

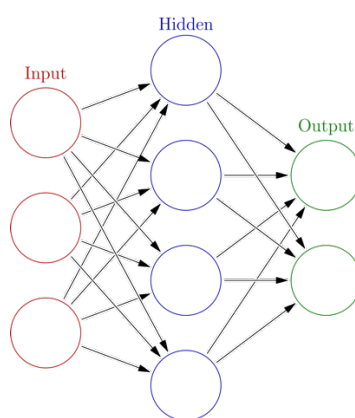


ΠΟΛΥΤΕΧΝΕΙΟ ΚΡΗΤΗΣ
ΣΧΟΛΗ ΜΗΧΑΝΙΚΩΝ ΟΡΥΚΤΩΝ ΠΟΡΩΝ

ΜΕΤΑΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ
ΓΕΩΤΕΧΝΟΛΟΓΙΑ ΚΑΙ ΠΕΡΙΒΑΛΛΟΝ
ΕΥΑΓΓΕΛΟΣ ΚΑΚΑΡΟΓΛΟΥ
A.M. 2016029001

ΘΕΜΑ:

«Ανάπτυξη μοντέλων τεχνητών νευρωνικών δικτύων για τον προσδιορισμό του χρόνου διαδρομής σεισμικών κυμάτων. Εφαρμογή σε δεδομένα σεισμικής τομογραφίας από το σταθμό Ανθούπολης του Αττικού Μετρώ.»



ΤΡΙΜΕΛΗΣ ΕΠΙΤΡΟΠΗ:

Επιβλέπων: **ΒΑΦΕΙΔΗΣ ΑΝΤΩΝΗΣ** (Καθηγητή Πολυτεχνείου Κρήτης)

Μέλη: **ΓΑΛΕΤΑΚΗΣ ΜΙΧΑΗΛ** (Καθηγητή Πολυτεχνείου Κρήτης)

ΣΟΥΠΙΟΣ ΠΑΝΤΕΛΗΣ (Καθηγητή King Fahd University of
Petroleum and Minerals, Saudi Arabia)

ΠΕΡΙΛΗΨΗ

Ο προσδιορισμός του χρόνου της πρώτης άφιξης σε σεισμικά δεδομένα είναι απαραίτητος για την εκτέλεση υπολογισμών στατικής διόρθωσης ή υπολογισμούς τομογραφίας καθοδικών κυμάτων (diving wave tomography). Η επιλογή των χρόνων πρώτων αφίξεων είναι ένα πρόβλημα αναγνώρισης μοτίβων και ως εκ τούτου συνεπάγεται σημαντική ανθρώπινη προσπάθεια.

Επιπλέον, καθώς αυξάνεται ο αριθμός των σεισμικών αισθητήρων, γίνεται όλο και πιο δύσκολο για τους αναλυτές να επιλέξουν σεισμικές φάσεις τόσο χειροκίνητα όσο και αναλυτικά, παρόλα αυτά οι προσπάθειες αυτές είναι θεμελιώδεις.

Παρά τις βελτιώσεις που έγιναν τα τελευταία χρόνια με τις μεθόδους αυτόματης επιλογής πρώτων αφίξεων, εξακολουθεί να είναι πολύ δύσκολο να προσομοιαστούν τα αποτελέσματα που προκύπτουν από έμπειρους αναλυτές, με ένα πιο λεπτό ζήτημα να είναι το γεγονός ότι διάφοροι γεωφυσικοί αναλυτές επιλέγουν πρώτες αφίξεις με διαφορετικούς τρόπους, κάτι που μπορεί να εισάγει μεροληψία (bias) στα δεδομένα. Από τα παραπάνω γίνεται φανερό ότι πρέπει να εισαχθεί ένας πιο αξιόπιστος και ακριβής τρόπος αυτόματης επιλογής χρόνου διαδρομής.

Με τη χρήση των νευρωνικών δικτύων που γίνονται όλο και πιο δημοφιλή στις γεωφυσικές εφαρμογές, λόγω των δυνατοτήτων τους ως καθολικών προσεγγιστών (global approximators), είναι δηλαδή εργαλεία που μπορούν να προσεγγίσουν οποιαδήποτε συνεχή συνάρτηση με ακρίβεια και ως εκ τούτου καθιστούν έναν πολύ ελπιδοφόρο τρόπο αυτόματης επιλογής πρώτων αφίξεων επεξεργασίας σεισμικών δεδομένων.

Σε αυτή την εργασία, εισάγονται και επεξεργάζονται σεισμικά δεδομένα και στη συνέχεια εκπαιδεύονται τεχνητά νευρωνικά δίκτυα. Επιλέγεται το δίκτυο με τις καλύτερες επιδόσεις και χρησιμοποιείται για την αυτόματη επιλογή των συμβάντων πρώτης άφιξης σε σεισμικά δεδομένα. Χρησιμοποιώντας την έξοδο του νευρωνικού δικτύου έναντι της εξόδου από έναν έμπειρο χρήστη, προκύπτουν χρήσιμα συμπεράσματα.

Έχοντας ένα μεγαλύτερο σύνολο δεδομένων εισόδου βελτιώνεται η απόδοση του δικτύου και ο αριθμός των κόμβων συμβάλλει θετικά στα αποτελέσματα, αλλά μετά από ένα συγκεκριμένο σημείο δεν υπάρχει περαιτέρω βελτίωση.

ABSTRACT

The time of the first-arrival in seismic data, is essential for performing refraction statics computations or diving-wave tomography calculations. Picking first-arrival times, is a pattern recognition problem and, as such, involves a substantial amount of human effort. Furthermore, with the number of seismic sensors growing, it is becoming increasingly difficult for analysts to pick seismic phases both, manually and comprehensively, yet such efforts are fundamental. Despite the improvements made in recent years with the automatic phase picking methods, it is still very difficult to match the performance of experienced analysts, with a subtler issue being the fact that, different seismic analysts may pick phases in different ways, which most of the times can introduce bias in the data. From the above it becomes obvious that a more reliable and accurate way of automated phase picking needs to be introduced.

With the use of neural networks becoming increasingly popular in geophysical applications, because of their abilities as universal approximators, they are tools that can approximate any continuous function with an arbitrary precision, and as such they make for a very promising way to automatically pick first break events and edit seismic trace data.

In this thesis, a number of seismic data is imported, edited and then a number of neural networks is trained. The best performing one is chosen and used to automatically pick first break events on seismic data. Using the neural network's output against the output from an experienced user, useful conclusions are derived.

Having a larger input data set improves the output of the network and the number of nodes positively contributes to the output, but after a certain point there is no improvement.

Ευχαριστίες:

Θα ήθελα να ευχαριστήσω ιδιαίτερα τον επιβλέποντα μου Αντώνη Βαφείδη για την εμπιστοσύνη του στην ανάθεση του θέματος και την στήριξη του σε καθ' όλη την διάρκεια του μεταπτυχιακού. Τον Δρ. Γιώργο Κρητικάκη του τμήματος μηχανικών ορυκτών πόρων του Πολυτεχνείου Κρήτης για τις πολύτιμες συμβουλές του στο κομμάτι της επεξεργασίας των δεδομένων και της μεθόδου της σεισμικής τομογραφίας. Τον Δρ. Ιωάννη Τριχάκη του τμήματος μηχανικών περιβάλλοντος του Πολυτεχνείου Κρήτης για την συνδρομή του στο αντικείμενο των νευρωνικών δικτύων και για την βοήθεια του που ήταν ανεκτίμητη για την περάτωση της εργασίας αυτής. Την οικογένεια μου για την στήριξη και την βοήθεια τους.

Τέλος την αγαπημένη μου Κατερίνα Βιρβιδάκη που ήταν και είναι πάντα δίπλα μου.

ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

ΠΕΡΙΛΗΨΗ	2
ABSTRACT	3
1 Η ΜΕΘΟΔΟΣ ΤΗΣ ΣΕΙΣΜΙΚΗΣ ΤΟΜΟΓΡΑΦΙΑΣ.....	8
1.1 ΑΡΧΗ ΤΗΣ ΣΕΙΣΜΙΚΗΣ ΤΟΜΟΓΡΑΦΙΑΣ	8
1.2 ΕΦΑΡΜΟΓΕΣ ΤΗΣ ΜΕΘΟΔΟΥ	9
1.3 ΤΟΜΟΓΡΑΦΙΑ ΠΕΡΙΘΛΑΣΗΣ ΚΑΙ ΤΟΜΟΓΡΑΦΙΑ ΣΕΙΣΜΙΚΗΣ ΑΚΤΙΝΑΣ.....	10
1.4 ΤΟΜΟΓΡΑΦΙΑ ΣΕΙΣΜΙΚΗΣ ΑΚΤΙΝΑΣ (SEISMIC RAY TOMOGRAPHY)	11
1.4.1 ΕΙΣΑΓΩΓΗ	11
1.4.2 ΜΕΘΟΔΟΙ ΜΕΤΑΣΧΗΜΑΤΙΣΜΟΥ	11
1.4.2.2 ΤΟΜΟΓΡΑΦΙΑ ΑΚΤΙΝΩΝ ΜΕ ΕΥΘΥ ΜΕΤΑΣΧΗΜΑΤΙΣΜΟ	15
1.4.2.3 ΤΟΜΟΓΡΑΦΙΑ ΑΚΤΙΝΑΣ ΜΕ ΟΠΙΣΘΟΠΡΟΒΟΛΗ	18
1.4.3 ΜΕΘΟΔΟΙ ΑΝΑΠΤΥΞΗΣ ΣΕΙΡΩΝ	19
1.4.3.1 ΤΟ ΠΡΟΒΛΗΜΑ ΤΗΣ ΜΟΝΤΕΛΟΠΟΙΗΣΗΣ	19
1.4.3.2 Η ΜΕΘΟΔΟΣ KACZMARZ	22
2. ΤΕΧΝΗΤΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ.....	27
2.1 ΕΙΣΑΓΩΓΗ	27
2.1.1 ΤΙ ΟΝΟΜΑΖΕΤΑΙ ΝΕΥΡΩΝΙΚΟ ΔΙΚΤΥΟ	27
2.1.2 ΜΟΝΑΔΕΣ ΣΤΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ	28
2.2 ΔΟΜΕΣ ΔΙΚΤΥΟΥ	29
2.2.1 ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ ΕΝΟΣ ΕΠΙΠΕΔΟΥ ΜΕ ΠΡΟΣ ΣΤΑ ΕΜΠΡΟΣ ΤΡΟΦΟΔΟΤΗΣΗ ΤΟΥ ΣΗΜΑΤΟΣ (PERCEPTONS)	30
2.2.2 ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ ΠΟΛΛΩΝ ΕΠΙΠΕΔΩΝ ΜΕ ΠΡΟΣ ΤΑ ΕΜΠΡΟΣ ΤΡΟΦΟΤΗΣΗ ΤΟΥ ΣΗΜΑΤΟΣ.....	33
2.3 ΜΑΘΗΣΗ ΔΟΜΩΝ ΝΕΥΡΩΝΙΚΩΝ.....	36
2.4 Ο ΑΛΓΟΡΙΘΜΟΣ ΕΚΠΑΙΔΕΥΣΗΣ LEVENBERG-MARQUARDT	37
2.4.1 ΠΡΟΕΛΕΥΣΗ.....	37
2.4.2 Ο ΑΛΓΟΡΙΘΜΟΣ	38
2.4.3 ΕΦΑΡΜΟΓΗ ΤΟΥ ΑΛΓΟΡΙΘΜΟΥ	38
2.4.4 ΥΠΟΛΟΓΙΣΜΟΣ ΤΟΥ ΙΑΚΩΒΙΑΝΟΥ ΠΙΝΑΚΑ	38
2.5 ΣΧΕΔΙΑΣΜΟΣ ΤΗΣ ΔΙΑΔΙΚΑΣΙΑΣ ΕΚΠΑΙΔΕΥΣΗΣ	44
2.6 ΣΥΝΟΨΗ ΓΙΑ ΤΟΝ ΑΛΓΟΡΙΘΜΟ ΕΚΠΑΙΔΕΥΣΗΣ LEVENBERG-MARQUARDT.....	45
ΚΕΦΑΛΑΙΟ 3: ΕΦΑΡΜΟΓΗ ΤΗΣ ΜΕΘΟΔΟΛΟΓΙΑΣ, ΓΙΑ ΤΗΝ ΑΥΤΟΜΑΤΗ ΕΠΙΛΟΓΗ ΠΡΩΤΩΝ ΑΦΙΞΕΩΝ.....	46
3.1 ΕΙΣΑΓΩΓΗ	46
3.2 ΠΡΟΕΛΕΥΣΗ ΤΩΝ ΔΕΔΟΜΕΝΩΝ	46
3.3 ΔΙΑΔΙΚΑΣΙΑ ΜΕΤΑΤΡΟΠΗΣ ΤΩΝ ΔΕΔΟΜΕΝΩΝ.....	46
3.4 ΔΗΜΙΟΥΡΓΙΑ ΤΟΥ ΠΙΝΑΚΑ ΤΩΝ ΔΕΔΟΜΕΝΩΝ ΕΚΠΑΙΔΕΥΣΗΣ (TRAINING DATA SET).....	49
3.5 ΕΠΙΛΟΓΗ ΤΩΝ ΠΡΩΤΩΝ ΑΦΙΞΕΩΝ ΚΑΙ ΔΗΜΙΟΥΡΓΙΑ ΤΟΥ ΣΥΝΟΛΟΥ ΣΤΟΧΩΝ (TARGETS).....	50
3.6 ΔΗΜΙΟΥΡΓΙΑ ΚΑΙ ΕΚΠΑΙΔΕΥΣΗ ΤΩΝ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ	52
3.7 ΣΥΝΟΨΗ	54
4: ΠΑΡΟΥΣΙΑΣΗ ΤΩΝ ΑΠΟΤΕΛΕΣΜΑΤΩΝ	54
4.1 ΕΙΣΑΓΩΓΗ	54
4.1.2 ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΔΙΚΤΥΩΝ.....	54

4.2.1 ΔΕΙΓΜΑ 1, 8 ΚΟΜΒΟΙ	55
4.2.2 ΔΕΙΓΜΑ 1, 9 ΚΟΜΒΟΙ	56
4.2.3 ΔΕΙΓΜΑ 1, 10 ΚΟΜΒΟΙ	58
4.2.4 ΔΕΙΓΜΑ 1, 11 ΚΟΜΒΟΙ	59
4.2.5 ΔΕΙΓΜΑ 2, 10 ΚΟΜΒΟΙ	60
4.2.6 ΔΕΙΓΜΑ 2, 11 ΚΟΜΒΟΙ	62
4.2.7 ΔΕΙΓΜΑ 2, 12 ΚΟΜΒΟΙ	63
4.2.8 ΔΕΙΓΜΑ 3, 8 ΚΟΜΒΟΙ	65
4.2.8 ΔΕΙΓΜΑ 3, 9 ΚΟΜΒΟΙ	67
4.2.8 ΔΕΙΓΜΑ 3, ΚΟΜΒΟΙ 10	68
4.2.8 ΔΕΙΓΜΑ 3, 11 ΚΟΜΒΟΙ	70
4.2.8 ΔΕΙΓΜΑ 3, 12 ΚΟΜΒΟΙ	71
4.2.13 ΔΕΙΓΜΑ 4, 8 ΚΟΜΒΟΙ	73
4.2.14 ΔΕΙΓΜΑ 4, 9 ΚΟΜΒΟΙ	74
4.2.15 ΔΕΙΓΜΑ 4, 10 ΚΟΜΒΟΙ	76
4.2.16 ΔΕΙΓΜΑ 4, 11 ΚΟΜΒΟΙ	77
4.2.17 ΔΕΙΓΜΑ 4, 20 ΚΟΜΒΟΙ	79
4.3 ΣΥΓΚΡΙΣΗ ΑΠΟΤΕΛΕΣΜΑΤΩΝ ΑΠΟ ΝΕΥΡΩΝΙΚΟ ΔΙΚΤΥΟ ΜΕ ΑΠΟΤΕΛΕΣΜΑΤΑ ΑΠΟ ΕΞΕΙΔΙΚΕΥΜΕΝΟ ΧΡΗΣΤΗ.....	80
4.3.1 ΔΙΑΔΙΚΑΣΙΑ ΣΥΓΚΡΙΣΗΣ.....	80
5 ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΜΕΛΛΟΝΤΙΚΕΣ ΠΡΟΤΑΣΕΙΣ	81
5.1 ΕΙΣΑΓΩΓΗ	81
5.2 ΣΥΜΠΕΡΑΣΜΑΤΑ	81
5.3 ΠΡΟΒΛΗΜΑΤΑ	82
5.4 ΜΕΛΛΟΝΤΙΚΕΣ ΠΡΟΤΑΣΕΙΣ	82
6 ΚΩΔΙΚΑΣ ΜΑΤΛΑΒ ΓΙΑ ΤΗΝ ΔΗΜΙΟΥΡΓΙΑ ΝΕΥΡΩΝΙΚΟΥ ΔΙΚΤΥΟΥ	83
7 ΠΑΡΑΡΤΗΜΑ Ι-ΑΛΓΟΡΙΘΜΟΙ: STEEPEST DESCENT, NEWTON, GAUSS- NEWTON	85
7.1 ΑΛΓΟΡΙΘΜΟΣ STEEPEST DESCENT.....	85
7.2 ΜΕΘΟΔΟΣ NEWTON.....	85
7.3 ΑΛΓΟΡΙΘΜΟΣ GAUSS-NEWTON.....	88
8 ΒΙΒΛΙΟΓΡΑΦΙΑ.....	89

1 Η ΜΕΘΟΔΟΣ ΤΗΣ ΣΕΙΣΜΙΚΗΣ ΤΟΜΟΓΡΑΦΙΑΣ

Στο κεφάλαιο αυτό θα γίνει μια συνοπτική αναφορά στην μέθοδο της σεισμικής τομογραφίας, η οποία ήταν και η κύρια μέθοδος που χρησιμοποιήθηκε για την απόκτηση των δεδομένων της παρούσας εργασίας.

1.1 ΑΡΧΗ ΤΗΣ ΣΕΙΣΜΙΚΗΣ ΤΟΜΟΓΡΑΦΙΑΣ

Ως τομογραφία ορίζεται η τεχνική απεικόνισης, η οποία παράγει την εικόνα της διατομής (τομόγραμμα) ενός αντικειμένου, με βάση την απόκριση του αντικείμενου αυτού στην μη-καταστροφική διερευνητική ενέργεια μιας εξωτερικής πηγής. (Lo)

Η σεισμική τομογραφία χρησιμοποιεί πηγές που παράγουν σεισμικά κύματα τα οποία διερευνούν ένα γεωλογικό στόχο ενδιαφέροντος.

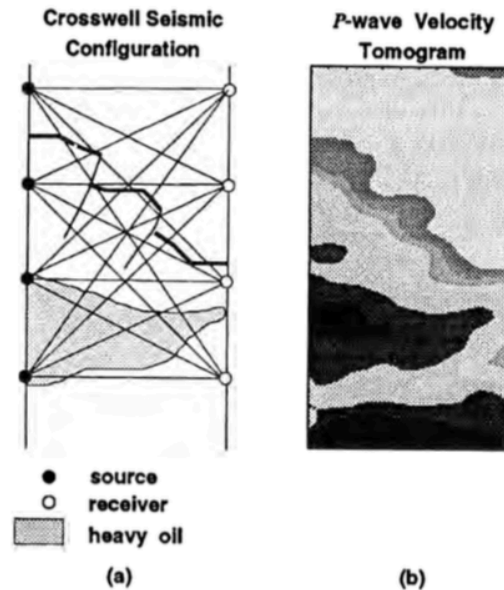
Η εικόνα 1.1(α) αποτελεί ένα παράδειγμα διάταξης για σεισμική τομογραφία σε γεώτρηση. Μια σεισμική πηγή τοποθετείται σε γεώτρηση και ένα σύστημα δέκτη τοποθετείται σε διπλανή γεώτρηση. Σεισμικά κύματα που παράγονται από την πηγή διερευνούν έναν στόχο που, στην προκειμένη περίπτωση είναι ένα ταμιευτήρας βαρέως πετρελαίου που βρίσκεται μεταξύ των δυο γεωτρήσεων.

Η απόκριση του ταμιευτήρα στην σεισμική ενέργεια καταγράφεται από αισθητήρες που τοποθετούνται σε διαφορετικά βάθη στη γεώτρηση (receiver well). Ο ταμιευτήρας μελετάται σε πολλές κατευθύνσεις με την καταγραφή της σεισμικής ενεργείας από την ίδια διάταξη δέκτη για διαφορετικές θέσεις πηγής.

Με τον τρόπο αυτό, δημιουργείται ένα δίκτυο διαδρομών σεισμικών ακτινών που ταξιδεύουν μέσα στον ταμιευτήρα. Η μετρούμενη απόκριση του ταμιευτήρα στο σεισμικό κύμα ονομάζεται, δεδομένο προβολής (projection data). (Lo)

Οι μέθοδοι τομογραφικής ανακατασκευής εικόνας λειτουργούν πάνω στα δεδομένα προβολής για να δημιουργήσουν ένα τομόγραμμα, όπως αυτό της εικόνα 1.1(β). Στην περίπτωση αυτή χρησιμοποιήθηκαν δεδομένα προβολής από χρόνους διαδρομής απευθείας κυμάτων (direct arrivals travel time) και τομογραφία σεισμικής ακτίνας για την κατασκευή τομογράμματος ταχύτητας P-κυμάτων.

Γενικά, διαφορετικά χρώματα ή αποχρώσεις του γκρι σε ένα τομόγραμμα παριστάνουν λιθολογικές δομές με διαφορετικά χαρακτηριστικά. Οι υψηλές τιμές ταχύτητας P κυμάτων (σκούρο γκρι/ μαύρο) στο τομόγραμμα της εικόνας 1.1(β) σχετίζονται με ταμιευτήρα με υψηλό βαθμό κορεσμού κορεσμό σε πετρέλαιο.



Εικόνα 1: 1 (α)Γεωμετρία για σεισμική τομογραφία σε πηγάδι γεώτρησης. Η ενέργεια των P κυμάτων ταξιδεύει στις ακτίνες διαδρομής της ανίχνευσης του γεωλογικού στόχου. (β)Τομόγραμμα ταχύτητας P κυμάτων όπως αυτό κατασκευάστηκε από δεδομένα παρατηρούμενων χρόνων διαδρομής. Διαφορετικές αποχρώσεις του γκριζού αντιστοιχούν σε διαφορετικές τιμές ταχύτητας (L_o).

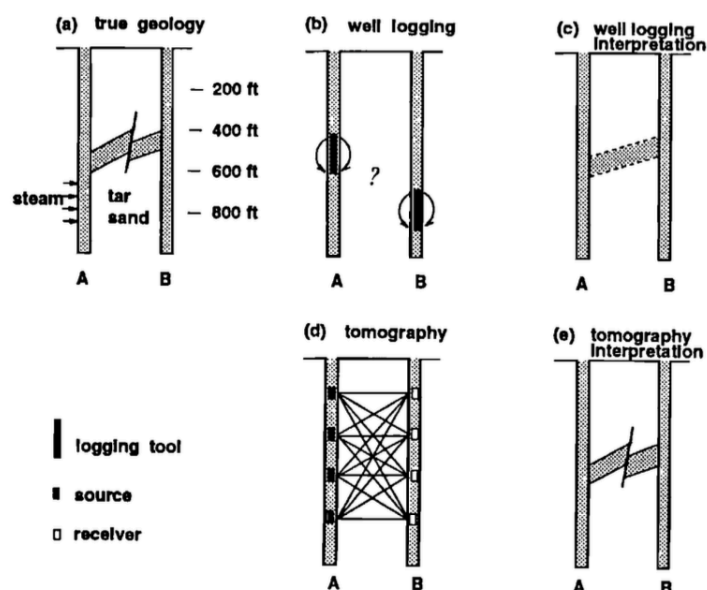
1.2 ΕΦΑΡΜΟΓΕΣ ΤΗΣ ΜΕΘΟΔΟΥ

Η σεισμική τομογραφία βρίσκει εφαρμογή σε ένα μεγάλο εύρος προβλημάτων στην βιομηχανία πετρελαίου, που κυμαίνονται από την εξερεύνηση μέχρι την ανάπτυξη και την παραγωγή. Πέρα από την βιομηχανία πετρελαίου, η εφαρμογή της μεθόδου της τομογραφίας σε επιφανειακά σεισμικά δεδομένα μπορεί να παράγει μοντέλα ταχύτητας χρήσιμα σε προβλήματα εξερεύνησης. (L_o)

Η σεισμική τομογραφία που χρησιμοποιείται σε προβλήματα ανάπτυξης και παραγωγής γενικά εφαρμόζεται με μια διάταξη μεταξύ γεωτρήσεων (crosswell configuration) όπως παρουσιάζεται στην εικόνα 1.1(α).

Η εικόνα 1.2 παρουσιάζει τα πλεονεκτήματα της σεισμικής τομογραφίας για την μελέτη ταμιευτήρων σε σχέση με συμβατικές μεθόδους.

Η εικόνα 1.2(α) παριστάνει την πραγματική γεωλογία μεταξύ δυο γεωτρήσεων σε ένα πεδίο παραγωγής, όπου ο παραγωγικός σχηματισμός είναι ένα στρώμα πίσσας με άμμο που βρίσκεται κάτω από ένα λεπτότερο, λιγότερο διαπερατό στρώμα.



Εικόνα 1: 2(α) Γεωλογικό μοντέλο (β) οι διαγραφές δειγματίζουν μόνο μια μικρή απόσταση μέσα στον ταμιευτήρα απαιτώντας κάποιου είδους παρεμβολής μεταξύ των γεωτρήσεων όπως φαίνεται στην εικόνα (γ). (δ) σεισμικές καταγραφές της απόκρισης της γης στην σεισμική ενέργεια μεταξύ των γεωτρήσεων, που επιτρέπει την ανακατασκευή του γεωλογικού μοντέλου όπως φαίνεται στο (ε).

1.3 ΤΟΜΟΓΡΑΦΙΑ ΠΕΡΙΘΛΑΣΗΣ ΚΑΙ ΤΟΜΟΓΡΑΦΙΑ ΣΕΙΣΜΙΚΗΣ ΑΚΤΙΝΑΣ

Για την εφαρμογή της μεθόδου της σεισμικής τομογραφίας πρέπει πρώτα να μοντελοποιηθεί το σεισμικό κύμα. Τόσο το θεωρητικό μοντέλο της τομογραφίας περίθλασης όσο και αυτό της τομογραφίας ακτίνας είναι διαθέσιμα για αυτό το σκοπό. Η επιλογή του μοντέλου εξαρτάται από τα σχετικά μεγέθη του μήκους κύματος και του στόχου που πρόκειται να εξεταστεί. Μια σωστή επιλογή του θεωρητικού μοντέλου για ένα δεδομένο σεισμικό κύμα είναι απαραίτητη για την επιτυχή εφαρμογή της μεθόδου της σεισμικής τομογραφίας.

Εάν το μέγεθος του στόχου είναι κατά πολύ μεγαλύτερο από το μήκος κύματος, τότε η διάδοσή του κύματος στο μέσο μπορεί να μοντελοποιηθεί με την χρήση της ακτινικής θεωρίας (ray theory), η οποία, περιγράφει τα κύματα με ακτίνες. Όταν το μέγεθος του στόχου είναι συγκρίσιμο με αυτό του μήκους του σεισμικού κύματος, τότε η μοντελοποίηση της διάδοσης του σεισμικού κύματος χρησιμοποιώντας την θεωρία της τομογραφίας περίθλασης. Τέτοιου είδους στόχοι διασκορπούν το σεισμικό κύμα σε πολλές κατευθύνσεις και μόνο η θεωρία περίθλασης μπορεί μοντελοποιήσει σωστά το φαινόμενο. (Lo)

1.4 ΤΟΜΟΓΡΑΦΙΑ ΣΕΙΣΜΙΚΗΣ ΑΚΤΙΝΑΣ (Seismic Ray Tomography)

1.4.1 ΕΙΣΑΓΩΓΗ

Στην εφαρμογή αυτής της μεθόδου γίνεται η υπόθεση ότι η σεισμική πηγή παράγει κύματα με μήκη κύματος πολύ μικρότερα από το μέγεθος των ανομοιογενειών που συναντιούνται στο μέσο. Μόνο όταν ικανοποιείται αυτή η υπόθεση η διάδοση των σεισμικών κυμάτων μπορεί να μοντελοποιηθεί με την χρήση ακτίνων. (Lo)

Υπάρχουν δυο διαφορετικές ομάδες τεχνικών για την ανακατασκευή εικόνων στην σεισμική τομογραφία. Η πρώτη ομάδα είναι αυτή των μεθόδων μετασχηματισμού (transform methods), οι εφαρμογές των μεθόδων αυτών έχουν τις ρίζες τους στην αστρονομία και την ιατρική. Είναι πολύ περιοριστικές όσον αφορά το πρόβλημα της σεισμικής απεικόνισης, καθώς γενικά οι διαδρομές διάδοσης των ακτίνων θεωρούνται ευθείες και είναι πλήρους σάρωσης (full scan aperture).

Παρόλα αυτά, οι μέθοδοι μετασχηματισμού κάνουν μια πολύ καλή εισαγωγή στις αρχές της σεισμικής τομογραφίας λόγω της απλότητάς τους και χρησιμεύουν σαν γέφυρα μεταξύ των εφαρμογών της τομογραφίας σε άλλα αντικείμενα και των εφαρμογών της σεισμικής τομογραφίας. Οι μέθοδοι ανάπτυξης σειρών (series expansion methods) είναι η δεύτερη κατηγορία μεθόδων ανακατασκευής εικόνας και αποτελούν τις πιο διαδομένες στην σεισμική τομογραφία.

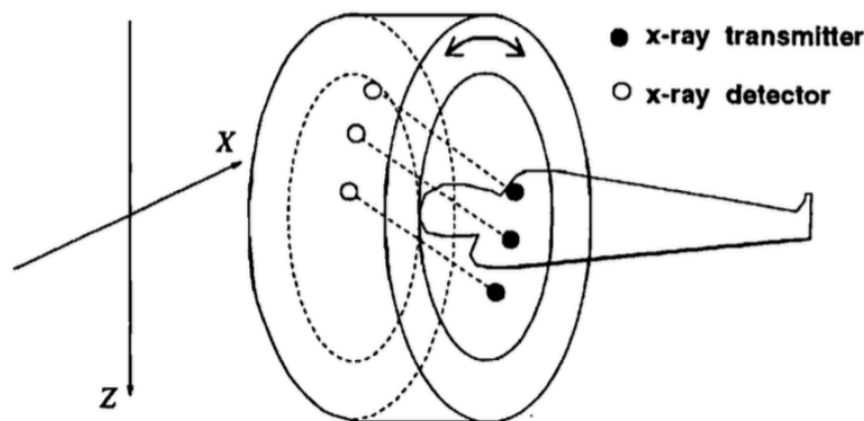
1.4.2 ΜΕΘΟΔΟΙ ΜΕΤΑΣΧΗΜΑΤΙΣΜΟΥ

Στο σημείο αυτό, παρουσιάζεται το θεώρημα προβολής (projection slice theorem) καθώς παρέχει το θεωρητικό υπόβαθρο για τις μεθόδους μετασχηματισμού.

Στην συνέχεια, από το θεώρημα τμήματος προβολής, προκύπτουν οι δυο μέθοδοι για την τομογραφία: η τομογραφία απευθείας ακτίνας και η τομογραφία οπισθοδιάδοσης ακτίνας.

1.4.2.1 ΘΕΩΡΗΜΑ ΠΡΟΒΟΛΗΣ

Η αρχή του θεωρήματος προβολής περιγράφεται από τυπικό πείραμα ιατρικής τομογραφίας και παρουσιάζεται στην εικόνα 1.3. Μια συνδεσμολογία πομπού-δέκτη ακτίνων X σε σχήμα ντόνατ περικυκλώνει τον στόχο, σε αυτήν την περίπτωση, το κεφάλι ενός ανθρώπου.



Εικόνα 1: 3 Γεωμετρία ιατρικού πειράματος τομογραφίας. Οι σαρώσεις των ακτίνων X λαμβάνονται από διαφορετικά σημεία του κεφαλιού του ατόμου με την περιστροφή της συνδεσμολογίας πομπού-δέκτη.

Η ένταση των ακτίνων X μετριέται για έναν συγκεκριμένο προσανατολισμό της συνδεσμολογίας. Στην συνέχεια η συνδεσμολογία περιστρέφεται γύρω από το άτομο έτσι ώστε οι ακτίνες X να περνούν μέσα από το κεφάλι με διαφορετική διεύθυνση. Το πείραμα ολοκληρώνεται όταν το κεφάλι του ατόμου έχει σαρωθεί από όλες τις κατευθύνσεις. Ο στόχος των μεθόδων μετασχηματισμού είναι να χρησιμοποιούν πληροφορίες εξασθένησης (attenuation information) από τις μετρημένες τιμές της έντασης των ακτίνων X για να ανακατασκευάσουν μια εικόνα της διατομής του κεφαλιού του ατόμου στο επίπεδο που περιλαμβάνει πομπό και δέκτη. Έτσι ένας εγκεφαλικός όγκος που εξασθενεί τις ακτίνες X διαφορετικά από ότι ο φυσιολογικός ιστός μπορεί να γίνει ορατός.

Η εικόνα 1.4 απεικονίζει μια διατομή του κεφαλιού. Η μεταβλητή αντίθεση μέσα στον στόχο αντιπροσωπεύει την ανομοιομορφή εξασθένηση των ακτίνων X που σχετίζεται με όγκο, κανονικό ιστό και το κρανίο.

Για τον σκοπό της παραγωγής του θεωρήματος προβολής, ορίζεται η συνάρτηση μοντέλου (model function) $M(x,z)$ ως η χωρική κατανομή της εξασθένησης. Γενικά η συνάρτηση μοντέλου αντιπροσωπεύει την άγνωστη κατανομή στο χώρο κάποιας φυσικής ιδιότητας του μέσου, η οποία επιδρά στην ενέργεια που διαδίδεται με κάποιον παρατηρήσιμο τρόπο.

Μια τυπική συνάρτηση μοντέλου στην σεισμική τομογραφία περιγράφει ταχύτητα ή επιβράδυνση του κύματος συμπίεσης (reciprocal compressional wave velocity or slowness), η οποία έχει άμεση επίδραση στον παρατηρούμενο χρόνο διαδρομής της διαδιδόμενης ενέργειας.

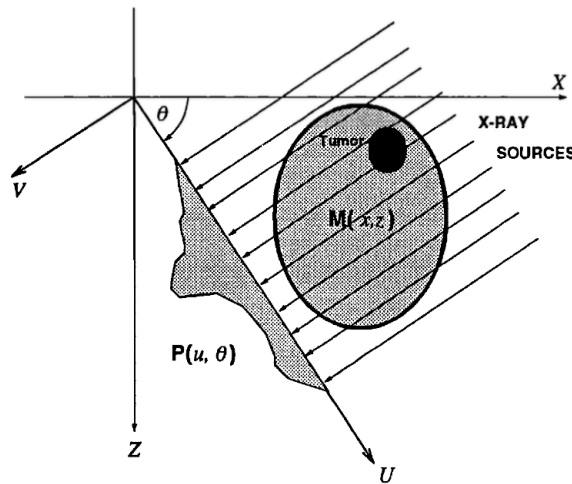
Το θεώρημα προβολής απαιτεί η παρατήρηση της διαδιδόμενης ενέργειας να λαμβάνεται κατά μήκος δοσμένης προβολής η οποία είναι κάθετη στην διαδρομή των ακτίνων. Η εικόνα 1.4 παρουσιάζει μια μοναδική προβολή στο πείραμα της ιατρικής τομογραφίας. Οι ακτίνες X που εκπέμπονται από τον πομπό ταξιδεύουν κατά μήκος των παράλληλων ακτίνων και καταγράφονται από ανιχνευτές τοποθετημένους κατά μήκος του u άξονα. Το περιστράμενο σύστημα χωρικών συντεταγμένων (u,v) εισάγεται για να περιγράψει όλους τους πιθανούς προσανατολισμούς για την συνδεσμολογία πομπού-δέκτη περί του στόχου. Ο άξονας v ορίζεται ως παράλληλος στην κατεύθυνση διάδοσης των ακτίνων X και ο άξονας u ορίζεται ως κάθετος στον v άξονα στην διεύθυνση κατά την οποία μετριέται η ένταση των ακτίνων X. Αν το σύστημα συντεταγμένων $u-v$ έχει την ίδια αρχή με το $x-z$, τότε η σχέση μεταξύ των δυο όταν το ένα περιστρέφεται με γωνία θ σε σχέση με το άλλο είναι:

$$\begin{bmatrix} x \\ z \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}. \quad (1.1)$$

Για δεδομένη ακτίνα στην εικόνα 4, το δεκαδικό ποσοστό πτώσης σε ένταση των ακτίνων X, ορίζεται ως $P(u, \theta)$,

$$P(u, \theta) = [I_0 - I(u, \theta)]/I_0,$$

Όπου $I(u, \theta)$, είναι η ένταση που μετρήθηκε από τον ανιχνευτή στο (u, θ) και I_0 είναι η ένταση των ακτίνων X στον πομπό.



