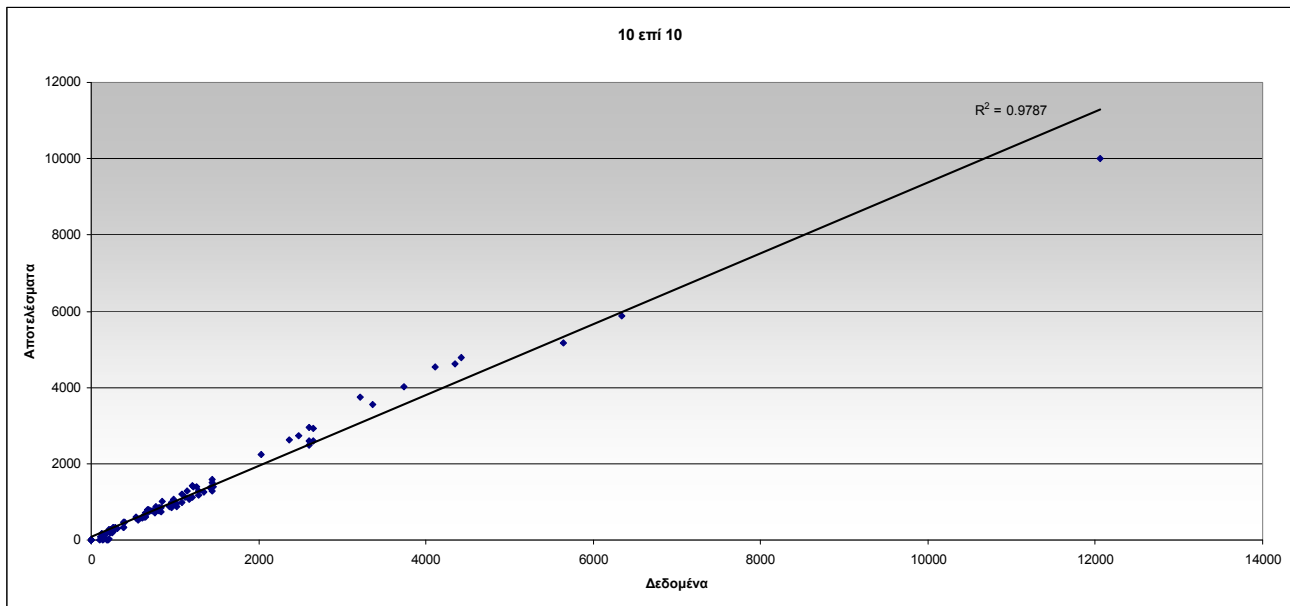
10 επί 10:

<i>Στατιστικά παλινδρόμησης</i>	
Πολλαπλό R	0.989
R Τετράγωνο	0.979
Προσαρμοσμένο R Τετράγωνο	0.979
Τυπικό σφάλμα	226.583
Μέγεθος δείγματος	100.000

<b>ΑΝΑΛΥΣΗ ΔΙΑΚΥΜΑΝΣΗΣ</b>					
	<i>βαθμοί ελευθερίας</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Σημαντικότητα F</i>
Παλινδρόμηση	1	231681328	231681328	4513	0
Υπόλοιπο	98	5031302	51340		
Σύνολο	99	236712630			

	<i>Συντελεστές</i>	<i>Τυπικό σφάλμα</i>	<i>t</i>	<i>πιμή- P</i>	<i>Κατώτερο 95%</i>	<i>Υψηλότερο 95%</i>	<i>Κατώτερο 95.0%</i>	<i>Υψηλότερο 95.0%</i>
Τεταγμένη επί την αρχή	80.576	27.842	2.894	0.005	25.325	135.827	25.325	135.827
Μεταβλητή X 1	0.929	0.014	67.177	0.000	0.901	0.956	0.901	0.956

**Πίνακας 5.19**



**Γράφημα 5.11**

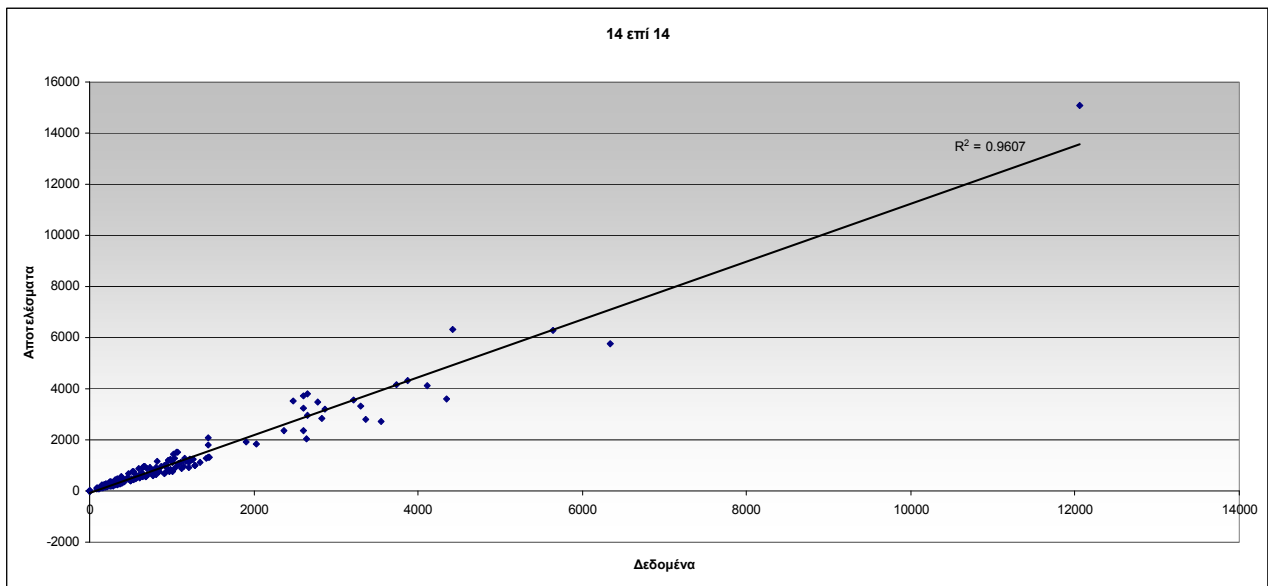
14 επί 14:

<i>Στατιστικά παλινδρόμησης</i>	
Πολλαπλό R	0.980
R Τετράγωνο Προσαρμοσμένο	0.961
R Τετράγωνο	0.960
Τυπικό σφάλμα Μέγεθος δείγματος	302.183 196.000

<b>ΑΝΑΛΥΣΗ ΔΙΑΚΥΜΑΝΣΗΣ</b>					
	<i>βαθμοί ελευθερίας</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Σημαντικότητα F</i>
Παλινδρόμηση	1	432827162	432827162	4740	0
Υπόλοιπο	194	17715020	91315		
Σύνολο	195	450542182			

	<i>Συντελεστές</i>	<i>Τυπικό σφάλμα</i>	<i>t</i>	<i>τιμή-P</i>	<i>Κατώτερο 95%</i>	<i>Υψηλότερο 95%</i>	<i>Κατώτερο 95.0%</i>	<i>Υψηλότερο 95.0%</i>
Τεταγμένη επί την αρχή	-65.154	26.337	-2.474	0.014	-117.097	-13.211	-117.097	-13.211
Μεταβλητή X 1	1.129	0.016	68.847	0.000	1.097	1.162	1.097	1.162

**Πίνακας 5.20**



Γράφημα 5.12

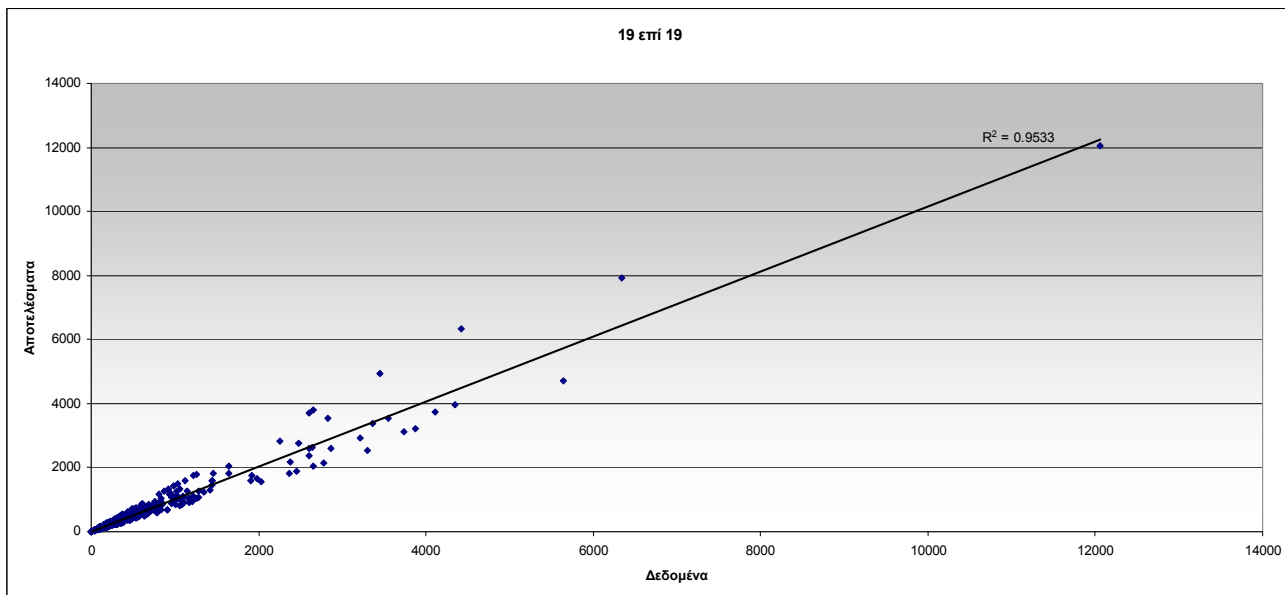
## 19 επί 19:

Στατιστικά παλινδρόμησης	
Πολλαπλό R	0.976
R Τετράγωνο Προσαρμοσμένο R Τετράγωνο	0.953
Τυπικό σφάλμα	238.984
Μέγεθος δείγματος	361.000

ΑΝΑΛΥΣΗ ΔΙΑΚΥΜΑΝΣΗΣ					
	βαθμοί ελευθερίας	SS	MS	F	Σημαντικότητα F
Παλινδρόμηση	1	418462699	418462699	7327	0
Υπόλοιπο	359	20503692	57113		
Σύνολο	360	438966392			

	Συντελεστές	Τυπικό σφάλμα	t	τιμή- P	Κατώτερο 95%	Υψηλότερο 95%	Κατώτερο 95.0%	Υψηλότερο 95.0%
Τεταγμένη επί την αρχή	7.875	14.847	0.530	0.596	-21.324	37.073	-21.324	37.073
Μεταβλητή X 1	1.016	0.012	85.597	0.000	0.993	1.039	0.993	1.039

Πίνακας 5.21



**Γράφημα 5.13**

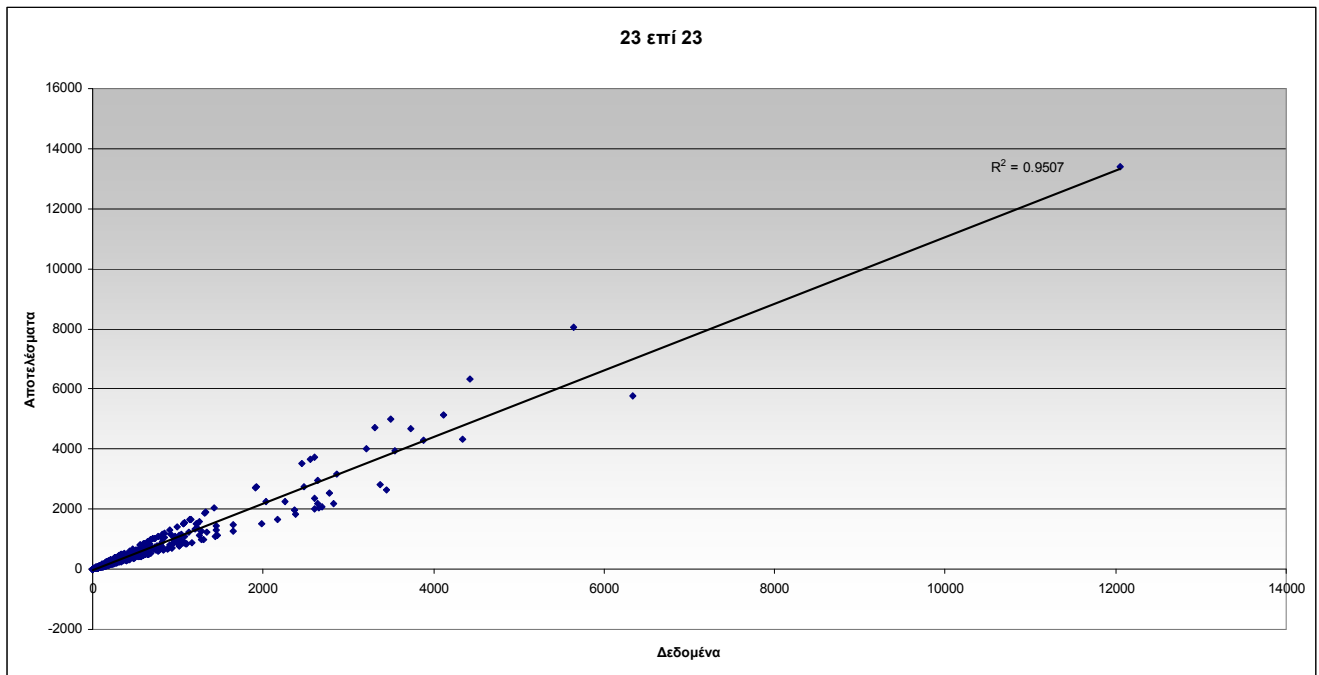
23 επί 23:

Στατιστικά παλινδρόμησης	
Πολλαπλό R	0.975
R Τετράγωνο	0.951
Προσαρμοσμένο R Τετράγωνο	0.951
Τυπικό σφάλμα	233.564
Μέγεθος δείγματος	529.000