Εικόνα 1: 4 Διατομή του κεφαλιού ενός ατόμου όπου οι διαφορετικές αντιθέσεις αντιπροσωπεύουν ανομοιόμορφες εξασθενίσεις των ακτίνων X. Η προβολή $P(u, \theta)$ είναι το δεκαδικό ποσοστό πτώσης σε ένταση των ακτίνων X μετρημένο κατά μήκος του περιστραμένου άξονα συντεταγμένων u . Ο v άξονας είναι κάθετος στον u που είναι πάντα παράλληλος με την διεύθυνση διάδοσης των ακτίνων X. Η συνάρτηση μοντέλου $M(x, z)$ δίνει μια αριθμητική τιμή για την εξασθένηση και είναι ένας άγνωστος που πρέπει να προσδιοριστεί από την παρατηρούμενη προβολή $P(u, \theta)$. (Lo)

Το $P(u, \theta)$ ορίζεται ως συνάρτηση δεδομένων (data function). Εάν η εξασθένηση είναι μικρή στην διαδρομή της ακτίνας τότε η συνάρτηση δεδομένων $P(u, \theta)$ συνδέεται γραμμικά με την εξασθένηση $M(x, z)$ σαν το γραμμικό ολοκλήρωμα (line integral):

$$P(u, \theta) = \int_{ray} M(x, z) dv, \quad (1.2)$$

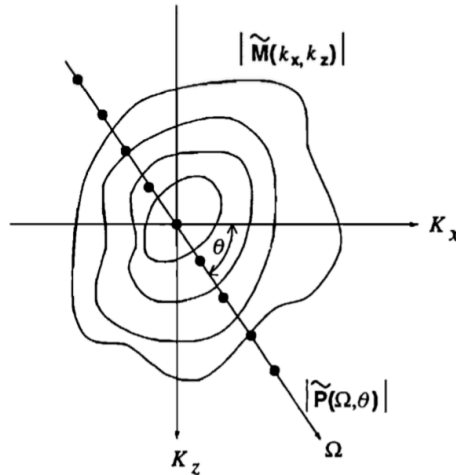
όπως αυτό λαμβάνεται για την διαδρομή της ακτίνας. Σημειώνεται ότι κάθε παρατήρηση της συνάρτησης δεδομένων $P(u, \theta)$ αποτελεί λύση στην εξίσωσης (1.2) κατά μήκος της δεδομένης διαδρομής ακτίνας χωρίς να υπάρχει γνώση της συνάρτησης μοντέλου $M(x, z)$. Σε περίπτωση που υπάρχει αμελητέα εξασθένηση των ακτίνων X εκτός του στόχου, η εξίσωση (1.2), δίνει την ίδια μετρημένη προβολή για οποιοδήποτε απόσταση ζεύγους πομπού-δέκτη όσο κάθε πομπός και δέκτης παραμένει εκτός του στόχου. Η υπόθεση αυτή χρησιμοποιείται για να ξαναγραφτεί η εξίσωση (1.2) με άπειρα όρια:

$$P(u, \theta) = \int_{-\infty}^{+\infty} M(x, z) dv. \quad (1.3)$$

Για την λήψη μιας πιο απλής και με περισσότερο νόημα σχέσης μεταξύ της συνάρτησης μοντέλου $M(x, z)$ και της συνάρτησης δεδομένων $P(u, \theta)$, αυτές μετασχηματίζονται στο πεδίο Fourier. Ο μετασχηματισμός Fourier σε δυο διαστάσεις της συνάρτησης μοντέλου $M(x, z)$ είναι:

$$\tilde{M}(k_x, k_z) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} M(x, z) e^{-j(k_x x + k_z z)} dx dz, \quad (1.4)$$

όπου k_x και k_z είναι χωρικές συχνότητες κατά μήκος του άξονα x και z αντίστοιχα. Η χωρική συχνότητα, ή, κυματαριθμός, ορίζεται ως $k=2\pi/\lambda$ όπου λ , το μήκος κύματος. Η εικόνα 1.5 αντιπροσωπεύει το εύρος του πλάτους του δισδιάστατου μετασχηματισμού Fourier μιας υποθετικής συνάρτησης μοντέλου $M(x, z)$.



Εικόνα 1: 5 Θεώρημα προβολής στο χώρο Fourier. Ο δισδιάστατος μετασχηματισμός Fourier μιας υποθετικής συνάρτησης μοντέλου $M(x, z)$ υπολογίζει το φάσμα πλάτους (amplitude spectrum) $[\tilde{M}(k_x, k_z)]$. Το φάσμα πλάτους στο επίπεδο k_x - k_z παριστάνεται με ισότιμες γραμμές. Το φάσμα πλάτους $|\tilde{P}(\Omega, \theta)|$ από τον μονοδιάστατο μετασχηματισμό Fourier της συνάρτησης δεδομένων $P(u, \theta)$ αντιπροσωπεύει ένα τμήμα του $[\tilde{M}(k_x, k_z)]$ κατά τον Ω άξονα στο επίπεδο k_x - k_z .

Σημειώνεται ότι εάν το $\tilde{M}(k_x, k_z)$ είναι γνωστό, τότε τη άγνωστη συνάρτηση μοντέλου $M(x, z)$ μπορεί να προσδιοριστεί από τον δισδιάστατο αντίστροφο μετασχηματισμό Fourier:

$$M(x, z) = \frac{1}{4\pi^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \tilde{M}(k_x, k_z) e^{j(k_x x + k_z z)} dk_x dk_z. \quad (1.5)$$

Τώρα, έστω ότι το $\tilde{P}(\Omega, \theta)$ αντιπροσωπεύει τον μονοδιάστατο μετασχηματισμό Fourier της συνάρτησης δεδομένων $P(u, \theta)$ κατά τον u -άξονα, όπως παρουσιάζεται στην εικόνα 1.4. Ο μονοδιάστατος μετασχηματισμός Fourier γράφεται:

$$\tilde{P}(\Omega, \theta) = \int_{-\infty}^{+\infty} P(u, \theta) e^{-j\Omega u} du, \quad (1.6)$$

Όπου Ω η χωρική συχνότητα κατά τον άξονα- u . Αντικαθιστώντας την εξίσωση (1.3) στην εξίσωση (1.6) προκύπτει:

$$\tilde{P}(\Omega, \theta) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} M(x, z) e^{-j\Omega u} dv du. \quad (1.7)$$

Τώρα, η εξίσωση (1.7) πρέπει να γραφτεί αποκλειστικά με x και z . Η μεταβλητή u αντικαθίσταται με x και z χρησιμοποιώντας την αντιστροφή της εξίσωσης (1.1) που δίνεται από:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix}. \quad (1.8)$$

Από την εξίσωση (8) και αντικαθιστώντας το $du du$ με $dx dz$ στην εξίσωση (1.7) προκύπτει:

$$\begin{aligned}\tilde{P}(\Omega, \theta) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} M(x, z) e^{-j\Omega(x \cos \theta + z \sin \theta)} dx dz \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} M(x, z) e^{-j[(\Omega \cos \theta)x + (\Omega \sin \theta)z]} dx dz\end{aligned}\quad (1.9)$$

Συγκρίνοντας τα ολοκληρωτέα μέλη (integrands) των εξισώσεων (1.9) και (1.4), φαίνεται ότι η εξίσωση (1.9) είναι απλά ο δισδιάστατος μετασχηματισμός Φουριέ του $M(x, z)$, όπου τα k_x και k_z περιορίζονται στον Ω άξονα θέτοντας:

$$\begin{aligned}k_x &= \Omega \cos \theta, \text{ and} \\ k_z &= \Omega \sin \theta.\end{aligned}\quad (1.10)$$

Η σχέση αυτή γίνεται εμφανής στην εικόνα (1.5). Αντικαθιστώντας τις (10) στην (9) προκύπτει:

$$\tilde{P}(\Omega, \theta) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} M(x, z) e^{-j(k_x x + k_z z)} dx dz, \quad (1.11)$$

Όπου τα k_x και k_z ορίζονται από την εξίσωση (1.10). Συγκρίνοντας τις εξισώσεις (1.11) και (4) φαίνεται ότι έχει επιτευχθεί μια πιο απλή σχέση μεταξύ της συνάρτησης δεδομένων $P(u, \theta)$ και της συνάρτησης μοντέλου $M(x, z)$ στο επίπεδο των χωρικών συχνοτήτων:

$$\tilde{P}(\Omega, \theta) = \tilde{M}(k_x, k_z). \quad (1.12)$$

Χρησιμοποιώντας λόγια, στην εξίσωση (1.12) διατυπώνεται ότι ο μονοδιάστατος μετασχηματισμός Fourier της προβολής που αναπαρίσταται από την εξίσωση δεδομένων $\tilde{P}(\Omega, \theta)$ είναι ίσος με ένα τμήμα του δισδιάστατου μετασχηματισμού Fourier της συνάρτησης μοντέλου $\tilde{M}(k_x, k_z)$ προσδιορισμένη στο γεωμετρικό τόπο των σημείων $k_x = \Omega \cos \theta$ και $k_z = \Omega \sin \theta$. Η εξίσωση (1.12) ονομάζεται θεώρημα προβολής. Το θεώρημα αυτό όπως φάνηκε και στην εικόνα 1.5 δίνει μόνο ένα τμήμα της συνάρτησης μοντέλου ανά προβολή. Στην επομένη ενότητα παρουσιάζεται, το πόσες προβολές σε διαφορετικές γωνίες θ , χρειάζονται για να κατασκευαστεί ολόκληρη η συνάρτηση μοντέλου μέσω του θεωρήματος προβολής.

Οι δύο αυτές τεχνικές είναι ο ευθύς μετασχηματισμός στην τομογραφία ακτίνας (direct transform ray tomography) και η οπισθοπροβολική τομογραφία ακτίνας (backprojection ray tomography).

1.4.2.2 ΤΟΜΟΓΡΑΦΙΑ ΑΚΤΙΝΩΝ ΜΕ ΕΥΘΥ ΜΕΤΑΣΧΗΜΑΤΙΣΜΟ

Η τομογραφία ακτίνων με ευθύ μετασχηματισμό χρησιμοποιεί το θεώρημα τμήματος προβολής με έναν σχετικά απλό τρόπο. Σε προηγούμενη ενότητα δείχτηκε ότι η εφαρμογή του θεωρήματος τμήματος προβολής στον μονοδιάστατο μετασχηματισμό Fourier μιας μοναδικής προβολής που αναπαρίσταται από την συνάρτηση δεδομένων $\tilde{P}(\Omega, \theta)$ προσδιορίζει μόνο τμήμα της συνάρτησης μοντέλου $\tilde{M}(k_x = \Omega \cos \theta, k_z = \Omega \sin \theta)$. Η εικόνα 1.5 δείχνει ένα τέτοιο τμήμα μέσω της συνάρτησης μοντέλου. Για την ανάκτηση

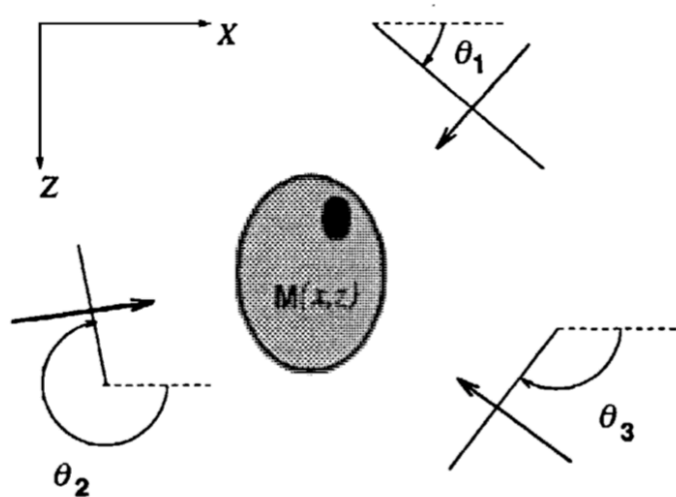
της πλήρους συνάρτησης μοντέλου, ο στόχος πρέπει να εξεταστεί από πολλές διαφορετικές κατευθύνσεις.

Η εικόνα 1.6 δείχνει τρεις διαφορετικές κατευθύνσεις κατά τις οποίες ακτίνες X εξετάζουν το κεφάλι του ασθενή. Οι παρατηρούμενες συναρτήσεις δεδομένων (observed data functions) για τις τρεις αυτές προβολές είναι οι $P(u_1\theta_1)$, $P(u_2\theta_2)$ και $P(u_3\theta_3)$.

Μετά την εφαρμογή του θεωρήματος προβολής στον μονοδιάστατο μετασχηματισμό Fourier αυτών των συναρτήσεων δεδομένων, λαμβάνονται τρία τμήματα μέσω του φάσματος πλάτους της συνάρτησης μοντέλου (Εικόνα 1.7).

Τώρα ο δισδιάστατος μετασχηματισμός Fourier της συνάρτησης μοντέλου $\tilde{M}(k_x k_z)$ είναι καλύτερα ορισμένος από ότι με το ένα τμήμα που παρουσιάζεται στην εικόνα 1.5, αλλά είναι ακόμα ανεπαρκής για την πλήρη ανακατασκευή της εικόνας.

Ο στόχος πρέπει να εξετασθεί με ακτίνες X από όλες τις κατευθύνσεις με τη γωνία θ να κυμαίνεται από 0-180 μοίρες. Μόνο τότε το επίπεδο $k_x k_z$ στην εικόνα 1.7 θα είναι εντελώς καλυμμένο με τμήματα $\tilde{M}(\Omega \cos \theta, \Omega \sin \theta)$, όπου το θ κυμαίνεται από 0 έως 180 μοίρες. Μετά από ένα τέτοιο πείραμα, ο πλήρης δισδιάστατος μετασχηματισμός Fourier της άγνωστης συνάρτησης μοντέλου $M(x,z)$ προσδιορίζεται κατά μήκος των ακτινωτών γραμμών στο επίπεδο $k_x k_z$.



Εικόνα 1: 6 Το κρανίο του ασθενή εξετάζεται με ακτίνες X σε τρεις διαφορετικές κατευθύνσεις θ_1, θ_2 και θ_3 . Εφαρμογή του θεωρήματος τμήματος προβολής στην συνάρτηση δεδομένων που προέρχεται από αυτές τις προβολές παράγει τα τμήματα που φαίνονται στην εικόνα 1.7

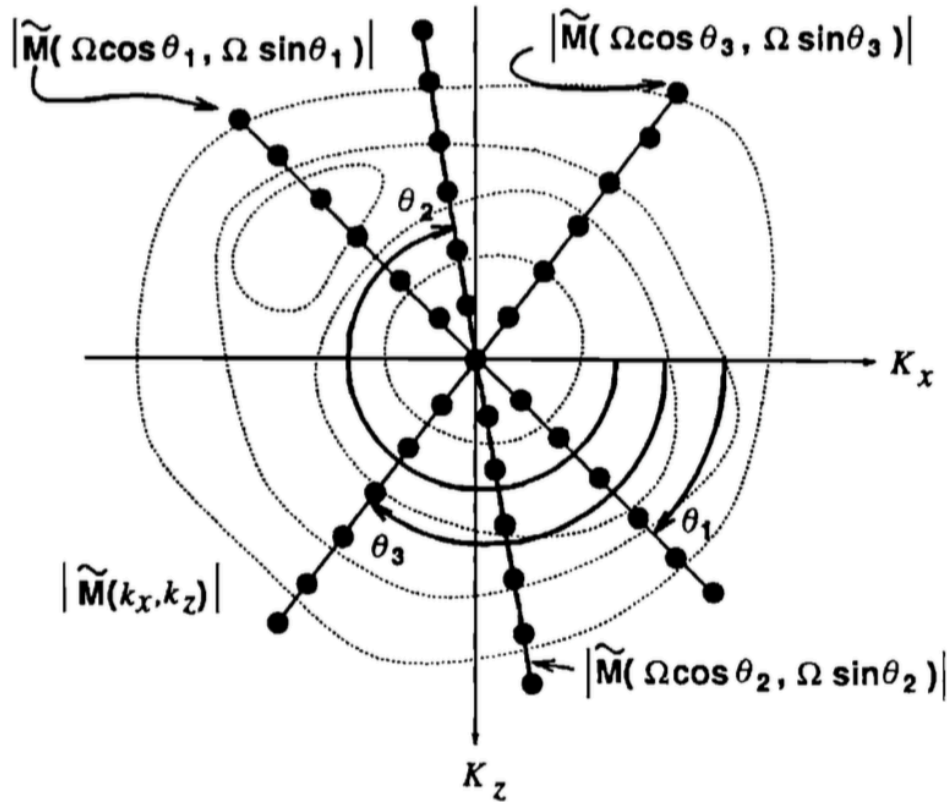
Για την λήψη του $M(x,z)$ από το $\tilde{M}(\Omega \cos \theta, \Omega \sin \theta)$, ένας αλγόριθμος που εφαρμόζει την μέθοδο τομογραφίας ακτίνας με ευθύ μετασχηματισμό πρέπει πρώτα να μετατρέψει τα δεδομένα από ένα πολικό πλέγμα $\tilde{M}(\Omega \cos \theta, \Omega \sin \theta)$ σε ένα καρτεσιανό $\tilde{M}(k_x, k_z)$ στο επίπεδο $k_x - k_z$ ή:

$$\tilde{M}(\Omega \cos \theta, \Omega \sin \theta) \xrightarrow{\text{interpolate}} \tilde{M}(k_x, k_z). \quad (1.13)$$

Τελικά, εφαρμόζεται ένας δισδιάστατος αντίστροφος μετασχηματισμός Fourier του $\tilde{M}(k_x k_z)$ για να προσδιοριστεί το $M(x,z)$:

$$M(x, z) = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \tilde{M}(k_x, k_z) e^{j(k_x x + k_z z)} dk_x dk_z. \quad (1.14)$$

Έτσι η ανακατασκευή της εικόνας ολοκληρώνεται.



Εικόνα 1: 7 Διαγραμματική απεικόνιση των τμημάτων μέσω του φάσματος πλάτους της άγνωστης συνάρτησης μοντέλου $[\tilde{M}(k_x, k_z)]$, το οποίο βρέθηκε με την εφαρμογή του θεωρήματος προβολής στην συνάρτηση δεδομένων που προσδιορίστηκε για την προβολή των ακτίνων X στις κατευθύνσεις θ_1, θ_2 και θ_3 (Lo)

Η μέθοδος τομογραφίας ακτίνας με ευθύ μετασχηματισμό μπορεί εύκολα να συνοψιστεί σε πέντε βήματα:

- 1) Προσδιορισμός της συνάρτησης δεδομένων $P(u, \theta)$ του στόχου με την διεύθυνση προβολής θ , να κυμαίνεται από 0 έως 180 μοίρες. Υπενθυμίζεται ότι η άγνωστη συνάρτηση μοντέλου $M(x, z)$ που αντιστοιχεί στον στόχο αναπαριστά μια φυσική ιδιότητα που επηρεάζει την διάδοση της ενέργειας με κάποιον τρόπο. Έτσι, μια συνάρτηση δεδομένων είναι απλά το ολοκλήρωμά της συνάρτησης μοντέλου κατά μήκος κάθε ακτίνας ή:

$$P(u, \theta) = \int_{-\infty}^{+\infty} M(x, z) dv.$$

- 2) Μονοδιάστατος μετασχηματισμός Fourier κατά τον άξονα u για κάθε συνάρτηση δεδομένων που δίνεται από:

$$\bar{P}(\Omega, \theta) = \int_{-\infty}^{+\infty} P(u, \theta) e^{-j\Omega u} du.$$

- 3) Χρήση του θεωρήματος προβολής για την λήψη τμημάτων από τον δισδιάστατο μετασχηματισμό Fourier της συνάρτησης μοντέλου. Κάθε τμήμα ορίζεται από:

$$\tilde{M}(\Omega \cos \theta, \Omega \sin \theta) = \bar{P}(\Omega, \theta).$$

4) Μετατροπή του δισδιάστατου μετασχηματισμού Fourier της συνάρτησης μοντέλου στο επίπεδο k_x-k_z από ένα πολικό πλέγμα $\tilde{M}(\Omega \cos \theta, \Omega \sin \theta)$ σε ένα καρτεσιανό $\tilde{M}(x, z)$

$$\tilde{M}(\Omega \cos \theta, \Omega \sin \theta) \xrightarrow{\text{interpolate}} \tilde{M}(k_x, k_z).$$

5) Αντίστροφος δισδιάστατος μετασχηματισμός Fourier στο $\tilde{M}(k_x, k_z)$ για να προκύψει το $M(x, z)$, η ανακατασκευασμένη εικόνα του στόχου. Ο αντίστροφος μετασχηματισμός δίνεται από:

$$M(x, z) = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \tilde{M}(k_x, k_z) e^{j(k_x x + k_z z)} dk_x dk_z.$$

Η μέθοδος τομογραφία ακτίνας με ευθύ μετασχηματισμό θα ήταν γρήγορη στην εφαρμογή αν δεν υπήρχε το τέταρτο βήμα όπως αυτό περιγράφηκε παραπάνω, που απαιτεί παρεμβολή της συνάρτησης μοντέλου στο πεδίο των συχνοτήτων.

1.4.2.3 ΤΟΜΟΓΡΑΦΙΑ ΑΚΤΙΝΑΣ ΜΕ ΟΠΙΣΘΟΠΡΟΒΟΛΗ

Η τομογραφία ακτίνας με οπισθοπροβολή χρησιμοποιεί το ίδιο σύστημα λήψης δεδομένων με την τομογραφία ακτίνας με ευθύ μετασχηματισμό: καταγραφή της συνάρτησης δεδομένων $P(u, \theta)$ με την πειραματική μέτρηση του ολοκληρώματος της άγνωστης συνάρτησης μοντέλου $M(x, z)$ κατά διαφορετικές διαδρομές ακτίνων ή προβολές. Η διαφορά μεταξύ των δυο μεθόδων είναι το πώς υπολογίζεται η συνάρτηση μοντέλου από την συνάρτηση δεδομένων. (Lo)

Για την εφαρμογή της μεθόδου της τομογραφίας ακτίνας με οπισθοπροβολή πρέπει πρώτα να γραφτεί ο δισδιάστατος αντίστροφος μετασχηματισμός Fourier της συνάρτησης μοντέλου $\tilde{M}(k_x, k_z)$:

$$M(x, z) = \frac{1}{4\pi^2} \int_{-\pi}^{+\pi} \int_{-\pi}^{+\pi} \tilde{M}(k_x, k_z) e^{j(k_x x + k_z z)} dk_x dk_z. \quad (1.15)$$

Στην συνέχεια γίνεται αλλαγή στις μεταβλητές της εξίσωσης (1.15) αντικαθιστώντας το k_x με $\Omega \cos \theta$ και το k_z με $\Omega \sin \theta$, και αλλάζοντας την ολοκλήρωση από dk_x, dk_z σε $|\Omega| d\Omega d\theta$ και έτσι:

$$M(x, z) = \frac{1}{4\pi^2} \int_0^{2\pi} \int_0^{\infty} \tilde{M}(\Omega \cos \theta, \Omega \sin \theta) \times e^{j\Omega(x \cos \theta + z \sin \theta)} |\Omega| d\Omega d\theta. \quad (1.16)$$

Οι ολοκληρώσεις ως προς θ στην εξίσωση (1.16) μπορούν να γραφτούν σαν δύο ολοκληρώματα:

$$\begin{aligned} M(x, z) &= \frac{1}{4\pi^2} \int_0^{\pi} \int_0^{\infty} \tilde{M}(\Omega \cos \theta, \Omega \sin \theta) \\ &\times e^{j\Omega(x \cos \theta + z \sin \theta)} |\Omega| d\Omega d\theta \\ &+ \frac{1}{4\pi^2} \int_{\pi}^{2\pi} \int_0^{\infty} \tilde{M}[\Omega \cos(\theta + \pi), \Omega \sin(\theta + \pi)] \\ &\times e^{j\Omega[x \cos(\theta + \pi) + z \sin(\theta + \pi)]} |\Omega| d\Omega d\theta. \end{aligned} \quad (1.17)$$

Το δεξί μέλος της εξίσωσης (17) μπορεί να γραφτεί ως:

$$\tilde{M}[\Omega \cos(\theta + \pi), \Omega \sin(\theta + \pi)] = \tilde{M}(-\Omega \cos \theta, -\Omega \sin \theta). \quad (1.18)$$

Το δεύτερο σέτ ολοκληρωμάτων στην δεξιά πλευρά της εξίσωσης (1.17) μπορεί να ξαναγραφτεί αντικαθιστώντας την συνάρτηση μοντέλου με την εξίσωση (1.18),

εφαρμόζοντας προσθέσεις γωνιών θέτοντας $\Omega = -\Omega$ και $d\Omega = -d\Omega$, και αντιστρέφοντας την φορά της ολοκλήρωσης ως προς Ω . Βάση αυτών η εξίσωση 1.17 γράφεται:

$$\begin{aligned} M(x, z) = & \frac{1}{4\pi^2} \int_0^\pi \int_0^\infty \tilde{M}(\Omega \cos \theta, \Omega \sin \theta) \\ & \times e^{j\Omega(x \cos \theta + z \sin \theta)} | \Omega | d\Omega d\theta \\ & + \frac{1}{4\pi^2} \int_0^\pi \int_{-\infty}^0 \tilde{M}(\Omega \cos \theta, \Omega \sin \theta) \\ & \times e^{j\Omega(x \cos \theta + z \sin \theta)} | \Omega | d\Omega d\theta. \end{aligned}$$

Συνδυάζοντας τα ολοκληρώματα ως προς Ω προκύπτει:

$$\begin{aligned} M(x, z) = & \frac{1}{4\pi^2} \int_0^\pi \int_{-\infty}^{+\infty} \tilde{M}(\Omega \cos \theta, \Omega \sin \theta) \\ & \times e^{j\Omega(x \cos \theta + z \sin \theta)} | \Omega | d\Omega d\theta. \end{aligned} \quad (1.19)$$

Χρησιμοποιώντας το θεώρημα προβολής αντικαθίσταται το $\tilde{M}(\Omega \cos \theta, \Omega \sin \theta)$ στην εξίσωση (1.19) με το $\tilde{P}(\Omega, \theta)$ και προκύπτει ο τύπος της τομογραφίας ακτίνας με οπισθοπροβολή:

$$\begin{aligned} M(x, z) = & \frac{1}{4\pi^2} \int_0^\pi \int_{-\infty}^{+\infty} \tilde{P}(\Omega, \theta) \\ & \times e^{j\Omega(x \cos \theta + z \sin \theta)} | \Omega | d\Omega d\theta. \end{aligned} \quad (1.20)$$

1.4.3 ΜΕΘΟΔΟΙ ΑΝΑΠΤΥΞΗΣ ΣΕΙΡΩΝ

Οι μέθοδοι ανάπτυξης σειρών αποτελούν μια ομάδα από υπολογιστικούς αλγορίθμους οι οποίοι, όπως οι μέθοδοι μετασχηματισμού προσδιορίζουν την συνάρτηση μοντέλου, $M(x, z)$ της περιοχής του στόχου. Όμως, σε αντίθεση με τις μεθόδους μετασχηματισμού, αυτοί οι αλγόριθμοι επιτρέπουν εύκολα καμπυλόγραμμες διαδρομές ακτίνων μέσα από τον στόχο και είναι έτσι, πολύ χρήσιμες για εφαρμογές της σεισμικής τομογραφίας. Όπως και πριν, η συζήτηση περιορίζεται για ένα πρόβλημα δυο διαστάσεων, έτσι ώστε η συνάρτηση μοντέλου $M(x, z)$ να προσδιορίζεται σε επίπεδο που περνάει διαμέσου του στόχου και περιέχει όλες τις πηγές και τους δέκτες.

1.4.3.1 ΤΟ ΠΡΟΒΛΗΜΑ ΤΗΣ ΜΟΝΤΕΛΟΠΟΙΗΣΗΣ

Όπως θα δειχθεί και παρακάτω, η μέθοδος ανάπτυξης σειρών ενημερώνει επαναληπτικά μια εκτιμώμενη συνάρτηση μοντέλου M^{est} έτσι ώστε αυτή να συγκλίνει στην πραγματική συνάρτηση μοντέλου M^{true} .

Οι ενημερώσεις αυτές προκύπτουν από την σύγκριση της παρατηρούμενης συνάρτησης δεδομένων P^{obs} με μία εκτιμώμενη συνάρτηση δεδομένων P^{pre} . Η εμπρός μοντελοποίηση είναι απαραίτητη για τον προσδιορισμό της εκτιμώμενης συνάρτησης δεδομένων και θα αποτελέσει το θέμα αυτού του κεφαλαίου.

Η εξίσωση (1.2) προσδιορίζει την πειραματική διαδικασία για τις μεθόδους μετασχηματισμού ως το γραμμικό ολοκλήρωμα (line integral) της συνάρτησης $M(x, z)$ κατά μήκος μιας ευθύγραμμης ακτίνας κατά την διεύθυνση του u-άξονα. Σύμφωνα με τα παραπάνω η εξίσωση (1.2) γράφεται:

$$P^{\text{obs}}(u, \theta) = \int_{\text{ray}} M^{\text{true}}(x, z) dv.$$

Σε προηγούμενη περίπτωση δεν απαιτούνταν μοντελοποίηση καθώς οι ακτίνες διαδρομής ήταν ευθύγραμμες και το θεώρημα προβολής μπορούσε να εφαρμοσθεί απευθείας για τον προσδιορισμό της πραγματικής συνάρτησης μοντέλου $M^{true}(x,z)$. Για την μέθοδο ανάπτυξης σειρών όμως, μπορούν να συμπεριληφθούν και καμπυλόγραμμες ακτίνες διαδρομής. Η εξίσωση (1.2) μπορεί εύκολα να μετασχηματιστεί για να περιγράψει καμπυλόγραμμες ακτίνες διαδρομής με την μορφή καμπυλών με το να ξαναγραφτεί η συνάρτηση μοντέλου σε όρους του διανύσματος θέσης \mathbf{r} . Έτσι για ένα δεδομένο ζευγάρι πηγής και δέκτη το γραμμικό ολοκλήρωμα της συνάρτησης μοντέλου $M(\mathbf{r})$ κατά μήκος της ακτίνας διαδρομής είναι:

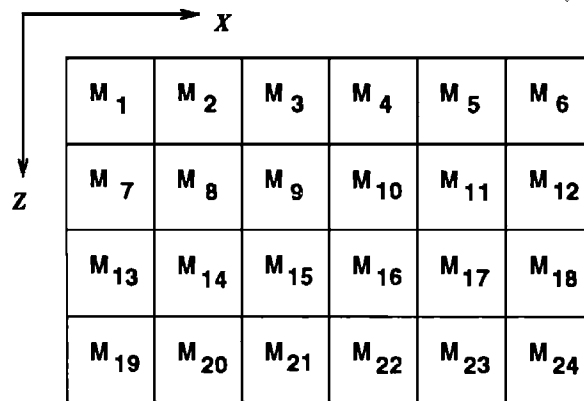
$$P^{obs} = \int_{ray} M^{true}(\mathbf{r}) d\mathbf{r},$$

όπου η παρατηρούμενη προβολή που δίνεται από την συνάρτηση δεδομένων P^{obs} αντιπροσωπεύει το μετρούμενο γραμμικό ολοκλήρωμα και η $M^{true}(\mathbf{r})$ είναι η πραγματική συνάρτηση μοντέλου που πρέπει να βρεθεί.

Η τελευταία εξίσωση χρησιμοποιείται για να διατυπωθεί το πρόβλημα της μοντελοποίησης θέτοντας:

$$P = \int_{ray} M(\mathbf{r}) d\mathbf{r}, \quad (1.21)$$

όπου P τώρα είναι η εκτιμώμενη συνάρτηση δεδομένων και $M(\mathbf{r})$ η εκτιμώμενη συνάρτηση μοντέλου.



M_1	M_2	M_3	M_4	M_5	M_6
M_7	M_8	M_9	M_{10}	M_{11}	M_{12}
M_{13}	M_{14}	M_{15}	M_{16}	M_{17}	M_{18}
M_{19}	M_{20}	M_{21}	M_{22}	M_{23}	M_{24}

Εικόνα 1: 8 Η μέθοδος ανάπτυξης σειρών χρησιμοποιεί μια διακριτή συνάρτηση μοντέλου M_j $j=1, \dots, J$, όπου M_j είναι η μέση τιμή της συνεχούς συνάρτησης μοντέλου $M(\mathbf{r})$ μέσα στο j κελί. Εδώ $j=24$.

Έτσι η μοντελοποίηση ορίζεται ως η διαδικασία προσδιορισμού της προβλεπόμενης συνάρτησης δεδομένων από το γραμμικό ολοκλήρωμα κατά μήκος της ακτίνας διαδρομής μέσω μιας γνωστής, αλλά εκτιμώμενης συνάρτησης μοντέλου.

Όπως και με τις μεθόδους μετασχηματισμού, η συνάρτηση μοντέλου στην τομογραφία ανάπτυξης σειρών διακριτοποιείται.

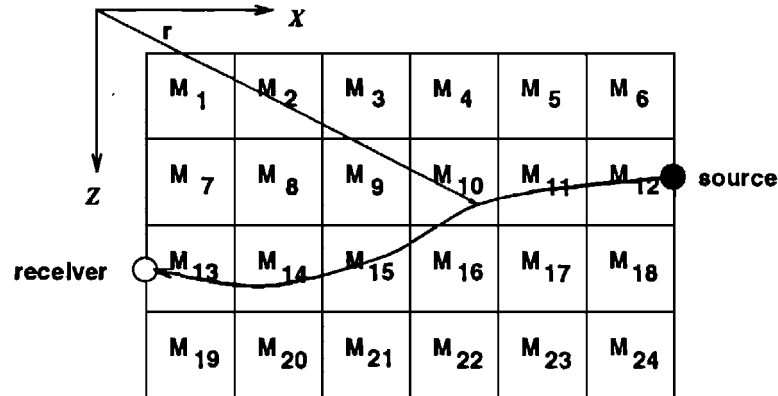
Η εικόνα 1.8 δείχνει την εικόνα περιοχής του στόχου χωρισμένη σε πολλά μικρά κελιά. Κάθε κελί λαμβάνει την μέση τιμή της φυσικής ιδιότητας (π.χ. εξασθένηση ακτίνων X , βραδύτητα, κλπ.) που αντιπροσωπεύεται από την συνεχή συνάρτηση μοντέλου $M(\mathbf{r})$ μέσα στο κελί. Η συνάρτηση μοντέλου στην εικόνα 1.8 χωρίζεται σε 24 κελιά και γράφεται διακριτά ως M_j όπου $j=1, \dots, 24$.

Το M_j αντιπροσωπεύει λοιπόν την μέση τιμή της $M(\mathbf{r})$ μέσα στο j -οστο κελί.

Η Εικόνα 1.9 δείχνει μια ακτίνα που ταξιδεύει μέσα από τη διακριτοποιημένη συνάρτηση μοντέλου. Η εξίσωση (1.21) ξαναγράφεται σε διακριτή μορφή, για να περιγράψει την ακτίνα που ταξιδεύει μέσα από την διακριτή συνάρτηση μοντέλου ως:

$$P = \sum_{j=1}^J M_j S_j, \quad (1.22)$$

όπου M_j η εκτιμώμενη συνάρτηση μοντέλου για το j -οστό κελί, S_j το μήκος της ακτίνας διαδρομής μέσα στο j -οστό κελί και J ο συνολικός αριθμός κελιών στον στόχο.



Εικόνα 1: 9 Ακτίνα ταξιδεύει μέσα από ένα διακριτό μοντέλο.

Το παράδειγμα της εικόνας 1.9 έχει $J=24$ κελιά, αλλά η ακτίνα διαπερνά μόνο επτά κελιά ($j=12,11,10,16,15,14$ και 13). Για να κρατηθεί η εξίσωση (1.22) σε συνοχή με την εξίσωση (21) θέτεται ότι $S_j=0$ για όλα τα κελιά που δεν διαπερνούνται από την ακτίνα. Η εικόνα 1.9 δείχνει 17 κελιά για τα οποία δεν υπάρχει πληροφορία καθώς η μοναδική ακτίνα δεν τα διαπέρασε.

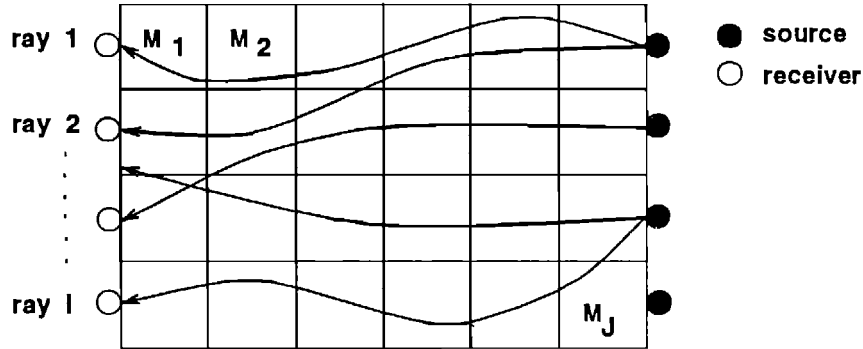
Προσθέτοντας περισσότερες πηγές και δέκτες γύρω από την περιοχή του άγνωστου στόχου, διαφορετικές ακτίνες δειγματίζουν τα 17 αδειγμάτιστα κελιά καθώς και μερικά από αυτά που είχαν ήδη δειγματιστεί. Η προσθήκη περισσότερων ακτίνων παρουσιάζεται στην εικόνα 1.10. Τώρα όλα τα κελιά μελετώνται από αυτό το δίκτυο ακτίνων.