ΑΝΑΛΥΣΗ ΔΙΑΚΥΜΑΝΣΗΣ					
	βαθμοί ελευθερίας	SS	MS	F	Σημαντικότητα F
Παλινδρόμηση	1	554398383	554398383	10163	0
Υπόλοιπο	527	28748935	54552		
Σύνολο	528	583147318			

	Συντελεστές	Τυπικό σφάλμα	t	τιμή-P	Κατώτερο 95%	Υψηλότερο 95%	Κατώτερο 95.0%	Υψηλότερο 95.0%
Τεταγμένη επί την αρχή	-29.969	11.917	-2.515	0.012	-53.380	-6.559	-53.380	-6.559
Μεταβλητή X 1	1.108	0.011	100.810	0.000	1.086	1.129	1.086	1.129

**Πίνακας 5.22**



Γράφημα 5.14

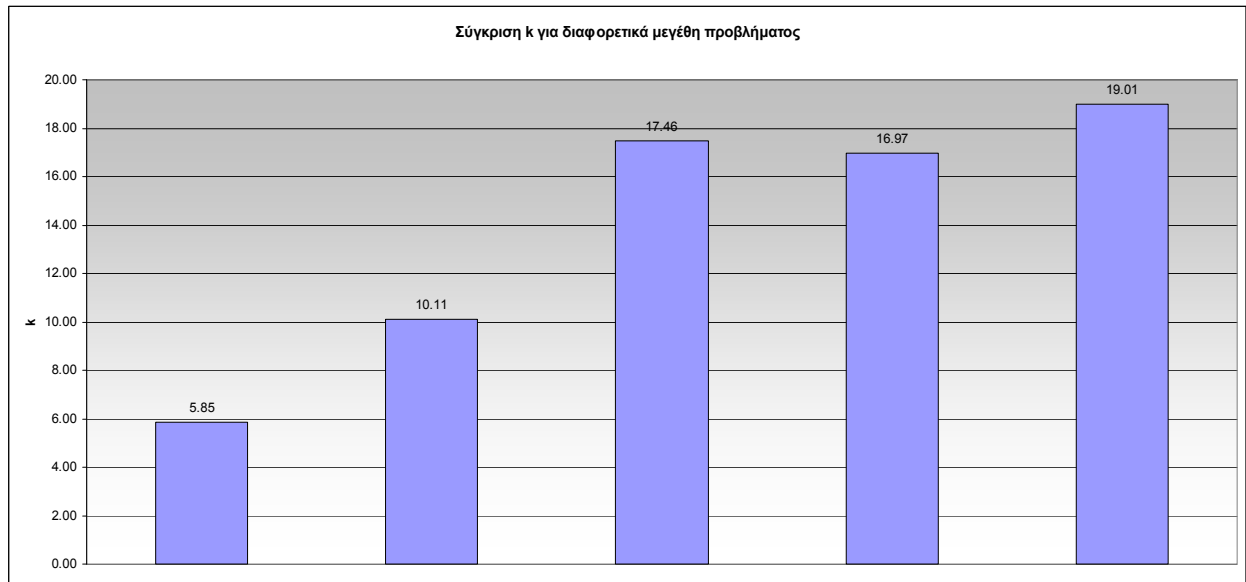
### 5.4.5. Συμπεράσματα

Έχοντας παρουσιάσει τα αποτελέσματα που προέκυψαν παραπάνω, μπορούμε να εξάγουμε κάποια χρήσιμα συμπεράσματα. Στο γράφημα 5.15 βλέπουμε την μεταβολή του συντελεστή  $k$  για τα διάφορα μεγέθη του προβλήματος. Είναι εμφανές ότι όσο αυξάνεται το μέγεθος τόσο η σταθμισμένη ποσοστιαία μεταβολή αυξάνεται. Επίσης στο γράφημα 5.16 βλέπουμε τη μεταβολή του συντελεστή συσχέτισης  $R^2$  για τα διάφορα μεγέθη του προβλήματος. Παρατηρούμε ότι όσο αυξάνεται το μέγεθος τόσο μειώνεται η τιμή του  $R^2$ . Στον πίνακα 5.23 φαίνεται το πλήθος των αγνώστων και των σχέσεων για κάθε μέγεθος του προβλήματος.

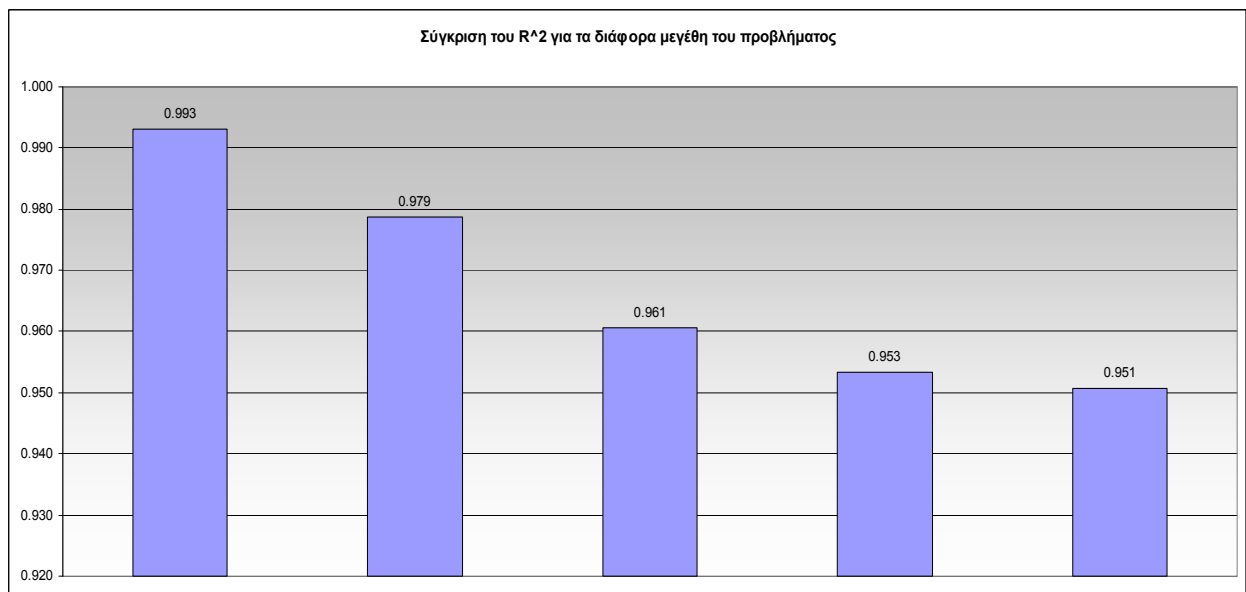
Μέγεθος	Αγνώστοι	Σχέσεις
6x6	30	12
10x10	90	20
14x14	182	28
19x19	342	38
23x23	506	46

Πίνακας 5.23

Το συμπέρασμα που προκύπτει είναι ότι η ακρίβεια που προσφέρουν τα μοντέλα μεγιστοποίησης της εντροπίας είναι αντιστρόφως ανάλογη του μεγέθους του προβλήματος.



Γράφημα 5.15

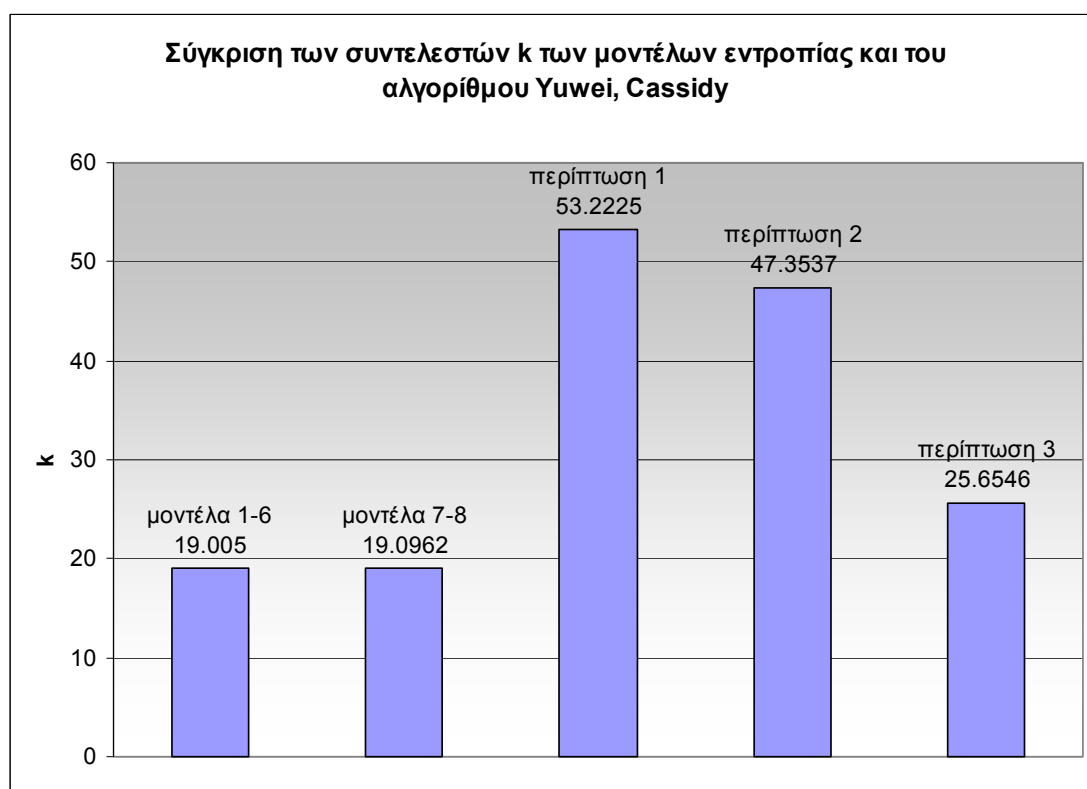


Γράφημα 5.16

## 5.5. Συμπεράσματα

Στα προηγούμενα κεφάλαια εφαρμόσαμε την μεθοδολογία που περιγράφει ο αλγόριθμος των Li, Cassidy (2006) αλλά και αυτή των μοντέλων μεγιστοποίησης της εντροπίας, για ένα πραγματικό πρόβλημα του οποίου γνωρίζουμε την λύση, ώστε να μπορούμε να προσδιορίσουμε την ακρίβεια του δίνουν.

Σε αυτό το κεφάλαιο θα συγκρίνουμε τα αποτελέσματα που προέκυψαν από τις δύο αυτές διαφορετικές προσεγγίσεις. Ακολουθεί γράφημα που συγκρίνει τον συντελεστή σταθμισμένης ποσοστιαίας μεταβολής  $k$ .



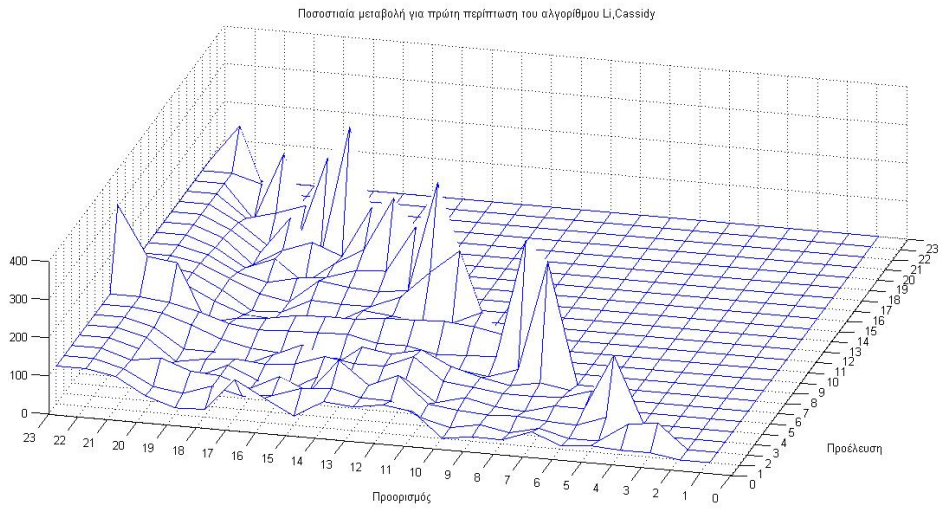
Γράφημα 5.17

Συμπεραίνουμε ότι η ακρίβεια των αποτελεσμάτων των μοντέλων μεγιστοποίησης της εντροπίας είναι σαφώς μεγαλύτερη από αυτή που δίνει ο αλγόριθμος των Li, Cassidy.