Πρέπει να τροποποιηθούν οι δείκτες στην εξίσωση (1.22) για να συμπεριλάβουν μια τιμή προβολής για κάθε ακτίνα. Αν το P_i αντιπροσωπεύει την προβολή τότε η εξίσωση (1.22) γράφεται ως:

$$P_i = \sum_{j=1}^J M_j S_{ij}, \text{ for } i = 1, \dots, I, \quad (1.23)$$

όπου I ο συνολικός αριθμός ακτίνων, S_{ij} το μήκος διαδρομής της i -οστής ακτίνας μέσω του j -οστού κελιού και όπως προηγουμένως M_j η διακριτή εκτίμηση της συνάρτησης μοντέλου για το j -οστό κελί και J ο συνολικός αριθμός κελιών. Η εξίσωση (1.23) είναι η

μαθηματική διατύπωση του προβλήματος της μοντελοποίησης που χρησιμοποιείται στην τομογραφία ακτίνας με ανάπτυξη σειρών



Εικόνα 1: 10 Γενικά μια μόνο ακτίνα δεν παρέχει πληροφορίες για όλα τα κελιά της συνάρτησης μοντέλου. Όμως χρησιμοποιώντας περισσότερες θέσεις πηγής και δέκτη γύρω από το στόχο όλα τα κελιά μπορούν τελικά να ελεγχθούν.

Η εξίσωση (1.23) μπορεί να μοντελοποιήσει αποτελεσματικά την διαδικασία απόκτησης δεδομένων αν οι προβολές P_i , $i=1, \dots, I$ είναι τα παρατηρούμενα δεδομένα (πχ ο χρόνος διαδρομής) και η συνάρτηση M_j , $j=1, \dots, J$ είναι η πραγματική αλλά άγνωστη συνάρτηση μοντέλου:

$$P_i^{obs} = \sum_{j=1}^J M_j^{true} S_{ij}, \text{ for } i = 1, \dots, I. \quad (1.24)$$

1.4.3.2 Η ΜΕΘΟΔΟΣ KACZMARZ

Όπως αναφέρθηκε παραπάνω η μοντελοποίηση απαιτεί τον προσδιορισμό της πραγματικής συνάρτησης μοντέλου. Έτσι η εξίσωση (1.23) θα γραφτεί με μορφή πινάκων για να απλοποιηθούν τα μαθηματικά. Σε μορφή πινάκων η εξίσωση (1.23) γράφεται:

$$\mathbf{P} = \mathbf{SM}, \quad (1.25)$$

όπου οι εκτιμώμενες προβολές στο διάνυσμα δεδομένων \mathbf{P} είναι:

$$\mathbf{P} = \begin{bmatrix} P_1 \\ P_2 \\ \vdots \\ P_I \end{bmatrix}, \quad (1.26)$$

οι τιμές της διακριτής εκτιμώμενης συνάρτησης μοντέλου στο διάνυσμα μοντέλου \mathbf{M} είναι:

$$\mathbf{M} = \begin{bmatrix} M_1 \\ M_2 \\ \vdots \\ M_J \end{bmatrix}, \quad (1.27)$$

και τα μήκη των ακτίνων διαδρομής για I ακτίνες και J κελιά στο \mathbf{S} είναι:

$$\mathbf{S} = \begin{bmatrix} S_{11} & S_{12} & \cdots & S_{1J} \\ S_{21} & S_{22} & \cdots & S_{2J} \\ \vdots & \vdots & \ddots & \vdots \\ S_{I1} & S_{I2} & \cdots & S_{IJ} \end{bmatrix} \quad (1.28)$$

Η εξίσωση (1.24) μπορεί να γραφτεί επίσης σε μορφή πινάκων σαν:

$$\mathbf{P}^{obs} = \mathbf{S}\mathbf{M}^{true} \quad (1.29)$$

Παρόλο που η (1.29) δεν θα λυθεί απευθείας, είναι επιθυμητό να προσδιοριστεί το πραγματικό διάνυσμα μοντέλου \mathbf{M}^{true} με δεδομένο το \mathbf{P}^{obs} και το \mathbf{S} . Το πρόβλημα τώρα μεταφέρεται στο να βρεθεί ένας γενικευμένος αντίστροφος μετασχηματισμός \mathbf{S}^{-g} . Τότε, θα μπορούσε να εφαρμοσθεί ο γενικευμένος αντίστροφος μετασχηματισμός \mathbf{S}^{-g} και στις δυο πλευρές τις εξίσωσης (1.29) για να βρεθεί το διάνυσμα του πραγματικού μοντέλου:

$$\begin{aligned} \mathbf{S}^{-g}\mathbf{P}^{obs} &= \mathbf{S}^{-g}\mathbf{S}\mathbf{M}^{true} \\ &= \mathbf{M}^{true} \end{aligned}$$

Θεωρητικά η τελευταία εξίσωση ισχύει αλλά πρακτικά είναι πολύ δύσκολο να προσδιοριστεί το \mathbf{S}^{-g} για δυο λόγους.

Πρώτον, ο \mathbf{S} είναι συνήθως αρκετά μεγάλος και αραιός το οποίο αυξάνει αρκετά το υπολογιστικό κόστος του \mathbf{S}^{-g} . Δεύτερον ο \mathbf{S} είναι συνήθως κακώς ορισμένος (ill conditioned) το οποίο κάνει την μέθοδο υπολογισμού του \mathbf{S}^{-g} πολύ ασταθή.

Η μέθοδος του Kaczmarz αποφεύγει τα προβλήματα που σχετίζονται με την αντιστροφή ενός μεγάλου και αραιού πίνακα και παρέχει αποτελεσματικά μέσα για τον προσδιορισμό μιας προσεγγιστικής λύσης για την εξίσωση (1.29) χρησιμοποιώντας μια επαναληπτική διαδικασία. Η εικόνα 1.11 δείχνει ένα διάγραμμα ροής για την μέθοδο όπου υπάρχουν τρία βασικά βήματα στο επαναληπτικό κομμάτι του αλγορίθμου.

Μια αρχική εκτίμηση του διανύσματος μοντέλου \mathbf{M}^{init} είναι η είσοδος στην επαναληπτική λούπα του αλγορίθμου και λειτουργεί ως πρώτη εκτίμηση \mathbf{M}^{est} της πραγματικής λύσης \mathbf{M}^{true} .

Με την πρώτη εκτίμηση του διανύσματος του μοντέλου \mathbf{M}^{est} να είναι γνωστή, το πρώτο βήμα είναι η χρήση της μοντελοποίησης όπως αυτή ορίζεται από την εξίσωση (1.25) για την πρόβλεψη ενός διανύσματος δεδομένων \mathbf{P}^{pre} . Αυτό το βήμα πραγματοποιείται πολλαπλασιάζοντας τον πίνακα \mathbf{S} (operator) με το διάνυσμα του εκτιμώμενου μοντέλου:

$$\mathbf{P}^{pre} = \mathbf{S}\mathbf{M}^{est} \quad (1.30)$$

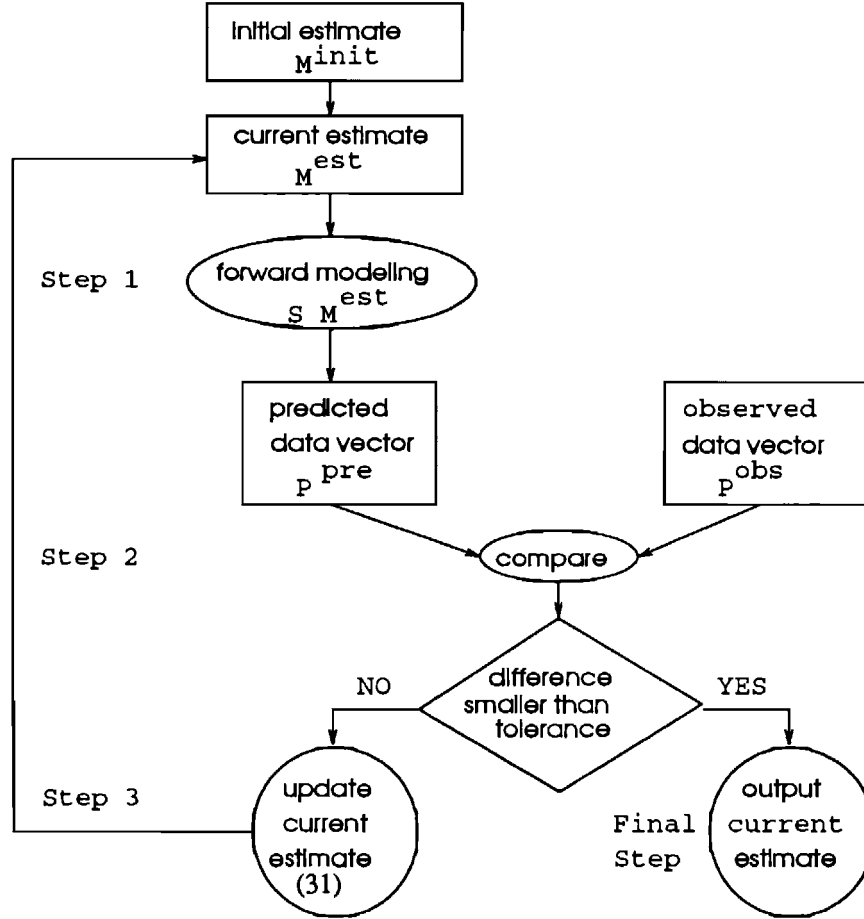
Στο δεύτερο βήμα το εκτιμώμενο διάνυσμα δεδομένων \mathbf{P}^{pre} συγκρίνεται με τις παρατηρούμενες τιμές του διανύσματος δεδομένων \mathbf{P}^{obs} με την λήψη των διαφορών μεταξύ των δυο.

Μια μικρή διαφορά ή καλή συμφωνία μεταξύ των εκτιμώμενων και παρατημένων δεδομένων υπονοεί συμφωνία μεταξύ του διανύσματος του εκτιμώμενου μοντέλου και του πραγματικού. Έτσι, εάν η διαφορά είναι μικρότερη μιας ορισμένης ανοχής τότε η τρέχουσα εκτίμηση του διανύσματος του μοντέλου είναι το αποτέλεσμα της λύσης της εξίσωσης (1.29) στο τελευταίο βήμα του αλγορίθμου.

Το τρίτο βήμα για το επαναληπτικό κομμάτι της μεθόδου Kaczmarz ξεκινάει όταν η διαφορά μεταξύ των εκτιμώμενων και των παρατηρούμενων διανυσμάτων δεδομένων είναι μεγαλύτερη της ορισμένης ανοχής.

Αυτό το σημαντικό βήμα ουσιαστικά χρησιμοποιεί την πληροφορία της διαφοράς $\mathbf{P}^{obs} - \mathbf{P}^{pre}$, για να ενημερώσει το τρέχον εκτιμώμενο διάνυσμα μοντέλου \mathbf{M}^{est} με μια νέα εκτίμηση του διανύσματος $\mathbf{M}^{(new)est}$, που βρίσκεται πιο κοντά στο διάνυσμα του πραγματικού μοντέλου \mathbf{M}^{true} . Αυτό το τρίτο βήμα γράφεται σε μορφή εξισώσεων ως:

$$\begin{aligned} M^{(new)est} &= M^{est} + \Delta^i M, \\ \text{for } i &= 1, \dots, I, \end{aligned} \quad (1.31)$$



Εικόνα 1: 11 Διάγραμμα ροής για την μέθοδο Kaczmarz (Lo)

Όπου το $\Delta^i M$ αντιπροσωπεύει την σταδιακή ενημέρωση στην τρέχουσα εκτίμηση του διανύσματος μοντέλου και ο δείκτης i εννοεί την εφαρμογή της (1.31) όταν συγκρίνονται οι σειρές των διανυσμάτων P^{obs} P^{pre} που αντιστοιχούν στην τελευταία τιμή του δείκτη.

Η νέα εκτίμηση $M^{(new)est}$ χρησιμοποιείται ως η τρέχουσα εκτίμηση για την επόμενη επανάληψη.

Η μέθοδος του υπολογισμού του $\Delta^i M$ στην εξίσωση (1.31) είναι προφανώς ένας κρίσιμος παράγοντας στην επιτυχία της μεθόδου του Kaczmarz. Η μέθοδος υπολογίζει το $\Delta^i M$ με την εξίσωση:

$$\Delta^i M = \begin{bmatrix} \Delta^i M_1 \\ \Delta^i M_2 \\ \vdots \\ \Delta^i M_J \end{bmatrix} \quad (1.32)$$

όπου:

$$\begin{aligned}
\Delta^i M_j &= S_{ij} \frac{P_i^{obs} - P_i^{pre}}{\sum_{j'=1}^J (S_{ij'})^2} \\
&= S_{ij} \frac{P_i - \sum_{j'=1}^J S_{ij'} M_{j'}^{est}}{\sum_{j'=1}^J (S_{ij'})^2}.
\end{aligned}
\quad (1.33)$$

Τώρα η εξίσωση (1.33) μπορεί να παραγωγιστεί γεωμετρικά και να δειχθεί η σύγκλιση της μεθόδου μέσω ενός απλού παραδείγματος.

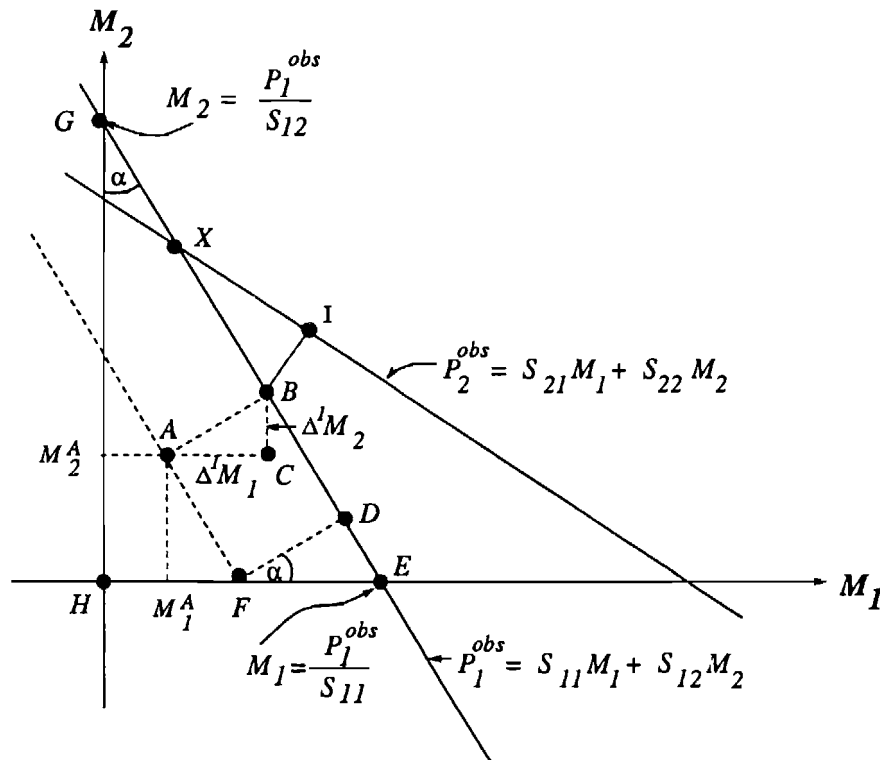
Έστω δύο ακτίνες ($I=2$) που ταξιδεύουν μέσα από ένα μοντέλο δύο κελιών, ($J=2$), έτσι ώστε η εξίσωση (1.29) για το πρόβλημα αυτό να γράφεται ως:

$$P_1^{obs} = S_{11}M_1 + S_{12}M_2, \text{ για την ακτίνα 1 } (1.34)$$

$$P_2^{obs} = S_{21}M_1 + S_{22}M_2, \text{ για την ακτίνα 2 } (1.35)$$

Όπου P_1^{obs} και P_2^{obs} είναι οι παρατηρήσεις και M_1 και M_2 οι άγνωστοι.

Σχεδιάζονται οι εξισώσεις (1.34) και (1.35) στην εικόνα 1.12 σαν γραμμές σε ένα δισδιάστατο χώρο με άξονες M_1 και M_2 . Η λύση της εξίσωσης προκύπτει στο σημείο X όπου οι δυο γραμμές τέμνονται.



Εικόνα 1: 12 Για δυο ακτίνες ($I=2$) σε ένα μοντέλο 2 κελιών ($J=2$), πιθανές τιμές για τα M_1 και M_2 ορίζονται από τις δυο γραμμές. (Lo)

Αφού η λύση στο σημείο X με συντεταγμένες (M_1^{true}, M_2^{true}) είναι άγνωστη, η διαδικασία πρέπει να ξεκινήσει με μια αρχική εκτίμηση στο σημείο A με συντεταγμένες (M_1^A, M_2^A) . Η αρχική εκτίμηση γίνεται η τρέχουσα εκτίμηση όπως φαίνεται στο διάγραμμα ροής (Εικόνα 1.11).

Το γεωμετρικό βήμα στην εικόνα 1.12 είναι να βρεθεί η κάθετη προβολή του σημείου A στην γραμμή που ορίζεται από την εξίσωση (1.34) στο σημείο B. Μαθηματικά αυτό το βήμα περιγράφεται από:

$$M_1^B = M_1^A + \Delta^1 M_1, \text{ και } (1.36)$$

$$M_2^B = M_2^A + \Delta^1 M_2, \quad (1.37)$$

όπου $\Delta^1 M_1$ και $\Delta^2 M_2$ οι διορθώσεις που αναζητούνται γεωμετρικά και ορίζονται από την εξίσωση (1.33) για $i=1$ και $j=1,2$.

Η πρώτη γεωμετρική σχέση από την εικόνα 1.12 είναι η ομοιότητα των τριγώνων ABC, FED και GEH, Χρησιμοποιώντας αυτές τις ομοιότητες:

$$\begin{aligned} \Delta^1 M_1 &= \overline{AB} \cos \alpha = \overline{DF} \cos \alpha = \overline{EF} \cos^2 \alpha \\ &= \overline{EF} \frac{\overline{GH}^2}{\overline{GE}^2}, \text{ και } (1.38) \end{aligned}$$

$$\begin{aligned} \Delta^1 M_2 &= \overline{AB} \sin \alpha = \overline{DF} \sin \alpha = \overline{EF} \cos \alpha \sin \alpha \\ &= \overline{EF} \frac{\overline{GH}}{\overline{GE}} \frac{\overline{EH}}{\overline{GE}}. \quad (1.39) \end{aligned}$$

Τώρα, πρέπει να προσδιοριστούν τα \overline{EF} , $\overline{GH}/\overline{GE}$ και $\overline{EH}/\overline{GE}$ στις εξισώσεις (1.38) και (1.39). Στο ευθύγραμμο τμήμα \overline{EF} , το σημείο E βρίσκεται στο $(M_1=P_1^{\text{obs}}/S_{11}, M_2=0)$. Το σημείο F είναι η τομή με τον άξονα M_1 και μιας ευθείας παράλληλης στην εξίσωση (1.34) και περιέχει τα σημεία (M_1^A, M_2^A) . Η εξίσωση για αυτή την ευθεία είναι:

$$S_{11} M_1^A + S_{12} M_2^A = S_{11} M_1 + S_{12} M_2. \quad (1.40)$$

Από την εξίσωση (1.40) προσδιορίζονται οι συντεταγμένες του σημείου F κατά μήκος του άξονα M_1 ως:

$$(M_1 = (S_{11} M_1^A + S_{12} M_2^A)/S_{11}, M_2 = 0).$$

Τελικά το μήκος του ευθύγραμμου τμήματος EF δίνεται από:

$$\overline{EF} = \frac{1}{S_{11}} [P_1^{\text{obs}} - (S_{11} M_1^A + S_{12} M_2^A)]. \quad (1.41)$$

Ο λόγος $\overline{GH}/\overline{GE}$ είναι απλά:

$$\begin{aligned} \frac{\overline{GH}}{\overline{GE}} &= \frac{P_1^{\text{obs}}/S_{12}}{\sqrt{(P_1^{\text{obs}}/S_{11})^2 + (P_1^{\text{obs}}/S_{12})^2}} \\ &= \frac{S_{11}}{\sqrt{S_{11}^2 + S_{12}^2}}. \quad (1.42) \end{aligned}$$

Παρόμοια το $\overline{EH}/\overline{GE}$ δίνεται από:

$$\begin{aligned} \frac{\overline{EH}}{\overline{GE}} &= \frac{P_1^{\text{obs}}/S_{11}}{\sqrt{(P_1^{\text{obs}}/S_{11})^2 + (P_1^{\text{obs}}/S_{12})^2}} \\ &= \frac{S_{12}}{\sqrt{S_{11}^2 + S_{12}^2}}. \quad (1.43) \end{aligned}$$

Αντικαθιστώντας τις εξισώσεις (1.41), (1.42) και (1.43) στις εξισώσεις (1.38) και (1.39) οδηγεί στις διορθώσεις του μοντέλου λόγου της πρώτης ακτίνας:

$$\Delta^1 M_1 = S_{11} \frac{P_1^{\text{obs}} - P_1^{\text{pre}}}{S_{11}^2 + S_{12}^2} \quad (1.44)$$

και

$$\Delta^1 M_2 = S_{12} \frac{P_1^{obs} - P_1^{pre}}{S_{11}^2 + S_{12}^2}, \quad (1.45)$$

όπου $P_1^{pr} = S_{11}M_1^A + S_{12}M_2^A$. Οι εξισώσεις (1.44) και (1.45) είναι οι ίδιες με την εξίσωση (1.33) όταν $i=1$ και $j=1,2$.

Έτσι φαίνεται ότι η εξίσωση (1.33) απλά αντιπροσωπεύει την προβολή της εκτίμησης του μοντέλου σε ένα από τα υπερ-επίπεδα που ορίστηκαν από την εξίσωση (1.29).

Συνεχίζοντας αυτό το παράδειγμα ένα βήμα παραπέρα, δύναται να προσδιοριστεί το σημείο I στην εικόνα 1.12, την προβολή του σημείου B στο υπερ-επίπεδο 2 που ορίζεται από την εξίσωση (1.35) για την δεύτερη ακτίνα χρησιμοποιώντας τους δείκτες $i=2$ και $j=1,2$ στην εξίσωση (1.33).

Εάν εναλλαχθούν οι προβολές των εκτιμήσεων του μοντέλου μεταξύ των δυο υπερ. επιπέδων, τότε οι ενημερωμένες εκτιμήσεις του μοντέλου πρέπει να συγκλίνουν στο σημείο X όπως φαίνεται στην εικόνα 1.13. Άρα, η μέθοδος Kaczmarz θα συγκλίνει στην λύση της εξίσωσης (1.29). (Lo)

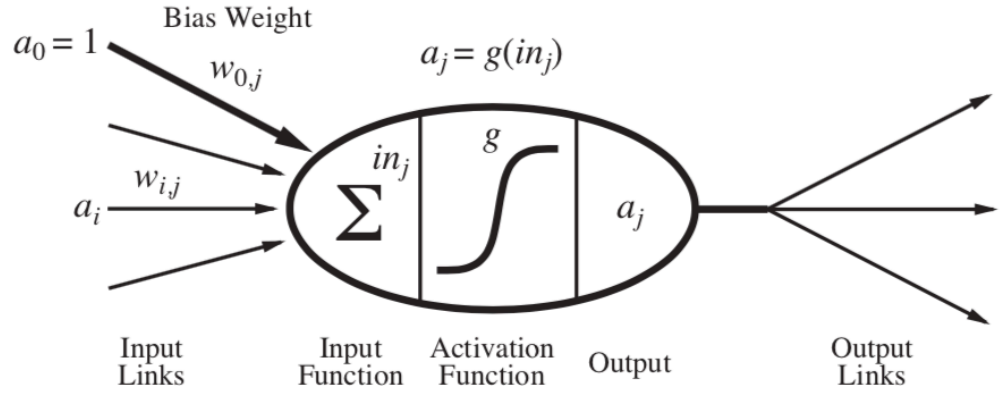
2. ΤΕΧΝΗΤΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ

2.1 ΕΙΣΑΓΩΓΗ

Όπως αναφέρθηκε και σε προηγούμενο κεφαλαίο το πρόβλημα της επιλογής των πρώτων αφίξεων είναι αρκετά σημαντικό καθώς είναι πολύ χρονοβόρο. Στον κλάδο της γεωφυσικής μπορεί κανείς να βρει αρκετές εφαρμογές των τεχνητών νευρωνικών δικτύων, σε περιορισμένη όμως έκταση καθώς είναι ένα σχετικά καινούργιο επιστημονικό αντικείμενο. Στο παρόν κεφάλαιο θα γίνει μια συνοπτική αναφορά στην θεωρία πίσω από την κατασκευή ενός νευρωνικού δικτύου καθώς και στον αλγόριθμο εκπαίδευσης αυτού.

2.1.1 ΤΙ ΟΝΟΜΑΖΕΤΑΙ ΝΕΥΡΩΝΙΚΟ ΔΙΚΤΥΟ

Ο νευρώνας (neuron) είναι ένα κύτταρο στον εγκέφαλο με κύρια λειτουργία του, την συγκέντρωση, επεξεργασία και διάδοση ηλεκτρικών σημάτων. Η δυνατότητα επεξεργασίας πληροφοριών του εγκέφαλου θεωρείται ότι απορρέει κυρίως από δίκτυα τέτοιων νευρώνων. Η Εικόνα 2.1 δείχνει ένα απλό μαθηματικό μοντέλο του νευρώνα, το οποίο επινοήθηκε από τους McCulloch και Pitts (1943). Γενικά μιλώντας, ο νευρώνας «πυροδοτείται» όταν ένας γραμμικός συνδυασμός εισόδων του, υπερβαίνει ένα κατώφλι. Τεχνητά νευρωνικά δίκτυα ή νευρωνικά δίκτυα ονομάζονται τα μαθηματικά μοντέλα που βασίζονται στην θεώρηση ότι ο τρόπος επεξεργασίας πληροφοριών του εγκεφάλου, μέσω των νευρώνων, είναι σαφώς ταχύτερος από ψηφιακούς υπολογιστές σε κομμάτια που αφορούν την αντίληψη και την αναγνώριση μοτίβων. (Russel Stuart, 2010)



Εικόνα 2. 1: Ένα απλό μαθηματικό μοντέλο νευρώνα

2.1.2 ΜΟΝΑΔΕΣ ΣΤΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ

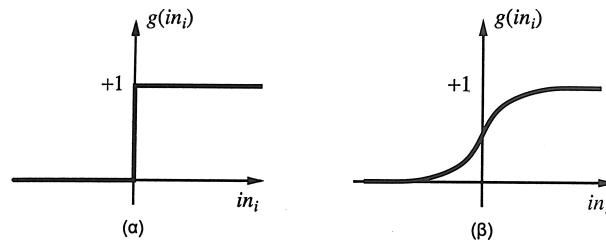
Τα νευρωνικά δίκτυα αποτελούνται από κόμβους ή μονάδες (units) που συνδέονται με κατευθυνόμενους συνδέσμους (links). Ο σύνδεσμος από μια μονάδα j στη μονάδα i εξυπηρετεί στη διάδοση ενεργοποίησης (activation) a_j από το j στο i . Με κάθε σύνδεσμο συσχετίζεται επίσης ένα αριθμητικό βάρος (weight) $W_{j,i}$, το οποίο προσδιορίζει την ισχύ και το πρόσημο της σύνδεσης. Κάθε μονάδα i υπολογίζει πρώτα ένα σταθμισμένο άθροισμα των εισόδων (inputs) της:

$$in_i = \sum_{j=0}^n W_{j,i} a_j \quad (2.1)$$

Στην συνέχεια εφαρμόζεται μια συνάρτηση ενεργοποίησης (activation function) g σε αυτό το άθροισμα για να παραχθεί η έξοδος (output):

$$a_i = g(in) = g\left(\sum_{j=0}^n W_{j,i} a_j\right) \quad (2.2)$$

Η συνάρτηση ενεργοποίησης g είναι σχεδιασμένη έτσι ώστε να εξυπηρετεί δυο βασικούς σκοπούς. Αρχικά η μονάδα να είναι «ενεργή» (κοντά στο +1) όταν δίνονται οι «σωστές» εισοδοί και «ανενεργή» όταν δίνονται οι λανθασμένες. Πρέπει επίσης η ενεργοποίηση να είναι μη γραμμική, αλλιώς το συνολικό νευρωνικό δίκτυο θα εκφυλίζεται σε μια απλή γραμμική συνάρτηση. Δυο επιλογές για την συνάρτηση g είναι η συνάρτηση κατωφλίου (threshold) και η σιγμοειδής συνάρτηση (sigmoid function). Η σιγμοειδής συνάρτηση έχει το πλεονέκτημα να είναι διαφορίσιμη, γεγονός σημαντικό για τον αλγόριθμο μάθησης των βαρών. Σημαντικό είναι επίσης, ότι, και οι δύο συναρτήσεις έχουν κατώφλι (είτε άκαμπτο είτε εύκαμπτο). Στο μηδέν το βάρος πόλωσης $W_{0,i}$, ορίζει το πραγματικό κατώφλι για την μονάδα, με την έννοια ότι η μονάδα ενεργοποιείται όταν το σταθμισμένο άθροισμα των εισόδων $\sum_{j=0}^n W_{j,i} a_j$ (με εξαίρεση της εισόδου πόλωσης) υπερβαίνει το $W_{0,i}$. (Russel Stuart, 2010)



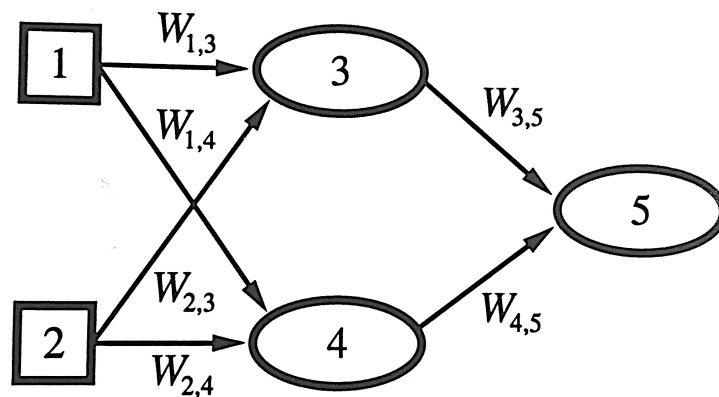
Εικόνα 2. 2: (α) Η συνάρτηση ενεργοποίησης κατωφλίου η οποία έχει έξοδο 1 όταν η είσοδος είναι θετική και 0 διαφορετικά (β) η σιγμοειδής συνάρτηση

2.2 ΔΟΜΕΣ ΔΙΚΤΥΟΥ

Υπάρχουν δυο κύριες κατηγορίες δομών νευρωνικών δικτύων: τα μη κυκλικά ή δίκτυα με προς τα εμπρός τροφοδότηση του σήματος (feed-forward networks), και τα κυκλικά ή αναδρομικά δίκτυα (recurrent networks). Ένα δίκτυο με προς τα εμπρός τροφοδότηση του σήματος αναπαριστά μια συνάρτηση των τρεχουσών εισόδων του, με συνέπεια, να μην έχει άλλη εσωτερική κατάσταση πέρα από τα ίδια τα βάρη. Για να γίνει αυτό πιο κατανοητό, θεωρείται το απλό δίκτυο που παρουσιάζεται στην Εικόνα 2.3. Το δίκτυο αυτό αποτελείται από δυο μονάδες εισόδου, δυο κρυφές μονάδες και μια μονάδα εξόδου (για χάρη απλότητας έχουν παραλειφθεί σε αυτό το παράδειγμα οι μονάδες πόλωσης). Με δεδομένο ένα διάνυσμα εισόδου $x=(x_1, x_2)$, οι ενεργοποιήσεις των μονάδων εισόδου ορίζονται στο $(a_1, a_2)=(x_1, x_2)$ και το δίκτυο υπολογίζει:

$$a_5 = g(W_{3,5}a_3 + W_{4,5}a_4) = g(W_{3,5}g(W_{1,3}a_1 + W_{2,3}a_2) + W_{4,5}g(W_{1,4}a_1 + W_{2,4}a_2)) \quad (2.1)$$

Αυτό σημαίνει ότι, με την έκφραση της εξόδου κάθε κρυφής μονάδας ως συνάρτηση των εισόδων της, δείχθηκε ότι η έξοδος του συνόλου του δικτύου, a_5 , είναι συνάρτηση των εισόδων του δικτύου. Επιπλέον γίνεται προφανές ότι τα βάρη στο δίκτυο λειτουργούν ως παράμετροι αυτής της συνάρτησης. Αν οι παράμετροι αυτοί μετονομαστούν σε W , το δίκτυο υπολογίζει μια συνάρτηση $h_w(x)$. Αν αλλαχθούν τα βάρη, αλλάζει με την σειρά της η συνάρτηση που αντιπροσωπεύει το δίκτυο. Αυτός είναι και ο τρόπος με τον οποίο συμβαίνει η μάθηση στα νευρωνικά δίκτυα.

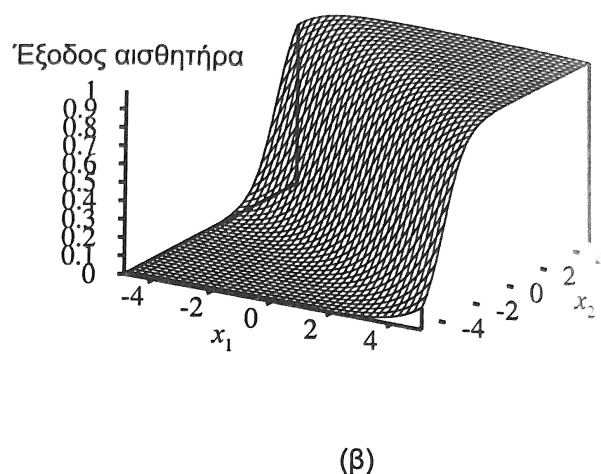
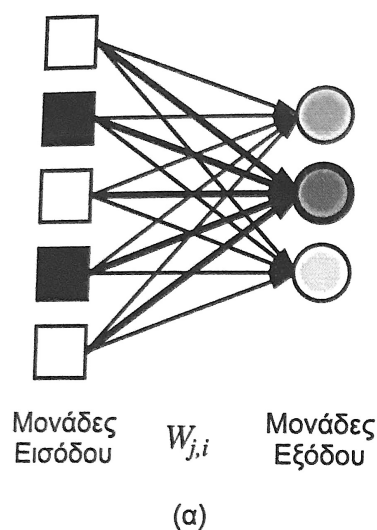


Εικόνα 2. 3 Ένα απλό νευρωνικό δίκτυο με δύο εισόδους, ένα κρυφό επίπεδο δύο μονάδων και μία έξοδο.

Το νευρωνικό δίκτυο μπορεί να χρησιμοποιηθεί για ταξινόμηση ή παλινδρόμηση. Τα δίκτυα με προς τα εμπρός τροφοδότηση του σήματος είναι συνήθως διατεταγμένα σε επίπεδα (layers), έτσι ώστε κάθε μονάδα να λαμβάνει είσοδο μόνο από τις μονάδες στο αμέσως προηγούμενο επίπεδο.

2.2.1 ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ ΕΝΟΣ ΕΠΙΠΕΔΟΥ ΜΕ ΠΡΟΣ ΤΑ ΕΜΠΡΟΣ ΤΡΟΦΟΔΟΤΗΣΗ ΤΟΥ ΣΗΜΑΤΟΣ (perceptrons)

Ένα δίκτυο του οποίου όλες οι εισόδοι είναι άμεσα συνδεδεμένες στις εξόδους ονομάζεται νευρωνικό δίκτυο ενός επιπέδου. (single layer neural network) ή δίκτυο αισθητήρα.



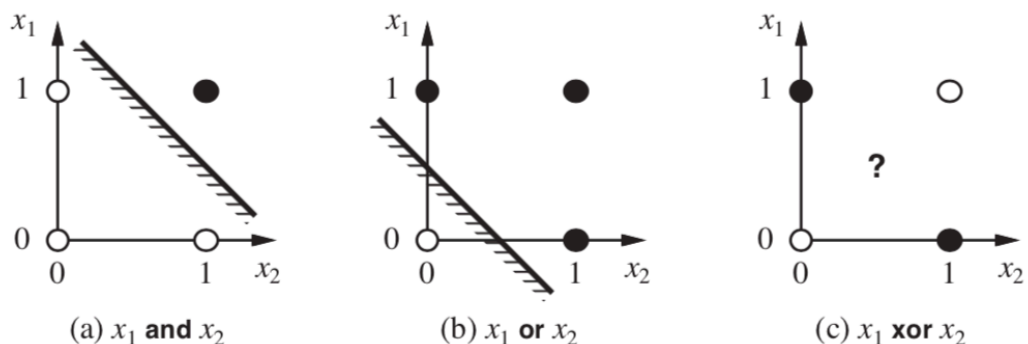
Εικόνα 2. 4(α) Ένα δίκτυο το οποίο αποτελείται από τρεις μονάδες εξόδου που μοιράζονται πέντε εισόδους. (β) Γράφημα της εξόδου μιας μονάδας αισθητήρα δύο εισόδων με σιγμοειδή ενεργοποίηση. (Russel Stuart, 2010)

Θα ακολουθήσει μια σύντομη περιγραφή του χώρου υποθέσεων που μπορεί να αναπαραστήσει ένας αισθητήρας. Με μία συνάρτηση ενεργοποίησης κατωφλίου, θεωρείται ότι ο αισθητήρας αναπαριστά μια «Boolean» συνάρτηση. Εκτός από τις στοιχειώδεις Boolean συναρτήσεις AND, OR, NOT, ένας αισθητήρας μπορεί να αναπαραστήσει με σαφήνεια και αρκετά «πολύπλοκες» συναρτήσεις. Για παράδειγμα, η συνάρτηση πλειοψηφίας, η οποία έχει έξοδο 1 μόνο εάν περισσότερες από τις μισές των n εισόδων της έχουν τιμή 1, δύναται να αναπαρασταθεί από έναν αισθητήρα με $W_i=1$ και κατώφλι $W_0=n/2$. Από την άλλη υπάρχουν και πολλές Boolean συναρτήσεις που δεν γίνεται να αναπαρασταθούν από τον αισθητήρα κατωφλίου. Εξετάζοντας την εξίσωση 2.2, γίνεται φανερό ότι ο αισθητήρας επιστρέφει την τιμή 1 αν και μόνο αν το

σταθμισμένο άθροισμα των εισόδων του (συμπεριλαμβανομένης και της πόλωσης) είναι θετικό:

$$\sum_{j=0}^n W_j x_j > 0 \text{ ή } W * x > 0 \quad (2.4)$$

Τώρα η εξίσωση $W * x = 0$ ορίζει ένα υπερεπίπεδο στο χώρο εισόδου, έτσι ώστε ο αισθητήρας επιστρέφει την τιμή 1 αν και μόνο αν η είσοδος βρίσκεται στην μία πλευρά αυτού του υπερεπιπέδου. Για τον λόγο αυτό ο αισθητήρας κατωφλίου ονομάζεται και γραμμικός διαχωριστής. Οι εικόνες 2.5(a) και 2.5(b) δείχνουν αυτό το υπερεπίπεδο (μια ευθεία γραμμή, στις δύο διαστάσεις) για τις αναπαραστάσεις αισθητήρα των συναρτήσεων AND και OR με δύο εισόδους. Οι μαύρες κουκκίδες δείχνουν ένα σημείο στο χώρο εισόδου όπου η τιμή της συνάρτησης είναι 1, ενώ οι λευκές κουκκίδες υποδηλώνουν ένα σημείο στο οποίο η τιμή είναι 0. Ο αισθητήρας μπορεί να αναπαραστήσει αυτές τις συναρτήσεις επειδή υπάρχει μια ευθεία στο χώρο που διαχωρίζει όλες τις λευκές από τις μαύρες κουκκίδες. Οι συναρτήσεις αυτές ονομάζονται και γραμμικά διαχωρίσιμες. (linearly separable) Η εικόνα 2.5(c) δείχνει ένα παράδειγμα συνάρτησης που δεν είναι γραμμικά διαχωρίσιμη τη συνάρτηση XOR. Προφανώς δεν υπάρχει τρόπος ώστε να μπορεί ένας αισθητήρας κατωφλίου να διδαχτεί αυτή την συνάρτηση. Γενικά, οι αισθητήρες κατωφλίου μπορούν να αναπαραστήσουν μόνο γραμμικά διαχωρίσιμες συναρτήσεις. Οι σιγμοειδείς αισθητήρες έχουν παρόμοιους περιορισμούς με την έννοια ότι αναπαριστούν μόνο «εύκαμπτους» γραμμικούς διαχωριστές.



Εικόνα 2. 5 Γραμμική διαχωρισιμότητα σε αισθητήρες κατωφλίου

Παρά την περιορισμένη τους εκφραστική δύναμη, οι αισθητήρες κατωφλίου διατηρούν ορισμένα σημαντικά πλεονεκτήματα. Πιο συγκεκριμένα, υπάρχει ένας απλός αλγόριθμος μάθησης που προσαρμόζει έναν αισθητήρα κατωφλίου σε οποιοδήποτε γραμμικά διαχωρίσιμο σύνολο εκπαίδευσης. Η κεντρική ιδέα πίσω από αυτόν τον αλγόριθμο όπως και από τους περισσότερους αλγορίθμους μάθησης στα νευρωνικά δίκτυα, είναι η προσαρμογή των βαρών του δικτύου ώστε να ελαχιστοποιηθεί κάποιο μέτρο του σφάλματος στο σύνολο εκπαίδευσης. Κατά συνέπεια η μάθηση διατυπώνεται ως αναζήτηση βελτιστοποίησης στο χώρο βαρών (weight space). Η «κλασσική» μέτρηση

του σφάλματος είναι το άθροισμα των τετραγώνων των σφαλμάτων (sum of squared errors). Το τετράγωνο του σφάλματος για ένα και μόνο παράδειγμα εκπαίδευσης με είσοδο x και έξοδο y γράφεται ως:

$$E = \frac{1}{2} Err^2 = \frac{1}{2} (y - h_w(x))^2 \quad (2.5)$$

όπου $h_w(x)$ είναι η έξοδος του αισθητήρα.

Για να μειωθεί το τετράγωνο των σφαλμάτων, χρησιμοποιείται η μέθοδος κατάβασης πλαγιάς (gradient descent), υπολογίζοντας την μερική παράγωγο του E ως προς κάθε βάρος:

$$\begin{aligned} \frac{\partial E}{\partial W_j} &= Err * \frac{\partial Err}{\partial W_j} = \\ Err * \frac{\partial}{\partial W_j} \left[y - g\left(\sum_{j=1}^n W_j x_j\right) \right] &= -Err * g'(in) * x_j \end{aligned} \quad (2.6)$$

όπου g' είναι η πρώτη παράγωγος της συνάρτησης ενεργοποίησης, που για την σιγμοειδή δίνεται από το $g' = g(1-g)$. Στον αλγόριθμο κατάβασης πλαγιάς που ζητούμενο είναι η μείωση του E , το βάρος μετατρέπεται όπως φαίνεται παρακάτω:

$$W_j \leftarrow (W_j + a * Err * g'(in) * x_j) \quad (2.7)$$

όπου a είναι ο ρυθμός μάθησης (learning rate). Ο πλήρης αλγόριθμος παρουσιάζεται στην Εικόνα 2.6. Διατρέχει ένα-ένα τα παραδείγματα εκπαίδευσης στο δίκτυο, ρυθμίζοντας ελαφρώς τα βάρη μετά από κάθε επανάληψη ώστε να μειωθεί το σφάλμα. Κάθε κύκλος επαναλήψεων ονομάζεται εποχή (epoch). Οι εποχές επαναλαμβάνονται μέχρι την επίτευξη κάποιου κριτηρίου παύσης-συνήθως, ότι οι αλλαγές βάρους έχουν γίνει πολύ μικρές. Άλλες μέθοδοι υπολογίζουν την κλίση για όλο το σύνολο εκπαίδευσης προσθέτοντας όλες τις συνεισφορές κλίσης στην εξίσωση 2.7 πριν από την ενημέρωση των βαρών. Η μέθοδος στοχαστικής κλίσης (stochastic gradient) επιλέγει με τυχαίο τρόπο παραδείγματα από το σύνολο εκπαίδευσης, αντί να κάνει κύκλους μέσα σε αυτά.

```

function PERCEPTRON-LEARNING (παράδειγματα, δίκτυο) returns μια υπόθεση αισθητήρα
inputs: παράδειγματα, ένα σύνολο παραδειγμάτων, το καθένα με είσοδο  $\mathbf{x} = x_1, \dots, x_n$  και έξοδο  $y$ 
           δίκτυο, ένας αισθητήρας με βάρη  $W_j, j = 0, \dots, n$ , και συνάρτηση ενεργοποίησης  $g$ 

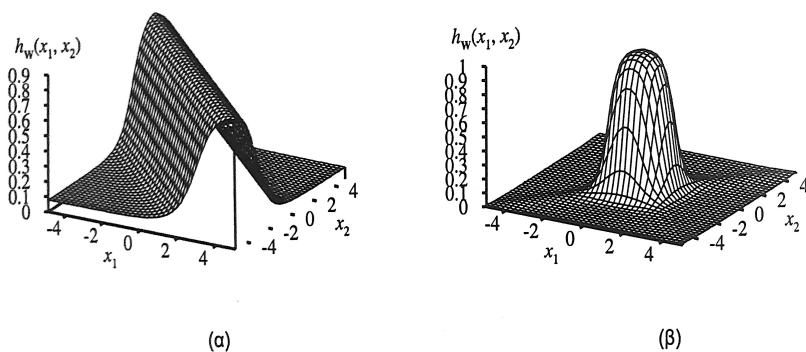
repeat
  for each  $e$  in παράδειγματα do
     $in \leftarrow \sum_{j=0}^n W_j x_j[e]$ 
     $Err \leftarrow y[e] - g(in)$ 
     $W_j \leftarrow W_j + \alpha \times Err \times g'(in) \times x_j[e]$ 
until να ικανοποιηθεί ένα κριτήριο παύσης
return NEURAL-NET-HYPOTHESIS(δίκτυο)

```

Εικόνα 2. 6 Ο αλγόριθμος μάθησης κατάβασης πλαγιάς για αισθητήρες, αν θεωρήσουμε μια διαφορίσιμη συνάρτηση ενεργοποίησης g .

2.2.2 ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ ΠΟΛΛΩΝ ΕΠΙΠΕΔΩΝ ΜΕ ΠΡΟΣ ΤΑ ΕΜΠΡΟΣ ΤΡΟΦΟΤΗΣΗ ΤΟΥ ΣΗΜΑΤΟΣ

Στο σημείο αυτό θα παρουσιαστούν συνοπτικά δίκτυα με κρυφές μονάδες. Η πιο συνηθισμένη περίπτωση περιλαμβάνει ένα και μόνο κρυφό επίπεδο, όπως στην Εικόνα 2.8. Το πλεονέκτημα της προσθήκης κρυφών επιπέδων είναι η αύξηση του μεγέθους του χώρου υποθέσεων που μπορεί να αναπαραστήσει το δίκτυο. (Russel Stuart, 2010) Έστω ότι κάθε κρυφή μονάδα είναι ένας αισθητήρας που αναπαριστά μια εύκαμπτη συνάρτηση κατωφλίου στο χώρο εισόδου όπως παρουσιάζεται στην εικόνα 2.4 (β). Στην συνέχεια, έστω ότι η μονάδα εξόδου είναι ένα εύκαμπτο κατώφλι, γραμμικός συνδυασμός πολλών τέτοιων συναρτήσεων. Για παράδειγμα εάν προστεθούν δυο συναρτήσεις εύκαμπτου κατωφλίου με αντίθετη κατεύθυνση και βρεθεί το κατώφλι του αποτελέσματος, είναι δυνατόν να εξαχθεί μια συνάρτηση «κορυφογραμμής» όπως φαίνεται στην Εικόνα 2.7 (α). Εάν συνδυαστούν δυο τέτοιες κορυφογραμμές σε ορθή γωνία μεταξύ τους (δηλαδή αν συνδυαστούν οι έξοδοι τεσσάρων κρυφών μονάδων), δημιουργείται μια προεξοχή όπως φαίνεται στην Εικόνα 2.7 (β).

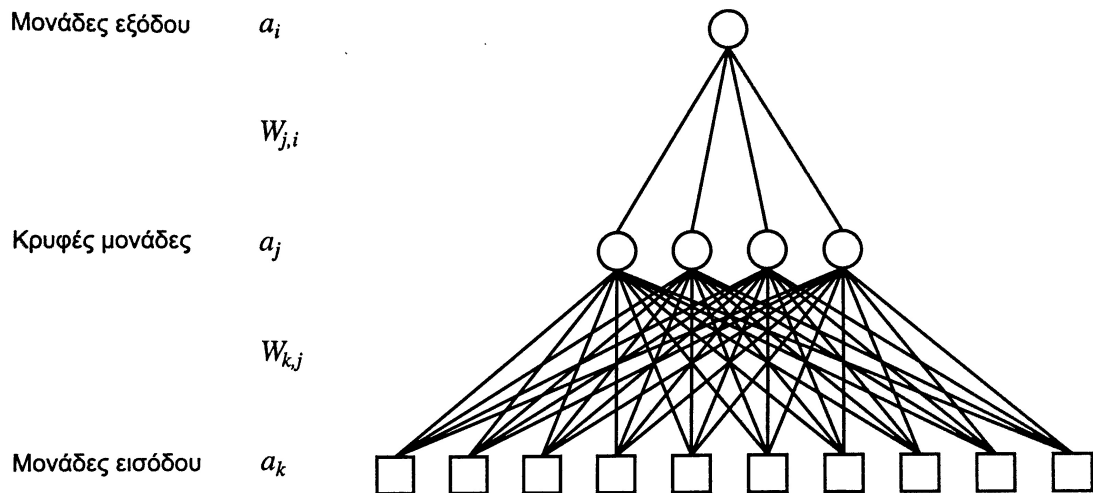


Εικόνα 2. 7 (α) Το αποτέλεσμα του συνδυασμού δυο συναρτήσεων εύκαμπτου κατωφλίου με αντίθετη κατεύθυνση για την δημιουργία μιας κορυφογραμμής. (β) Το αποτέλεσμα συνδυασμού δυο κορυφογραμμών για την δημιουργία μιας προεξοχής. (Russel Stuart, 2010)

Σε περίπτωση που υπάρχουν περισσότερες κρυφές μονάδες, είναι δυνατόν να δημιουργηθούν περισσότερες προεξοχές διαφορετικού μεγέθους και σε διαφορετικές θέσεις. Στην πραγματικότητα, με ένα και μόνο επαρκώς μεγάλο κρυφό επίπεδο είναι δυνατή η αναπαράσταση οποιασδήποτε συνεχής συνάρτησης των εισόδων με αυθαίρετη

ακρίβεια, με δύο επίπεδα μπορούν να αναπαρασταθούν ακόμα και οι ασυνεχείς συναρτήσεις.

Οι αλγόριθμοι μάθησης για δίκτυα πολλών επιπέδων είναι παρόμοιοι με τον αλγόριθμο μάθησης αισθητήρα όπως παρουσιάστηκε στην Εικόνα 2.6. Μια μικρή διαφορά είναι η δυνατότητα πολλών εξόδων, λαμβάνοντας έτσι το διάνυσμα $h_w(x)$ αντί για μια και μόνο τιμή και κάθε παράδειγμα έχει διάνυσμα εξόδου y . Η κύρια διαφορά είναι ότι, ενώ το σφάλμα $y-h_w$ στο επίπεδο εξόδου είναι ξεκάθαρο, το σφάλμα στο κρυφό επίπεδο είναι ασαφές καθώς τα δεδομένα εκπαίδευσης δεν δείχνουν την τιμή που θα πρέπει να έχουν οι κρυφοί κόμβοι. Αποδεικνύεται ότι το σφάλμα εξόδου μπορεί να οπισθοδιαδοθεί (back-propagate) από το επίπεδο εξόδου στα κρυφά επίπεδα. Η διαδικασία αυτή, της οπισθοδιάδοσης προκύπτει άμεσα από την διαφορίση της κλίσης του συνολικού σφάλματος. (Russel Stuart, 2010)



Εικόνα 2. 8 Ένα νευρωνικό δίκτυο πολλών επιπέδων με ένα κρυφό επίπεδο και 10 εισόδους. (Russel Stuart, 2010)

Στο επίπεδο εξόδου ο κανόνας ενημέρωσης του βάρους είναι πανομοιότυπος με την Εξίσωση 2.7. Έχοντας πολλές μονάδες εξόδου, έστω E_{out_i} η i -οστή συνιστώσα του διανύσματος σφάλματος $y-h_w$. Είναι επίσης χρήσιμο να οριστεί ένα τροποποιημένο σφάλμα $\Delta_i = E_{\text{out}_i} * g'(in_i)$, έτσι ώστε ο κανόνας ενημέρωσης βάρους να γίνει:

$$W_{j,i} \leftarrow W_{j,i} + (a * a_j \Delta_i) \quad (2.8)$$

Για την ενημέρωση των συνδέσεων μεταξύ των μονάδων εισόδου και των κρυφών μονάδων, είναι απαραίτητο να οριστεί μια μεταβλητή αντίστοιχη με τον όρο σφάλματος των κόμβων εξόδου. Εδώ είναι που θα γίνει και η οπισθοδιάδοση του σφάλματος. Η βασική αρχή είναι ότι ο κρυφός κόμβος j είναι «υπεύθυνος» για κλάσμα του σφάλματος Δ_i σε κάθε έναν από τους κόμβους με τους οποίους συνδέεται. Έτσι, οι τιμές Δ_i διαιρούνται ανάλογα με την ισχύ της σύνδεσης μεταξύ του κρυφού κόμβου και του κόμβου εξόδου και στην συνέχεια διαδίδονται προς τα πίσω για να δώσουν τις τιμές του Δ_i που αντιστοιχούν στο κρυφό επίπεδο. Ο κανόνας διάδοσης για τις τιμές του Δ είναι ο:

$$\Delta_j = g'(in_j) \sum_i W_{j,i} \Delta_i \quad (2.9)$$

Έτσι ο κανόνας ενημέρωσης βάρους για τα βάρη μεταξύ των εισόδων και του κρυφού επιπέδου είναι σχεδόν πανομοιότυπος με τον κανόνα ενημέρωσης για το επίπεδο εξόδου

$$W_{k,j} \leftarrow W_{k,j} + (a * a_k \Delta_j) \quad (2.10)$$

Η διαδικασία της οπισθοδιάδοσης δύναται να συνοψισθεί ως εξής: (Russel Stuart, 2010)

- Υπολογίζονται οι τιμές του Δ για τις μονάδες εξόδου χρησιμοποιώντας το παρατηρηθέν σφάλμα.

- Ξεκινώντας με το επίπεδο εξόδου, επαναλαμβάνεται η παρακάτω διαδικασία για κάθε επίπεδο του δικτύου μέχρι το πρώτο κρυφό επίπεδο.
 - Οι τιμές του Δ διαδίδονται πίσω στο προηγούμενο επίπεδο
 - Τα βάρη μεταξύ των δύο επιπέδων ενημερώνονται

```

function BACK-PROP-LEARNING(παραδείγματα, δίκτυο) returns ένα νευρωνικό δίκτυο
  inputs: παραδείγματα, ένα σύνολο παραδειγμάτων, κάθε ένα με διάνυσμα εισόδου  $x$  και διάνυσμα εξόδου  $y$ 
           δίκτυο, ένα δίκτυο πολλών επιπέδων με  $L$  επίπεδα, βάρη  $W_{j,i}$ , συνάρτηση ενεργοποίησης  $g$ 

  repeat
    for each  $e$  in παραδείγματα do
      for each κόμβο  $j$  στο επίπεδο εισόδου do  $a_j \leftarrow x_j[e]$ 
      for  $l = 2$  to  $L$  do
         $in_i \leftarrow \sum_j W_{j,i} a_j$ 
         $a_i \leftarrow g(in_i)$ 
      for each κόμβο  $i$  στο επίπεδο εξόδου do
         $\Delta_i \leftarrow g'(in_i) \times (y_i[e] - a_i)$ 
      for  $l = L - 1$  to  $1$  do
        for each κόμβο  $j$  στο επίπεδο  $l$  do
           $\Delta_j \leftarrow g'(in_j) \sum_i W_{j,i} \Delta_i$ 
          for each κόμβο  $i$  στο επίπεδο  $l + 1$  do
             $W_{j,i} \leftarrow W_{j,i} + \alpha \times a_j \times \Delta_i$ 
      until να ικανοποιηθεί κάποιο κριτήριο παύσης
  return NEURAL-NET-HYPOTHESIS(δίκτυο)

```

Εικόνα 2. 9 Ο αλγόριθμος οπισθοδιάδοσης για μάθηση σε δίκτυα πολλών επιπέδων

Στην συνέχεια θα παρατεθούν οι εξισώσεις οπισθοδιάδοσης από τις βασικές αρχές. Το τετράγωνο των σφαλμάτων σε ένα και μόνο παράδειγμα ορίζεται ως:

$$E = \frac{1}{2} \sum_i (y_i - a_i)^2 \quad (2.11)$$

όπου το άθροισμα εφαρμόζεται στους κόμβους του επιπέδου εξόδου. Για να προσδιοριστεί η κλίση ως προς ένα συγκεκριμένο βάρος $W_{j,i}$ στο επίπεδο εξόδου, χρειάζεται μόνο η ανάπτυξη της ενεργοποίησης a_i καθώς οι υπόλοιποι του αθροίσματος δεν επηρεάζονται από το $W_{j,i}$.

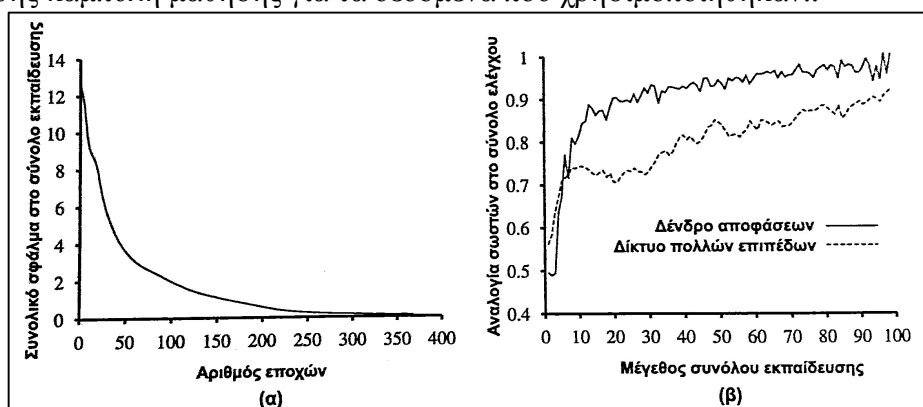
$$\begin{aligned}
 \frac{\partial E}{\partial W_{j,i}} &= -(y_i - a_i) \frac{\partial a_i}{\partial W_{j,i}} = -(y_i - a_i) \frac{\partial g(in_i)}{\partial W_{j,i}} \\
 &= -(y_i - a_i) g'(in_i) \frac{\partial (in_i)}{\partial W_{j,i}} = -(y_i - a_i) g'(in_i) \frac{\partial}{\partial W_{j,i}} \left(\sum_j W_{j,i} a_j \right) \\
 &= -(y_i - a_i) g'(in_i) a_j = -a_j \Delta_i.
 \end{aligned} \quad (2.12)$$

Με το Δ_i να ορίζεται όπως και προηγουμένως. Για να προσδιοριστεί η κλίση ως προς τα βάρη $W_{k,j}$ που συνδέουν το επίπεδο εισόδου με το κρυφό επίπεδο, πρέπει να διατηρηθεί ολόκληρη η άθροιση πάνω στο i καθώς η κάθε τιμή της εξόδου a_i μπορεί να επηρεαστεί από τις αλλαγές στο $W_{k,j}$. Πρέπει επίσης να αναπτυχθούν οι ενεργοποιήσεις a_j .

$$\begin{aligned}
\frac{\partial E}{\partial W_{kj}} &= -\sum_i (y_i - a_i) \frac{\partial a_i}{\partial W_{k,j}} = -\sum_i (y_i - a_i) \frac{\partial g(in_i)}{\partial W_{k,j}} \\
&= -\sum_i (y_i - a_i) g'(in_i) \frac{\partial in_i}{\partial W_{k,j}} = -\sum_i \Delta_i \frac{\partial}{\partial W_{k,j}} \left(\sum_j W_{j,i} a_j \right) \\
&= -\sum_i \Delta_i W_{j,i} \frac{\partial a_j}{\partial W_{k,j}} = -\sum_i \Delta_i W_{j,i} \frac{\partial g(in_j)}{\partial W_{k,j}} \\
&= -\sum_i \Delta_i W_{j,i} g'(in_j) \frac{\partial in_j}{\partial W_{k,j}} \\
&= -\sum_i \Delta_i W_{j,i} g'(in_j) \frac{\partial}{\partial W_{k,j}} \left(\sum_k W_{k,j} a_k \right) \\
&= -\sum_i \Delta_i W_{j,i} g'(in_j) a_k = -a_k \Delta_j,
\end{aligned}
\tag{2.13}$$

Όπου και πάλι το Δ_j ορίζεται όπως και προηγουμένως. Έτσι προσδιορίζονται οι κανόνες ενημέρωσης. Είναι σαφές ότι η διαδικασία μπορεί να επεκταθεί σε δίκτυα με περισσότερα από ένα κρυφά επίπεδα, γεγονός που αιτιολογεί τον γενικό αλγόριθμο που παρουσιάστηκε στην Εικόνα 2.9

Στην εικόνα 2.10 παρουσιάζονται δυο καμπύλες. Η πρώτη είναι μια καμπύλη εκπαίδευσης (training curve), η οποία δείχνει το μέσο τετράγωνο των σφαλμάτων κατά τη διάρκεια της διαδικασίας ενημέρωσης βάρους σε ένα δεδομένο σύνολο εκπαίδευσης με 100 παραδείγματα. Το συμπέρασμα που εξάγεται είναι ότι το δίκτυο πραγματικά συγκλίνει σε τέλεια προσαρμογή στα δεδομένα εκπαίδευσης. Η δεύτερη καμπύλη είναι η συνήθης καμπύλη μάθησης για τα δεδομένα που χρησιμοποιήθηκαν..



Εικόνα 2. 10 Καμπύλη εκπαίδευσης που δείχνει τη σταδιακή μείωση του σφάλματος, (β) Συγκριτικές καμπύλες μάθησης (Russel Stuart, 2010)

Τα νευρωνικά έχουν φυσικά την δυνατότητα εκτέλεσης πιο πολύπλοκων εργασιών μάθησης, παρόλο που απαιτούνται αρκετές μικρό ρυθμίσεις προκειμένου να είναι σωστή η δομή του δικτύου και να επιτευχθεί σύγκλιση σε κάτι κοντινό του καθολικά βέλτιστου στο χώρο των βαρών.

2.3 ΜΑΘΗΣΗ ΔΟΜΩΝ ΝΕΥΡΩΝΙΚΩΝ

Μέχρι τώρα εξετάστηκαν προβλήματα μάθησης και προσαρμογής βαρών με δεδομένη μια σταθερή δομή δικτύου, θα πρέπει λοιπόν να βρεθεί ένας τρόπος για τον προσδιορισμό της βέλτιστης δομής δικτύου. Αν επιλεγθεί ένα υπερβολικά μεγάλο δίκτυο, θα μπορεί μεν να αποστηθίσει όλα τα παραδείγματα σχηματίζοντας έναν μεγάλο πίνακα αναζήτησης ,αλλά αυτό δεν σημαίνει κατ' ανάγκην ότι θα γενικεύει ικανοποιητικά για

εισόδους που δεν έχει ξανασυναντήσει. Έχει παρατηρηθεί ότι τα πολύ μεγάλα δίκτυα πράγματι γενικεύουν καλά εφόσον τα βάρη διατηρούνται μικρά. Αυτός ο περιορισμός διατηρεί τις τιμές ενεργοποίησης στη γραμμική περιοχή της σιγμοειδούς συνάρτησης $g(x)$ όπου το x είναι πολύ κοντά στο μηδέν. Αυτό με τη σειρά του, σημαίνει ότι το δίκτυο συμπεριφέρεται σαν γραμμική συνάρτηση με πολύ λίγες παραμέτρους. Με άλλα λόγια, όπως όλα τα στατιστικά μοντέλα, τα νευρωνικά δίκτυα υπόκεινται σε υπερπροσαρμογή (overfitting) όταν υπάρχουν υπερβολικά πολλές παράμετροι στο μοντέλο. Δίνοντας έμφαση στα πλήρως συνδεδεμένα δίκτυα, οι μόνες επιλογές που πρέπει να γίνουν αφορούν τον αριθμό των κρυφών επιπέδων και τα μεγέθη τους. Η πιο συνηθισμένη προσέγγιση είναι η δοκιμή πολλών και η επιλογή του βέλτιστου. Έτσι οι τεχνικές της διασταυρωμένης επικύρωσης (cross validation) είναι απαραίτητες για την αποφυγή του κρυφοκοιτάγματος (peeking) στο σύνολο ελέγχου. Επιλέγεται δηλαδή η αρχιτεκτονική δικτύου που δίνει την μεγαλύτερη ακρίβεια πρόβλεψης στα σύνολα επικύρωσης.

2.4 Ο ΑΛΓΟΡΙΘΜΟΣ ΕΚΠΑΙΔΕΥΣΗΣ LEVENBERG-MARQUARDT

Ο αλγόριθμος Levenberg–Marquardt, αναπτύχθηκε από τους Kenneth Levenberg και Donald Marquardt και παρέχει μια αριθμητική λύση στο πρόβλημα της ελαχιστοποίησης του E το οποίο περιγράφεται από μια μη-γραμμική συνάρτηση. Ο αλγόριθμος αυτός είναι γρήγορος και παρουσιάζει σταθερή σύγκλιση. Ο αλγόριθμος είναι κατάλληλος για εκπαίδευση μεσαίου μεγέθους δικτύων. (Hao Yu, 2011)

2.4.1 ΠΡΟΕΛΕΥΣΗ

Ο αλγόριθμος Levenberg–Marquardt είναι αποτέλεσμα του συνδυασμού της μεθόδου Steepest Descent και της μεθόδου Gauss–Newton (παράρτημα I κεφάλαιο 7). Διαθέτει το πλεονέκτημα της ταχύτητας από την Gauss–Newton και την σταθερότητα της Steepest Descent. Είναι πιο εύρωστος (robust) από την Gauss–Newton, επειδή, σε πολλές περιπτώσεις συγκλίνει καλά ακόμα και αν η επιφάνεια του σφάλματος περιγράφεται από, πιο πολύπλοκη της τετραγωνικής, μη γραμμική συνάρτηση. Παρόλο που η Levenberg–Marquardt τείνει να είναι πιο αργή από την Gauss–Newton (σε περιπτώσεις σύγκλισης), συγκλίνει πολύ πιο γρήγορα από την Steepest Descent.

Η βασική αρχή της Levenberg–Marquardt είναι ότι εκτελεί μια συνδυασμένη εκπαίδευση. (Hao Yu, 2011) Γύρω από την περιοχή με πολύπλοκη καμπυλότητα, ο αλγόριθμος Levenberg–Marquardt αλλάζει σε Steepest Descent μέχρι η τοπική καμπυλότητα να είναι αρκετή ώστε να γίνει μια τετραγωνική εκτίμηση και τελικά μετατρέπεται σε σχεδόν αλγόριθμο Gauss–Newton (σε περιπτώσεις σύγκλισης), το οποίο επιταχύνει σημαντικά την διαδικασία της σύγκλισης.

2.4.2 Ο ΑΛΓΟΡΙΘΜΟΣ

Για να εξασφαλιστεί ότι ο Εσσιανός πίνακας που προσεγγίζεται (approximated Hessian matrix) $J^T J$ είναι αντιστρέψιμος, στα πλαίσια του αλγορίθμου Levenberg–Marquardt εισάγεται μια καινούργια προσέγγιση στον Εσσιανό πίνακα:

$$H := J^T J + \mu I \quad (2.14)$$

όπου: μ πάντα θετικό και καλείται συντελεστής συνδυασμού (combination coefficient) και I ο ταυτοτικός πίνακας)

Από την εξίσωση 2.14 μπορεί να παρατηρηθεί ότι τα στοιχεία της κύριας διαγώνιου του προς προσέγγιση Εσσιανού πίνακα θα είναι μεγαλύτερα του μηδενός. Έτσι με αυτήν την προσέγγιση, γίνεται σίγουρο ότι ο πίνακας H θα είναι πάντα αντιστρέψιμος.

Συνδυάζοντας τον βασικό κανόνα για τον αλγόριθμο Gauss–Newton:

$$\mathbf{w}_{k+1} = \mathbf{w}_k - \left(J_k^T J_k \right)^{-1} J_k \mathbf{e}_k \quad (2.15)$$

και την 2.14 ο κανόνας για την Levenberg–Marquardt γράφεται ως :

$$\mathbf{w}_{k+1} = \mathbf{w}_k - \left(J_k^T J_k + \mu I \right)^{-1} J_k \mathbf{e}_k \quad (2.16)$$

Όπως αναφέρθηκε παραπάνω, σαν συνδυασμός του Steepest Descent και του Gauss–Newton, ο Levenberg–Marquardt αλλάζει μεταξύ των δυο αλγορίθμων κατά την διάρκεια της εκπαίδευσης. Όταν ο συντελεστής συνδυασμού μ είναι πολύ μικρός (σχεδόν μηδέν), χρησιμοποιείται η Gauss–Newton. Αντίστοιχα όταν ο συντελεστής συνδυασμού είναι πολύ μεγάλος η 2.16 τείνει στην:

$$\mathbf{w}_{k+1} = \mathbf{w}_k - \alpha \mathbf{g}_k \quad (2.17)$$

και χρησιμοποιείται η Steepest Descent με τον μ να ερμηνεύεται ως συντελεστής μάθησης στην Steepest Descent $\alpha=1/\mu$ (Hao Yu, 2011)

2.4.3 ΕΦΑΡΜΟΓΗ ΤΟΥ ΑΛΓΟΡΙΘΜΟΥ

Για να εφαρμοσθεί ο αλγόριθμος στην εκπαίδευση νευρωνικού δικτύου πρέπει να αντιμετωπιστούν δυο προβλήματα: πώς μπορεί να υπολογιστεί ο Ιακωβιανός πίνακας και πώς μπορεί να οργανωθεί επαναληπτικά η διαδικασία εκπαίδευσης για την ενημέρωση των βαρών. Στην συνέχεια αυτού του κεφαλαίου η εφαρμογή του αλγορίθμου Levenberg–Marquardt θα γίνει σε δύο μέρη, αρχικά θα παρουσιαστεί ο τρόπος υπολογισμού του Ιακωβιανού πίνακα και στην συνέχεια ο σχεδιασμός της διαδικασίας εκπαίδευσης.

2.4.4 ΥΠΟΛΟΓΙΣΜΟΣ ΤΟΥ ΙΑΚΩΒΙΑΝΟΥ ΠΙΝΑΚΑ

Στον υπολογισμό που ακολουθεί, τα j και k χρησιμοποιούνται ως δείκτες των νευρώνων, από το 1 έως nn , όπου nn είναι ο αριθμός των νευρώνων που περιέχονται σε μια

τοπολογία δικτύου. Ως i ορίζεται ο δείκτης των νευρώνων εισόδου, από 1 έως n_i , όπου n_i είναι ο αριθμός των εισόδων και μπορεί να διαφέρει για διαφορετικούς νευρώνες.

Σαν εισαγωγή στις βασικές έννοιες της εκπαίδευσης νευρωνικών δικτύων, θεωρείται ένας νευρώνας j με είσοδο n_i , όπως φαίνεται στην Εικόνα 2.11. Εάν ο νευρώνας j βρίσκεται στο πρώτο επίπεδο, όλες οι εισόδους του θα συνδεθούν με τις εξόδους του δικτύου, αλλιώς οι εισόδους του μπορούν να συνδεθούν με τις εξόδους άλλων νευρώνων ή με τις εισόδους των δικτύων εάν επιτρέπονται συνδέσεις μεταξύ των επιπέδων.

Ο κόμβος y είναι μια σημαντική και ευέλικτη ιδέα. Μπορεί να γραφτεί ως $y_{j,i}$, που αναφέρεται στην είσοδο i του νευρώνα j . Μπορεί επίσης να γραφεί ως y_j για να καθορίσει την έξοδο του νευρώνα j .

Στην συνέχεια θα χρησιμοποιείται με τρόπο τέτοιο ώστε, εάν ο κόμβος y έχει ένα δείκτη τότε θα χρησιμοποιείται ως κόμβος εξόδου νευρώνων, αλλά αν έχει δύο δείκτες (νευρώνας και είσοδος), είναι ένας κόμβος εισόδου νευρώνων. (Hao Yu, 2011)

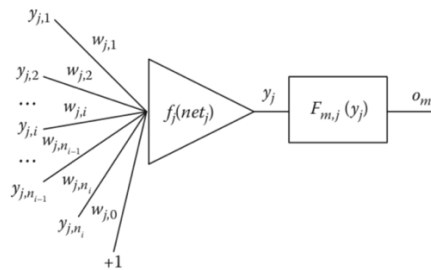
Ο κόμβος εξόδου του νευρώνα j υπολογίζεται χρησιμοποιώντας την

$$y_i = f_i(net_j) \quad (2.18)$$

όπου f η συνάρτηση ενεργοποίησης του νευρώνα j και net_j το άθροισμα των σταθμισμένων εισόδων των κόμβων (sum of weighted inputs nodes of neuron j) του νευρώνα j :

$$net_j = \sum_{i=1}^{n_i} w_{j,i} y_{j,i} + w_{j,0} \quad (2.19)$$

όπου $y_{j,i}$ το i -οστή είσοδος του κόμβου του νευρώνα j σταθμισμένη κατά $w_{j,i}$ και $w_{j,0}$ η τάση στάθμισης του νευρώνα j .



Εικόνα 2. 11 Σύνδεση ενός νευρώνα j με το υπόλοιπο δίκτυο

Χρησιμοποιώντας την 2.19 προκύπτει ότι η παράγωγος του net_j ισούται με:

$$\frac{\partial net_j}{\partial w_{j,i}} = y_{j,i} \quad (2.20)$$

και η κλίση s_j της συνάρτησης ενεργοποίησης f_j είναι

$$s_j = \frac{\partial y_j}{\partial net_j} = \frac{\partial f_j(net_j)}{\partial net_j} \quad (2.21)$$

Μεταξύ του κόμβου εξόδου y_i ενός κρυφού νευρώνα j και της εξόδου o_m του δικτύου υπάρχει μια πολύπλοκη μη-γραμμική σχέση (Εικόνα 2.11)

$$o_m = F_{m,j}(y_i) \quad (2.22)$$

όπου o_m η m -ιοστή έξοδος του δικτύου.