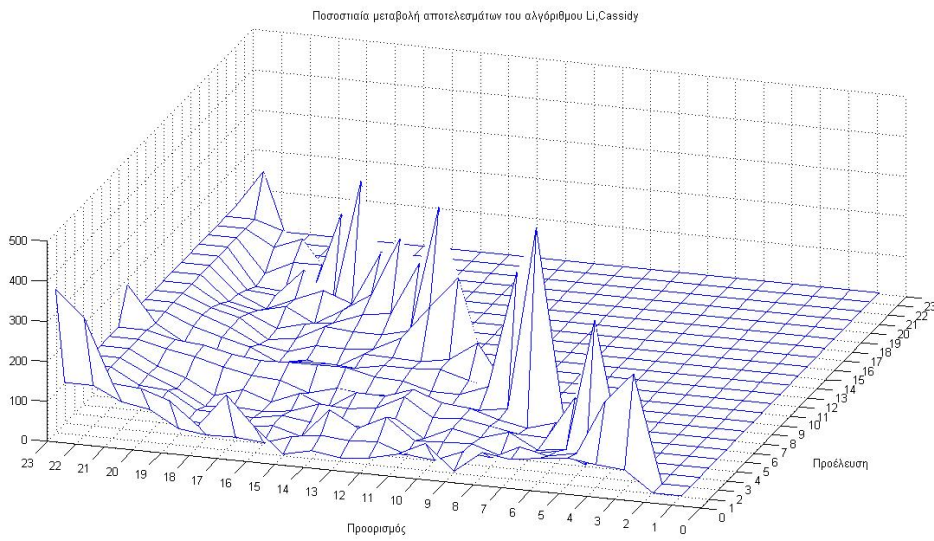
Στη συνέχεια θα παρουσιάσουμε τρισδιάστατα γραφήματα των ποσοστιαίων μεταβολών των αποτελεσμάτων σε σχέση με τα δεδομένα για όλες τις περιπτώσεις



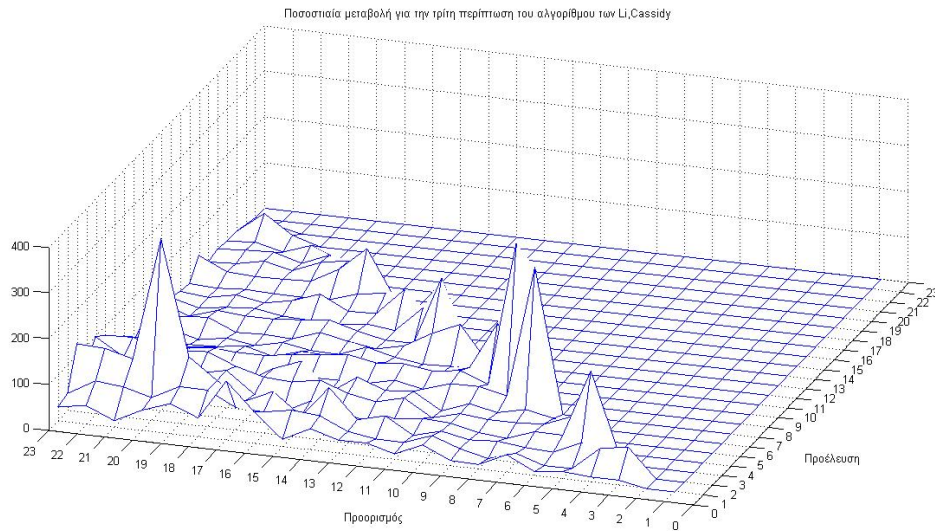
που εξετάσαμε, με σκοπό να εντοπίσουμε τυχόν συστηματικά σφάλματα ή γενικά κάποια συστηματική συμπεριφορά των αποτελεσμάτων.



**Γράφημα 5.18**



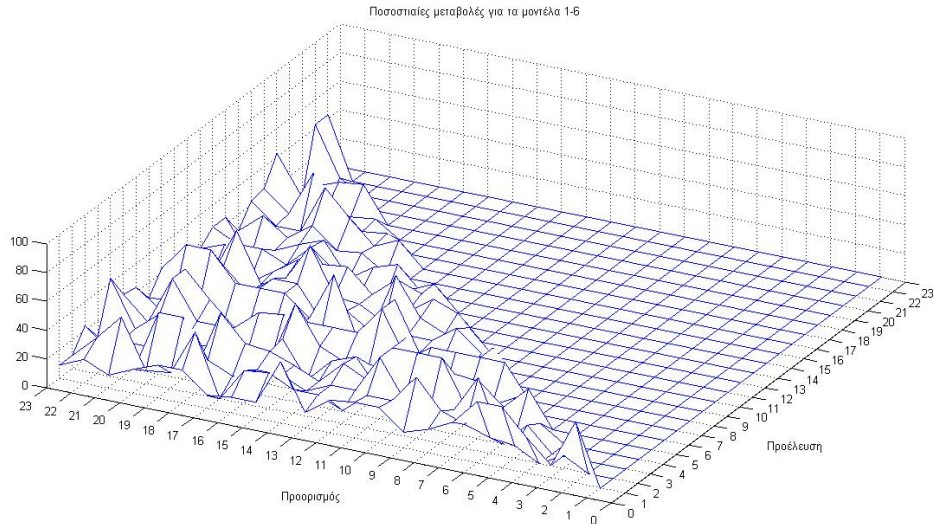
**Γράφημα 5.19**



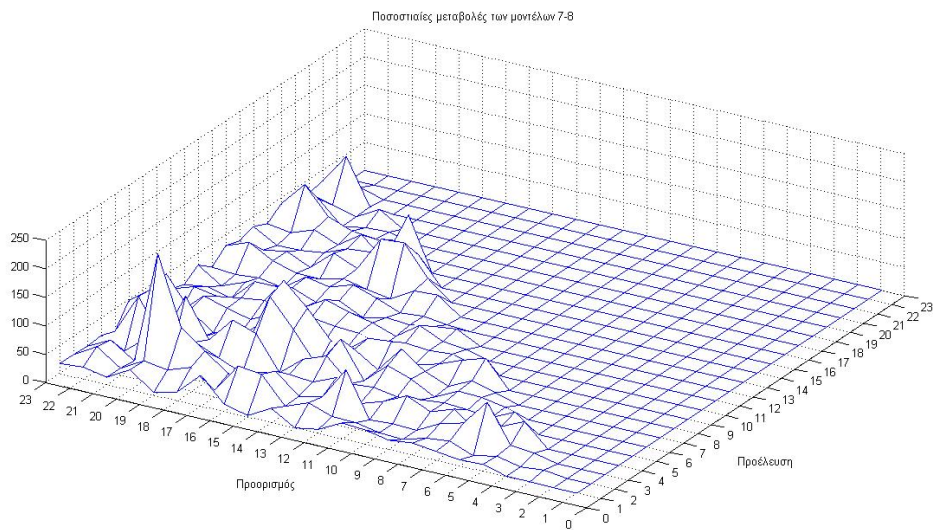
**Γράφημα 5.20**

Αυτό που παρατηρούμε για τον αλγόριθμο των Li, Cassidy είναι ένα συστηματικό σφάλμα που παρουσιάζεται κυρίως στα πρώτα κελιά της κάθε γραμμής του πίνακα. Η ποσοστιαία μεταβολή είναι πολύ μεγάλη σε σχέση με αυτή των κελιών του υπόλοιπου πίνακα και παρατηρείται και στις τρεις περιπτώσεις. Επίσης για τις δύο πρώτες περιπτώσεις παρατηρούμε ότι οι δύο τελευταίες στήλες έχουν μια ομοιόμορφη ποσοστιαία μεταβολή. Αυτό οφείλεται στο γεγονός ότι οι τιμές των κελιών τους είναι μηδέν. Ο μηδενισμός των κελιών παρατηρείται σε όλες τις περιπτώσεις που έχουμε ορίσει σημαντικές στάσεις. Στην τρίτη περίπτωση όπου δεν λαμβάνουμε υπόψη μας τις βασικές παραδοχές (ύπαρξη σημαντικών στάσεων) του αλγορίθμου δεν μηδενίζεται η τιμή κανενός κελιού. Αυτό σημαίνει ότι υπάρχει κάποια δυσλειτουργία στον αλγόριθμο που προκαλεί αυτό το συστηματικό σφάλμα στα αποτελέσματα.

Στο επόμενο γράφημα θα δούμε τις ποσοστιαίες μεταβολές για τα μοντέλα της εντροπίας.



**Γράφημα 5.21**



**Γράφημα 5.22**

Αυτό που παρατηρούμε για τα μοντέλα μεγιστοποίησης της εντροπίας είναι ότι δεν υπάρχει κάποια συστηματική ποσοστιαία μεταβολή. Επίσης βλέπουμε πως οι ποσοστιαίες μεταβολές για τα μοντέλα αυτά, είναι κατά πού μικρότερες από αυτές του αλγορίθμου των Li, Cassidy κάτι που επιβεβαιώνει και το γεγονός ότι τα μοντέλα μεγιστοποίησης της εντροπίας προσφέρουν πολύ καλύτερη ακρίβεια.

## 6. Επαναληπτικός αλγόριθμος μεγιστοποίησης της εντροπίας

### 6.1. Εισαγωγή

Στο προηγούμενο κεφάλαιο αξιολογήσαμε την ακρίβεια που δίνουν τα μοντέλα μεγιστοποίησης της εντροπίας και ο γενικευμένος αλγόριθμος εκτίμησης πινάκων Π-Π του Li, Cassidy. Σε αυτό το κεφάλαιο αναπτύσσεται ένας νέος αλγόριθμος προσέγγισης του προβλήματος με απώτερο σκοπό την βελτίωση της ακρίβειας του εκτιμώμενου πίνακα Π-Π.

Ως εφελκτήριο για την ανάπτυξη του νέου αλγορίθμου ήταν η συμπεριφορά που παρουσίασαν τα μοντέλα για διάφορα μεγέθη του πίνακα των μετακινήσεων. Γνωρίζουμε ότι τα μοντέλα αυτά υπολογίζουν τον πίνακα στο σύνολο του, δηλαδή τις τιμές όλων των κελιών του πίνακα Π-Π. Επίσης είδαμε ότι τα μοντέλα αυτά όσο μεγαλύτερο είναι το πρόβλημα που πρέπει να επιλύσουν, τόσο δυσκολότερο είναι να ικανοποιήσουν πλήρως τους περιορισμούς. Παρόλα αυτά όμως τα προβλήματα που εντοπίσαμε, παρατηρήσαμε ότι όσο μειώνεται η κλίμακα του προβλήματος, τόσο καλύτερη είναι η ακρίβεια δίνουν τα μοντέλα. Αυτό μας έκανε να σκεφτούμε πως μπορούμε να εκμεταλλευτούμε αυτή τους την ιδιότητα ώστε να επιτύχουμε καλύτερη ακρίβεια από αυτή που τώρα προσφέρουν.

### 6.2. Ο αλγόριθμος

Αυτό που σκεφτήκαμε λοιπόν είναι, αντί να αντιμετωπίζουμε το πρόβλημα ενιαία, να το χωρίσουμε σε επιμέρους υποπροβλήματα και να τα επιλύουμε με την σειρά κατά την φορά που έχει το δίκτυο. Με αυτό τον τρόπο προκύπτουν περισσότερα προβλήματα αλλά μικρά σε μέγεθος. Ένα μειονέκτημα αυτής της μεθόδου είναι ότι το σφάλμα του πρώτου υποπροβλήματος που επιλύεται, θα μεταφέρεται σε όλα τα επόμενα. Αλλά όπως δείξαμε όσο μικρότερη η κλίμακα του προβλήματος, τόσο μικρότερο το σφάλμα.

Η επίλυση του προβλήματος επιτυγχάνεται, εφαρμόζοντας μια επαναληπτική διαδικασία όπου σε κάθε επανάληψη περιλαμβάνεται η εφαρμογή ενός οποιουδήποτε μοντέλου μεγιστοποίησης της εντροπίας. Δηλαδή σε κάθε επανάληψη γίνεται βελτιστοποίηση του κάθε κόμβου. Εμείς για τη επίλυση του προβλήματος χρησιμοποιήσαμε το μοντέλο (3.6.1) γιατί έχει την πιο απλή μορφή.

Ο αλγόριθμος αυτός μπορεί να εφαρμοστεί σε γραμμικά δίκτυα μονής κατεύθυνσης με σκοπό τον προσδιορισμό του πίνακα Π-Π.

Ως δεδομένα χρειαζόμαστε ένα αρχικό πίνακα και τον αριθμό αυτών που εισέρχονται ( $O_i$ ) και εξέρχονται ( $D_i$ ) από το δίκτυο σε κάθε κόμβο.

Γενικά, η μορφή του πίνακα που θέλουμε να προσδιορίσουμε φαίνεται στον παρακάτω πίνακα.

**ΑΓΝΩΣΤΟΙ**

	1	2	3	.....	.....	κ	.....	v	sum
1	0	$T_{1,2}$	$T_{1,3}$			$T_{1,\kappa}$		$T_{1,v}$	$O_1$
2	0	0	$T_{2,3}$			$T_{2,\kappa}$		$T_{2,v}$	$O_2$
3	0	0	0						$O_3$
	0	0	0						
	0	0	0						
	0	0	0						
	0	0	0						
λ	0	0	0			$T_{\lambda,\kappa}$		$T_{\lambda,v}$	$O_\lambda$
	0	0	0						
v	0	0	0	0	0	0	0	0	$O_v$
sum	0	$D_2$	$D_3$			$D_\kappa$		$D_v$	

**ΠΕΡΙΟΡΙΣΜΟΙ**

**Πίνακας 6.1**

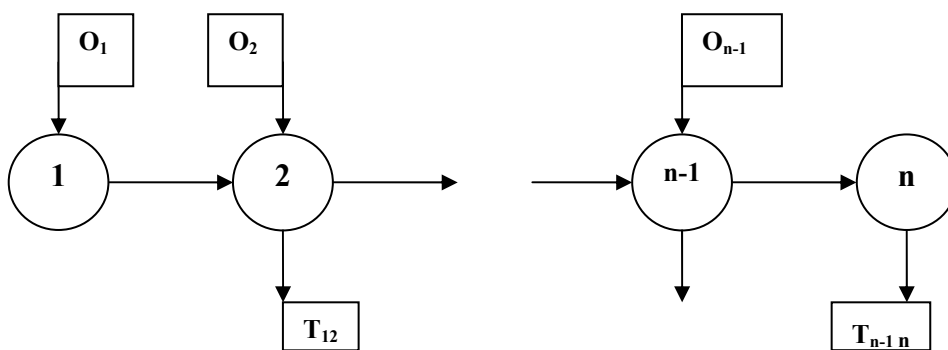
Δηλαδή σε κάθε επανάληψη θα υπολογίζουμε τους αγνώστους για τον κάθε κόμβο υπό τους αντίστοιχους περιορισμούς.

Επανάληψη	Άγνωστοι					
Αρχικά	$T_{1,2}$					
1	$T_{1,3}$	$T_{2,3}$				
2	$T_{1,4}$	$T_{2,4}$	$T_{3,4}$			
3	$T_{1,5}$	$T_{2,5}$	$T_{3,5}$	$T_{4,5}$		
.....	.....	.....	.....	.....	.....	
.....	.....	.....	.....	.....	.....	
κ	$T_{1,v}$	$T_{2,v}$	$T_{3,v}$	$T_{4,v}$	.....	$T_{v-1,v}$

Πίνακας 6.2

Με τον επαναληπτικό αλγόριθμο προσδιορίζουμε για κάθε κόμβο, από πού προέρχονται αυτοί που αποβιβάζονται στον κόμβο αυτό και πόσοι συνολικά συνεχίζουν το ταξίδι τους. Δηλαδή το μέγεθος του προβλήματος θα είναι κάθε φορά ένας πίνακας  $(n-1) \times 2$  όπου το  $n$  θα εξαρτάται από την σειρά της στάσης στο δίκτυο. Η πρώτη στήλη του πίνακα θα δίνει αυτούς που αποβιβάζονται και η δεύτερη στήλη θα δίνει αυτούς που συνεχίζουν. Η πρώτη στήλη θα είναι η λύση για τον κάθε κόμβο ενώ η δεύτερη στήλη χρησιμοποιείται στον έλεγχο των περιορισμών.

Λόγο της γραμμικότητας του δικτύου μπορούμε να υπολογίσουμε κάποιους αγνώστους πριν αρχίσουμε την επίλυση. Όπως φαίνεται και στο παρακάτω σχήμα, είναι προφανές ότι όσοι εξέρχονται από το δίκτυο στον δεύτερο κόμβο θα έχουν εισέλθει στον πρώτο και όσοι εισέρχονται στο δίκτυο από τον προτελευταίο κόμβο είναι σίγουρο ότι θα εξέλθουν στον τελευταίο.



Σχήμα 6.1

Οπότε ξεκινάμε γνωρίζοντας

	1	2
1	0	$T_{12}$
2	0	0

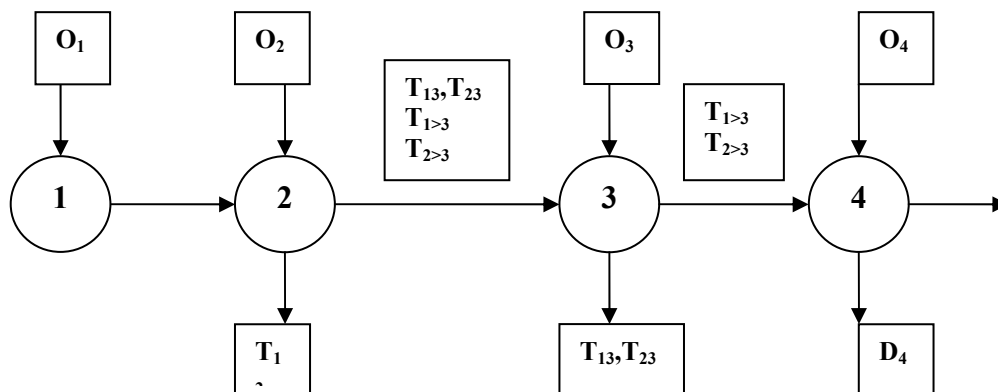
Πίνακας 6.3

Συνεχίζουμε με τον κόμβο 3, όπου θέλουμε να προσδιορίσουμε το από ποιους κόμβους προέρχονται αυτοί που αποβιβάζονται. Δηλαδή, έχοντας υπολογίσει την τιμή του  $T_{12}$ , στο βήμα αυτό θα εφαρμόσουμε τον αλγόριθμο μεγιστοποίησης της εντροπίας για να υπολογίσουμε τις τιμές των  $T_{13}, T_{23}$ . Θέλουμε δηλαδή να προσδιορίσουμε τα πράσινα κελία.

	1	2	3
1	0	$T_{12}$	$T_{13}$
2	0	0	$T_{23}$
3	0	0	0

Πίνακας 6.4

Γενικά για τον κάθε κόμβο γνωρίζουμε πόσοι ακριβώς ταξιδιώτες υπάρχουν στο δίκτυο ακριβώς πριν από αυτόν, πόσοι εισέρχονται, πόσοι εξέρχονται αλλά και πόσοι συνεχίζουν από αυτόν.



Σχήμα 6.2

Συνεπώς από το σχήμα βλέπουμε ότι για τον κόμβο 3 ισχύει:

$$T_{13} + T_{1>3} = O_1 - (T_{12})$$

$$T_{23} + T_{2>3} = O_2$$

$$T_{13} + T_{23} = D_3$$

Όπου  $T_{i>j}$  είναι ο αριθμός των μετακινήσεων που έχουν ξεκινήσει το ταξίδι τους από τον κόμβο  $i$  και έχουν τελικό προορισμό κατάντη του κόμβου  $j$ .

Παρατηρούμε ότι έχουμε 3 σχέσεις για να προσδιορίσουμε 4 αγνώστους. Συνεπώς θα βελτιστοποιήσουμε το σύστημα αυτό για να βρούμε την πιθανότερη λύση του. Ο πίνακας που θα προσδιορίσουμε θα έχει την μορφή

$T_{13}$	$T_{1>3}$
$T_{23}$	$T_{2>3}$

**Πίνακας 6.5**

και γενικά θα ισχύει

$$T_{i>j} = \sum_{k=j+1}^n T_{i,k}$$

Τώρα θα προσδιορίσουμε και τον αρχικό πίνακα που θα χρησιμοποιήσουμε για την λύση του προβλήματος. Υποχρεωτικά ο αρχικός πίνακας θα έχει ίδια μορφή με τον πίνακα 6.5. Άρα θα είναι της μορφής

$t_{13}$	$t_{1>3}$
$t_{23}$	$t_{2>3}$

**Πίνακας 6.6**

όπου γενικά θα ισχύει και για αυτόν

$$t_{i>j} = \sum_{k=j+1}^n t_{i,k}$$



όπου στην αριστερή στήλη θα έχουμε τα αντίστοιχα στοιχεία του αρχικού πίνακα για τον κόμβο που εξετάζουμε κάθε φορά και στην δεξιά στήλη θα είναι το άθροισμα των αντίστοιχων μετακινήσεων που συνεχίζουν κατάντη του κόμβου. Δηλαδή το  $t_{1>3}$  θα ισούται με το άθροισμα των έγχρωμων κελιών της πρώτης γραμμής του αρχικού πίνακα και το  $t_{2>3}$  θα ισούται με το άθροισμα των έγχρωμων κελιών της δεύτερης γραμμής του αρχικού πίνακα.

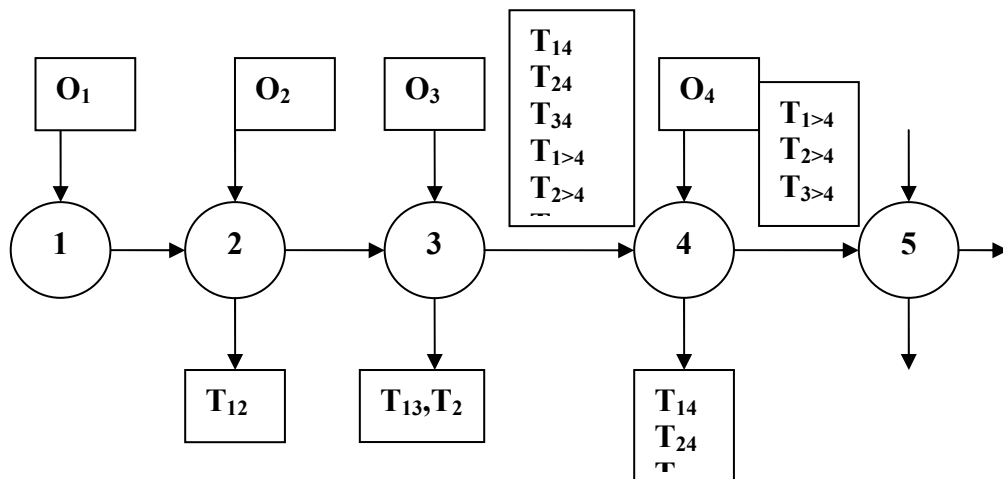
	1	2	3	4	.....	n-1	n
1	0	$t_{12}$	$t_{13}$	$t_{14}$	.....	$t_{1,n-1}$	$t_{1n}$
2	0	0	$t_{23}$	$t_{24}$	.....	$t_{2,n-1}$	$t_{2n}$
3	0	0	0	$t_{34}$			
4	0	0	0	0			
5	0	0	0	0			
6	0	0	0	0		0	
7	0	0	0	0		0	0

**Πίνακας 6.7**

Η διαδικασία τώρα συνεχίζεται στον κόμβο 4. Έχοντας υπολογίσει τις τιμές των  $T_{12}, T_{13}, T_{23}$ , στο βήμα αυτό εφαρμόζουμε τον αλγόριθμο της μέγιστης εντροπίας για να υπολογίσουμε τις τιμές των  $T_{14}, T_{24}, T_{34}$ .

	1	2	3	4
1	0	$T_{12}$	$T_{13}$	$T_{14}$
2	0	0	$T_{23}$	$T_{24}$
3	0	0	0	$T_{34}$
4	0	0	0	0

**Πίνακας 6.8**



**Σχήμα 6.3**

Συνεπώς από το σχήμα βλέπουμε πως στον κόμβο 4 ισχύει ότι:

$$T_{14} + T_{1>4} = O_1 - (T_{12} + T_{13})$$

$$T_{24} + T_{2>4} = O_2 - (T_{23})$$

$$T_{34} + T_{3>4} = O_3$$

$$T_{14} + T_{24} + T_{34} = D_4$$

Αυτό που παρατηρούμε εδώ, είναι ότι σε αυτές τις σχέσεις, που αποτελούν και τους περιορισμούς μας υπάρχουν στοιχεία που τα θεωρούμε γνωστά αλλά προέρχονται από τα αποτελέσματα που έδωσε η επίλυση του ανάντη κόμβου. Οπότε ότι σφάλμα περιέχεται στα αποτελέσματα αυτά θα μεταφερθεί και στα επόμενα. Ο πίνακας που θέλουμε να προσδιορίσουμε τώρα και θα ισχύουν οι παραπάνω περιορισμοί είναι

$T_{14}$	$T_{1>4}$
$T_{24}$	$T_{2>4}$
$T_{34}$	$T_{3>4}$

**Πίνακας 6.9**

και ο αντίστοιχος αρχικός πίνακας θα είναι:

$t_{13}$	$t_{1>4}$
$t_{23}$	$t_{2>4}$
$t_{34}$	$t_{3>4}$

**Πίνακας 6.10**

όπου στην αριστερή στήλη θα έχουμε τα αντίστοιχα στοιχεία του αρχικού πίνακα για τον κόμβο που εξετάζουμε κάθε φορά και στην δεξιά στήλη θα είναι το άθροισμα των αντίστοιχων μετακινήσεων που συνεχίζουν κατάντη του κόμβου. Δηλαδή το  $t_{1>4}$  θα ισούται με το άθροισμα των έγχρωμων κελιών της πρώτης γραμμής του αρχικού πίνακα, το  $t_{2>4}$  θα ισούται με το άθροισμα των έγχρωμων κελιών της δεύτερης γραμμής του αρχικού πίνακα, το  $t_{3>4}$  θα ισούται με το άθροισμα των έγχρωμων κελιών της τρίτης γραμμής του αρχικού πίνακα και ούτω καθεξής.

	1	2	3	4	5	.....	n-1	n
1	0	$t_{12}$	$t_{13}$	$t_{14}$	$t_{15}$	.....	$t_{1,n-1}$	$t_{1n}$
2	0	0	$t_{23}$	$t_{24}$	$t_{25}$	.....	$t_{2,n-1}$	$t_{2n}$
3	0	0	0	$t_{34}$	$t_{35}$	.....	$t_{3,n-1}$	$t_{3n}$
4	0	0	0	0	$t_{45}$			
5	0	0	0	0	0			
6	0	0	0	0	0		0	
7	0	0	0	0	0		0	0

Η συγκεκριμένη μεθοδολογία συνεχίζεται μέχρι και τον προτελευταίο κόμβο του δικτύου μας. Αν λοιπόν μετά την επίλυση του προβλήματος σε κάθε κόμβο, κρατάμε την αριστερή στήλη του πίνακα των αποτελεσμάτων στο τέλος θα δημιουργήσουμε τον συνολικό εκτιμώμενο πίνακα Π-Π.

### 6.3. Γενικευμένη μορφή του επαναληπτικού αλγορίθμου

Έστω ότι έχουμε ένα γραμμικό δίκτυο μονής κατεύθυνσης με  $n$  κόμβους. Ως δεδομένα έχουμε τον αρχικό πίνακα  $t$  και τον αριθμό αυτών που εισέρχονται ( $O_i$ ) και εξέρχονται ( $D_i$ ) από το δίκτυο σε κάθε κόμβο. Εμείς θέλουμε να υπολογίσουμε τον πιο πιθανό πίνακα Π-Π  $T$  ακολουθώντας μια επαναληπτική διαδικασία

βελτιστοποίησης σε κάθε κόμβο. Για την βελτιστοποίηση μπορεί να χρησιμοποιηθεί οποιοδήποτε μοντέλο μεγιστοποίησης της εντροπίας.

Στον πρώτο κόμβο δεν αποβιβάζεται κανείς. Λόγο της γεωμετρίας του δικτύου, για τον δεύτερο κόμβο ισχύει ότι  $T_{12} = D_2$ .

Γενικά τώρα για τους κόμβους 3 έως  $(n-1)$  ισχύει:

Για ένα κόμβο  $j$  θέλουμε να προσδιορίσουμε τα  $T_{1,j}, T_{2,j}, K, T_{i,j}$ .