Η πολυπλοκότητα αυτής της μη γραμμικής συνάρτησης $F_{m,j}(y_i)$ εξαρτάται από τον αριθμό των νευρώνων που βρίσκονται μεταξύ του νευρώνα j και της εξόδου του δικτύου m . Τα στοιχεία του Ιακωβιανού πίνακα στην:

$$J = \begin{bmatrix} \frac{\partial e_{1,1}}{\partial w_1} & \frac{\partial e_{1,1}}{\partial w_2} & \dots & \frac{\partial e_{1,1}}{\partial w_N} \\ \frac{\partial e_{1,2}}{\partial w_1} & \frac{\partial e_{1,2}}{\partial w_2} & \dots & \frac{\partial e_{1,2}}{\partial w_N} \\ \dots & \dots & \dots & \dots \\ \frac{\partial e_{1,M}}{\partial w_1} & \frac{\partial e_{1,M}}{\partial w_2} & \dots & \frac{\partial e_{1,M}}{\partial w_N} \\ \dots & \dots & \dots & \dots \\ \frac{\partial e_{p,1}}{\partial w_1} & \frac{\partial e_{p,1}}{\partial w_2} & \dots & \frac{\partial e_{p,1}}{\partial w_N} \\ \frac{\partial e_{p,2}}{\partial w_1} & \frac{\partial e_{p,2}}{\partial w_2} & \dots & \frac{\partial e_{p,2}}{\partial w_N} \\ \dots & \dots & \dots & \dots \\ \frac{\partial e_{p,M}}{\partial w_1} & \frac{\partial e_{p,M}}{\partial w_2} & \dots & \frac{\partial e_{p,M}}{\partial w_N} \end{bmatrix}$$

μπορούν να υπολογιστούν ως:

$$\frac{\partial e_{p,m}}{\partial w_{j,i}} = \frac{\partial (d_{p,m} - o_{p,m})}{\partial w_{j,i}} = -\frac{\partial o_{p,m}}{\partial w_{j,i}} = -\frac{\partial o_{p,m}}{\partial y_j} \frac{\partial y_j}{\partial net_j} \frac{\partial net_j}{\partial w_{j,i}} \quad (2.23)$$

Συνδυάζοντας τις εξισώσεις 2.18 2.19 2.20 2.21 η 2.22 γράφεται ως :

$$\frac{\partial e_{p,m}}{\partial w_{j,i}} = -F'_{mj} s_j y_{j,i} \quad (2.24)$$

όπου F'_{mj} η πρώτη παράγωγος της συνάρτησης που συνδέει τον νευρώνα j και την έξοδο m . Η υπολογιστική διαδικασία για τον προσδιορισμό του Ιακωβιανού πίνακα μπορεί να οργανωθεί σύμφωνα με τις αρχές της οπισθοδιάδοσης των υπολογισμών για πρώτης τάξης αλγόριθμους (όπως στον EBP). Υπάρχουν όμως και σημαντικές διαφορές.

Κατ' αρχήν για κάθε παράδειγμα στον EBP, χρειάζεται μόνο μία διαδικασία οπισθοδιάδοσης, ενώ, για τον Levenberg–Marquardt η διαδικασία αυτή πρέπει να επαναληφθεί για κάθε έξοδο ξεχωριστά ώστε να εξαχθούν συνεχόμενες στήλες του Ιακωβιανού πίνακα.

Μια άλλη διαφορά είναι ότι, η αρχή της οπισθοδιάδοσης της παραμέτρου δ πρέπει να τροποποιηθεί. Στον αλγόριθμο EPB, τα σφάλματα εξόδου συμπεριλαμβάνονται στην δ :

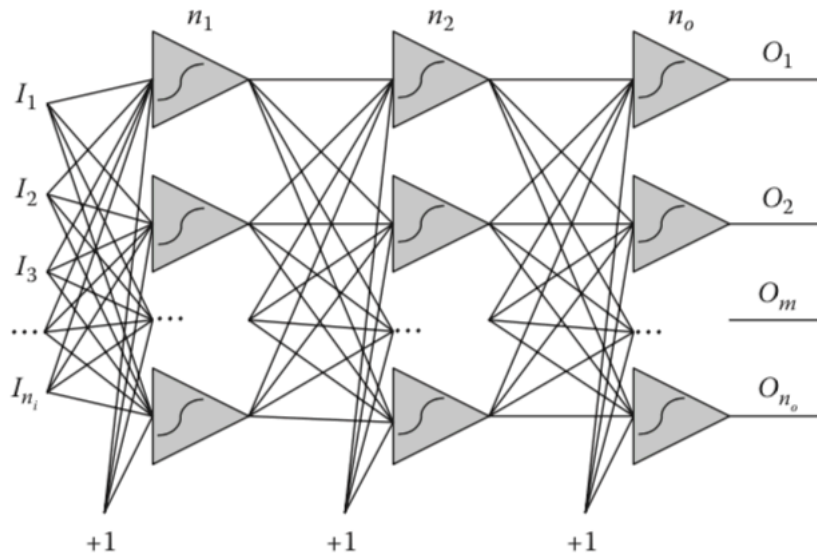
$$\delta_j = s_j \sum_{m=1}^M F'_{mj} e_m \quad (2.25)$$

Στην Levenberg–Marquardt, οι παράμετροι δ υπολογίζονται για κάθε νευρώνα και έξοδο m , ξεχωριστά. Επίσης κατά την οπισθοδιάδοση, τα σφάλματα αντικαθίστανται από μια τιμή μονάδας (unit value).

$$\delta_{m,j} = s_j F'_{mj} \quad (2.26)$$

Συνδυάζοντας τις εξισώσεις 2.24 και 2.25 τα στοιχεία του Ιακωβιανού πίνακα μπορούν να υπολογιστούν από το:

$$\frac{\partial e_{p,m}}{\partial w_{i,j}} = -\delta_{m,j} y_{i,j} \quad (2.27)$$



Εικόνα 2. 12 Δίκτυο αισθητήρων τριών επιπέδων

Στην εξίσωση 2.27 για τον υπολογισμό του Ιακωβιανού υπάρχουν δυο άγνωστοι. Ο κόμβος εισόδου $y_{j,i}$, μπορεί να υπολογιστεί στον προς τα εμπρός υπολογισμό (forward computation) (το σήμα διαδίδεται από τις εισόδους στις εξόδους), ενώ το $\delta_{m,j}$ λαμβάνεται από τον προς τα πίσω υπολογισμό (backward computation), ο οποίος ορίζεται ως σφάλματα που οπισθοδιαδίδονται από νευρώνες εξόδου (επίπεδο εξόδου) σε νευρώνες εισόδου (επίπεδο εισόδου). Στον νευρώνα εξόδου $m(j=m)$, $\delta_{m,j}=s_m$

Για καλύτερη ερμηνεία των προς τα εμπρός και προς τα πίσω υπολογισμών, θεωρείται ένα πολύ επίπεδο δίκτυο τριών επιπέδων (Εικόνα 2.12) σαν παράδειγμα. Για ένα δοσμένο μοτίβο (pattern) ο προς τα εμπρός υπολογισμός μπορεί να συνοψισθεί στα ακόλουθα βήματα: (Hao Yu, 2011)

α) Υπολογίζονται οι τιμές, κλίσεις και όλες οι εξόδους (net values, slopes outputs) για όλους τους νευρώνες του πρώτου επιπέδου:

$$net_j^1 = \sum_{i=1}^{n_i} I_i w_{j,i}^1 + w_{j,0}^1$$

$$y_j^1 = f_j^1(net_j^1)$$

$$s_j^1 = \frac{\partial f_j^1}{\partial net_j^1}$$

(2.28)

όπου:

I_i οι είσοδοι του δικτύου, ο εκθέτης «1» σημαίνει επίπεδο 1 και, j ο δείκτης των νευρών του πρώτου επιπέδου

β) Χρησιμοποιώντας τις εξόδους των νευρώνων του πρώτου επιπέδου σαν εισόδους όλων των νευρώνων του δεύτερου επιπέδου, γίνεται ένας παρόμοιος υπολογισμός των τιμών κλίσεων και εξόδων:

$$net_j^2 = \sum_{i=1}^{n_1} y_i^1 w_{j,i}^2 + w_{j,0}^2$$

$$y_j^2 = f_j^2(net_j^2)$$

$$s_j^2 = \frac{\partial f_j^2}{\partial net_j^2}$$

(2.29)

γ) Χρησιμοποιώντας τις εξόδους των νευρώνων του δευτέρου επιπέδου σαν εισόδους όλων των νευρώνων του τρίτου επιπέδου, γίνεται ένας παρόμοιος υπολογισμός των τιμών κλίσεων και εξόδων:

$$net_j^3 = \sum_{i=1}^{n_2} y_i^2 w_{j,i}^3 + w_{j,0}^3$$

$$o_j = f_j^3(net_j^3)$$

$$s_j^3 = \frac{\partial f_j^3}{\partial net_j^3}$$

(2.30)

Μετά τον προς τα εμπρός υπολογισμό, ο πίνακας των κόμβων (node array) και ο πίνακας των κλίσεων (slope array) μπορούν να προσδιοριστούν για όλους τους νευρώνες με το συγκεκριμένο μοτίβο(given pattern)

Με τα αποτελέσματα από τον προς τα εμπρός υπολογισμό, για μία συγκεκριμένη έξοδο j, ο προς τα πίσω υπολογισμός μπορεί να οργανωθεί ως:

δ) Υπολογίζεται το σφάλμα στην έξοδο j και το αρχικό δ σαν η κλίση της εξόδου j:

$$e_j = d_j - o_j \quad (2.31)$$

$$\delta_{j,j}^3 = s_j^3 \quad (2.32)$$

$$\delta_{j,k}^3 = 0 \quad (2.33)$$

όπου :

d_j : η επιθυμητή έξοδος στην έξοδο j

o_j : η πραγματική έξοδος στην έξοδο j, από τον προς τα εμπρός υπολογισμό

$\delta_{j,j}^3$: η αυτό-οπισθοδιάδοση (self-backpropagation)

$\delta_{j,k}^3$: η οπισθοδιάδοση από άλλους νευρώνες στο ίδιο επίπεδο (επίπεδο εξόδου)

ε) Το δ οπισθοδιαδίδεται από το τις εισόδους του τρίτου επιπέδου στις εξόδους του δευτέρου επιπέδου:

$$\delta_{j,j}^2 = w_{j,k}^3 \delta_{j,j}^3 \quad (2.34)$$

όπου k ο δείκτης των νευρώνων του δευτέρου επιπέδου, από 1 έως n_2 .

στ) Το δ οπισθοδιαδίδεται από το τις εξόδους του δευτέρου επιπέδου στις εισόδους του δευτέρου επιπέδου:

$$\delta_{j,k}^2 = \delta_{j,k}^2 s_k^2 \quad (2.35)$$

όπου k ο δείκτης των νευρώνων του δευτέρου επιπέδου, από 1 έως n_2 .

ζ) Το δ οπισθοδιαδίδεται από το τις εισόδους του δευτέρου επιπέδου στις εξόδους του πρώτου επιπέδου:

$$\delta_{j,k}^1 = \sum_{i=1}^{n_2} w_{j,i}^2 \delta_{j,i}^2 \quad (2.36)$$

όπου k ο δείκτης των νευρώνων του πρώτου επιπέδου, από 1 έως n_1 .

η) Το δ οπισθοδιαδίδεται από το τις εξόδους του πρώτου επιπέδου στις εισόδους του πρώτου επιπέδου:

$$\delta_{j,k}^1 = \delta_{j,k}^1 s_k^1 \quad (2.37)$$

όπου k ο δείκτης των νευρώνων του δευτέρου επιπέδου, από 1 έως n_1 .

Για την οπισθοδιάδοση άλλων εξόδων, επαναλαμβάνονται τα βήματα (δ)-(η). Με το να εφαρμόζεται προς τα εμπρός και προς τα πίσω υπολογισμός, μπορούν να υπολογιστούν ολόκληροι οι πίνακες δ και y για το δοσμένο μοτίβο. Για άλλα μοτίβα, ολόκληρος ο Ιακωβιανός πίνακας μπορεί να προσδιοριστεί, επαναλαμβάνοντας τον προς τα εμπρός και προς τα πίσω υπολογισμό.

```
for all patterns
% Forward computation
for all layers
    for all neurons in the layer
        calculate net;      % Equation (12.28)
        calculate output;  % Equation (12.27)
        calculate slope;   % Equation (12.30)
    end;
end;
%Backward computation
initial delta as slope;
for all outputs
    calculate error;
    for all layers
        for all neurons in the previous layer
            for all neurons in the current layer
                multiply delta through weights
                sum the backpropagated delta at proper nodes
            end;
            multiply delta by slope;
        end;
    end;
end;
```

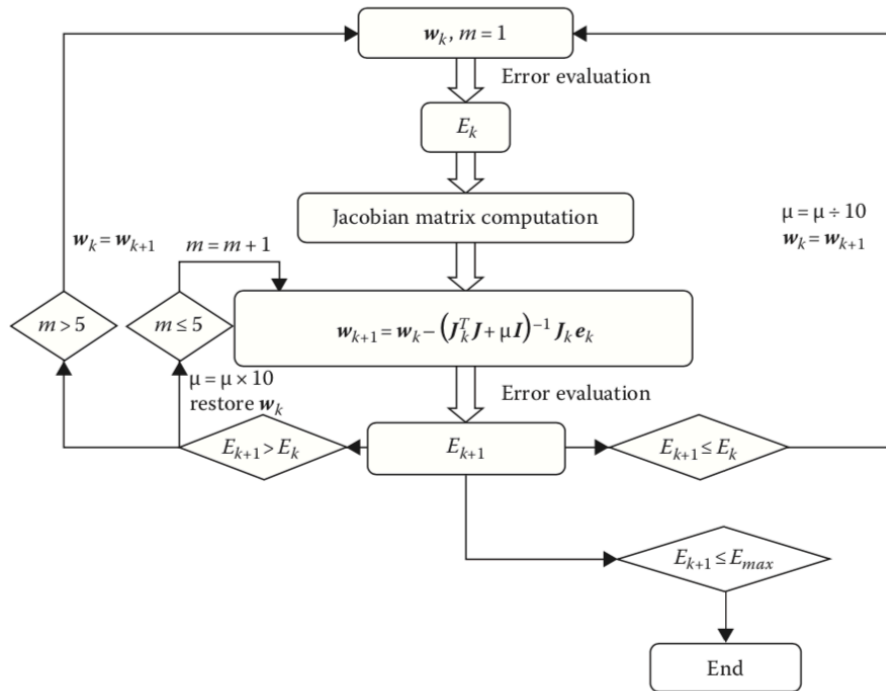
Εικόνα 2. 13 Ψευδο κώδικας για τον προς τα εμπρός και πίσω υπολογισμό

2.5 ΣΧΕΔΙΑΣΜΟΣ ΤΗΣ ΔΙΑΔΙΚΑΣΙΑΣ ΕΚΠΑΙΔΕΥΣΗΣ

Μετά την αναβάθμιση του κανόνα για την Levenberg–Marquardt και τον υπολογισμό του Ιακωβιανού πίνακα, το επόμενο βήμα είναι ο σχεδιασμός της διαδικασίας εκπαίδευσης. Σύμφωνα με τον καινούργιο κανόνα, εάν το σφάλμα μειωθεί, που σημαίνει τιμή σφάλματος μικρότερη από την πιο πρόσφατη τιμή σφάλματος, η τετραγωνική προσέγγισή στην συνάρτηση ολικού σφάλματος λειτουργεί και η σταθερά συνδυασμού μ μπορεί να λάβει μικρότερη τιμή για να μειωθεί η επήρεια της μεθόδου Gauss-Newton (αύξηση στην ταχύτητα). Από την άλλη, εάν το σφάλμα αυξηθεί, που σημαίνει τιμή σφάλματος μεγαλύτερη από την πιο πρόσφατη τιμή σφάλματος, κρίνεται απαραίτητο να βρεθεί περιοχή με καλύτερη καμπυλότητα για την εφαρμογή της τετραγωνικής προσέγγισης, έτσι η σταθερά συνδυασμού λαμβάνει μεγαλύτερες τιμές.

Τελικά, η διαδικασία εκπαίδευσης με την χρήση του αλγορίθμου Levenberg-Marquardt μπορεί να συνοψισθεί ως εξής: (Hao Yu, 2011)

1. Με τα αρχικά βάρη (τυχαία κατασκευασμένα (randomly generated)) εκτιμάται το συνολικό σφάλμα (SSE)
2. Γίνεται ενημέρωση σύμφωνα με την εξίσωση 2.16 για να προσαρμοστούν τα βάρη.
3. Με τα καινούργια βάρη εκτιμάται ξανά το συνολικό σφάλμα
4. Εάν το τρέχον ολικό σφάλμα αυξηθεί λόγω της ενημέρωσης, τότε γίνεται ένα βήμα πίσω (όπως η επιστροφή των τιμών του διανύσματος βαρών στις αμέσως προηγούμενες) και αυξάνεται η τιμή της σταθεράς συνδυασμού μ , με ένα παράγοντα της τάξης του 10 ή κάποιου άλλου παράγοντα. Στην συνέχεια γίνεται επιστροφή στο βήμα 2 και ξαναγίνεται δοκιμή ενημέρωσης των βαρών.
5. Εάν το συνολικό σφάλμα μειωθεί σαν αποτέλεσμα της αναβάθμισης, το βήμα γίνεται δεκτό (οι καινούργιες τιμές του διανύσματος των βαρών γίνονται δεκτές) και η τιμή της σταθεράς συνδυασμού μειώνεται με παράγοντα ανάλογο αυτού που χρησιμοποιήθηκε στο βήμα 4.
6. Επιστροφή στο βήμα 2 με τις νέες τιμές των βαρών και επανάληψη της διαδικασίας μέχρι η τιμή του συνολικού σφάλματος να είναι η επιθυμητή.



Εικόνα 2. 14 Διάγραμμα εκπαίδευσης για τον αλγόριθμο Levenberg–Marquardt

2.6 ΣΥΝΟΨΗ ΓΙΑ ΤΟΝ ΑΛΓΟΡΙΘΜΟ ΕΚΠΑΙΔΕΥΣΗΣ Levenberg–Marquardt

Καταληκτικά, ο αλγόριθμος Levenberg–Marquardt λύνει υπάρχοντα προβλήματα των μεθόδων Gauss-Newton και Steepest Descent για εκπαίδευση νευρωνικών δικτύων, με το συνδυασμό των δύο αλγορίθμων. Θεωρείται ένας από τους πιο αποδοτικούς αλγόριθμους για εκπαίδευση (Hao Yu, 2011).

Ωστόσο υπάρχουν και κάποια ελαττώματα. Ένα από αυτά τα προβλήματα είναι ότι η αντιστροφή του Εσσιανού πίνακα πρέπει να υπολογίζεται κάθε φορά μετά την αναβάθμιση των βαρών, και μπορεί να υπάρξουν αρκετές αναβαθμίσεις σε κάθε επανάληψη.

Για δίκτυα μικρού μεγέθους, ο υπολογισμός είναι αποτελεσματικός, αλλά, για μεγάλα δίκτυα, όπως αυτά που κατασκευάζονται για την αναγνώριση εικόνων η αντιστροφή θα είναι προβληματική με συνέπεια ο χρόνος που εξοικονομείται με την αύξηση της ταχύτητας να χάνεται, στην περίπτωση αυτή ο Levenberg–Marquardt ίσως να είναι πιο αργός ακόμα και από άλλες μεθόδους.

Ακόμα ένα πρόβλημα είναι ότι ο Ιακωβιανός πίνακας πρέπει να αποθηκεύεται για τον υπολογισμό και το μέγεθος του είναι $P \times M \times N$, όπου P ο αριθμός των μοτίβων (patterns) M ο αριθμός των εξόδων και N ο αριθμός των βαρών. Για μεγάλου μεγέθους μοτίβα εκπαίδευσης το κόστος σε μνήμη για την αποθήκευση του Ιακωβιανού πίνακα είναι τόσο μεγάλο που δεν καθιστά τη μέθοδο πρακτική.

Τελικά παρόλα τα προβλήματα που αναφέρθηκαν η Levenberg–Marquardt, για δίκτυα μικρού- μεσαίου μεγέθους, αποτελεί μια πολύ καλή και αποδοτική μέθοδο εκπαίδευσης και για αυτό επιλέχθηκε για την εκπαίδευση των δικτύων που κατασκευάστηκαν για την εργασία αυτή.

ΚΕΦΑΛΑΙΟ 3: ΕΦΑΡΜΟΓΗ ΤΗΣ ΜΕΘΟΔΟΛΟΓΙΑΣ, ΓΙΑ ΤΗΝ ΑΥΤΟΜΑΤΗ ΕΠΙΛΟΓΗ ΠΡΩΤΩΝ ΑΦΙΞΕΩΝ.

3.1 ΕΙΣΑΓΩΓΗ

Για τους σκοπούς αυτής της εργασίας, επεξεργάστηκαν δεδομένα που αντιπροσωπεύουν σεισμικά ίχνη σε μορφή «.dat», μετατράπηκαν σε μορφή πινάκων ώστε να είναι συμβατά για επεξεργασία σε περιβάλλον Matlab και, βάσει θεωρίας, ερμηνεύθηκαν και επιλέχθηκαν τα κελιά των πινάκων που αντιστοιχούν στις θέσεις των πρώτων αφίξεων. Στην συνέχεια χρησιμοποιώντας τις θέσεις των πρώτων αφίξεων (όπως αυτές επιλέχθηκαν από των συγγραφέα), δημιουργήθηκε και εκπαιδεύτηκε νευρωνικό δίκτυο ικανό να αναγνωρίζει με σχετικά μεγάλη ακρίβεια τη θέση και τη χρονική στιγμή των πρώτων αφίξεων. Στη συνέχεια του κεφαλαίου που ακολουθεί παρουσιάζεται μια αναλυτική περιγραφή της μεθοδολογίας για τη διαδικασία που ακολουθήθηκε

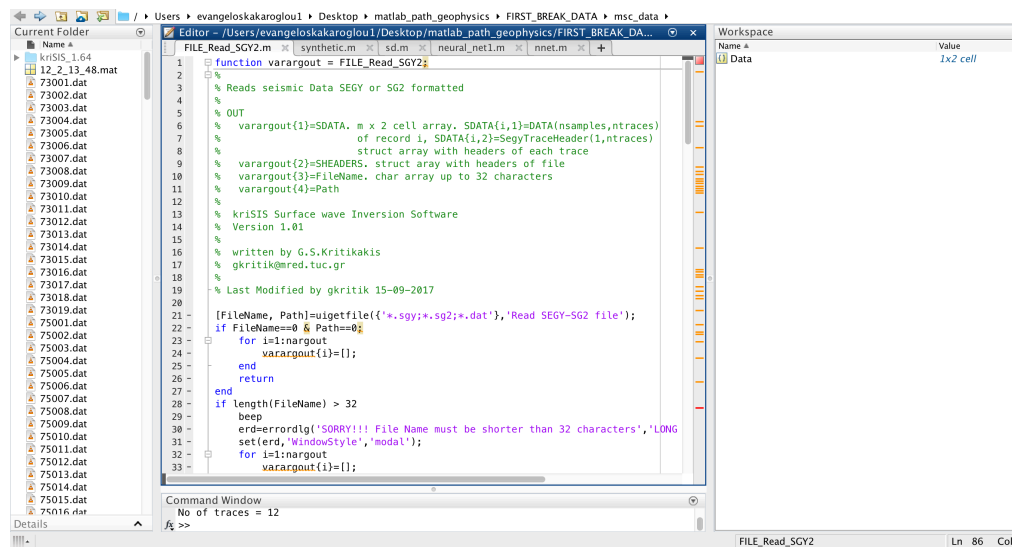
3.2 ΠΡΟΕΛΕΥΣΗ ΤΩΝ ΔΕΔΟΜΕΝΩΝ

Τα δεδομένα σεισμικής τομογραφίας μεταξύ γεωτρήσεων που χρησιμοποιήθηκαν, αποτελούν προϊόν συνεργασίας του Εργαστηρίου Εφαρμοσμένης Γεωφυσικής του Πολυτεχνείου Κρήτης υπό τον καθηγητή Αντώνη Βαφείδη σε συνεργασία με την εταιρία ΓΕΩΤΕΚ Α.Τ.Ε για την επέκταση της γραμμής 2 του αττικού μετρό στην περιοχή Ίλιον-Αργυρούπολη. Τα δεδομένα αυτά προκύπτουν από το έργο για το οποίο σχεδιάστηκε σεισμική γεωφυσική διασκόπηση μεταξύ των προκαθορισμένων σημείων των γεωτρήσεων σε περιοχές που υποδείχθηκαν κατά μήκος της νέας χάραξης για την επέκταση της γραμμής 2 του Αττικού Μετρό στο τμήμα «Ανθούπολη- Ίλιον». Η γεωφυσική αυτή διασκόπηση περιλαμβάνει συνολικά, την διεξαγωγή έντεκα σεισμικών τομογραφιών μεταξύ εννιά γεωτρήσεων. Για κάθε εφαρμογή της μεθόδου της σεισμικής τομογραφίας χρησιμοποιήθηκε ηλεκτρομηχανική σφυρά αμφίδρομης οριζόντιας κρούσης σε ισαποστάσεις των 2 m ως πηγή για την σεισμική διέγερση. Συστοιχία δώδεκα υδροφώνων σε ισαποστάσεις του ενός μέτρου για την καταγραφή των σεισμικών κυμάτων καθώς και σειсмоγράφος δώδεκα καναλιών με εύρος μετατροπέα αναλογικού σε ψηφιακό σήμα (A/D converter) 24 bit. Συνολικά, διασκοπήθηκαν τομές μήκους 225.3 μέτρων (m) και βάθους 520 μέτρων (m), οριοθετώντας μια έκταση σεισμικών τομογραμμάτων που υπερβαίνει τα 117000 m². Η αρχική μορφή των δεδομένων όπως αναφέρθηκε και στην εισαγωγή του κεφαλαίου ήταν .dat. Για την μετατροπή τους χρησιμοποιήθηκε ο αλγόριθμος krisis, ο οποίος αναπτύχθηκε από τον Δρ. Γ. Κρητικάκη.

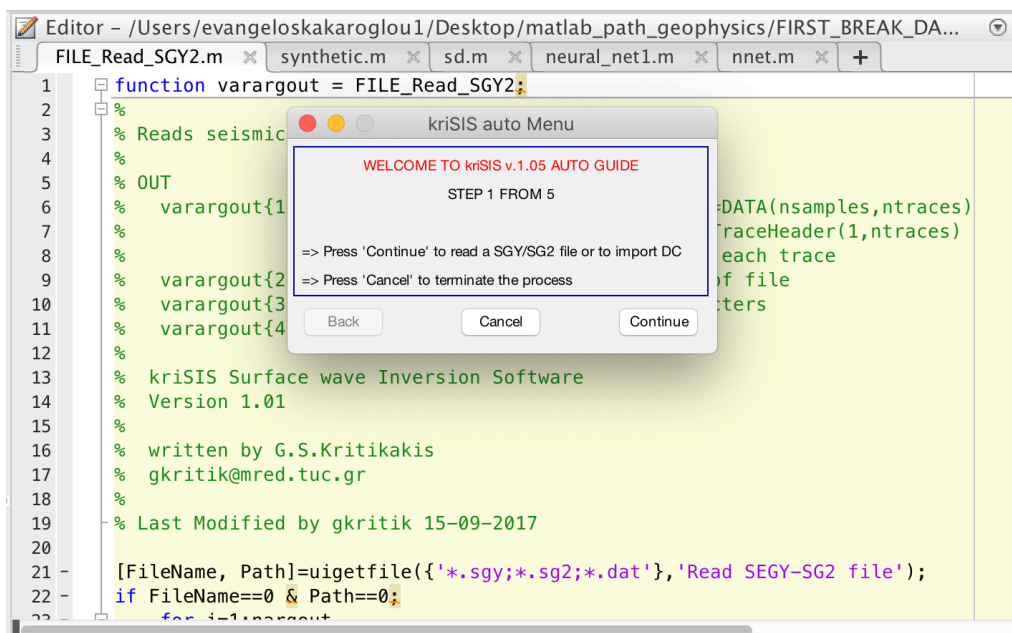
3.3 ΔΙΑΔΙΚΑΣΙΑ ΜΕΤΑΤΡΟΠΗΣ ΤΩΝ ΔΕΔΟΜΕΝΩΝ

Η διαδικασία που ακολουθήθηκε σε αυτό το κομμάτι είναι σχετικά απλή καθώς, αρχικά επιλέγεται το είδος των, προς μετατροπή, δεδομένων (.dat) και στην συνέχεια «τρέχοντας» και ακολουθώντας το interface του αλγορίθμου Krisis (Εικόνα 3.1, Εικόνα 3.2) εισάγονται τα προς μετατροπή δεδομένα, επεξεργάζονται (Εικόνα 3.3, Εικόνα 3.4) και τελικά μετατρέπονται στην επιθυμητή μορφή εξόδου (output), στην προκειμένη

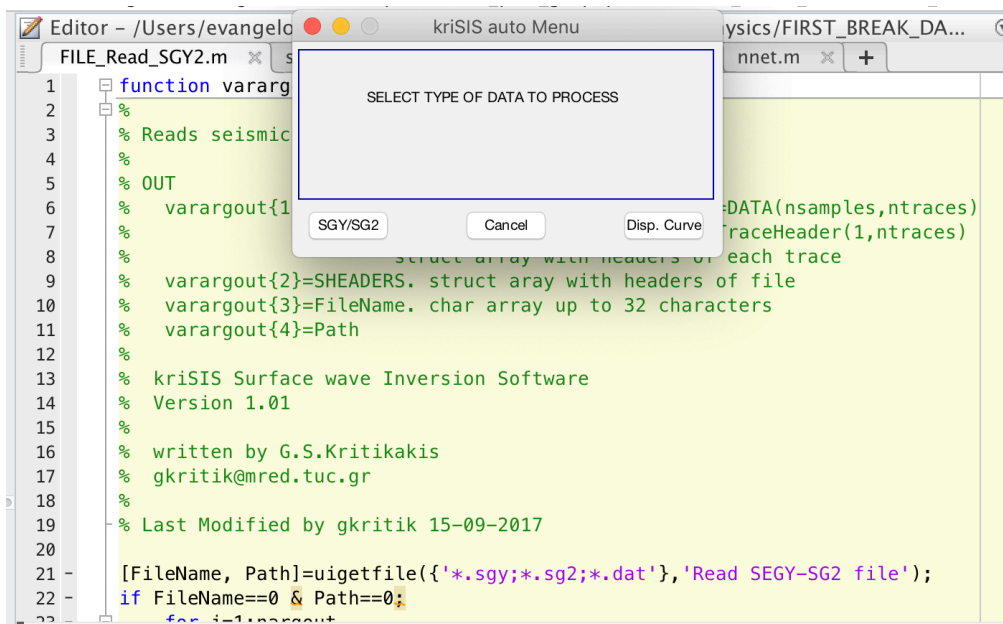
περίπτωση, σε μορφή πινάκων της Matlab (cell array). Κάθε στήλη του πίνακα αντιστοιχεί σε μία θέση πηγής (Εικόνα 3.5).



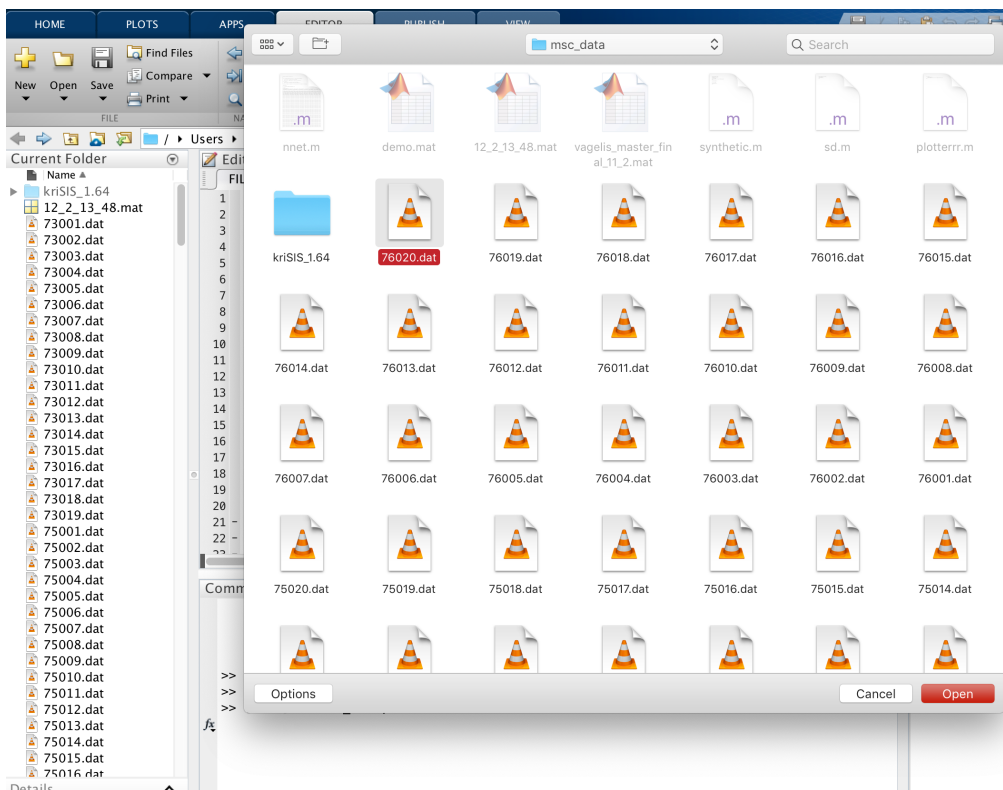
Εικόνα 3. 1 Ο αλγόριθμος Krisis για την μετατροπή των δεδομένων



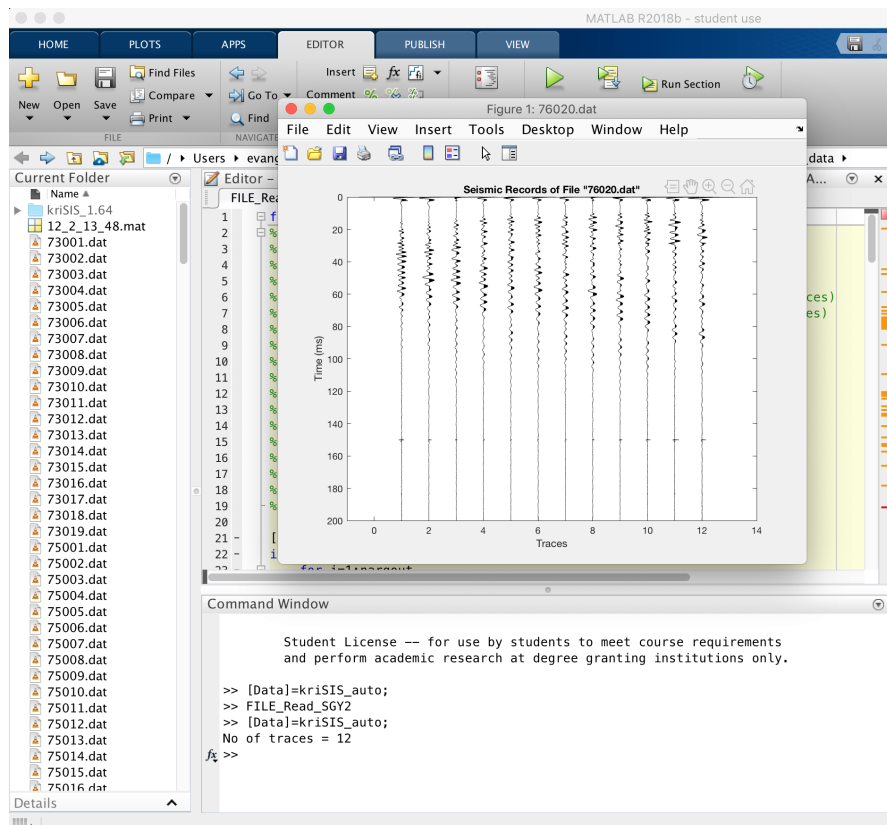
Εικόνα 3. 2 Γραφικό interface(GUI) του αλγορίθμου Krisis



Εικόνα 3. 3 Επιλογή του είδους των δεδομένων για μετατροπή από τον αλγόριθμο Krisis



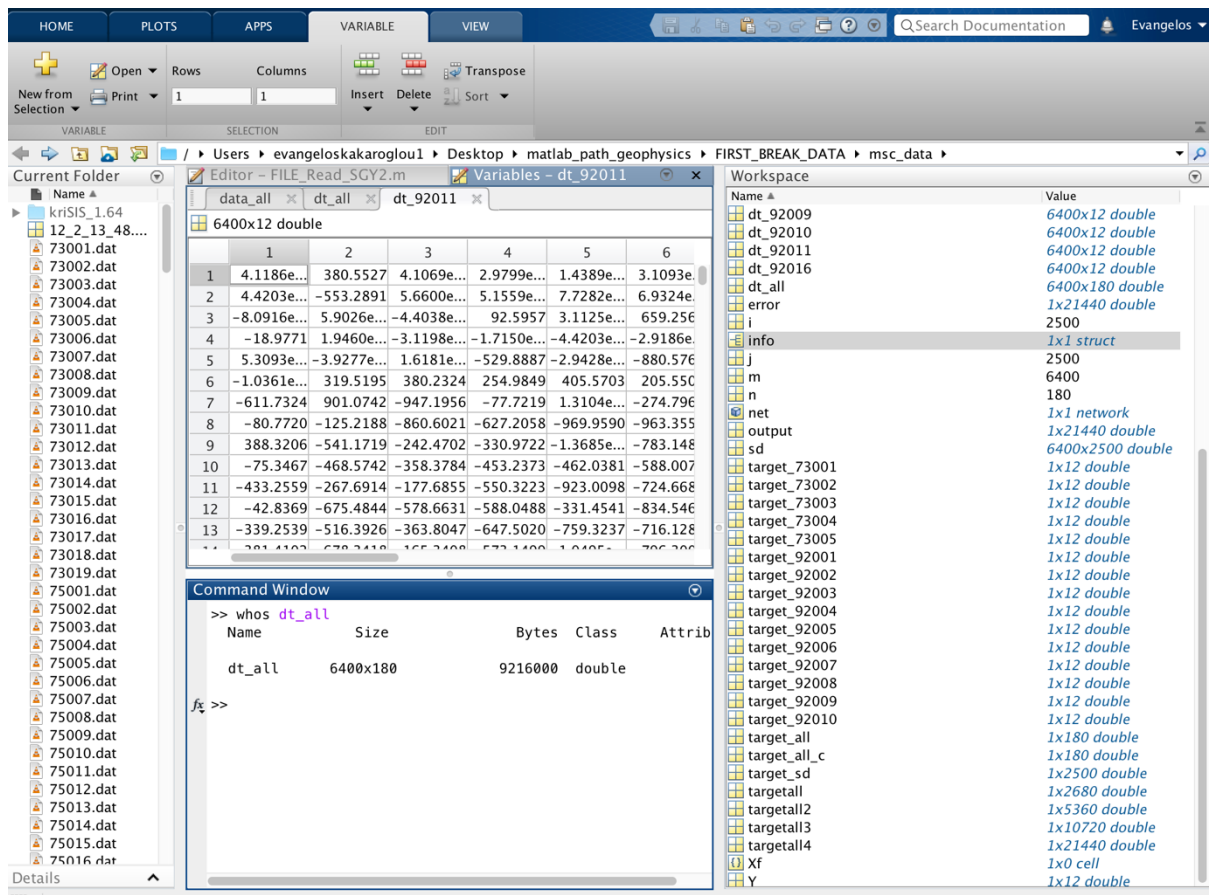
Εικόνα 3. 4 Επιλογή των προς μετατροπή .dat αρχείων



Εικόνα 3.5 Τελικό αποτέλεσμα του αλγορίθμου, πίνακες και σεισμόγραμμα

3.4 ΔΗΜΙΟΥΡΓΙΑ ΤΟΥ ΠΙΝΑΚΑ ΤΩΝ ΔΕΔΟΜΕΝΩΝ ΕΚΠΑΙΔΕΥΣΗΣ (training data set)

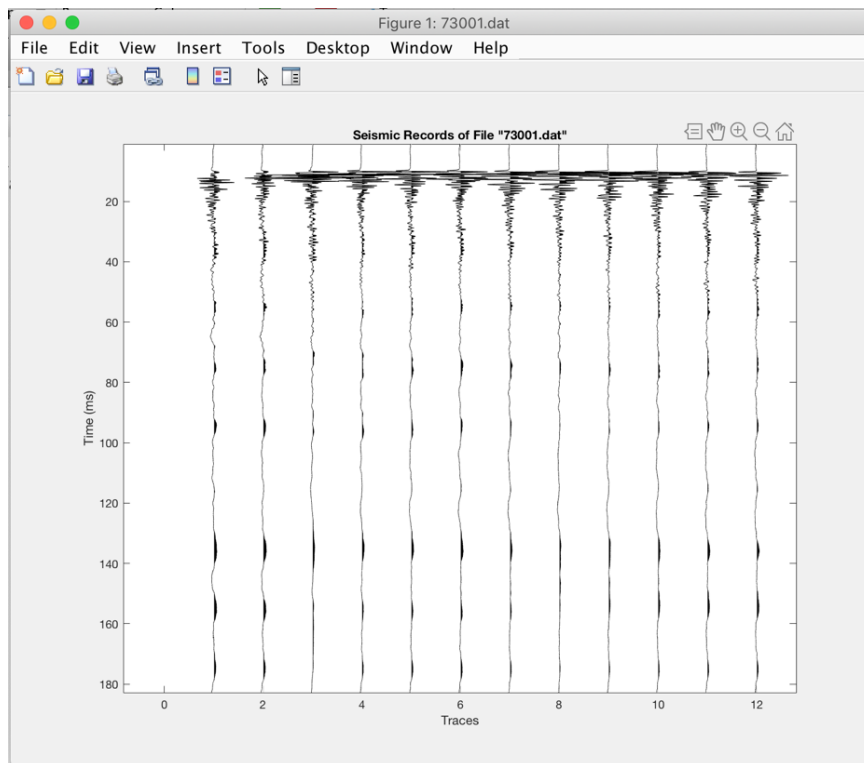
Για την δημιουργία του συνόλου εκπαίδευσης (training data set) επιλέγονται τα δεδομένα όπως αυτά μετασχηματίστηκαν παραπάνω, και στην συνέχεια τα δεδομένα από τις σεισμικές καταγραφές, ομαδοποιούνται σε έναν μεγαλύτερο πίνακα, που περιέχει συγκεντρωτικά τα δεδομένα από όλες τις προς μελέτη καταγραφές (Εικόνα 3.6). Δημιουργείται έτσι ένας πίνακας που περιέχει πληροφορίες σχετικά με τις πρώτες αφίξεις σε όλα τα σημεία της προς μελέτη περιοχής.