Για κάθε κόμβο  $j$  του δικτύου, θα επιλύουμε ένα πρόβλημα βελτιστοποίησης του οποίου η διάσταση θα είναι ένας πίνακας  $R$ , μεγέθους  $(j-1) \times 2$ . Στην αριστερή στήλη του  $R$  θα είναι ο αριθμός αυτών που αποβιβάζονται στον κόμβο  $(R_{i,1})$  και στην δεξιά θα είναι το άθροισμα των αντίστοιχων μετακινήσεων που συνεχίζουν κατάντη του κόμβου  $(R_{i,2})$ . Η αντικειμενική συνάρτηση του μοντέλου που θα βελτιστοποιείται κάθε φορά είναι:

$$\max Z_j = \frac{R!}{\prod_{i=1}^{j-1} \prod_{k=1}^2 (R_{i,k})} \quad (6.3.1)$$

όπου

$$R = \sum_{i=1}^{j-1} \sum_{k=1}^2 (R_{i,k}) \quad (6.3.2)$$

Οι περιορισμοί που θα ισχύουν σε κάθε κόμβο εκφράζονται από τις ακόλουθες σχέσεις και βασίζονται στην αρχή, ότι όσοι εισέρχονται σε ένα κόμβο, τόσοι και θα εξέρχονται από αυτόν.

$$\sum_{k=1}^2 (R_{i,k}) = O_i - \sum_{k=1}^{j-1} (T_{i,k}) \quad (6.3.3)$$

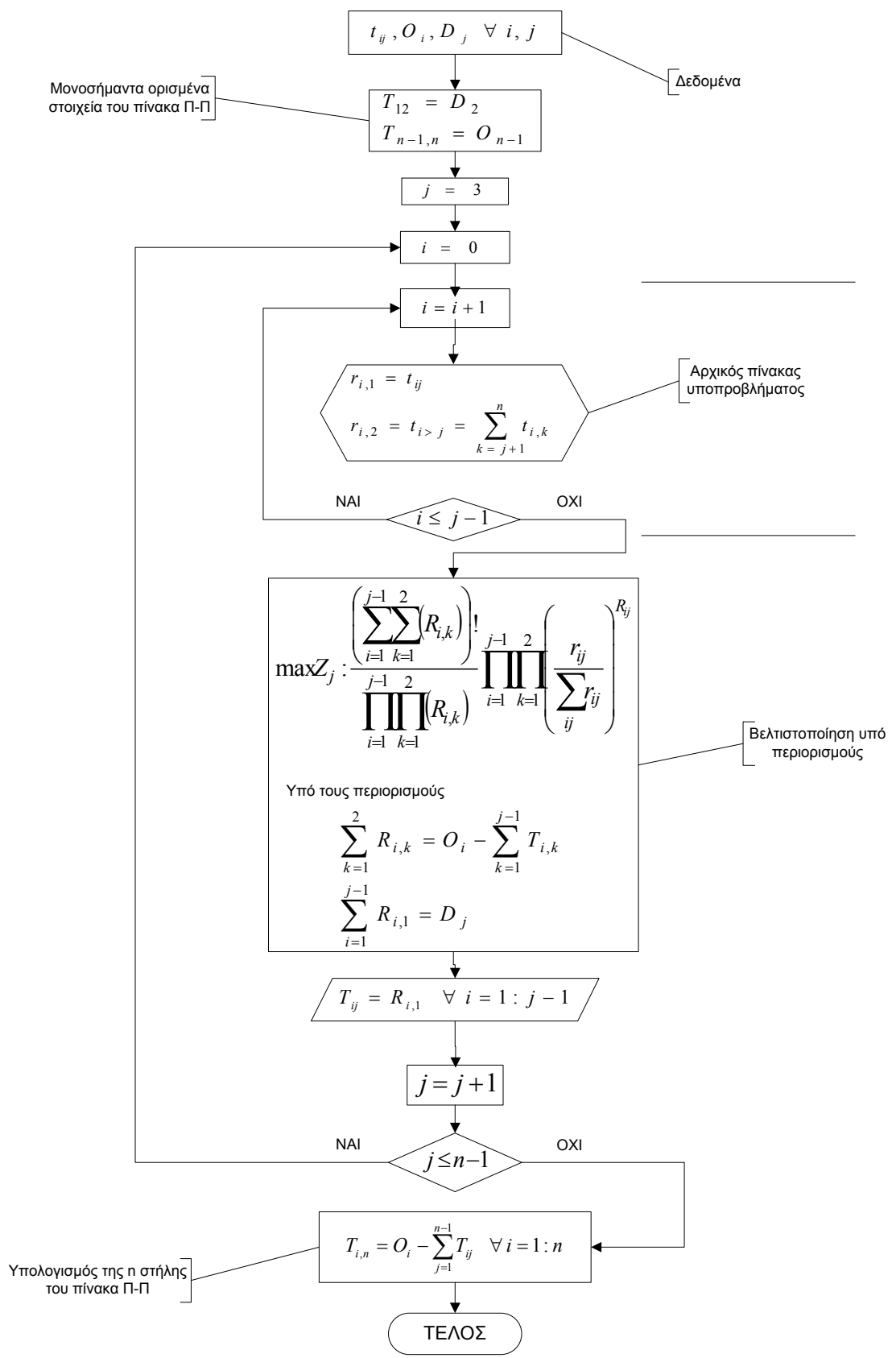
$$\sum_{i=1}^{j-1} (R_{i,1}) = D_j \quad (6.3.4)$$

Οι παρακάτω σχέσεις δείχνουν την αντιστοιχία των τιμών των κελιών του πίνακα  $R$  με αυτά του πίνακα  $T$  που θέλουμε να προσδιορίσουμε.

$$R_{i,1} = T_{i,j} \quad (6.3.5)$$

$$R_{i,2} = T_{i>j} \quad (6.3.6)$$

# ΔΙΑΓΡΑΜΜΑ ΡΟΗΣ



Στο παραπάνω διάγραμμα ροής παρουσιάζουμε την σειρά των βημάτων του επαναληπτικού αλγορίθμου μεγιστοποίησης της εντροπίας.

Ξεκινάει με την εισαγωγή των δεδομένων, που είναι ο αρχικός πίνακας  $t_{ij}$ , ο αριθμός αυτόν που εισέρχονται σε κάθε κόμβο ( $O_i$ ) και ο αριθμός αυτόν που εξέρχονται από κάθε κόμβο ( $D_i$ ).

Στην συνέχεια υπολογίζονται τα μονοσήμαντα ορισμένα στοιχεία του πίνακα Π-Π. Αυτά είναι τα στοιχεία  $T_{12}, T_{n-1,n}$  που τα προσδιορίζουμε λόγο της γεωμετρίας του δικτύου.

Ο αλγόριθμος τώρα συνεχίζει στον επόμενο κόμβο. Πρώτα δημιουργεί τον αρχικό πίνακα  $r$  για το υποπροβλημα που εξετάζεται. Αφού ολοκληρωθεί αυτό το βήμα ακολουθεί η βελτιστοποίηση του υποπροβλήματος υπό τους περιορισμούς που ισχύουν. Για την βελτιστοποίηση έχει επιλεγεί το βασικό μοντέλο μεγιστοποίησης της εντροπίας αλλά μπορεί να αντικατασταθεί με οποιοδήποτε άλλο μοντέλο. Τα αποτελέσματα της βελτιστοποίησης δίνονται στον πίνακα  $R$ . Τα κελιά της πρώτης στήλης του  $R$  είναι αυτά που αντιστοιχούν στα κελιά του πίνακα Π-Π που θέλουμε να εκτιμήσουμε.

Ο αλγόριθμος συνεχίζει στον επόμενο κόμβο μέχρι να επιλυθεί ολόκληρο το δίκτυο.

#### **6.4. Συμπεράσματα για τον επαναληπτικό αλγόριθμο**

Σε αυτό το κεφάλαιο παρουσιάσαμε αναλυτικά την μορφή του επαναληπτικού αλγορίθμου και περιγράψαμε την μέθοδο που ακολουθείται για τον προσδιορισμό του πίνακα Π-Π. Ουσιαστικά συνδυάσαμε την σταδιακή προσέγγιση του προβλήματος που περιέγραφε ο αλγόριθμος των Li, Cassidy (2006) και τα μοντέλα μεγιστοποίησης της εντροπίας. Έτσι προέκυψε μια μέθοδος που βελτιστοποιεί το πρόβλημα σταδιακά, ακλουθώντας την κατεύθυνση του δικτύου.

## 7. Εφαρμογή του επαναληπτικού αλγορίθμου μεγιστοποίησης της εντροπίας

### 7.1. Εφαρμογή

Στο κεφάλαιο αυτό θα εφαρμόσουμε τον επαναληπτικό αλγόριθμο στο ίδιο πρόβλημα που εφαρμόστηκαν και τα μοντέλα μεγιστοποίησης της εντροπίας και ο αλγόριθμος των Li, Cassidy. Χρησιμοποιήθηκε ο ίδιος αρχικός πίνακας που χρησιμοποιήθηκε και στην εφαρμογή των άλλων μεθόδων.

Τα αποτελέσματα που προέκυψαν παρουσιάζονται στον πίνακα 7.2. Επίσης ακολουθεί ένας πίνακας με διαφορά στατιστικά μεγέθη και ένα γράφημα για τα αποτελέσματα.

Στατιστικά παλινδρόμησης	
Πολλαπλό R	0.992
R Τετράγωνο	0.983
Προσαρμοσμένο R Τετράγωνο	0.983
Τυπικό σφάλμα	109.185
Μέγεθος δείγματος	529.000

ΑΝΑΛΥΣΗ ΔΙΑΚΥΜΑΝΣΗΣ					
	βαθμοί ελευθερίας	SS	MS	F	Σημαντικότητα F
Παλινδρόμηση	1	367727701	367727701	30846	0
Υπόλοιπο	527	6282604	11921		
Σύνολο	528	374010305			

	Συντελεστές	Τυπικό σφάλμα	t	τιμή-P	Κατώτερο 95%	Υψηλότερο 95%	Κατώτερο 95.0%	Υψηλότερο 95.0%
Τεταγμένη επί την αρχή	-0.190	5.071	-0.037	0.970	-10.151	9.771	-10.151	9.771
Μεταβλητή X 1	1.001	0.006	175.630	0.000	0.989	1.012	0.989	1.012

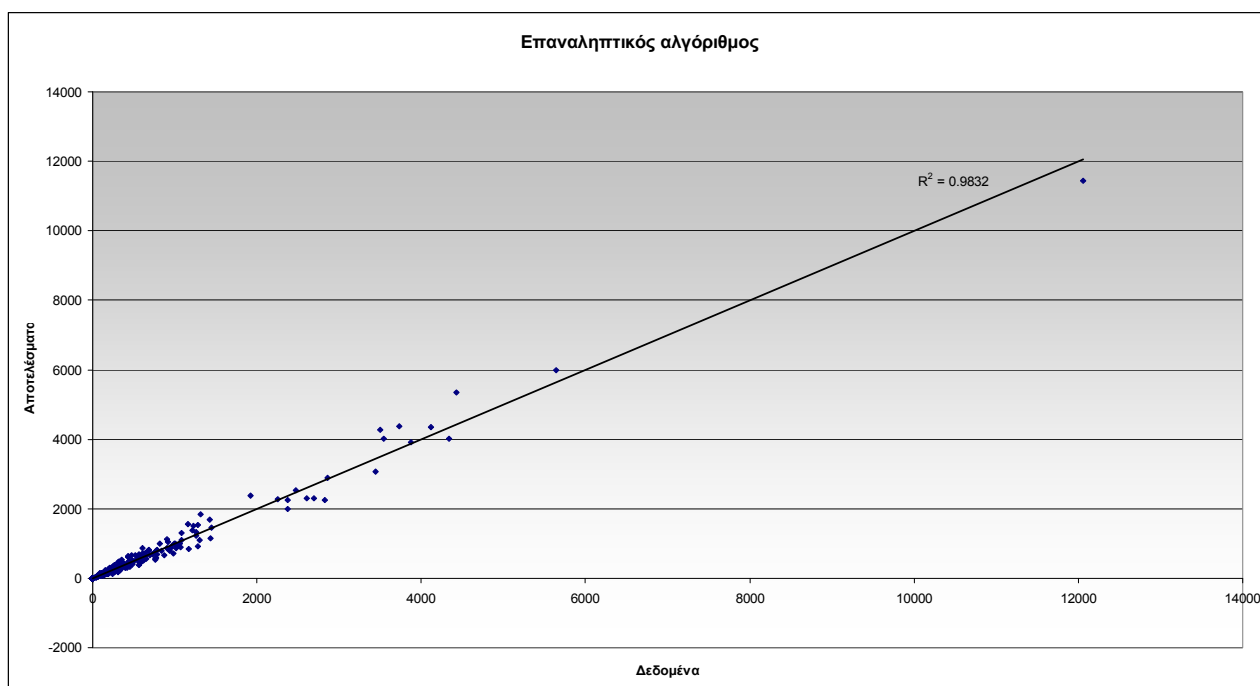
Πίνακας 7.1



**Αποτελέσματα του επαναληπτικού αλγόριθμου (Πειραιάς-Κηφισιά) – σύνολο ημέρας**

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
		Πειρ	Ν.Φάλ	Μοσχ	Καλλ	Ταύρ	Πετρ	Θησ	Μοναστ	Ομόν	Βικτ	Αττ	Αγ.Νικ	Κ.Πατ	Αγ.Ελευθ	Α.Πατ	Περισ	Πευκ	Νέα Ιωνία	Ν.Ηράκ	Ειρ	Μαρ	Κ.Α.Τ	Κηφ
1	Πειραιάς	0	990	788	1250	705	1109	585	1910	13284	2953	964	864	1053	456	662	297	193	693	617	203	948	348	910
2	Ν.Φάληρο	0	0	135	253	187	421	319	819	4339	1325	667	307	261	223	117	197	79	119	190	96	245	77	304
3	Μοσχάτο	0	0	0	214	161	139	263	727	2504	586	327	169	155	105	104	57	35	85	93	10	87	44	94
4	Καλλιθέα	0	0	0	0	108	835	602	1575	4547	1939	783	425	804	266	250	299	97	201	209	132	394	124	354
5	Ταύρος	0	0	0	0	0	255	347	846	4398	1064	473	225	298	134	150	123	51	104	184	55	182	61	185
6	Πετράλωνα	0	0	0	0	0	0	410	1548	3783	1103	629	296	310	291	263	121	91	221	148	59	281	97	261
7	Θησείο	0	0	0	0	0	0	0	141	1620	610	303	224	264	186	200	67	63	98	200	69	271	96	121
8	Μοναστηράκι	0	0	0	0	0	0	0	0	1190	1771	973	839	1507	869	648	497	171	341	520	228	1042	323	579
9	Ομόνοια	0	0	0	0	0	0	0	0	2435	3284	3518	3206	2756	2440	1752	850	2548	4204	1598	2903	955	3100	
10	Βικτώρια	0	0	0	0	0	0	0	0	0	0	1120	1223	1374	580	429	356	202	624	469	219	876	453	1012
11	Αττική	0	0	0	0	0	0	0	0	0	0	0	314	372	337	450	239	136	395	609	176	999	254	560
12	Αγ.Νικόλαος	0	0	0	0	0	0	0	0	0	0	0	0	177	310	308	201	83	546	277	200	553	129	608
13	Κ.Πατήσια	0	0	0	0	0	0	0	0	0	0	0	0	0	306	256	205	124	316	487	195	696	261	365
14	Αγ.Ελευθέριος	0	0	0	0	0	0	0	0	0	0	0	0	0	0	115	294	164	362	211	155	391	143	296
15	Α.Πατήσια	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	233	247	500	343	144	406	171	325
16	Περισσός	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	53	498	194	120	361	97	170
17	Πευκάκια	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	128	289	78	171	99	132
18	Νέα Ιωνία	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	558	189	374	209	369
19	Ν.Ηράκλειο	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	429	811	233	421
20	Ειρήνη	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	306	91	162
21	Μαρούσι	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	290	618
22	Κ.Α.Τ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	162
23	Κηφισιά	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

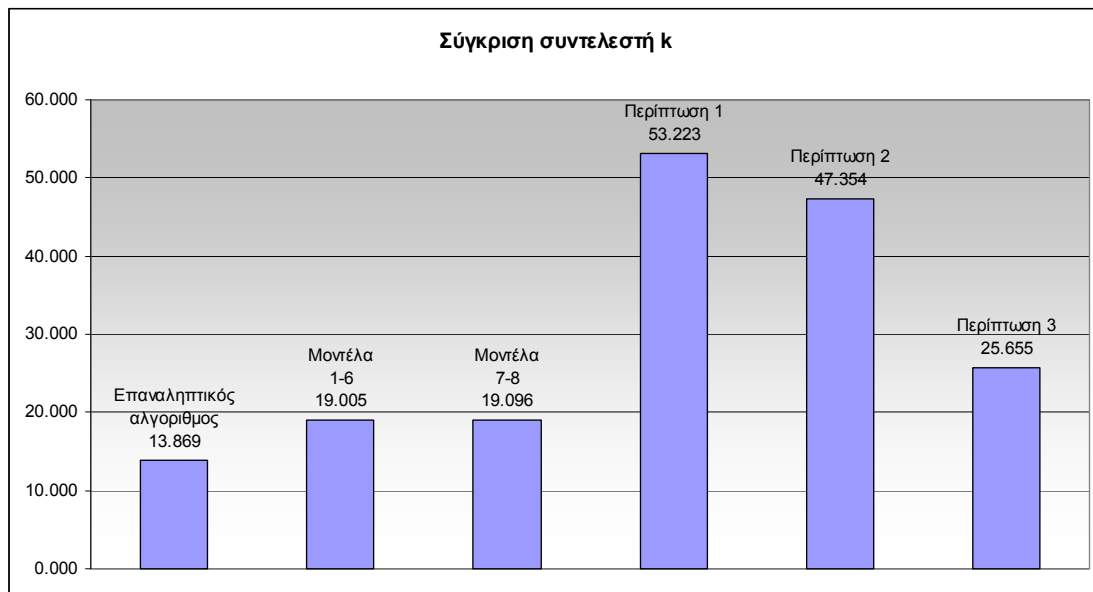
**Πίνακας 7.2**



**Γράφημα 7.1**

## 7.2. Συγκρίσεις

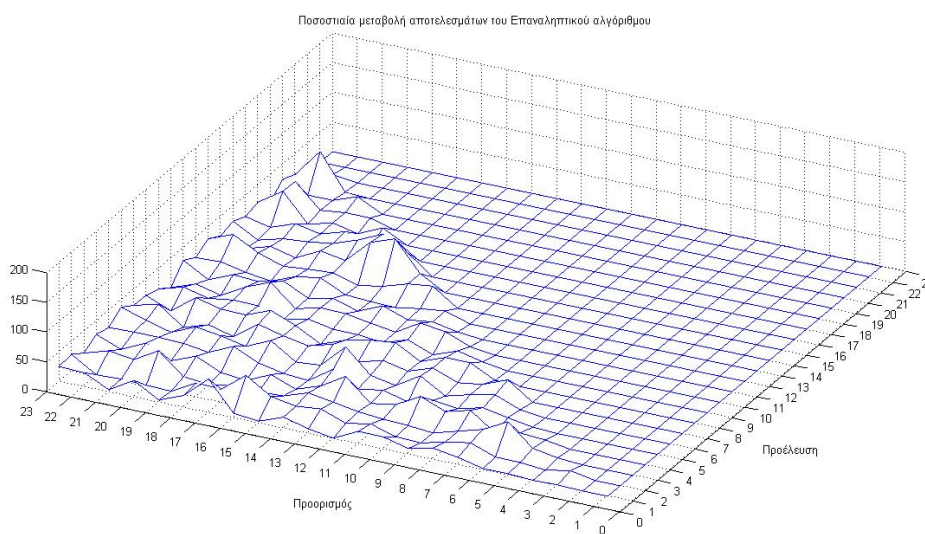
Σε αυτό το κεφάλαιο θα συγκρίνουμε την ακρίβεια των αποτελεσμάτων του επαναληπτικού αλγορίθμου με αυτή των άλλων μεθόδων. Στο γράφημα 7.2 γίνεται σύγκριση του συντελεστή σταθμισμένης ποσοστιαίας μεταβολής  $k$  για τον αλγόριθμό μας, τα μοντέλα εντροπίας και τον αλγόριθμο των Li, Cassidy (2006).



**Γράφημα 7.2**

Από το γράφημα 7.2 βλέπουμε ότι ο επαναληπτικός αλγόριθμος δίνει μεγαλύτερη ακρίβεια από οποιαδήποτε άλλη προσέγγιση που εξετάσαμε για την εκτίμηση του πίνακα Π-Π. Η βελτίωση της ακρίβειας των αποτελεσμάτων είναι της τάξεως του 5% από αυτή που προσφέρει η αμέσως καλύτερη προσέγγιση.

Ακολουθεί ένα τρισδιάστατο γράφημα της ποσοστιαίας μεταβολής των αποτελεσμάτων



**Γράφημα 7.3**

Στο γράφημα 7.3 παρατηρούμε ότι δεν υπάρχει κάποια έντονη ή συστηματική διακύμανση, αντιθέτως βλέπουμε ότι οι ποσοστιαίες μεταβολές είναι ομοιόμορφα κατανεμημένες.

### **7.3. Εφαρμογή σε διαφορετικό πρόβλημα**

Για να θεωρήσουμε τα αποτελέσματα που προέκυψαν από τον επαναληπτικό αλγόριθμο έγκυρα και τα συμπεράσματα που εξάγαμε αξιόπιστα θα πρέπει να εφαρμόσουμε τον αλγόριθμο και σε ένα διαφορετικό πρόβλημα. Εμείς εφαρμόσαμε τον επαναληπτικό αλγόριθμο για να εκτιμήσουμε τον πίνακα Π-Π της γραμμής του ΗΣΑΠ για την πρωινή αιχμή. Επίσης εκτιμήσαμε τον πίνακα Π-Π για την πρωινή αιχμή και για τα μοντέλα 1 έως 6 μεγιστοποίησης της εντροπίας ώστε να μπορούμε να κάνουμε σύγκριση των αποτελεσμάτων. Επιλέχθηκαν τα μοντέλα αυτά γιατί δίνουν την μεγαλύτερη ακρίβεια στα αποτελέσματα τους από τις υπόλοιπες μεθόδους επίλυσης που εξετάσαμε. Τα αποτελέσματα που προέκυψαν παρουσιάζονται στους παρακάτω πίνακες.

**Αποτελέσματα του επαναληπτικού αλγόριθμου (Πειραιάς-Κηφισιά) – πρωινή αιχμή**

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
		Πειρ	Ν.Φάλ	Μοσχ	Καλλ	Ταύρ	Πετρ	Θησ	Μοναστ	Ομόν	Βικτ	Αττ	Αγ.Νικ	Κ.Πατ	Αγ.Ελευθ	Α.Πατ	Περισ	Πευκ	Νέα Ιωνία	Ν.Ηράκ	Ειρ	Μαρ	Κ.Α.Τ	Κηφ
1	Πειραιάς	0	38	80	64	45	24	55	309	1449	340	40	24	35	25	22	20	3	22	33	15	43	21	80
2	Ν.Φάληρο	0	0	20	38	11	10	34	216	634	132	32	17	28	11	16	8	3	14	10	10	17	4	24
3	Μοσχάτο	0	0	0	16	15	15	31	87	333	67	13	8	11	9	8	6	1	7	3	2	9	3	13
4	Καλλιθέα	0	0	0	0	4	17	71	245	660	110	26	11	37	18	14	13	4	11	9	3	17	4	43
5	Ταύρος	0	0	0	0	0	14	44	169	720	120	31	9	27	16	7	8	7	29	14	11	23	7	23
6	Πετράλωνα	0	0	0	0	0	0	37	160	866	87	25	9	26	23	13	10	3	28	8	6	21	8	24
7	Θησείο	0	0	0	0	0	0	0	6	190	39	10	4	12	8	6	3	1	19	2	2	11	3	10
8	Μοναστηράκι	0	0	0	0	0	0	0	0	15	26	9	5	11	6	8	2	3	12	4	2	5	2	6
9	Ομόνοια	0	0	0	0	0	0	0	0	0	115	64	70	74	49	53	63	18	91	61	82	88	43	166
10	Βικτώρια	0	0	0	0	0	0	0	0	0	0	27	33	44	33	23	18	7	21	46	18	67	21	120
11	Αττική	0	0	0	0	0	0	0	0	0	0	0	24	27	29	25	17	9	46	37	23	131	35	119
12	Αγ.Νικόλαος	0	0	0	0	0	0	0	0	0	0	0	0	10	24	14	25	7	39	40	19	71	19	72
13	Κ.Πατήσια	0	0	0	0	0	0	0	0	0	0	0	0	0	12	19	17	13	30	40	40	56	33	95
14	Αγ.Ελευθέριος	0	0	0	0	0	0	0	0	0	0	0	0	0	0	21	11	5	32	28	16	51	22	55
15	Α.Πατήσια	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7	8	30	20	10	39	14	65
16	Περισσός	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	57	12	16	29	7	24
17	Πευκάκια	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	16	8	22	9	23
18	Νέα Ιωνία	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	17	9	26	17	45
19	Ν.Ηράκλειο	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	48	65	26	50
20	Ειρήνη	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7	6	7
21	Μαρούσι	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11	36
22	Κ.Α.Τ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11
23	Κηφισιά	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

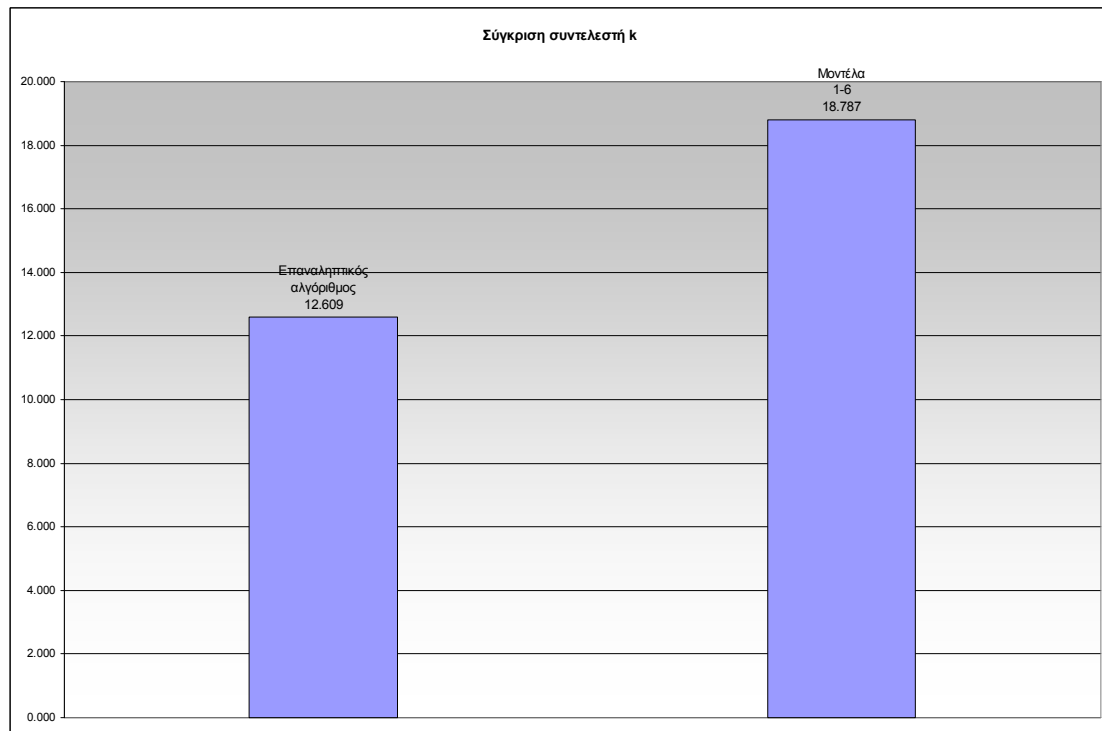
**Πίνακας 7.3**

**Αποτελέσματα των μοντέλων 1-6 (Πειραιάς-Κηφισιά) – πρωινή αιχμή**

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
		Πειρ	Ν.Φάλ	Μοσχ	Καλλ	Ταύρ	Πετρ	Θησ	Μοναστ	Ομόν	Βικτ	Αττ	Αγ.Νικ	Κ.Πατ	Αγ.Ελευθ	Α.Πατ	Περισ	Πευκ	Νέα Ιωνία	Ν.Ηράκ	Ειρ	Μαρ	Κ.Α.Τ	Κηφ
1	Πειραιάς	0	29	78	67	44	31	60	352	1206	274	40	40	33	23	20	17	4	27	30	10	44	30	49
2	Ν.Φάληρο	0	0	18	23	8	10	28	131	543	124	23	21	17	7	18	7	3	14	14	13	23	5	39
3	Μοσχάτο	0	0	0	13	11	13	25	123	461	44	11	7	13	8	10	4	1	13	6	3	9	4	8
4	Καλλιθέα	0	0	0	0	6	24	57	263	574	141	29	15	27	26	20	21	4	15	17	4	17	5	28
5	Ταύρος	0	0	0	0	0	13	73	169	596	108	35	13	23	18	5	8	5	30	11	9	17	9	27
6	Πετράλωνα	0	0	0	0	0	0	43	254	573	174	36	12	26	17	26	9	4	27	8	6	24	9	25
7	Θησείο	0	0	0	0	0	0	0	9	136	38	15	7	13	11	3	7	1	19	4	4	8	3	10
8	Μοναστηράκι	0	0	0	0	0	0	0	0	22	23	11	4	8	6	8	4	3	9	4	2	6	3	7
9	Ομόνοια	0	0	0	0	0	0	0	0	0	174	80	53	66	57	55	51	22	98	107	63	107	64	144
10	Βικτώρια	0	0	0	0	0	0	0	0	0	0	25	22	38	21	15	17	7	22	47	23	58	25	194
11	Αττική	0	0	0	0	0	0	0	0	0	0	0	33	25	25	27	29	13	38	36	29	146	22	138
12	Αγ.Νικόλαος	0	0	0	0	0	0	0	0	0	0	0	0	6	23	20	26	9	35	43	30	53	24	84
13	Κ.Πατήσια	0	0	0	0	0	0	0	0	0	0	0	0	0	21	18	34	11	27	25	25	77	24	81
14	Αγ.Ελευθέριος	0	0	0	0	0	0	0	0	0	0	0	0	0	0	18	23	7	41	29	16	48	15	81
15	Α.Πατήσια	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11	11	29	16	12	41	14	73
16	Περισσός	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	48	13	17	27	7	41
17	Πευκάκια	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11	12	16	15	10	26
18	Νέα Ιωνία	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30	17	44	15	31
19	Ν.Ηράκλειο	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	79	58	24	41
20	Ειρήνη	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11	3	10
21	Μαρούσι	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	12	39
22	Κ.Α.Τ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	14
23	Κηφισιά	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