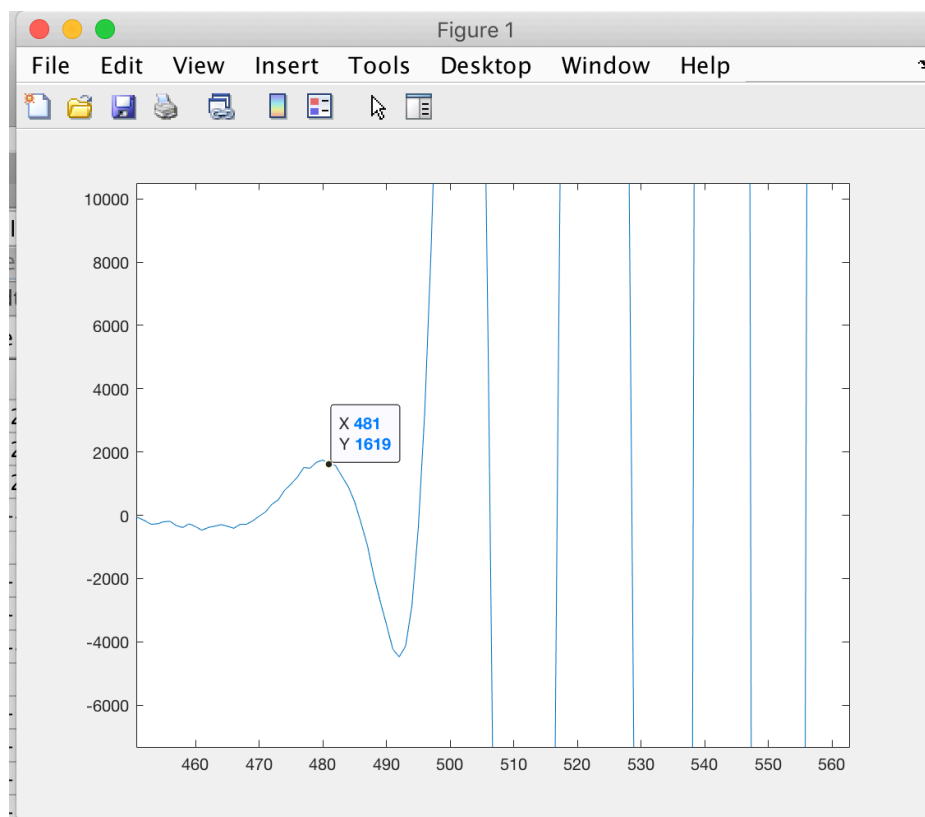
Εικόνα 3. 6 Δημιουργία πίνακα με το σύνολο εκπαίδευσης

3.5 ΕΠΙΛΟΓΗ ΤΩΝ ΠΡΩΤΩΝ ΑΦΙΞΕΩΝ ΚΑΙ ΔΗΜΙΟΥΡΓΙΑ ΤΟΥ ΣΥΝΟΛΟΥ ΣΤΟΧΩΝ (targets)

Η επιλογή των πρώτων αφίξεων έγινε εμπειρικά σύμφωνα με την κρίση του συγγραφέα. Τα δεδομένα αναπαρίστανται με την μορφή σημάτων με την εντολή plot της Matlab (Εικόνα 3.7) και οπτικά επιλέγονται τα σημεία των πρώτων αφίξεων (Εικόνα 3.8). Τελικά εντοπίζεται και αποθηκεύεται το κελί του πίνακα στο οποίο αντιστοιχεί η κάθε διαφορετική πρώτη άφιξη.



Εικόνα 3. 7 Παράδειγμα σεισμικής καταγραφής



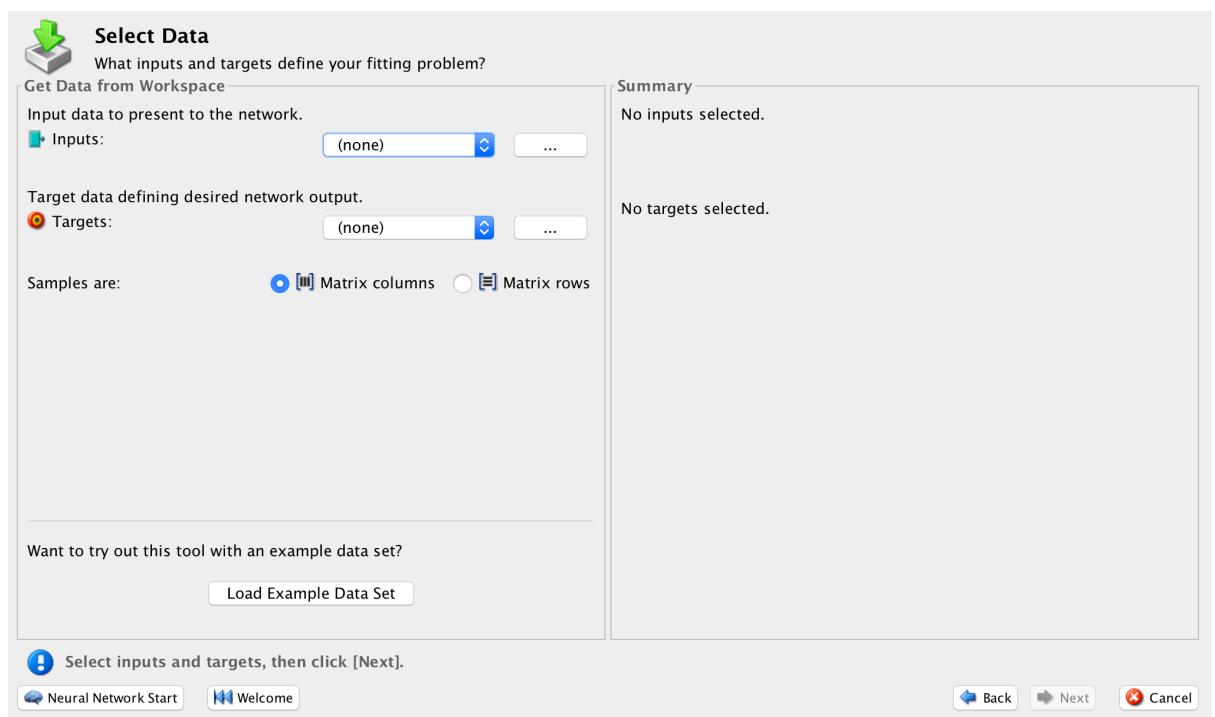
Εικόνα 3. 8 Οπτική επιλογή πρώτης άφιξης

Αφού επιλεγθούν οι θέσεις των πρώτων αφίξεων για τα δεδομένα μελέτης δημιουργείται πίνακας ο οποίος περιέχει τα κελιά που βρίσκονται οι πρώτες αφίξεις για κάθε γραμμή μελέτης. Στην συνέχεια οι επί μέρους πίνακες ενώνονται σε έναν που περιέχει τις θέσεις όλων των πρώτων αφίξεων, αυτός θα είναι και ο πίνακας που αντιστοιχεί στους στόχους για την εκπαίδευση του τεχνητού νευρωνικού δικτύου.

3.6 ΔΗΜΙΟΥΡΓΙΑ ΚΑΙ ΕΚΠΑΙΔΕΥΣΗ ΤΩΝ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ

Για την επιλογή της γεωμετρίας του δικτύου και του τρόπου εκπαίδευσης καθώς και για την τυχαία διαίρεση των δειγμάτων χρησιμοποιήθηκε το nftool του πακέτου machine learning and deep learning της Matlab. Ακολουθεί μια συνοπτική περιγραφή του τρόπου εργασίας στο πακέτο αυτό με τελικό αποτέλεσμα την κατασκευή ενός δικτύου ενδιαφέροντος.

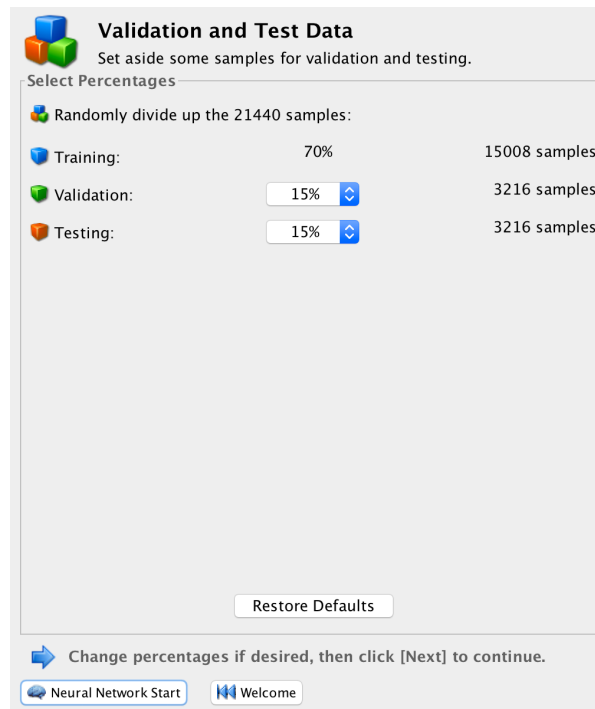
Αρχικά επιλέγεται το σετ δεδομένων εκπαίδευσης το οποίο κατασκευάστηκε με τον τρόπο που παρουσιάστηκε παραπάνω. Το σύνολο αυτό αποτελείται από όλα τα δεδομένα από όλες τις καταγραφές και λειτουργεί ως input του δικτύου. Στην συνέχεια, εισάγεται ο πίνακας που περιέχει τα targets, η κατασκευή του οποίου περιγράφεται επίσης παραπάνω (Εικόνα 3.9). Ο πίνακας αυτός είναι και ο πίνακας που περιέχει τις θέσεις των πρώτων αφίξεων όπως αυτές επιλέχθηκαν από τον συγγραφέα.



Εικόνα 3. 9 Επιλογή εισόδων και στόχων του νευρωνικού δικτύου

Το επόμενο βήμα στην διαδικασία είναι ο διαχωρισμός του αρχικού σετ δεδομένων σε τρία, αυτό της εκπαίδευσης (training), της επαλήθευσης (validation) και του ελέγχου (testing), ο διαχωρισμός αυτός γίνεται τυχαία, και μετά από δοκιμές. Με το σετ εκπαίδευσης να αντιστοιχεί στο 70%, το επαλήθευσης στο 15% του αρχικού και το ελέγχου σε 15% του αρχικού. (Εικόνα 3.10)

Στη συνέχεια ορίζεται ο αριθμός των νευρώνων για το κρυφό επίπεδο και καθορίζεται η γεωμετρία του δικτύου (Εικόνα 3.11).



Validation and Test Data
Set aside some samples for validation and testing.

Select Percentages

Randomly divide up the 21440 samples:

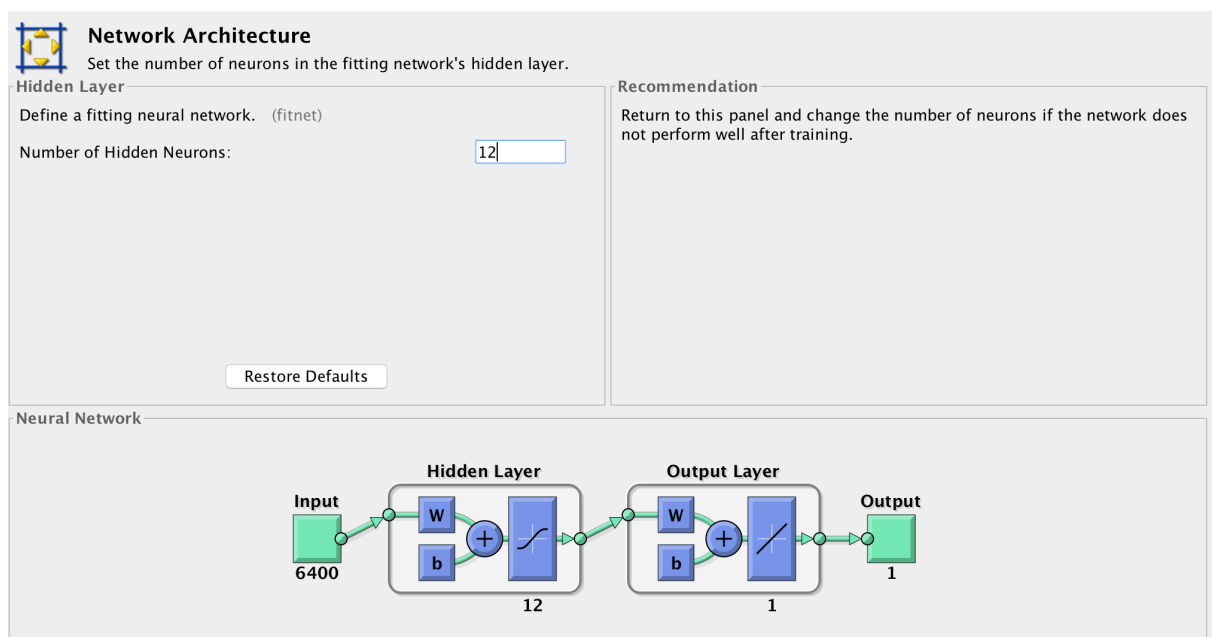
Training:	70%	15008 samples
Validation:	15%	3216 samples
Testing:	15%	3216 samples

Restore Defaults

Change percentages if desired, then click [Next] to continue.

Neural Network Start Welcome

Εικόνα 3. 10 Τυχαιος διαχωρισμός του αρχικού σετ δεδομένων εκπαίδευσης



Network Architecture
Set the number of neurons in the fitting network's hidden layer.

Hidden Layer

Define a fitting neural network. (fitnet)

Number of Hidden Neurons: 12

Restore Defaults

Recommendation

Return to this panel and change the number of neurons if the network does not perform well after training.

Neural Network

Input: 6400

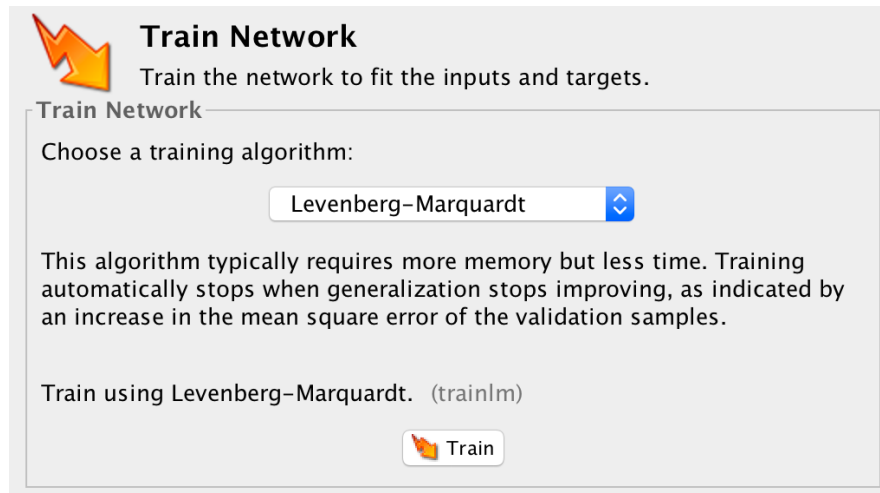
Hidden Layer: 12

Output Layer: 1

Output: 1

Εικόνα 3. 11 Επιλογή της γεωμετρίας του δικτύου

Τελικά επιλέγεται ο τρόπος εκπαίδευσης, για τον οποίο χρησιμοποιείται ο αλγόριθμος Levenberg-Marquardt του οποίου τα χαρακτηριστικά περιγράφονται αναλυτικά στο Κεφάλαιο 2 (Εικόνα 3.12).



Εικόνα 3. 12 Επιλογή του αλγορίθμου Levenberg-Marquardt ως αλγόριθμο εκπαίδευσης

3.7 ΣΥΝΟΨΗ

Στα πλαίσια αυτής εργασίας δημιουργήθηκαν σύμφωνα με την μεθοδολογία που παρουσιάστηκε παραπάνω, δίκτυα με διαφορετικές παραμέτρους και γεωμετρίες, με σκοπό την σύγκριση των αποδόσεων αυτών μεταξύ τους.

Τα χαρακτηριστικά των δικτύων συγκρίνονται μεταξύ τους και εξάγονται συμπεράσματα σχετικά με το βέλτιστο μέγεθος δείγματος και γεωμετρίας. Τα αποτελέσματα που προκύπτουν από τα πιο σταθερά δίκτυα συγκρίνονται και με αποτελέσματα που επεξεργάστηκαν χειροκίνητα από διαφορετικούς ανθρώπους.

Τα αποτελέσματα και τα χαρακτηριστικά των δικτύων που αναπτύχθηκαν παρουσιάζονται αναλυτικά στο επόμενο κεφάλαιο.

4: ΠΑΡΟΥΣΙΑΣΗ ΤΩΝ ΑΠΟΤΕΛΕΣΜΑΤΩΝ

4.1 ΕΙΣΑΓΩΓΗ

Στο κεφάλαιο αυτό θα γίνει η παρουσίαση και ο σχολιασμός των χαρακτηριστικών των δικτύων που προέκυψαν κατά την διαδικασία της εκπαίδευσης με την χρήση διαφορετικών γεωμετριών, και διαφορετικών μεγεθών δειγμάτων, τα οποία δημιουργήθηκαν σύμφωνα με την μεθοδολογία που αναφέρθηκε στο κεφάλαιο 3. Ακολουθεί δηλαδή η διαδικασία για την σταδιακή βελτίωση των χαρακτηριστικών των μοντέλων που κατασκευάστηκαν, με την επανάληψη της διαδικασίας για τα διάφορα δίκτυα και για διαφορετικούς αριθμούς κόμβων.

4.1.2 ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΔΙΚΤΥΩΝ

Τα στατιστικά χαρακτηριστικά τα οποία παρουσιάζονται και θα μελετηθούν στην συνέχεια, τα οποία χρησιμοποιούνται ως κριτήρια για την ευστάθεια και την αποτελεσματικότητα των τεχνητών νευρωνικών δικτύων που εκπαιδεύτηκαν είναι το μέσο τετραγωνικό σφάλμα (Mean Squared Error, MSE) και ο συντελεστής παλινδρόμηση (Regression, R).

Το μέσο τετραγωνικό σφάλμα ορίζεται ως η διαφορά μεταξύ των αποτελεσμάτων εξόδου του δικτύου (outputs) και των στόχων που είχαν τεθεί κατά την εκπαίδευση (targets),

υψωμένη στο τετράγωνο. Χαμηλότερες τιμές του μέσου τετραγωνικού σφάλματος είναι επιθυμητές καθώς όσο πιο πολύ τείνει στο μηδέν τόσο μικρότερη είναι η απόκλιση των τιμών εξόδου από τις επιθυμητές.

Ο συντελεστής παλινδρόμησης αναφέρεται στην συσχέτιση μεταξύ των αποτελεσμάτων και των στόχων. Τιμές του συντελεστή κοντά στην μονάδα υποδεικνύουν μια καλή συσχέτιση, ενώ κοντά στο μηδέν μια τυχαία συσχέτιση.

Παρακάτω παρουσιάζονται οι τιμές μέσου τετραγωνικού σφάλματος, συντελεστή παλινδρόμησης και τα ιστογράμματα σφαλμάτων κατά την διαδικασία της επιλογής του βέλτιστου νευρωνικού δικτύου.

Οι τιμές που παρουσιάζονται στους πίνακες που ακολουθούν αναφέρονται σε θέσεις κελιών πινάκων όπου βρίσκονται τα δεδομένα που επεξεργάστηκαν και για αυτό δεν αναφέρονται συγκεκριμένες μονάδες.

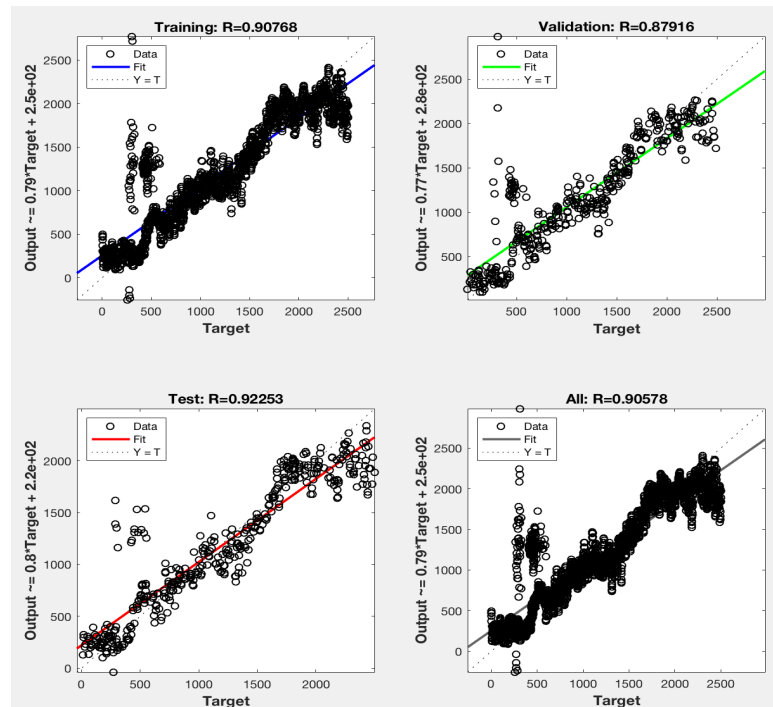
4.2.1 ΔΕΙΓΜΑ 1, 8 ΚΟΜΒΟΙ

Για το πρώτο δίκτυο που εκπαιδεύτηκε, χρησιμοποιήθηκε δείγμα με πληθυσμό 2680 και για την γεωμετρία επιλέχθηκαν 8 κόμβοι, δηλαδή το κρυφό επίπεδο περιέχει 8 νευρώνες. Παρατηρούνται πολύ υψηλές τιμές μέσου τετραγωνικού σφάλματος και τιμές συντελεστή παλινδρόμησης που απέχουν από την μονάδα. (Πίνακας 4.1). Επίσης παρατηρείται μεγάλη απόκλιση των αποτελεσμάτων (outputs) σε σχέση με την ευθεία $y=x$ στο διάγραμμα αποτελεσμάτων ως προς τις επιθυμητές τιμές (outputs-targets) (Εικόνα 4.1), πράγμα που επιβεβαιώνεται και από την μορφή του ιστογράμματος σφαλμάτων (Εικόνα 4.2).

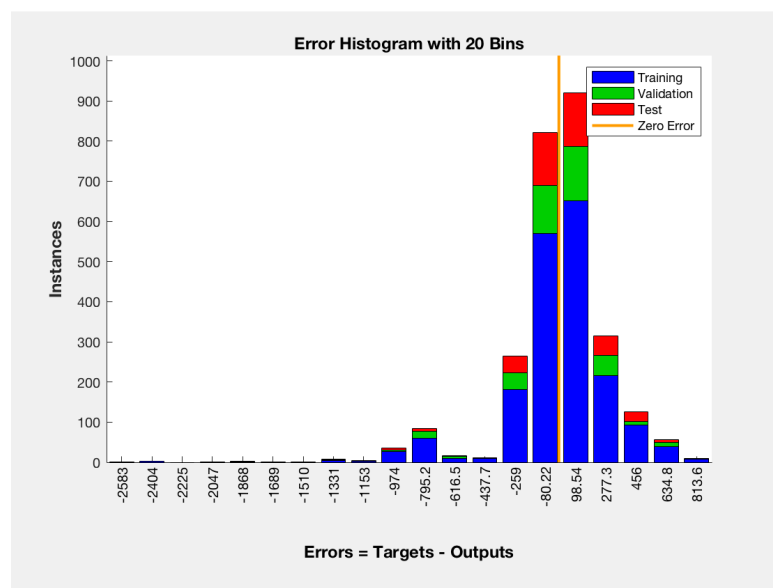
Βάσει των παρατηρήσεων αυτών προκύπτει ότι η συγκεκριμένη γεωμετρία δικτύου δεν είναι αποτελεσματική και απορρίπτεται.

SAMPLE 1 number of nodes 8	Sample size	MSE (mean squared error)	R(regression)
Training	1876	95866.57e-0	9.07e-1
Validation	402	112804.54e-0	8.79e-1
Testing	402	79468.15e-0	9.22e-1

Πίνακας 4. 1 Χαρακτηριστικά στοιχεία δικτύου για πληθυσμό 2680 και 8 κόμβους



Εικόνα 4. 1 Γραφική απεικόνιση των αποτελεσμάτων δικτύου με 2680 δείγματα και 8 κόμβους σε σχέση με τους στόχους



Εικόνα 4. 2 Ιστογράμμο σφαλμάτων δικτύου με 2680 δείγματα και 8 κόμβους

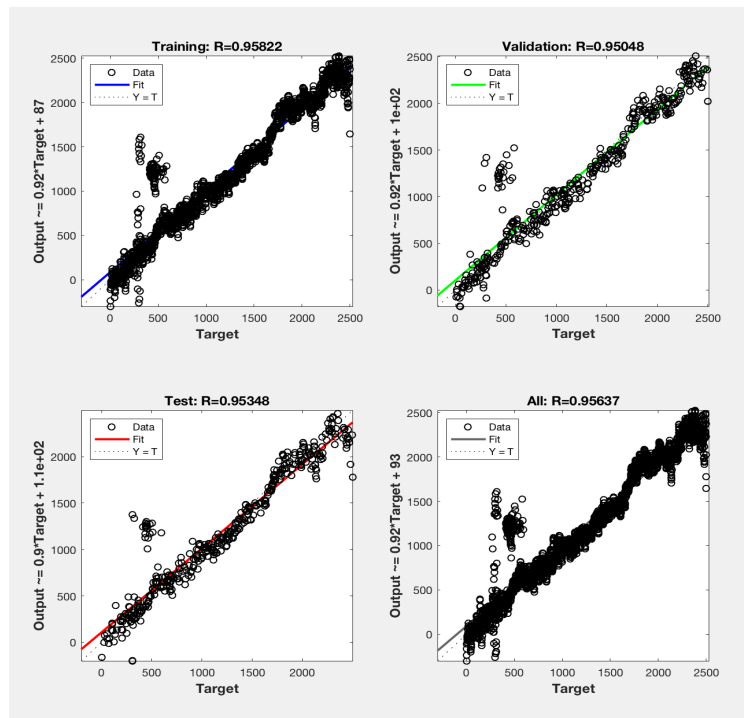
4.2.2 ΔΕΙΓΜΑ 1, 9 ΚΟΜΒΟΙ

Για το δεύτερο δίκτυο που εκπαιδεύτηκε, χρησιμοποιήθηκε ξανά δείγμα με πληθυσμό 2680 για την γεωμετρία αυτή την φορά επιλέχθηκαν 9 κόμβοι. Παρατηρούνται πολύ υψηλές τιμές μέσου τετραγωνικού σφάλματος και τιμές συντελεστή παλινδρόμησης που απέχουν από την μονάδα (Πίνακας 4.2). Επίσης παρατηρείται απόκλιση των αποτελεσμάτων (outputs) σε σχέση με την ευθεία $y=x$ στο διάγραμμα αποτελεσμάτων ως προς τις επιθυμητές τιμές (outputs-targets) (Εικόνα 4.3), πράγμα που επιβεβαιώνεται και από την μορφή του ιστογράμματος σφαλμάτων (Εικόνα 4.4).

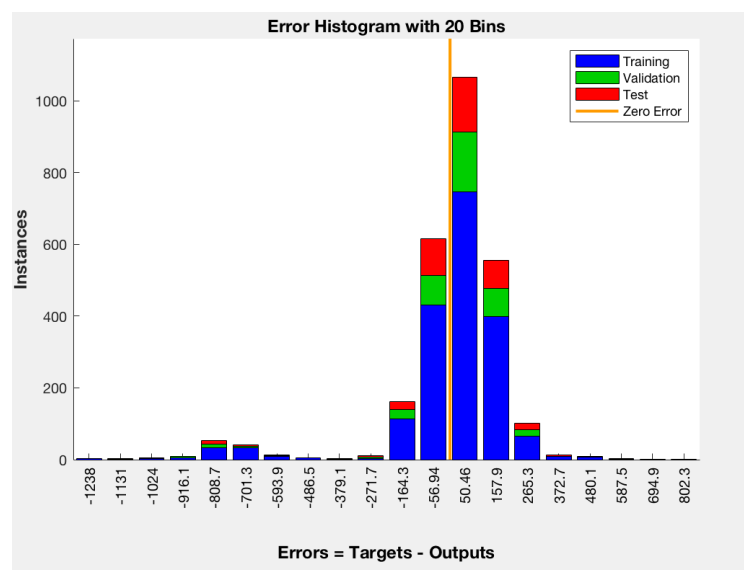
Βάσει των παρατηρήσεων αυτών προκύπτει ότι η συγκεκριμένη γεωμετρία δικτύου δεν είναι αποτελεσματική και απορρίπτεται.

SAMPLE 1 number of nodes 9	Sample size	MSE (mean squared error)	R(regression)
Training	1876	44072.24e-0	9.58e-1
Validation	402	52091.06e-0	9.50e-1
testing	402	44498.42e-0	9.53e-1

Πίνακας 4. 2 Χαρακτηριστικά στοιχεία δικτύου για πληθυσμό 2680 και 9 κόμβους



Εικόνα 4. 3 Γραφική απεικόνιση των αποτελεσμάτων δικτύου με 2680 δείγματα και 9 κόμβους σε σχέση με τους στόχους



Εικόνα 4. 4 Ιστόγραμμα σφαλμάτων δικτύου με 2680 δείγματα και 9 κόμβους

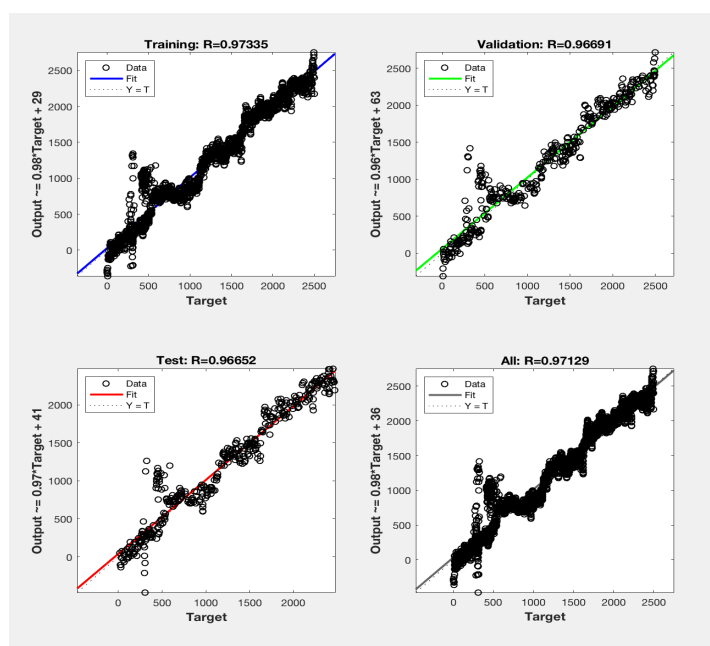
4.2.3 ΔΕΙΓΜΑ 1, 10 ΚΟΜΒΟΙ

Για το τρίτο δίκτυο που εκπαιδεύτηκε, χρησιμοποιήθηκε ξανά δείγμα με πληθυσμό 2680 για τη γεωμετρία αυτή την φορά επιλέχθηκαν 10 κόμβοι. Για ακόμη μια φορά παρατηρούνται πολύ υψηλές τιμές μέσου τετραγωνικού σφάλματος και τιμές συντελεστή παλινδρόμησης που απέχουν από την μονάδα (πίνακας 4.3). Επίσης παρατηρείται απόκλιση των αποτελεσμάτων (outputs) σε σχέση με την ευθεία $y=x$ στο διάγραμμα αποτελεσμάτων ως προς τις επιθυμητές τιμές (outputs-targets) (Εικόνα 4.5), πράγμα που επιβεβαιώνεται και από την μορφή του ιστογράμματος σφαλμάτων (Εικόνα 4.6).