**Πίνακας 7.4**

Κάνοντας την σύγκριση του συντελεστή σταθμισμένης ποσοστιαίας μεταβολής  $k$ , που φαίνεται στο γράφημα 7.4, βλέπουμε ότι η βελτίωση της ακρίβειας των αποτελεσμάτων ανάμεσα στα μοντέλα μεγιστοποίησης της εντροπίας και του επαναληπτικού αλγορίθμου είναι της τάξεως του 6%. Στην εφαρμογή για το σύνολο της ημέρας είχε προκύψει μια βελτίωση της τάξεως του 5%.



Γράφημα 7.4

## 7.4. Συμπεράσματα

Στο κεφάλαιο αυτό έγινε η σύγκριση της μεθόδου που προτείνεται από αυτή την εργασία σε σχέση με υπάρχουσες μεθόδους εκτίμησης πινάκων Π-Π από μετρήσεις φόρτων σε οδικούς συνδέσμους. Τα αποτελέσματα ήταν σαφή και διαπιστώσαμε ότι η βελτίωση της ακρίβειας που προσφέρει ο επαναληπτικός αλγόριθμος μεγιστοποίησης της εντροπίας είναι της τάξης του 5%.

## 8. Συμπεράσματα και προτάσεις

### 8.1. Συμπεράσματα

Η διπλωματική αυτή εργασία είχε σκοπό να αξιολογήσει τις μεθόδους προσδιορισμού πινάκων Π-Π από μετρήσεις φόρτων σε οδικούς συνδέσμους ώστε να εντοπίσει τα πλεονεκτήματα και μειονεκτήματα των μεθόδων αυτών. Με βάση τα συμπεράσματα αυτής της αξιολόγησης, σκοπός μας ήταν να προτείνουμε μια εναλλακτική μορφή μοντέλου που να παρέχει μεγαλύτερη ακρίβεια.

Οι προσεγγίσεις που εξετάστηκαν ήταν αυτή της μεγιστοποίησης της εντροπίας και του γενικευμένου αλγόριθμου εκτίμησης πινάκων Π-Π για γραμμικά δίκτυα.

Από τα αποτελέσματα των εφαρμογών που εξετάσαμε για τις δύο αυτές προσεγγίσεις, παρατηρήσαμε ότι η ακρίβεια των αποτελεσμάτων των μοντέλων μεγιστοποίησης της εντροπίας ήταν σαφώς μεγαλύτερη από αυτή που έδινε ο αλγόριθμος των Li, Cassidy (2006). Επίσης για τα μοντέλα μεγιστοποίησης της εντροπίας διαπιστώσαμε ότι η ακρίβεια που προσφέρουν είναι αντιστρόφως ανάλογη του μεγέθους του προβλήματος που επιλύουν. Για τον αλγόριθμο των Li, Cassidy (2006) διαπιστώσαμε ένα συστηματικό σφάλμα, που παρουσιάστηκε σε όλες τις περιπτώσεις που εξετάσαμε και έχει να κάνει με την υπερέτιμηση των πρώτων κελιών κάθε γραμμής του εκτιμώμενου πίνακα Π-Π και τον μηδενισμό των τελευταίων κελιών της γραμμής αυτής.

Στη συνέχεια αναπτύχθηκε μια εναλλακτική μορφή μοντέλου για την εκτίμηση του πίνακα Π-Π. Είναι ένας επαναληπτικός αλγόριθμος μεγιστοποίησης της εντροπίας, ο οποίος αντιμετωπίζει το πρόβλημα σταδιακά. Η ακρίβεια που προσφέρει ο αλγόριθμος αυτός, είναι συστηματικά καλύτερη από αυτή που δίνουν τα μοντέλα μεγιστοποίησης της εντροπίας και ο αλγόριθμος των Li, Cassidy (2006) και αυτό αποδείχθηκε μέσα από μία σειρά εφαρμογών σε πραγματικά προβλήματα.



## 8.2. Προτάσεις

Στην εργασία αυτή, ο επαναληπτικός αλγόριθμος μεγιστοποίησης της εντροπίας εφαρμόστηκε για ένα γραμμικό δίκτυο μονής κατεύθυνσης. Σε μελλοντική έρευνα θα μπορούσε να αναπτυχθεί ώστε να μπορεί να εφαρμοστεί και σε πιο πολύπλοκα δίκτυα.

Θα πρέπει να γίνουν περαιτέρω συγκρίσεις μεταξύ του επαναληπτικού αλγορίθμου μεγιστοποίησης της εντροπίας και των μοντέλων μεγιστοποίησης της εντροπίας. Μέσα από αυτές τις συγκρίσεις και τα αποτελέσματα που θα προκύψουν, θα μπορέσουν να εντοπιστούν τα πλεονεκτήματα και μειονεκτήματα της κάθε μεθόδου.

Ο επαναληπτικός αλγόριθμος θα μπορούσε να τροποποιηθεί, αν είναι δυνατό, ώστε να εκτιμά δυναμικούς πίνακες Π-Π.

## **Βιβλιογραφικές αναφορές**

Wilson A.G. (1967) “A statistical theory of spatial distribution models”,  
Transportation Research Part A, 253-269.

Tsygalnitsky S. (1977) “Simplified methods for transportation planning”, M.S.  
Thesis, Department of Civil Engineering, Massachusetts Institute of Technology,  
Cambridge.

Van Zuylen H. and Willumsen L.G. (1980) “The most likely trip matrix estimated  
from traffic counts”, Transportation Research Part B 14B, 281-293.

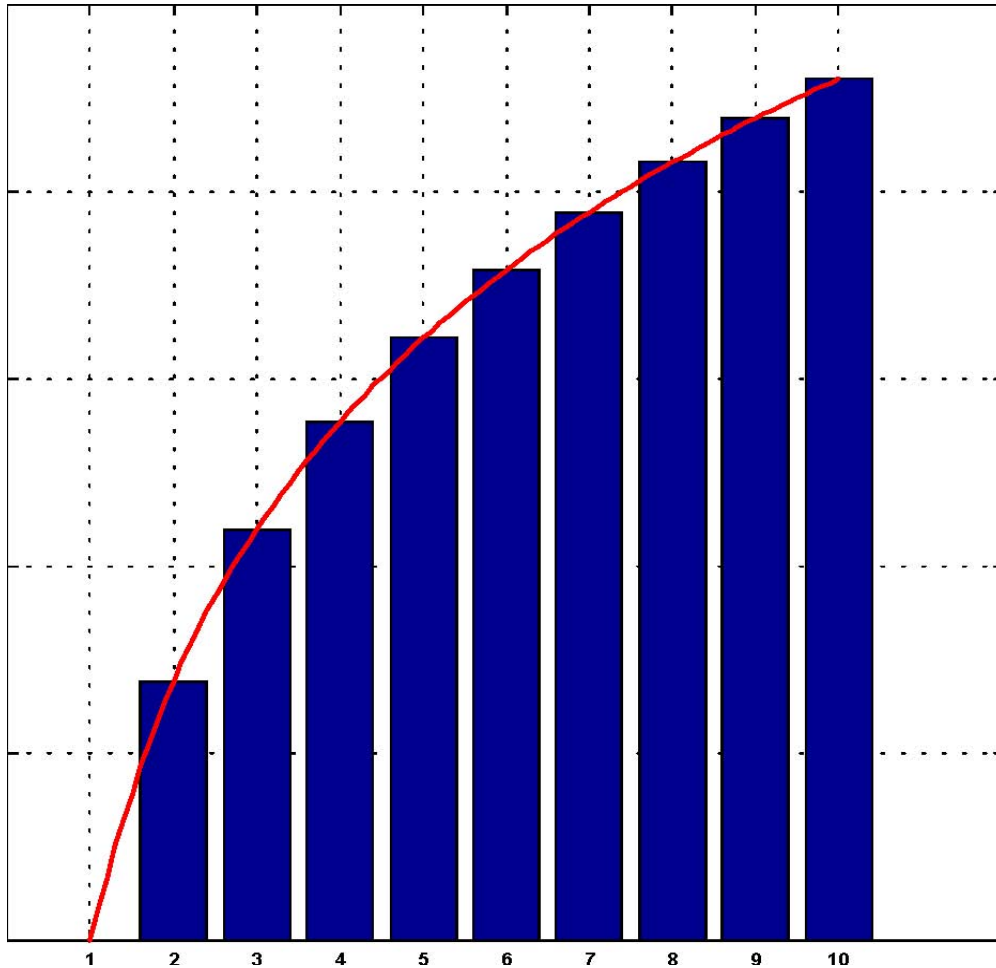
Willumsen L.G. (1981) “Simplified transport models based on traffic counts”,  
Transportation 10, 257-278.

Paramahamsan H. (1999) “Fundamental properties of synthetic o-d generation  
formulations and solutions”, Blacksburg, Virginia.

Li Y., Cassidy J. (2006) “A generalized and efficient algorithm for estimating transit  
route ODs from passenger counts”, Transportation Research Part B 41 (2007),  
114-125.

## Παράρτημα Α

Εδώ θα εξηγήσουμε τον μετασχηματισμό του Stirling. Ας θεωρήσουμε την εξίσωση  $y = \text{Ln}(x)$  όπως φαίνεται στο παρακάτω γράφημα.



Προσπαθώντας να προσεγγίσουμε την περιοχή κάτω από την καμπύλη, βρίσκουμε ότι:

Α μέθοδος: Η περιοχή κάτω από την καμπύλη μεταξύ του  $x=1$  και  $x=N$  είναι

$$\approx (1)\text{Ln}(1) + (1)\text{Ln}(2) + (1)\text{Ln}(3) + (1)\text{Ln}(4) + \dots + (1)\text{Ln}(N) = \text{Ln}(N!)$$

Β μέθοδος: Η περιοχή κάτω από την καμπύλη μπορεί να προσδιοριστεί ολοκληρώνοντας για  $x=1$  έως  $x=N$ .

$$\int_1^N \ln(x) dx$$

Γνωρίζουμε ότι

$$\int \ln(x) dx = x \ln(x) - x$$

Άρα ισχύει

$$\int_1^N \ln(x) dx = N \ln(N) - N - 0 + 1$$

Εάν το  $N$  είναι μεγάλος αριθμός, τότε το '1' στην τελευταία σχέση έχει αμελητέα επίδραση και μπορεί να παραληφθεί.

Συνδυάζοντας τις μεθόδους A και B καταλήγουμε στην παρακάτω σχέση:

$$\ln(N!) = N \ln(N) - N$$

## Παράρτημα Β

Ο αλγόριθμος Gauss-Newton είναι μια μέθοδος η οποία χρησιμοποιείται για την λύση προβλημάτων μη γραμμικών ελαχίστων τετραγώνων. Μπορεί να αναφερθεί και ως μια τροποποίηση της μεθόδου του Newton για την εύρεση της ελάχιστης τιμής μιας εξίσωσης. Σε αντίθεση με την μέθοδο του Newton, ο αλγόριθμος Gauss-Newton μπορεί να χρησιμοποιηθεί μόνο για να ελαχιστοποιήσει ένα άθροισμα των τιμών μιας εξίσωσης στο τετράγωνο, αλλά έχει το πλεονέκτημα ότι οι δεύτερες παράγωγοι, που είναι πρόκληση να υπολογιστούν, δεν χρειάζονται.

### Ο αλγόριθμος

Ας θεωρήσουμε  $m$  εξισώσεις  $r_i (i = 1, \dots, m)$  με  $n$  μεταβλητές  $\beta = (\beta_1, \beta_2, \dots, \beta_n)$ , με  $m \geq n$ , ο αλγόριθμος Gauss-Newton βρίσκει το ελάχιστο του αθροίσματος των τετραγώνων  $S(\beta) = \sum_{i=1}^m r_i^2(\beta)$ . Αρχίζοντας με μια αρχική εκτίμηση  $\beta^0$  για το ελάχιστο, η μέθοδος προχωράει με τις επαναλήψεις  $\beta^{s+1} = \beta^s + \delta\beta$ , όπου το  $\delta\beta$  ικανοποιεί τις εξισώσεις  $(J_r^T J_r) \delta\beta = -J_r^T r$ . Εδώ το  $r$  είναι το διάνυσμα της εξίσωσης  $r_i$  και το  $J_r$  είναι ο  $m \times n$  Ιακωβιανός πίνακας του  $r$  σε σχέση με το  $\beta$ . Ο υπέρτιτλος  $T$  υποδηλώνει ότι ο πίνακας έχει αντιμετατεθεί. Στην προσαρμογή των δεδομένων, όπου σκοπός είναι να βρούμε τις παραμέτρους  $\beta$  έτσι ώστε ένα μοντέλο λειτουργίας  $y = f(x, \beta)$  να ταιριάζει καλύτερα σε κάποια δεδομένα σημεία  $(x_i, y_i)$ , οι εξισώσεις  $r_i$  είναι τα υπόλοιπα  $r_i(\beta) = y_i - f(x_i, \beta)$ . Τότε το  $\delta\beta$  μπορεί να εκφραστεί σε σχέση με τον Ιακωβιανό πίνακα και την εξίσωση  $f$  όπως φαίνεται  $(J_f^T J_f) \delta\beta = -J_f^T r$ .