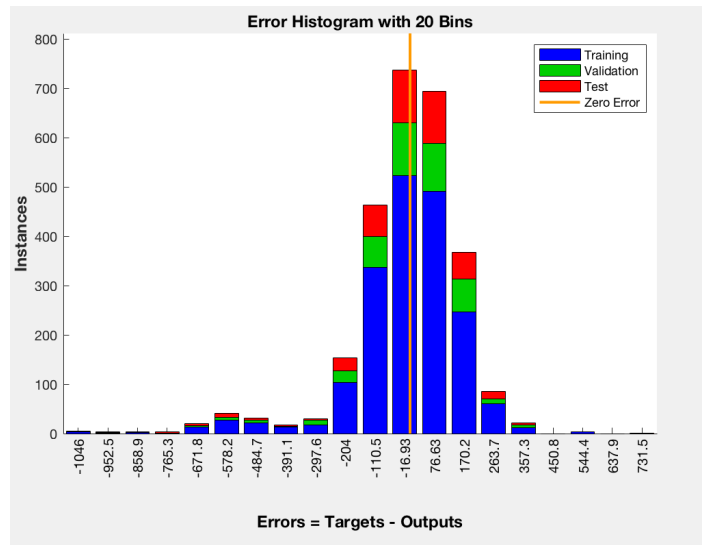
Βάσει των παρατηρήσεων αυτών προκύπτει ότι η συγκεκριμένη γεωμετρία δικτύου δεν είναι αποτελεσματική και απορρίπτεται.

SAMPLE 1 number of nodes 10	Sample size	MSE (mean squared error)	R(regression)
Training	1876	28235.40e-0	9.73e-1
Validation	402	37624.16e-0	9.669e-1
testing	402	35874.31e-0	9.665e-1

Πίνακας 4. 3 Χαρακτηριστικά στοιχεία δικτύου για πληθυσμό 2680 και 10 κόμβους



Εικόνα 4. 5 Γραφική απεικόνιση των αποτελεσμάτων δικτύου με 2680 δείγματα και 10 κόμβους σε σχέση με τους στόχους



Εικόνα 4. 6 Ιστογράμμο σφαλμάτων δικτύου με 2680 δείγματα και 10 κόμβους

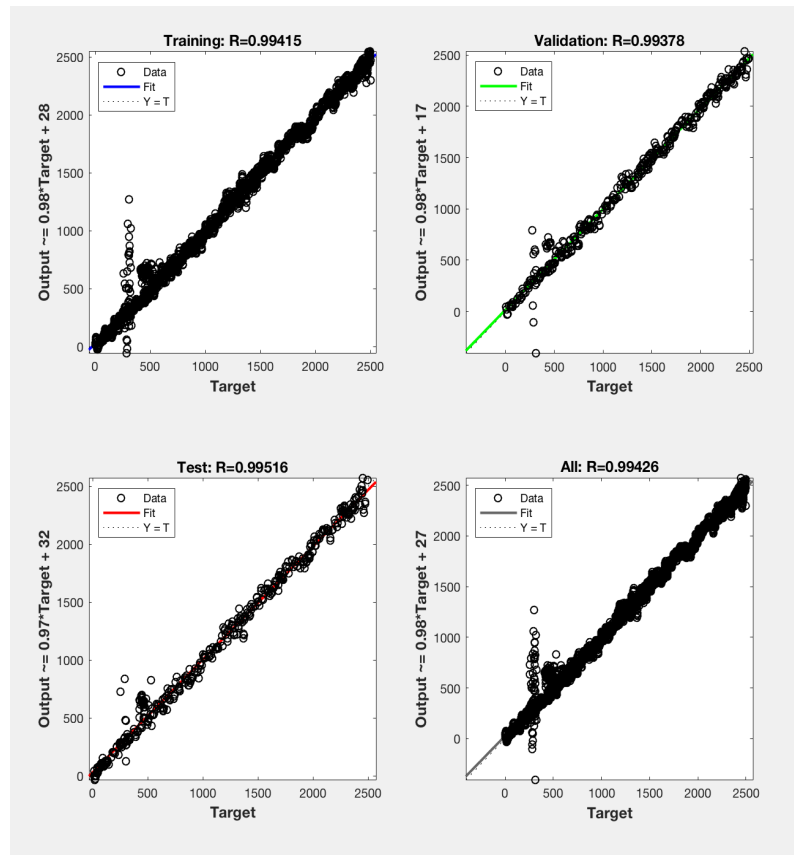
4.2.4 ΔΕΙΓΜΑ 1, 11 ΚΟΜΒΟΙ

Για το τέταρτο δίκτυο που εκπαιδεύτηκε, χρησιμοποιήθηκε ξανά δείγμα με πληθυσμό 2680 για την γεωμετρία αυτήν τη φορά επιλέχθηκαν 11 κόμβοι. Για ακόμη μια φορά παρατηρούνται πολύ υψηλές τιμές μέσου τετραγωνικού σφάλματος και τιμές συντελεστή παλινδρόμησης που απέχουν από την μονάδα (Πίνακας 4.4), πρέπει όμως να σημειωθεί η σημαντική βελτίωση σε σχέση με το πρώτο δίκτυο (4.2.1). Επίσης παρατηρείται απόκλιση των αποτελεσμάτων (outputs) σε σχέση με την ευθεία $y=x$ στο διάγραμμα αποτελεσμάτων ως προς τις επιθυμητές τιμές (outputs-targets) (εικόνα 4.7), πράγμα που επιβεβαιώνεται και από την μορφή του ιστογράμματος σφαλμάτων (εικόνα 4.8).

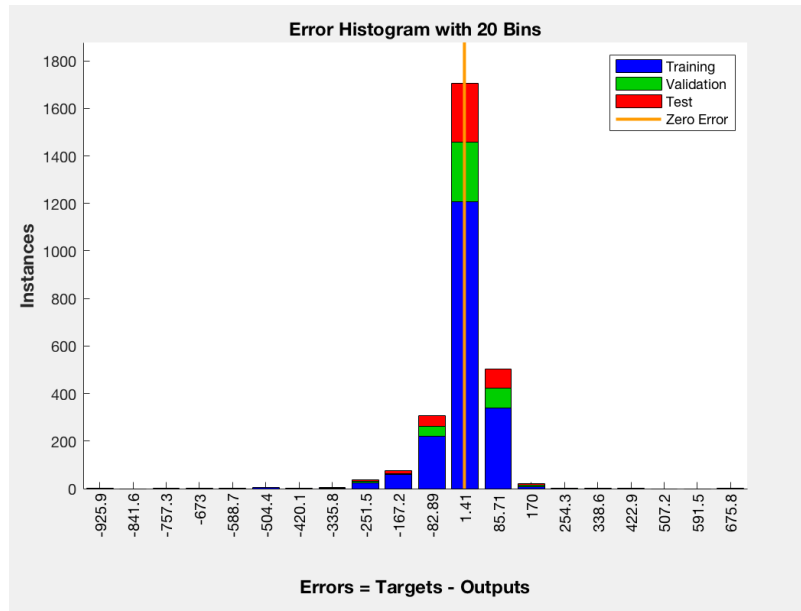
Βάσει των παρατηρήσεων αυτών προκύπτει ότι η συγκεκριμένη γεωμετρία δικτύου δεν είναι αποτελεσματική και απορρίπτεται. Στην συνέχεια δοκιμάζεται ο διπλασιασμός του μεγέθους του δείγματος, καθώς από ότι φαίνεται η αύξηση του αριθμού των κόμβων δεν μειώνει το σφάλμα με έναν ικανοποιητικά υψηλό ρυθμό.

SAMPLE 1 number of nodes 11	Sample size	MSE (mean squared error)	R(regression)
Training	1876	6133.01e-0	9.94e-1
Validation	402	6570.79e-0	9.93e-1
testing	402	5700.84e-0	9.95e-1

Πίνακας 4. 4 Χαρακτηριστικά στοιχεία δικτύου για πληθυσμό 2680 και 11 κόμβους



Εικόνα 4. 7 Γραφική απεικόνιση των αποτελεσμάτων του δικτύου με 2680 δείγματα και 11 κόμβους σε σχέση με τους στόχους



Εικόνα 4. 8 Ιστόγραμμα σφαλμάτων δικτύου με 2680 δείγματα και 11 κόμβους

4.2.5 ΔΕΙΓΜΑ 2, 10 ΚΟΜΒΟΙ

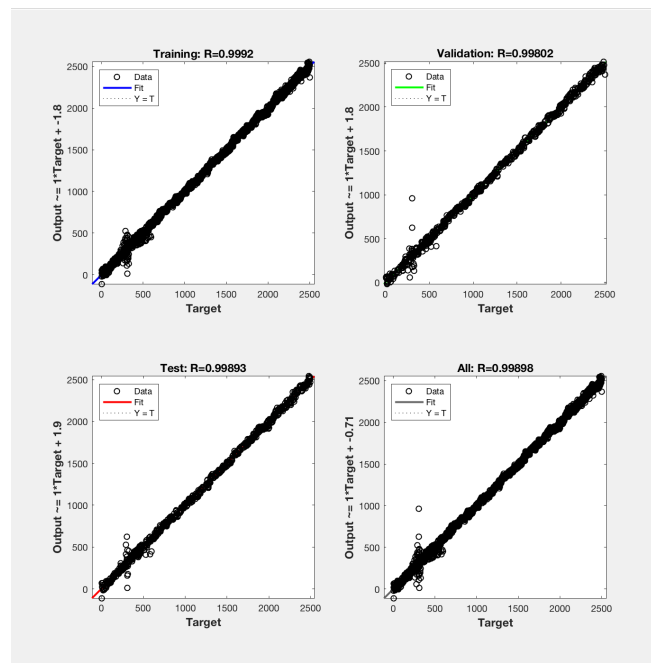
Για το πέμπτο δίκτυο που εκπαιδεύτηκε, χρησιμοποιήθηκε δείγμα με πληθυσμό 5360 και για την γεωμετρία επιλέχθηκαν 10 κόμβοι. Παρατηρούνται υψηλές τιμές μέσου τετραγωνικού σφάλματος, αλλά επίσης ότι, τα αποτελέσματα είναι πολύ βελτιωμένα σε σχέση με το προηγούμενο αντίστοιχο δίκτυο (4.2.3). Ο συντελεστής συσχέτισης αρχίζει

να πλησιάζει την μονάδα (τέλεια συσχέτιση) (πίνακας 4.5). Επίσης παρατηρείται σχετική απόκλιση της κατανομής των αποτελεσμάτων (outputs) σε σχέση με την ευθεία σε σχέση με την ευθεία $y=x$ στο διάγραμμα αποτελεσμάτων ως προς τις επιθυμητές τιμές (outputs-targets) (εικόνα 4.9), πράγμα που επιβεβαιώνεται και από την μορφή του ιστογράμματος σφαλμάτων(εικόνα:4.10).

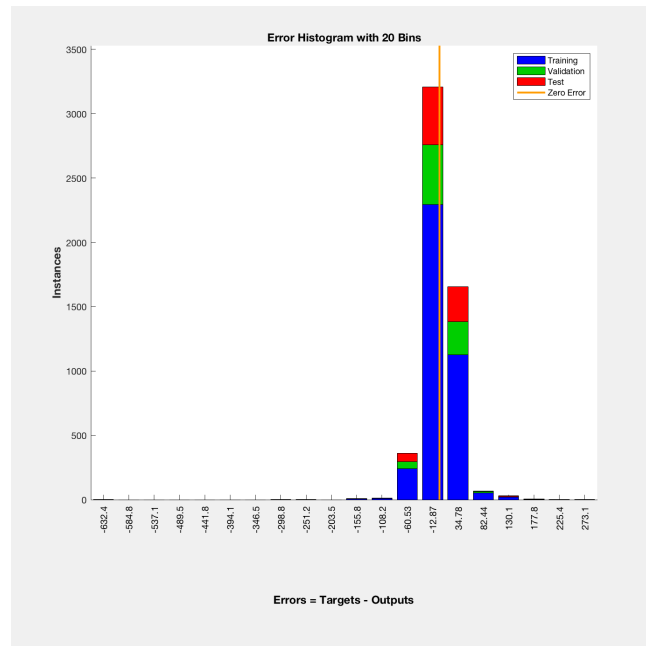
Βάσει των παρατηρήσεων αυτών προκύπτει ότι η συγκεκριμένη γεωμετρία δικτύου δεν είναι αποτελεσματική και απορρίπτεται.

SAMPLE 2 of number of nodes 10	Sample size	MSE (mean squared error)	R(regression)
Training	3752	833.41e-0	9.99e-1
Validation	804	2191.37e-0	9.98e-1
testing	804	1181.22e-0	9.98e-1

Πίνακας 4. 5 Χαρακτηριστικά στοιχεία δικτύου για πληθυσμό 5360 και 10 κόμβους



Εικόνα 4. 9 Γραφική απεικόνιση των αποτελεσμάτων δικτύου με 5360 δείγματα και 10 κόμβους σε σχέση με τους στόχους



Εικόνα 4. 10 Ιστόγραμμα σφαλμάτων δικτύου με 5360 δείγματα και 10 κόμβους

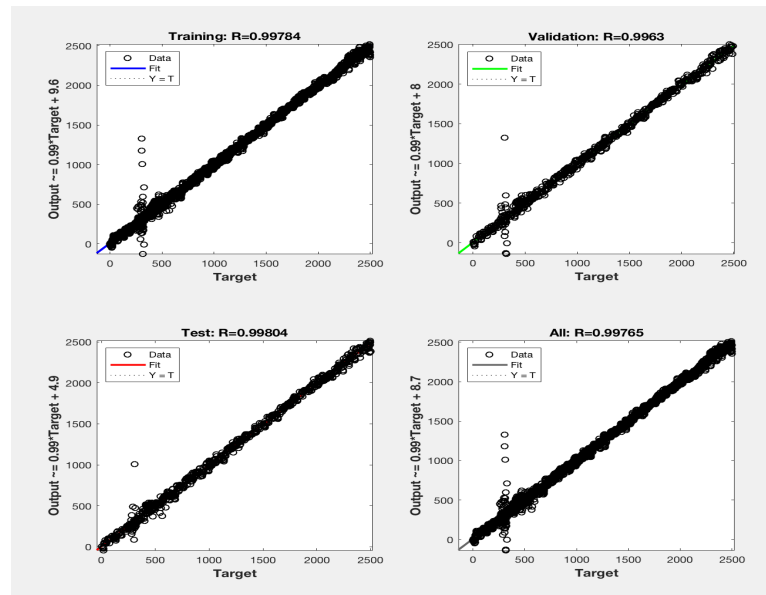
4.2.6 ΔΕΙΓΜΑ 2, 11 ΚΟΜΒΟΙ

Για το έκτο δίκτυο που εκπαιδεύτηκε, χρησιμοποιήθηκε δείγμα με πληθυσμό 5360 και για την γεωμετρία επιλέχθηκαν 11 κόμβοι. Παρατηρούνται ξανά υψηλές τιμές μέσου τετραγωνικού σφάλματος μεγαλύτερες απ' ότι στο προηγούμενο δίκτυο και ταυτόχρονα οι τιμές του συντελεστή παλινδρόμησης ,μειώθηκαν (Πίνακας 4.6). Επίσης παρατηρείται σχετική απόκλιση των αποτελεσμάτων (outputs) σε σχέση με την ευθεία $y=x$ στο διάγραμμα αποτελεσμάτων ως προς τις επιθυμητές τιμές (outputs-targets) (Εικόνα 4.11), πράγμα που επιβεβαιώνεται και από την μορφή του ιστογράμματος σφαλμάτων (εικόνα 4.12).

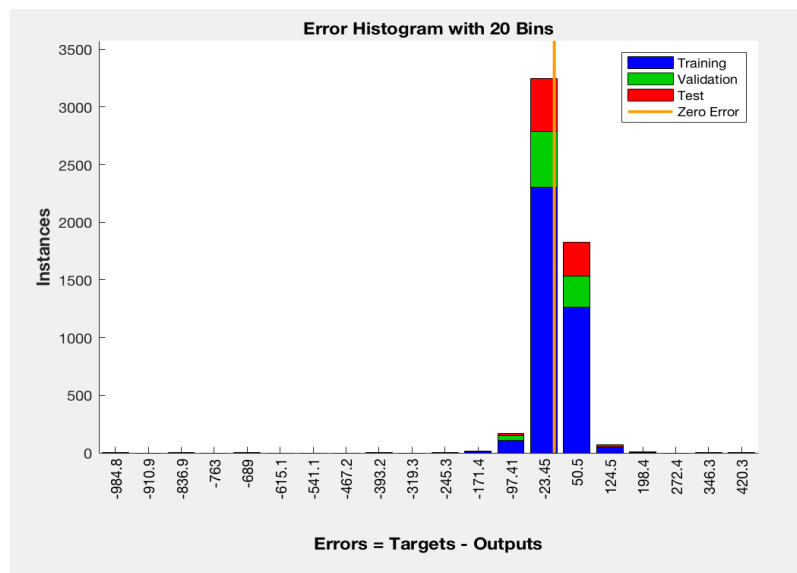
Βάσει των παρατηρήσεων αυτών προκύπτει ότι η συγκεκριμένη γεωμετρία δικτύου δεν είναι αποτελεσματική και απορρίπτεται.

SAMPLE 2 number of nodes 11	Sample size	MSE (mean squared error)	R(regression)
Training	3725	2297.66e-0	9.97e-1
Validation	804	3824.78e-0	9.96e-1
testing	804	2168.36e-0	9.98e-1

Πίνακας 4. 6 Χαρακτηριστικά στοιχεία δικτύου για πληθυσμό 5360 και 11 κόμβους



Εικόνα 4. 11 Γραφική απεικόνιση των αποτελεσμάτων δικτύου με 5360 δείγματα και 11 κόμβους σε σχέση με τους στόχους



Εικόνα 4. 12 Ιστογράμμο σφαλμάτων δικτύου με 5360 δείγματα και 11 κόμβους

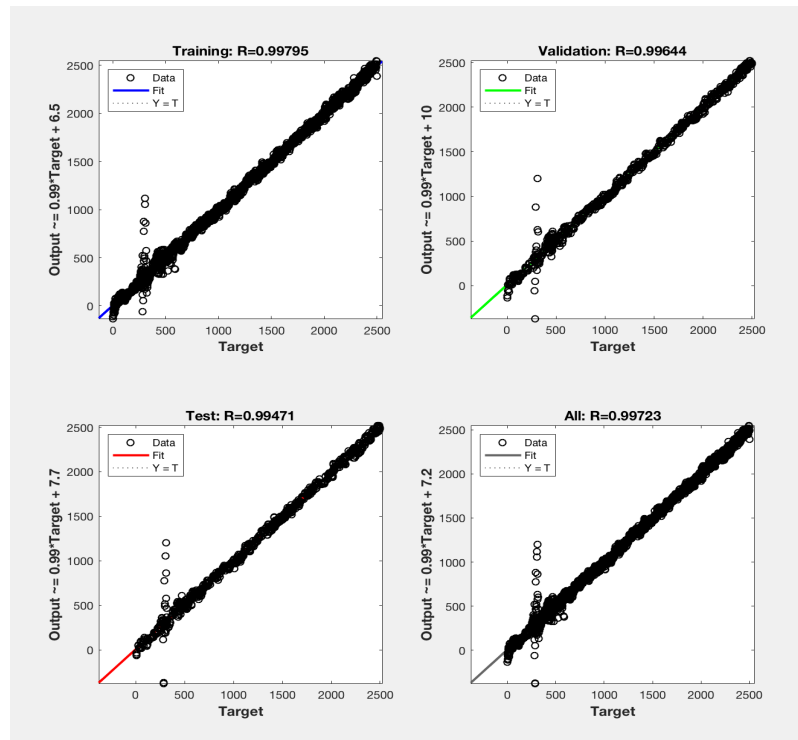
4.2.7 ΔΕΙΓΜΑ 2, 12 ΚΟΜΒΟΙ

Για το έβδομο δίκτυο που εκπαιδεύτηκε, χρησιμοποιήθηκε δείγμα με πληθυσμό 5360 και για την γεωμετρία επιλέχθηκαν 12 κόμβοι. Παρατηρούνται υψηλές τιμές μέσου τετραγωνικού σφάλματος και τιμές συντελεστή παλινδρόμησης μικρότερες από το προηγούμενο δίκτυο (Πίνακας 4.7). Επίσης παρατηρείται απόκλιση των αποτελεσμάτων (outputs) σε σχέση με την ευθεία $y=x$ στο διάγραμμα αποτελεσμάτων ως προς τις επιθυμητές τιμές (outputs-targets) (Εικόνα 4.13), πράγμα που επιβεβαιώνεται και από την μορφή του ιστογράμματος σφαλμάτων (σχήμα 4.14). Βάσει των παρατηρήσεων αυτών προκύπτει ότι η συγκεκριμένη γεωμετρία δικτύου δεν είναι αποτελεσματική και απορρίπτεται.

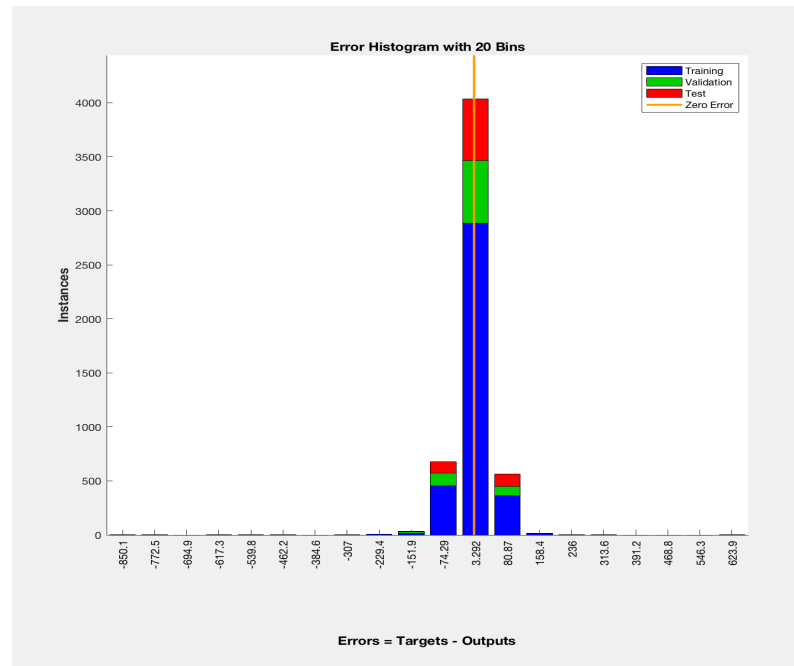
Στην συνέχεια θα γίνει διπλασιασμός του δείγματος για ακόμη μια φορά καθώς φαίνεται ότι η αύξησή του μεγέθους συνεφέρει θετικά σε αντίθεση με την αύξηση των κόμβων σε αυτό το σημείο που δείχνει να μην προσφέρει μεγάλη βελτίωση στα αποτελέσματα.

SAMPLE 2 number of nodes 12	Sample size	MSE (mean squared error)	R(regression)
Training	3752	2158.71e-0	9.97e-1
Validation	804	4009.46e-0	9.96e-1
testing	804	5507.88e-0	9.94e-1

Πίνακας 4.7 Χαρακτηριστικά στοιχεία δικτύου για πληθυσμό 5360 και 12 κόμβους



Εικόνα 4. 13 Γραφική απεικόνιση των αποτελεσμάτων δικτύου με 5360 δείγματα και 12 κόμβους σε σχέση με τους στόχους



Εικόνα 4. 14 Ιστόγραμμα σφαλμάτων δικτύου με 5360 δείγματα και 12 κόμβους

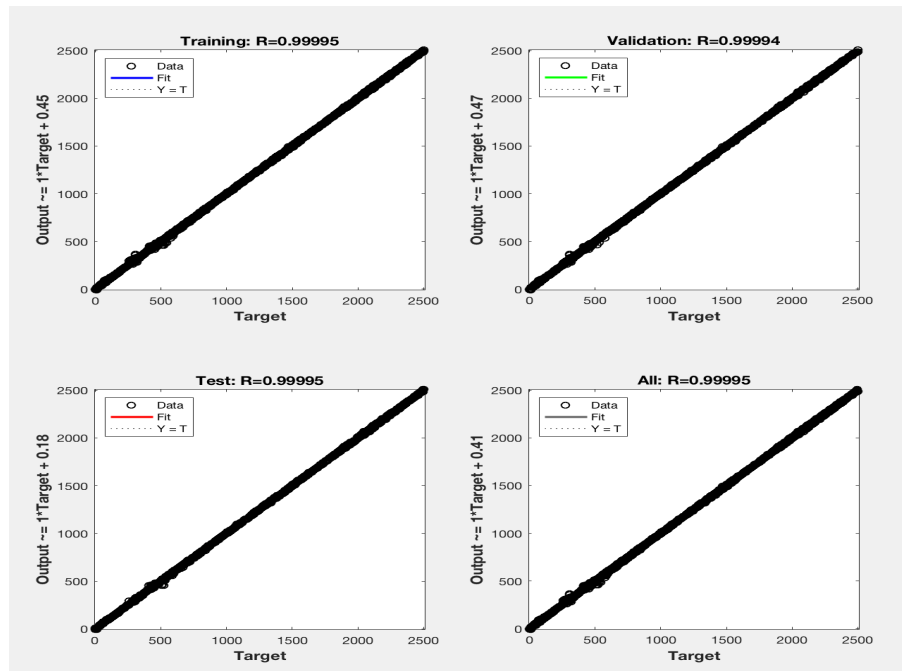
4.2.8 ΔΕΙΓΜΑ 3, 8 ΚΟΜΒΟΙ

Για το όγδοο δίκτυο που εκπαιδεύτηκε, χρησιμοποιήθηκε δείγμα με πληθυσμό 10720 και για την γεωμετρία επιλέχθηκαν 8 κόμβοι. Παρατηρούνται τιμές μέσου τετραγωνικού σφάλματος αρκετά μειωμένες σε σχέση με προηγούμενα δίκτυα και τιμές συντελεστή παλινδρόμησης που δεν απέχουν από την μονάδα. (Πίνακας 4.8). Επίσης παρατηρείται κατανομή των αποτελεσμάτων (outputs) σε ταύτιση με την ευθεία $y=x$ στο διάγραμμα αποτελεσμάτων ως προς τις επιθυμητές τιμές (outputs-targets) (Εικόνα 4.15), πράγμα που επιβεβαιώνεται μερικώς και από την μορφή του ιστογράμματος σφαλμάτων (Εικόνα 4.16).

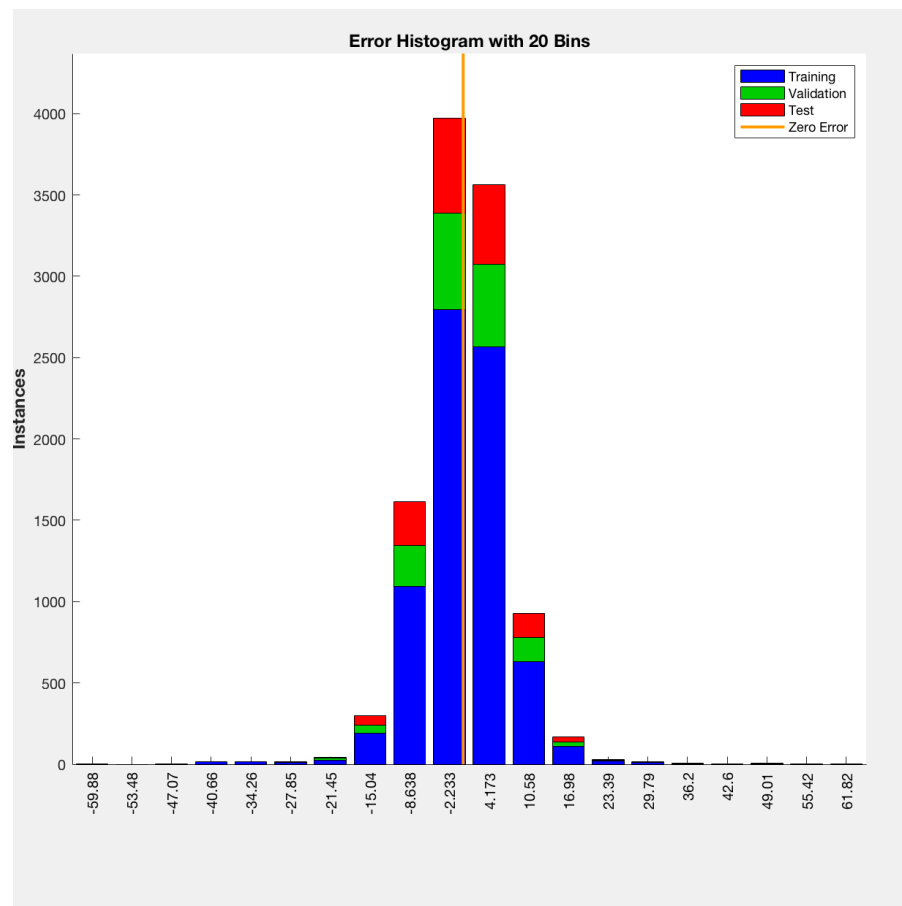
Βάσει των παρατηρήσεων αυτών προκύπτει ότι η συγκεκριμένη γεωμετρία δικτύου δίνει σχετικά καλά αποτελέσματα αλλά φαίνεται ότι μπορεί να βελτιωθεί και άλλο και έτσι συνεχίζεται η διαδικασία.

SAMPLE 3 of number of nodes 8	Sample size	MSE (mean squared error)	R(regression)
Training	7504	52.92e-0	9.99950e-1
Validation	1608	61.91e-0	9.99942e-1
testing	1608	57.97e-0	9.99945e-1

Πίνακας 4. 8 Χαρακτηριστικά στοιχεία δικτύου για πληθυσμό 10720 και 8 κόμβους



Εικόνα 4. 15 Γραφική απεικόνιση των αποτελεσμάτων δικτύου με 10720 δείγματα και 8 κόμβους σε σχέση με τους στόχους



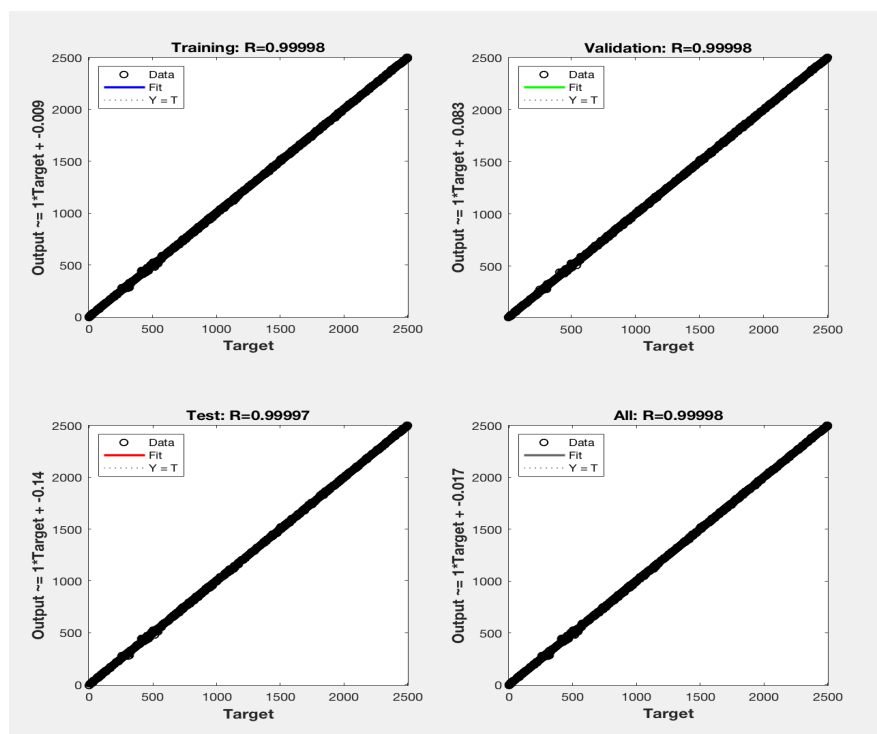
Εικόνα 4. 16 Ιστόγραμμα σφαλμάτων δικτύου με 10720 δείγματα και 8 κόμβους

4.2.8 ΔΕΙΓΜΑ 3, 9 ΚΟΜΒΟΙ

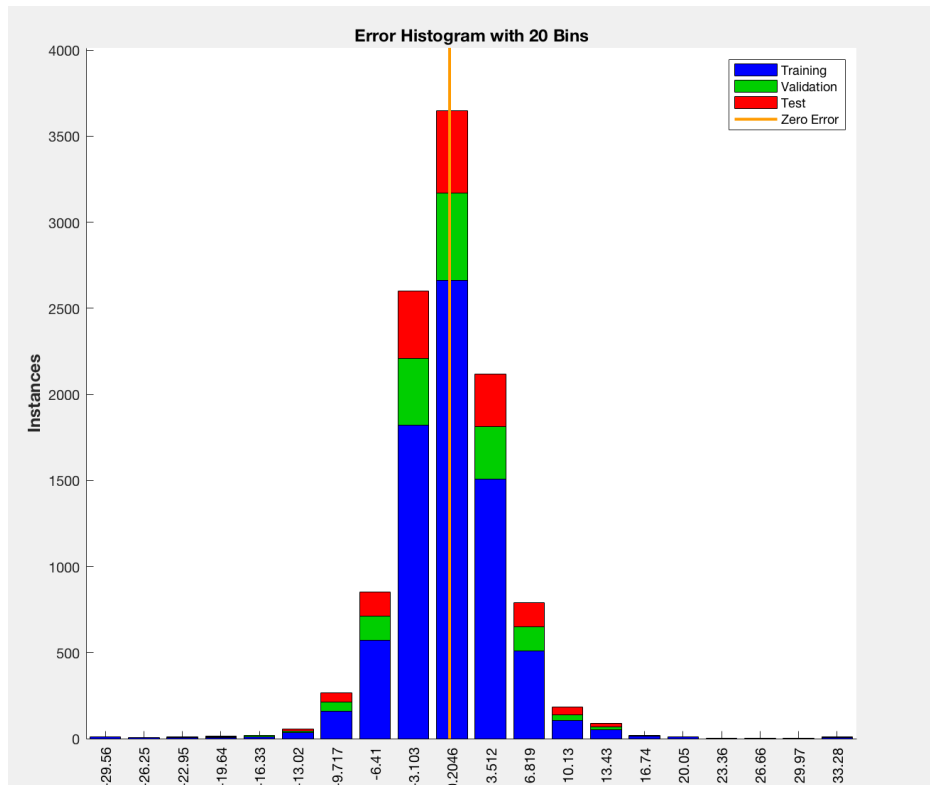
Για το ένατο δίκτυο που εκπαιδεύτηκε, χρησιμοποιήθηκε δείγμα με πληθυσμό 10720 και για την γεωμετρία επιλέχθηκαν 9 κόμβοι. Παρατηρούνται τιμές μέσου τετραγωνικού σφάλματος αρκετά μειωμένες σε σχέση με προηγούμενα δίκτυα και τιμές συντελεστή παλινδρόμησης που δεν απέχουν σημαντικά από την μονάδα (Πίνακας 4.9). Επίσης παρατηρείται κατανομή των αποτελεσμάτων (outputs) σε σχεδόν ταύτιση με την ευθεία $y=x$ στο διάγραμμα αποτελεσμάτων ως προς τις επιθυμητές τιμές (outputs-targets) (Εικόνα 4.17), πράγμα που επιβεβαιώνεται μερικώς και από την μορφή του ιστογράμματος σφαλμάτων (εικόνα 4.18). Βάσει των παρατηρήσεων αυτών προκύπτει ότι η συγκεκριμένη γεωμετρία δικτύου δίνει ακόμα καλύτερα αποτελέσματα αλλά φαίνεται επίσης ότι μπορεί να βελτιωθεί περισσότερο και έτσι συνεχίζεται η διαδικασία.