### Παρατηρήσεις

Η υπόθεση του αλγορίθμου ότι  $m \geq n$  είναι απαραίτητη, γιατί αλλιώς ο πίνακας  $J_r^T J_r$  δεν είναι αντιστρέψιμος και οι εξισώσεις δεν μπορούν να λυθούν. Ο

αλγόριθμος Gauss-Newton μπορεί να παραχθεί με γραμμική προσέγγιση της λειτουργίας  $r_i$ . Χρησιμοποιώντας το θεώρημα του Taylor, σε κάθε επανάληψη μπορούμε να γράφουμε  $r(\beta) \approx r(\beta^s) + J_r(\beta^s)\delta\beta$  όπου  $\delta\beta = \beta - \beta^s$ . Η αποστολή του να βρούμε το  $\delta\beta$  ελαχιστοποιώντας το άθροισμα των τετραγώνων της δεξιάς πλευράς είναι ένα γραμμικό πρόβλημα ελαχίστων τετραγώνων. Οι εξισώσεις είναι  $n$  γραμμικές εξισώσεις του αγνώστου  $\delta\beta$ . Μπορεί να λυθούν σε ένα βήμα, χρησιμοποιώντας τον παραγοντισμό του Cholesky ή καλύτερα τον παραγοντισμό QR του  $J_r$ . Για μεγάλα συστήματα η επαναληπτική μέθοδος είναι ίσως πιο αποδοτική. Αν υπάρχει γραμμική εξάρτηση μεταξύ των στηλών του  $J_r$ , οι επαναλήψεις θα αποτύχουν καθώς  $J_r^T J_r$  γίνεται μοναδιαίος.

### Ιδιότητες σύγκλισης

Μπορεί να αποδειχθεί ότι το  $\delta\beta$  δείχνει την κατεύθυνση προς ένα τοπικό ελάχιστο  $S$  και αν ο αλγόριθμος συγκλίνει τότε το όριο είναι ένα σταθερό σημείο του  $S$ . Παρόλα αυτά η σύγκλιση δεν είναι εγγυημένη, ούτε καν τοπικά σε αντίθεση με την μέθοδο του Newton. Ο αλγόριθμος μπορεί να συγκλίνει αργά ή και καθόλου αν η αρχική υπόθεση είναι μακριά από το ελάχιστο ή ο πίνακας  $J_r^T J_r$  είναι λάθος εξαρτημένος.

Ο αλγόριθμος του Gauss-Newton μπορεί να αποτύχει να φτάσει σε σύγκλιση. Για παράδειγμα, ας υποθέσουμε ένα πρόβλημα όπου  $m=2$  εξισώσεις και  $n=1$  μεταβλητές όπου ισχύει

$$r_1(\beta) = \beta + 1$$

$$r_2(\beta) = \lambda\beta^2 + \beta - 1$$

Το βέλτιστο είναι όταν  $\beta = 0$ . Αν  $\lambda = 0$  τότε το πρόβλημα είναι γραμμικό και η μέθοδος βρίσκει το βέλτιστο σε μία επανάληψη. Αν  $|\lambda| < 1$  τότε η μέθοδος συγκλίνει γραμμικά και το σφάλμα μειώνεται ασυμπτωτικά κατά ένα παράγοντα  $|\lambda|$  σε κάθε επανάληψη. Όταν όμως  $|\lambda| > 1$  τότε η μέθοδος δεν συγκλίνει ούτε τοπικά.

## Προέλευση από την μέθοδο του Newton

Η επανεμφανιζόμενη σχέση για την μέθοδο του Newton για την ελαχιστοποίηση μιας λειτουργίας  $S$  με παράμετρο  $\beta$  είναι  $\beta^{s+1} = \beta^s - H^{-1}g$ , όπου  $g$  υποδηλώνει το διάνυσμα κλίσης του  $S$  και  $H$  υποδηλώνει του πίνακα του Hessian για το  $S$ . Αφού

$$S = \sum_{i=1}^m r_i^2 \quad \text{τότε η κλίση δίνεται από } g_j = 2 \sum_{i=1}^m r_i \frac{\partial r_i}{\partial \beta_j}.$$

Τα στοιχεία του πίνακα Hessian υπολογίζονται από την παρακάτω σχέση:

$$H_{jk} = 2 \sum_{i=1}^m \left( \frac{\partial r_i}{\partial \beta_j} \frac{\partial r_i}{\partial \beta_k} + r_i \frac{\partial^2 r_i}{\partial \beta_j \partial \beta_k} \right).$$

Η μέθοδος του Gauss-Newton προκύπτει αν παραλείψουμε το δεύτερο μέλος της παραπάνω εξίσωσης. Τότε προκύπτει ότι ο προσεγγιστικός πίνακας Hessian

προσεγγίζεται από την ακόλουθη σχέση:  $H_{jk} \approx 2 \sum_{i=1}^m J_{ij} J_{ik}$  όπου  $J_{ij} = \frac{\partial r_i}{\partial \beta_j}$  είναι από

τον Ιακωβιανό πίνακα  $J_r$ . Η κλίση και ο προσεγγιστικός πίνακας Hessian μπορούν να γραφούν σε μορφή πίνακα σαν:  $g = 2J_r^T r$  και  $H \approx 2J_r^T J_r$ . Αυτές οι εκφράσεις υποβάλλονται στις παραπάνω σχέσεις για να προκύψουν οι λειτουργικές εξισώσεις:

$$\beta^{s+1} = \beta^s + \delta\beta$$

$$\delta\beta = -(J_r^T J_r)^{-1} J_r^T r$$

Η σύγκλιση του αλγορίθμου Gauss-Newton δεν είναι εγγυημένη για κάθε περίπτωση.

Η προσέγγιση  $\left| r_i \frac{\partial^2 r_i}{\partial \beta_j \partial \beta_k} \right| \ll \left| \frac{\partial r_i}{\partial \beta_j} \frac{\partial r_i}{\partial \beta_k} \right|$  που πρέπει να ισχύει ώστε να μπορούμε να

κάνουμε τις παραπάνω προσεγγίσεις, μπορεί να είναι έγκυρη σε δύο περιπτώσεις, στις οποίες μπορούμε να περιμένουμε σύγκλιση

- Οι τιμές του  $r_i$  να είναι μικρές, τουλάχιστον γύρο από την περιοχή του τοπικού ελάχιστου.

- Οι λειτουργίες να είναι «ελαφρός» μη γραμμικές, ώστε η τιμή του όρου

$$\frac{\partial^2 r_i}{\partial \beta_j \partial \beta_k} \text{ να είναι μικρή.}$$



## Παράρτημα Γ

Στα μαθηματικά ο αλγόριθμος Levenberg-Marquardt (LM) δίνει μια αριθμητική λύση στο πρόβλημα ελαχιστοποίησης μιας λειτουργίας, γενικά σε μη γραμμικές, για ένα εύρος παραμέτρων της λειτουργίας. Ο (LM) αλγόριθμος είναι κάτι ανάμεσα στον αλγόριθμο Gauss-Newton και της μεθόδου gradient descent. Ο (LM) αλγόριθμος είναι πιο ισχυρός από τον Gauss-Newton, το οποίο σημαίνει ότι σε πολλές περιπτώσεις βρίσκει λύση ακόμα και αν ξεκινήσει πολύ μακριά από το τελικό ελάχιστο. Από την άλλη όμως, για λειτουργίες που συμπεριφέρονται καλά και για λογικές αρχικές παραμέτρους, τείνει να είναι λίγο πιο αργός από τον Gauss-Newton.

### Το πρόβλημα

Έστω ότι έχουμε ένα σετ εμπειρικών ζευγών δεδομένων με ανεξάρτητες και εξαρτημένες μεταβλητές  $(x_i, y_i)$ , θέλουμε να βελτιστοποιήσουμε την παράμετρο  $\beta$  του μοντέλου  $f(x, \beta)$  έτσι ώστε το άθροισμα των τετραγώνων της εξίσωσης

$$S(\beta) = \sum_{i=1}^m [y_i - f(x_i, \beta)]^2 \text{ να γίνει ελάχιστο.}$$

### Η λύση

Ο αλγόριθμος Levenberg-Marquardt είναι μια επαναληπτική διαδικασία. Για να ξεκινήσει η ελαχιστοποίηση πρέπει δοθεί μια αρχική εκτίμηση της παραμέτρου  $\beta$ . Σε πολλές περιπτώσεις μια ανενημέρωτη εκτίμηση όπως  $\beta^T = (1, 1, \dots, 1)$  λειτουργεί μια χαρά, ενώ σε άλλες περιπτώσεις ο αλγόριθμος συγκλίνει μόνο όταν η αρχική εκτίμηση είναι κοντά στην τελική λύση.

Σε κάθε επανάληψη, η παράμετρος  $\beta$  αντικαθίσταται από μια καινούργια εκτίμηση  $\beta + \delta$ . Για να προσδιοριστεί το  $\delta$ , η λειτουργία  $f(\beta + \delta)$  προσεγγίζεται από την σχέση  $f(\beta + \delta) \approx f(\beta) + J\delta$  όπου  $J$  είναι ο Ιακωβιανός πίνακας της  $f$  στο  $\beta$ . Στο ελάχιστο του αθροίσματος των τετραγώνων του  $S$ , η κλίση του  $S$  για το  $\delta$  είναι μηδέν. Διαφορίζοντας το δεξιό μέρος της εξίσωσης οδηγεί στην σχέση

$(J^T J)\delta = J^T [y - f(\beta)]$  που είναι ένα σει γραμμικών εξισώσεων που μπορεί να λυθεί για  $\delta$ .

Η συνεισφορά του Levenberg είναι η αντικατάσταση της προηγούμενης εξίσωσης με την ακόλουθη:  $(J^T J + \lambda I)\delta = J^T [y - f(\beta)]$  όπου  $I$  είναι ο ταυτοτικός πίνακας. Ο μη αρνητικός παράγοντας  $\lambda$  ρυθμίζεται σε κάθε επανάληψη. Αν η μείωση της τιμής του  $S$  είναι ραγδαία μπορεί να χρησιμοποιηθεί μια μικρότερη τιμή του  $\lambda$  για να φέρει τον αλγόριθμο πιο κοντά στον Gauss-Newton, όπου αν η μείωση της τιμής είναι μη επαρκής, τότε το  $\lambda$  μπορεί να αυξηθεί οδηγώντας τον αλγόριθμο πιο κοντά στον gradient descent. Αν είτε το υπολογισμένο  $\delta$  ή η μείωση του αθροίσματος των τετραγώνων που αντιστοιχεί στην παράμετρο  $\beta + \delta$ , βγει από τα προκαθορισμένα όρια, η επανάληψη απορρίπτεται και η τελευταία τιμή της παραμέτρου  $\beta$  θεωρείται ότι είναι η λύση.

Ο αλγόριθμος του Levenberg έχει το μειονέκτημα ότι αν η τιμή παράγοντα  $\lambda$  είναι μεγάλη, τότε το  $J^T J + \lambda I$  δεν χρησιμοποιείται. Ο Marquardt αντικατέστησε τον ταυτοτικό πίνακα  $I$  με την διαγώνιο του πίνακα Hessian  $J^T J$  και ως αποτέλεσμα προκύπτει ο αλγόριθμος Levenberg-Marquardt:

$$(J^T J + \lambda \text{diag}(J^T J))\delta = J^T [y - f(\beta)].$$

### **Επιλογή της παραμέτρου $\lambda$**

Πολλά επιχειρήματα έχουν λεχθεί για την καλύτερη επιλογή της παραμέτρου  $\lambda$ . Υπάρχουν θεωρητικά επιχειρήματα που δείχνουν γιατί μερικές επιλογές εγγυώνται τοπική σύγκλιση του αλγορίθμου. Παρόλα αυτά αυτές οι επιλογές μπορούν να δημιουργήσουν πρόβλημα στην ολική σύγκλιση του αλγορίθμου, πιο συγκεκριμένα να είναι πολύ αργή η σύγκλιση κοντά στο βέλτιστο σημείο.

Οι απόλυτες τιμές της κάθε επιλογής εξαρτώνται από το πόσο καλά βαθμονομημένο είναι το αρχικό πρόβλημα. Ο Marquardt πρότεινε να γίνεται η αρχή με μια τιμή  $\lambda_0$  και ένα παράγοντα  $\nu > 1$ . Αρχικά θέτοντας  $\lambda = \lambda_0$  και υπολογίζοντας το υπόλοιπο

άθροισμα των τετραγώνων  $S(\beta)$  μετά από ένα βήμα από το αρχικό σημείο πρώτα με παράγοντα  $\lambda = \lambda_0$  και μετά με  $\lambda/\nu$ . Αν και οι δύο αυτοί είναι χειρότεροι από το αρχικό σημείο τότε ο παράγοντας  $\lambda$  αυξάνεται με αλληπάλληλο πολλαπλασιασμό με  $\nu$  μέχρι να βρεθεί ένα καλύτερο σημείο με ένα καινούργιο παράγοντα  $\lambda\nu^k$  για κάποιο  $k$ .

Αν η χρήση του παράγοντα  $\lambda/\nu$  οδηγήσει στην μείωση του αθροίσματος των τετραγώνων τότε αυτή θεωρείται η νέα τιμή του  $\lambda$  και η διαδικασία συνεχίζεται. Αν όμως οδηγήσει σε χειρότερο αποτέλεσμα και η χρήση του  $\lambda$  οδηγεί σε καλύτερο, τότε δεν αλλάζουμε το  $\lambda$  και το βέλτιστο σημείο το παίρνουμε για με τον παράγοντα  $\lambda$ .