SAMPLE 3 number of nodes 9	Sample size	MSE (mean squared error)	R(regression)
Training	7504	21.02e-0	9.99980e-1
Validation	1608	26.52e-0	9.99975e-1
testing	1608	30.68e-0	9.99971e-1

Πίνακας 4. 9 Χαρακτηριστικά στοιχεία δικτύου για πληθυσμό 10720 και 9 κόμβους



Εικόνα 4. 17 Γραφική απεικόνιση των αποτελεσμάτων δικτύου με 10720 δείγματα και 9 κόμβους σε σχέση με τους στόχους



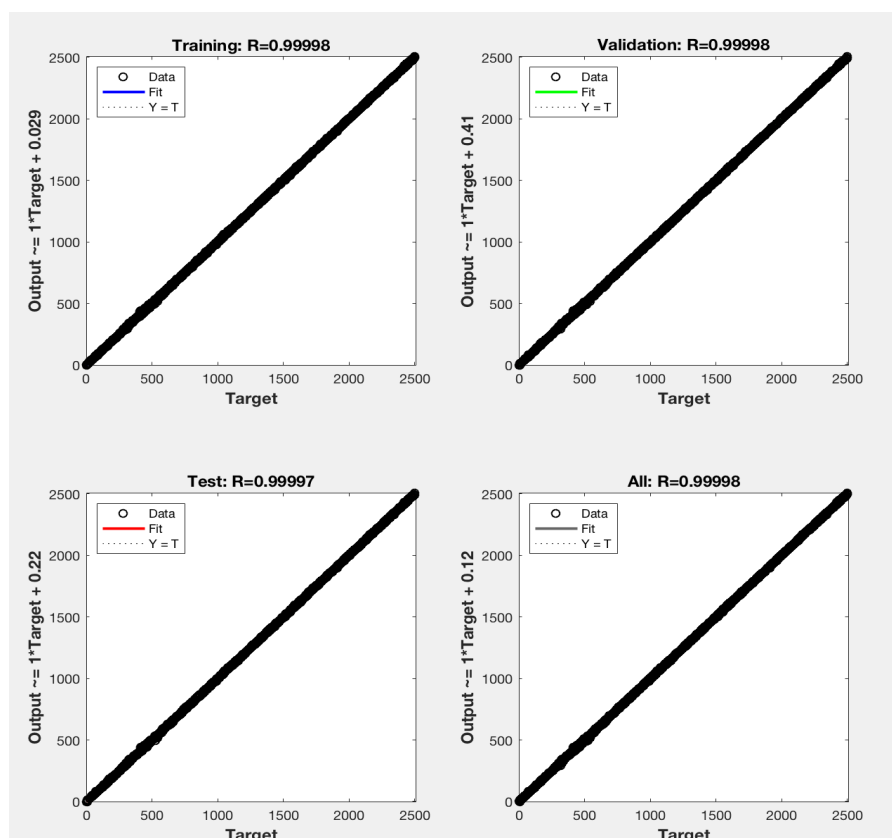
Εικόνα 4. 18 Ιστογράμμο σφαλμάτων δικτύου με 10720 δείγματα και 9 κόμβους

4.2.8 ΔΕΙΓΜΑ3, KOMBOI 10

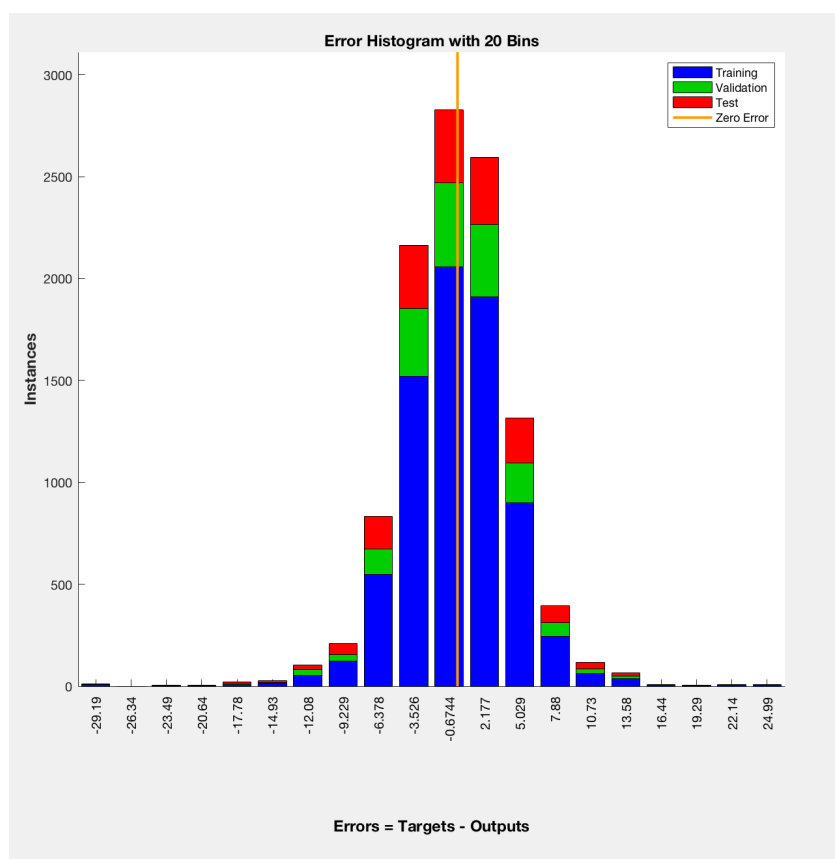
Για το δέκατο δίκτυο που εκπαιδεύτηκε, χρησιμοποιήθηκε δείγμα με πληθυσμό 10720 και για την γεωμετρία επιλέχθηκαν 10 κόμβοι. Παρατηρούνται τιμές μέσου τετραγωνικού σφάλματος αρκετά μειωμένες σε σχέση με προηγούμενα δίκτυα και τιμές συντελεστή παλινδρόμησης που δεν απέχουν σημαντικά από την μονάδα. (Πίνακας 4.10). Επίσης παρατηρείται κατανομή των αποτελεσμάτων (outputs) σε σχεδόν ταύτιση με την ευθεία $y=x$ στο διάγραμμα αποτελεσμάτων ως προς τις επιθυμητές τιμές (outputs-targets) (Εικόνα 4.19), πράγμα που επιβεβαιώνεται μερικώς και από την μορφή του ιστογράμματος σφαλμάτων (Εικόνα 4.20). Βάσει των παρατηρήσεων αυτών προκύπτει ότι η συγκεκριμένη γεωμετρία δικτύου δίνει ακόμα καλύτερα αποτελέσματα αλλά φαίνεται επίσης ότι μπορεί να βελτιωθεί περισσότερο και έτσι συνεχίζεται η διαδικασία προχωρώντας σε ακόμα μια αύξηση κόμβων.

SAMPLE 3 number of nodes 10	Sample size	MSE (mean squared error)	R(regression)
Training	7504	18.11e-0	9.99982e-1
Validation	1608	24.92e-0	9.99976e-1
testing	1608	27.28e-0	9.99974e-1

Πίνακας 4. 10 Χαρακτηριστικά στοιχεία δικτύου για πληθυσμό 10720 και 10 κόμβους



Εικόνα 4. 19 Γραφική απεικόνιση των αποτελεσμάτων δικτύου με 10720 δείγματα και 10 κόμβους σε σχέση με τους στόχους



Εικόνα 4. 20 Ιστόγραμμα σφαλμάτων δικτύου με 10720 δείγματα και 10 κόμβους

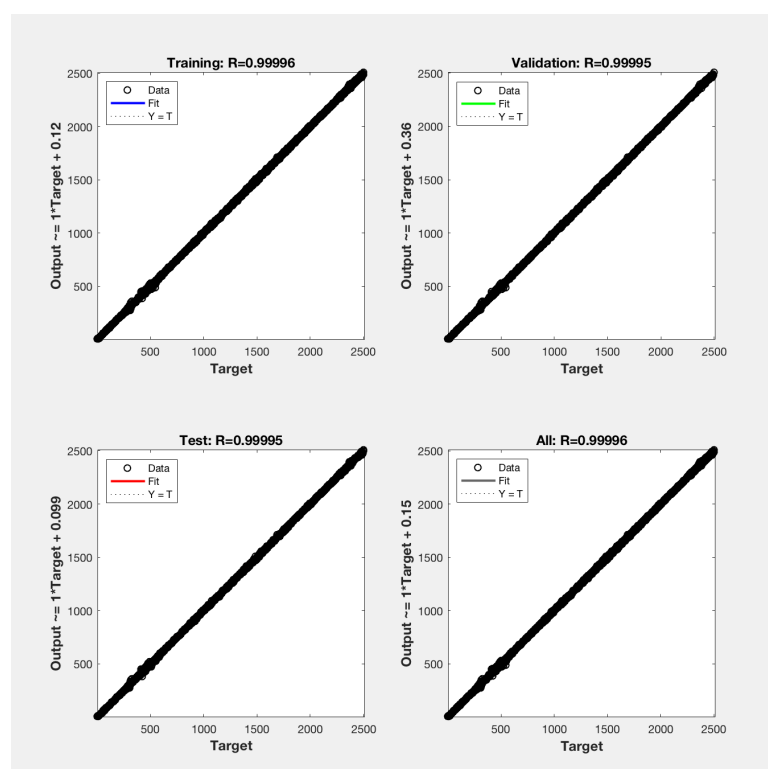
4.2.8 ΔΕΙΓΜΑ 3, 11 ΚΟΜΒΟΙ

Για το ενδέκατο δίκτυο που εκπαιδεύτηκε, χρησιμοποιήθηκε δείγμα με πληθυσμό 10720 και για την γεωμετρία επιλέχθηκαν 11 κόμβοι. Παρατηρούνται τιμές μέσου τετραγωνικού σφάλματος αυξημένες σε σχέση με το προηγούμενο δίκτυο και τιμές συντελεστή παλινδρόμησης που δεν απέχουν σημαντικά από την μονάδα (Πίνακας 4.11). Επίσης παρατηρούνται αποτελέσματα (outputs) σε σχεδόν ταύτιση με την ευθεία $y=x$ στο διάγραμμα αποτελεσμάτων ως προς τις επιθυμητές τιμές (outputs-targets) (Εικόνα 4.21), πράγμα που επιβεβαιώνεται μερικώς και από την μορφή του ιστογράμματος σφαλμάτων (Εικόνα 4.22).

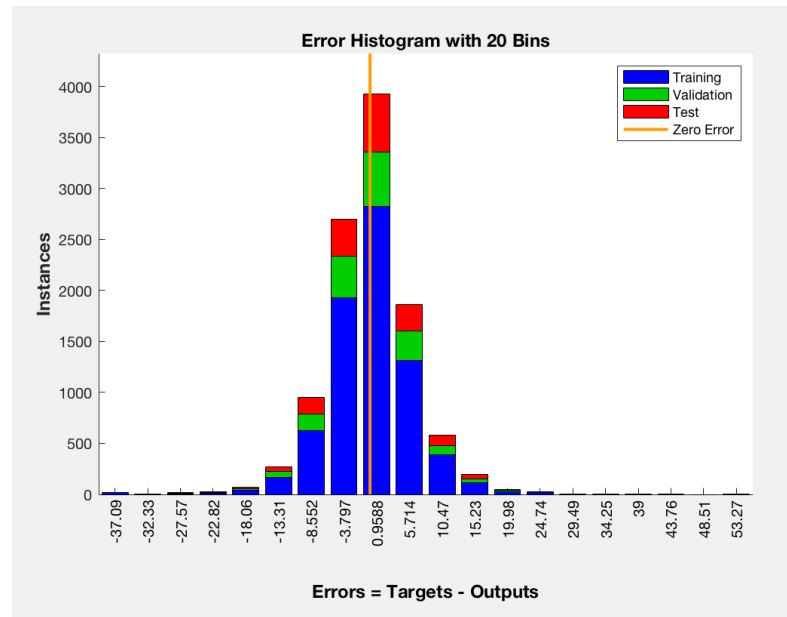
Βάσει των παρατηρήσεων αυτών προκύπτει ότι η συγκεκριμένη γεωμετρία δικτύου δεν δίνει καλύτερα αποτελέσματα και φαίνεται επίσης ότι η αύξηση των κόμβων δεν συνείσφερε θετικά στα αποτελέσματα. Ακολουθεί ακόμα μια αύξηση του αριθμού των κόμβων για επιβεβαίωση αυτού του συμπεράσματος.

SAMPLE 3 number of nodes 11	Sample size	MSE (mean squared error)	R(regression)
Training	7504	39.34e-0	9.99962e-1
Validation	1608	49.90e-0	9.99952e-1
testing	1608	50.33e-0	9.99953e-1

Πίνακας 4. 11 Χαρακτηριστικά στοιχεία δικτύου για πληθυσμό 10720 και 11 κόμβους



Εικόνα 4. 21 Γραφική απεικόνιση των αποτελεσμάτων δικτύου με 10720 δείγματα και 11 κόμβους σε σχέση με τους στόχους



Εικόνα 4. 22 Ιστόγραμμα σφαλμάτων δικτύου με 10720 δείγματα και 11 κόμβους

4.2.8 ΔΕΙΓΜΑ 3, 12 ΚΟΜΒΟΙ

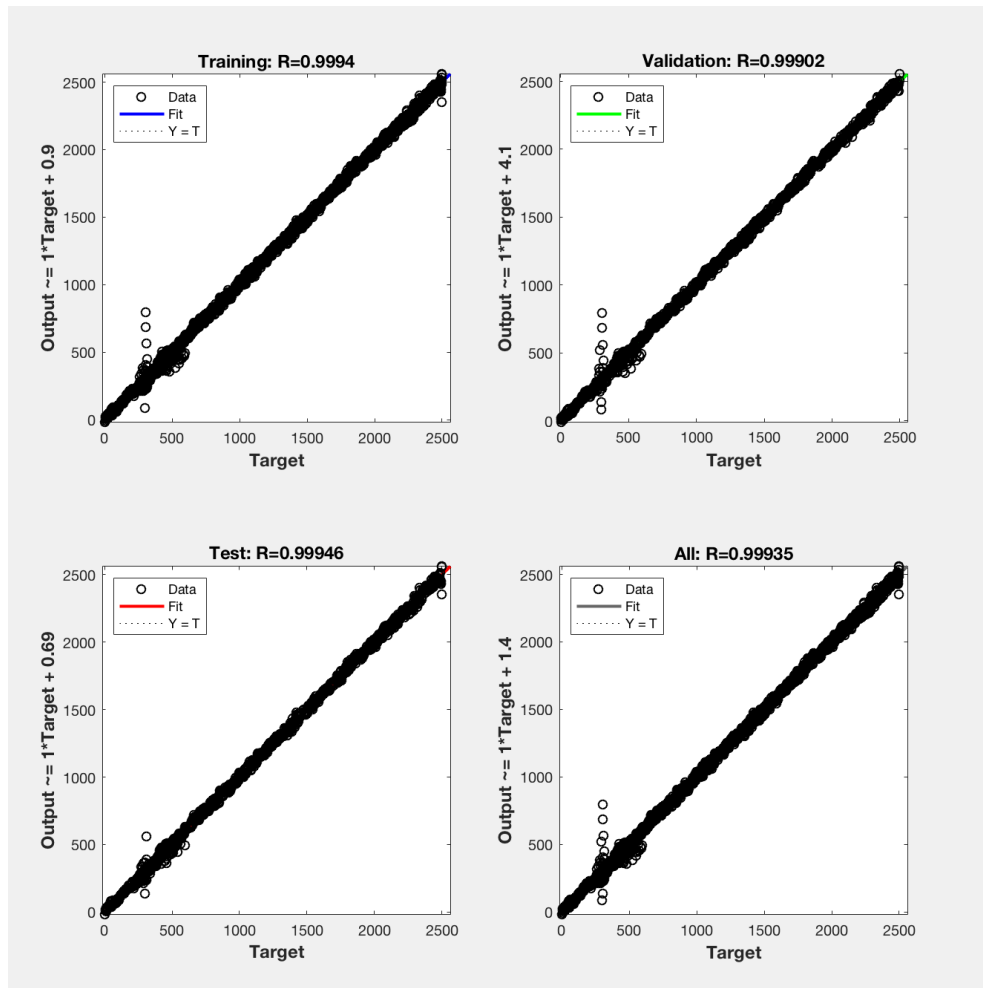
Για το δωδέκατο δίκτυο που εκπαιδεύτηκε, χρησιμοποιήθηκε δείγμα με πληθυσμό 10720 και για την γεωμετρία επιλέχθηκαν 11 κόμβοι. Παρατηρούνται μεγάλες τιμές μέσου τετραγωνικού σφάλματος και αυξημένες σε σχέση με προηγούμενα δίκτυα και τιμές συντελεστή παλινδρόμησης που δεν απέχουν σημαντικά από την μονάδα. (Πίνακας 4.12).

Επίσης παρατηρείται κατανομή των αποτελεσμάτων (outputs) που αρχίζει να απέχει από την ευθεία $y=x$ στο διάγραμμα αποτελεσμάτων ως προς τις επιθυμητές τιμές (outputs-targets) (Εικόνα 4.23), πράγμα που επιβεβαιώνεται μερικώς και από την μορφή του ιστογράμματος σφαλμάτων (Εικόνα 4.24).

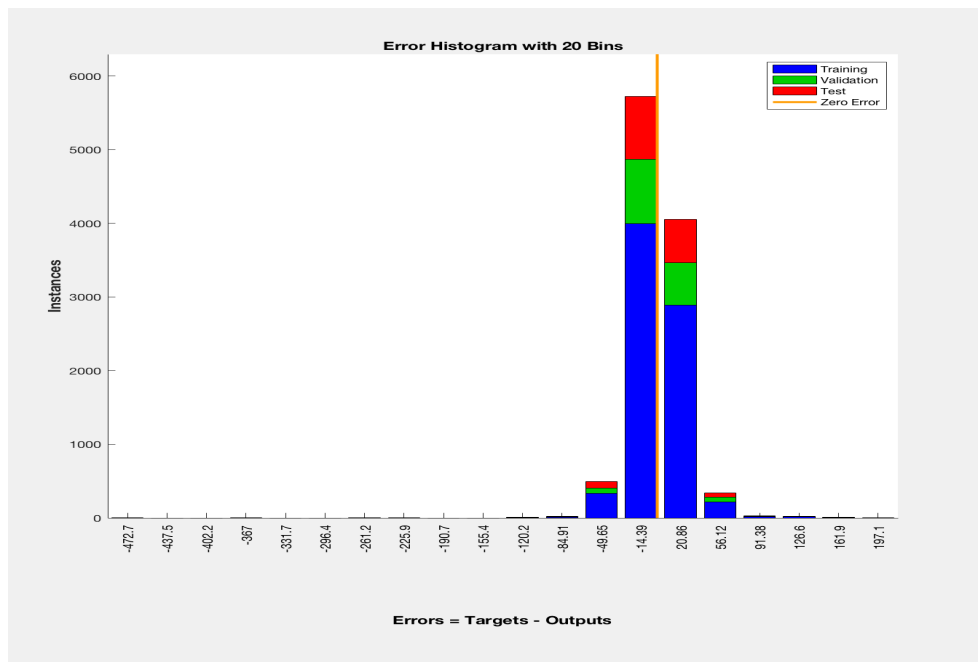
Βάσει των παρατηρήσεων αυτών προκύπτει ότι η συγκεκριμένη γεωμετρία δικτύου δεν δίνει καλύτερα αποτελέσματα και φαίνεται επίσης ότι η αύξηση των κόμβων τελικά δεν συνείσφερε θετικά στα αποτελέσματα. Ακολουθεί ένας τελευταίος διπλασιασμός του μεγέθους του δείγματος και επί μέρους δοκιμές με διαφορετικές γεωμετρίες δικτύων.

SAMPLE 3 number of nodes 12	Sample size	MSE (mean squared error)	R(regression)
Training	7504	640.17e-0	9.99399e-1
Validation	1608	1061.90e-0	9.99015e-1
testing	1608	550.79e-0	9.99463e-1

Πίνακας 4. 12 Χαρακτηριστικά στοιχεία δικτύου για πληθυσμό 10720 και 12 κόμβους



Εικόνα 4. 23 Γραφική απεικόνιση των αποτελεσμάτων δικτύου με 10720 δείγματα και 12 κόμβους σε σχέση με τους στόχους



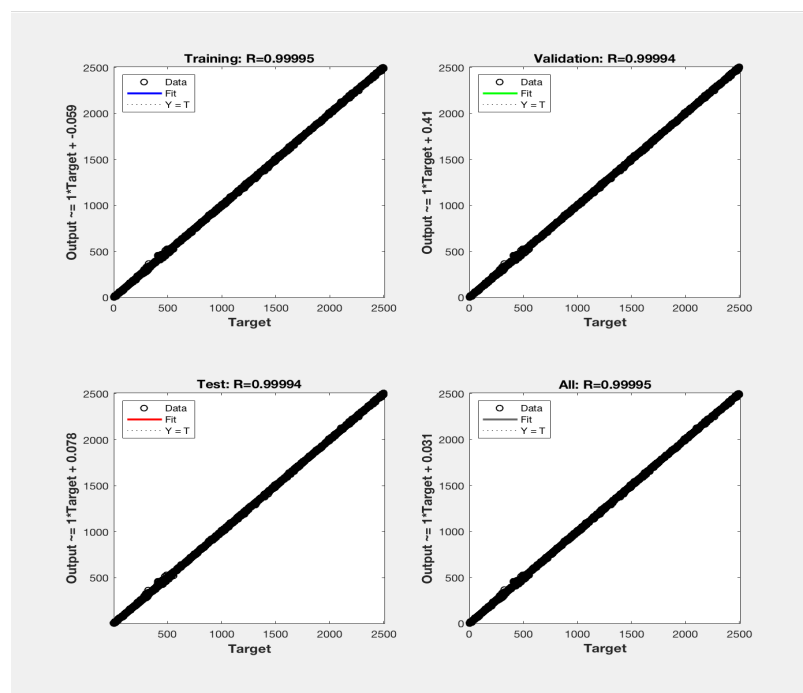
Εικόνα 4. 24 Ιστόγραμμα σφαλμάτων δικτύου με 10720 δείγματα και 12 κόμβους

4.2.13 ΔΕΙΓΜΑ 4,8 ΚΟΜΒΟΙ

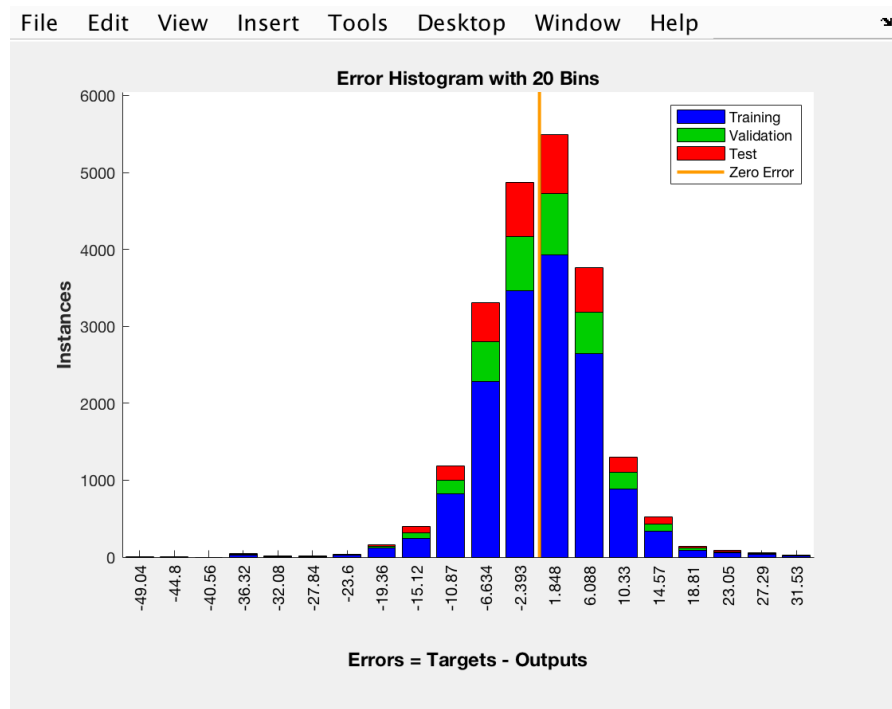
Για το δέκατο τρίτο δίκτυο που εκπαιδεύτηκε, χρησιμοποιήθηκε δείγμα με πληθυσμό 21440 και για την γεωμετρία επιλέχθηκαν 8 κόμβοι. Παρατηρούνται μειωμένες τιμές μέσου τετραγωνικού σφάλματος σε σχέση με το προηγούμενο δίκτυο και τιμές συντελεστή παλινδρόμησης που δεν απέχουν σημαντικά από την μονάδα. (Πίνακας 4.13). Επίσης παρατηρείται κατανομή των αποτελεσμάτων (outputs) που δεν απέχει από την ευθεία $y=x$ στο διάγραμμα αποτελεσμάτων ως προς τις επιθυμητές τιμές (outputs-targets) (Εικόνα 4.8), πράγμα που επιβεβαιώνεται μερικώς και από την μορφή του ιστογράμματος σφαλμάτων (Εικόνα 4.9). Βάσει των παρατηρήσεων αυτών προκύπτει ότι ο τελευταίος διπλασιασμός επηρέασε θετικά τα αποτελέσματα και τα έφερε σε αποδεκτά επίπεδα. Ακολουθεί ακόμα μια αύξηση του αριθμού των κόμβων για να ελεγχθεί πως αυτό θα επηρεάσει τα αποτελέσματα.

SAMPLE 4 number of nodes 8	Sample size	MSE (mean squared error)	R(regression)
Training	15008	51.09e-0	9.99951e-1
Validation	3216	63.70e-0	9.99939e-1
testing	3216	59.69e-0	9.99944e-1

Πίνακας 4. 13 Χαρακτηριστικά στοιχεία δικτύου για πληθυσμό 21440 και 8 κόμβους



Εικόνα 4. 25 Γραφική απεικόνιση των αποτελεσμάτων δικτύου με 21440 δείγματα και 8 κόμβους σε σχέση με τους στόχους



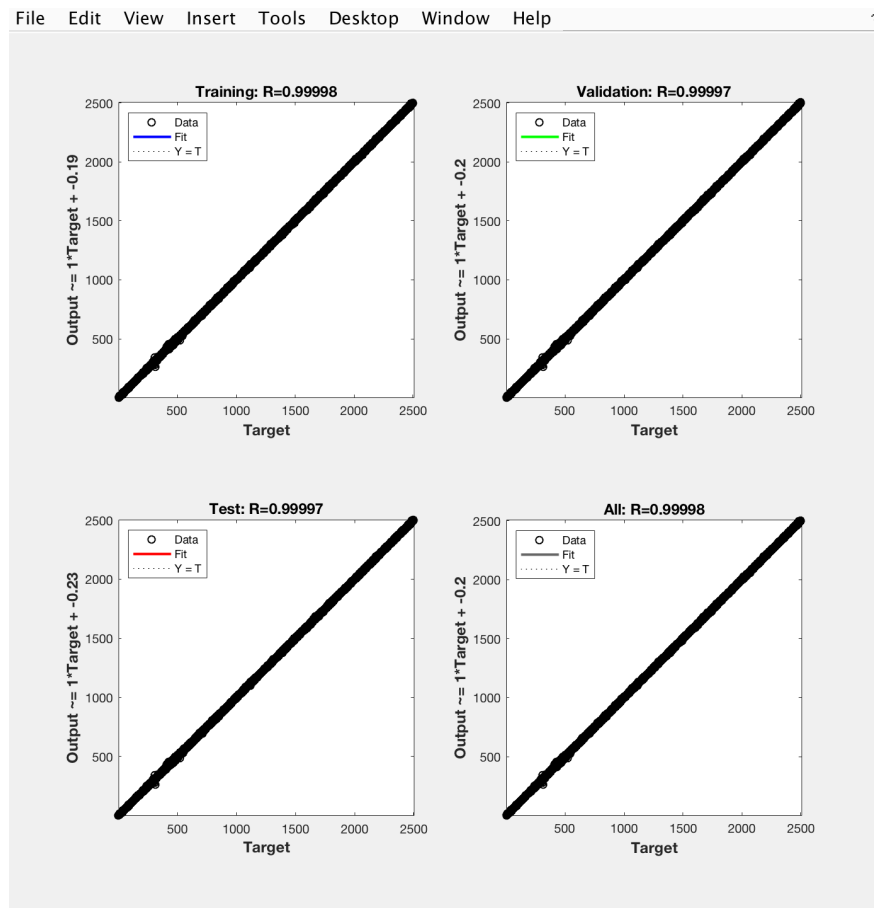
Εικόνα 4. 26 Ιστόγραμμα σφαλμάτων δικτύου με 21440 δείγματα και 8 κόμβους

4.2.14 ΔΕΙΓΜΑ 4, 9 ΚΟΜΒΟΙ

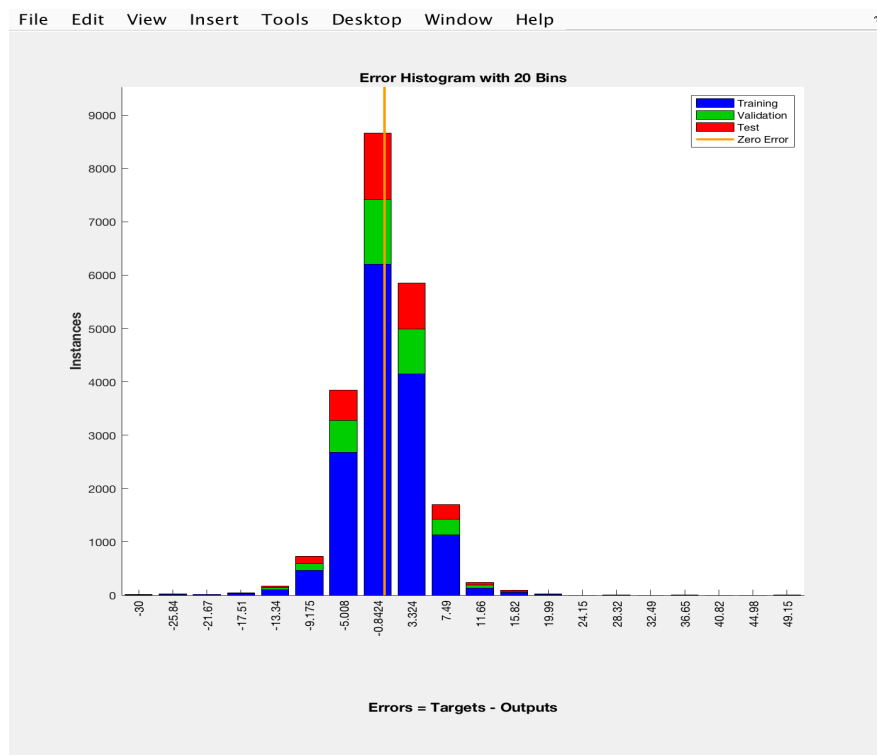
Για το δέκατο τέταρτο δίκτυο που εκπαιδεύτηκε, χρησιμοποιήθηκε δείγμα με πληθυσμό 21440 και για την γεωμετρία επιλέχθηκαν 9 κόμβοι. Παρατηρούνται μειωμένες τιμές μέσου τετραγωνικού σφάλματος σε σχέση με προηγούμενα δίκτυα και τιμές συντελεστή παλινδρόμησης που δεν απέχουν σημαντικά από την μονάδα. (Πίνακας 4.14). Επίσης παρατηρείται κατανομή των αποτελεσμάτων (outputs) που δεν απέχει από την ευθεία $y=x$ στο διάγραμμα αποτελεσμάτων ως προς τις επιθυμητές τιμές (outputs-targets) (Εικόνα 4.27), πράγμα που επιβεβαιώνεται μερικώς και από την μορφή του ιστογράμματος σφαλμάτων (Εικόνα 4.28). Βάσει των παρατηρήσεων αυτών προκύπτει ότι η αύξηση των κόμβων έφερε πολύ καλά αποτελέσματα, με το μικρότερο έως τώρα μέσο τετραγωνικό σφάλμα. Ακολουθούν και άλλες αυξήσεις στον αριθμό των κόμβων για να εξεταστεί το κατά πόσο ακόμα μπορούν να βελτιωθούν τα αποτελέσματα.

SAMPLE 4 number of nodes 9	Sample size	MSE (mean squared error)	R(regression)
Training	15008	19.90e-0	9.99981e-1
Validation	3216	27.14e-0	9.99974e-1
Testing	3216	27.10e-0	9.99974e-1

Πίνακας 4. 14 Χαρακτηριστικά στοιχεία δικτύου για πληθυσμό 21440 και 9 κόμβους



Εικόνα 4. 27 Γραφική απεικόνιση των αποτελεσμάτων δικτύου με 21440 δείγματα και 9 κόμβους σε σχέση με τους στόχους



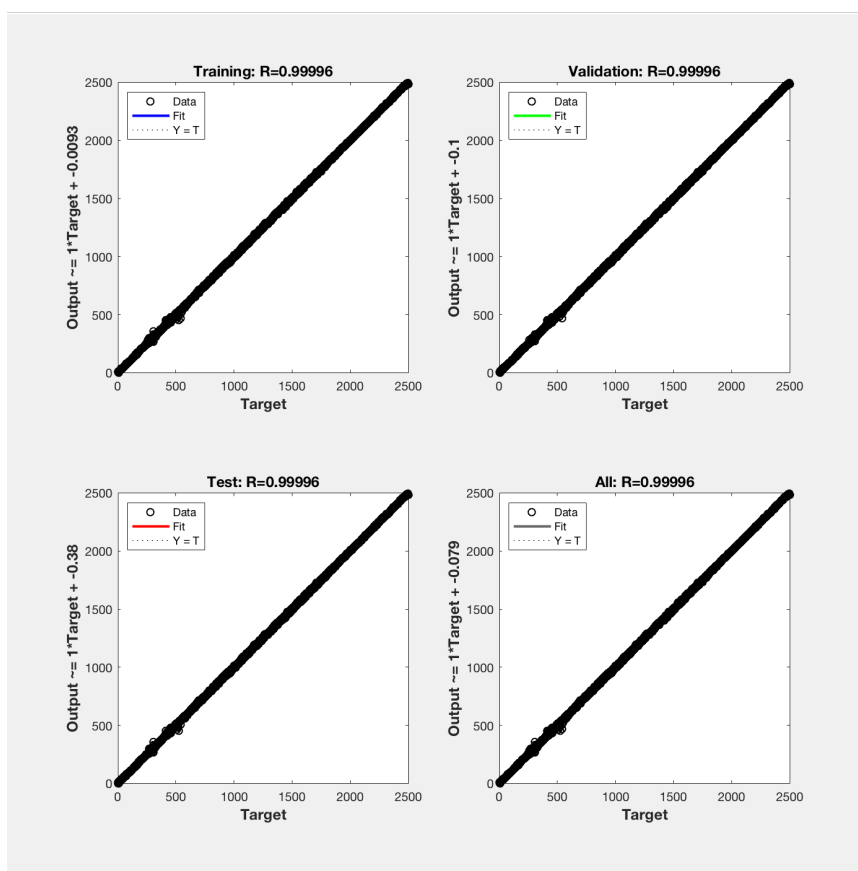
Εικόνα 4. 28 Ιστόγραμμα σφαλμάτων δικτύου με 21440 δείγματα και 9 κόμβους

4.2.15 ΔΕΙΓΜΑ 4, 10 ΚΟΜΒΟΙ

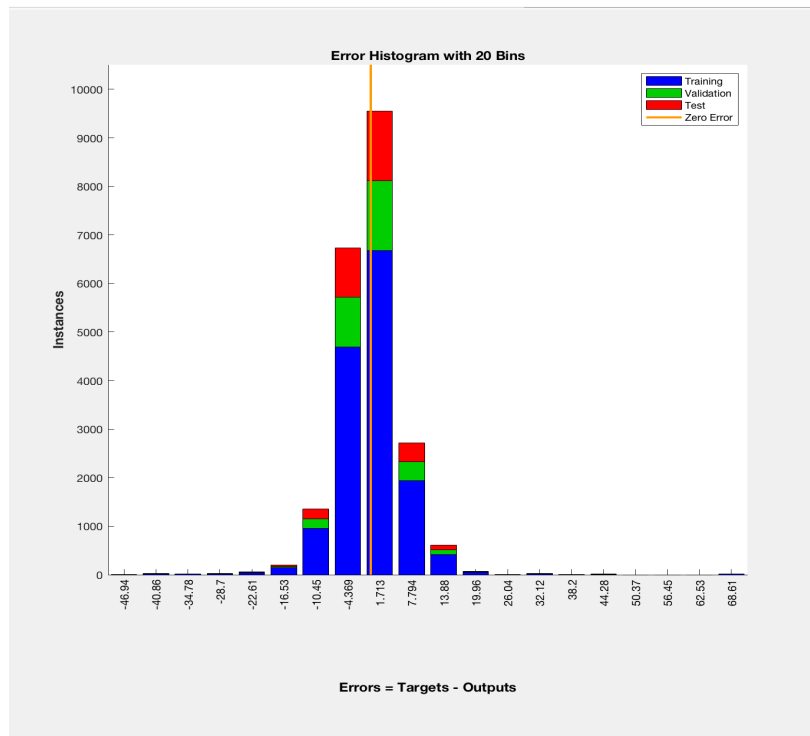
Για το δέκατο πέμπτο δίκτυο που εκπαιδεύτηκε, χρησιμοποιήθηκε δείγμα με πληθυσμό 21440 και για την γεωμετρία επιλέχθηκαν 10 κόμβοι. Παρατηρούνται αυξημένες τιμές μέσου τετραγωνικού σφάλματος σε σχέση με προηγούμενα δίκτυα και τιμές συντελεστή παλινδρόμησης που δεν απέχουν σημαντικά από την μονάδα. (πίνακας 4.15). Επίσης παρατηρείται κατανομή των αποτελεσμάτων (outputs) που δεν απέχει από την ευθεία $y=x$ στο διάγραμμα αποτελεσμάτων ως προς τις επιθυμητές τιμές (outputs-targets) (Εικόνα 4.29), πράγμα που επιβεβαιώνεται μερικώς και από την μορφή του ιστογράμματος σφαλμάτων (Εικόνα 4.30). Βάσει των παρατηρήσεων αυτών προκύπτει ότι η αύξηση των κόμβων δεν έφερε καλύτερα αποτελέσματα, σε σχέση με την προηγούμενη γεωμετρία, ακολουθεί ακόμα μια αύξηση στον αριθμό των κόμβων για την μελέτη των αποτελεσμάτων.

SAMPLE 4 number of nodes 10	Sample size	MSE (mean squared error)	R(regression)
Training	15008	40.15e-0	9.99961e-1
Validation	3216	41.00e-0	9.99962e-1
testing	3216	46.00e-0	9.99957e-1

Πίνακας 4. 15 Χαρακτηριστικά στοιχεία δικτύου για πληθυσμό 21440 και 10 κόμβους



Εικόνα 4. 29 Γραφική απεικόνιση των αποτελεσμάτων δικτύου με 21440 δείγματα και 10 κόμβους σε σχέση με τους στόχους



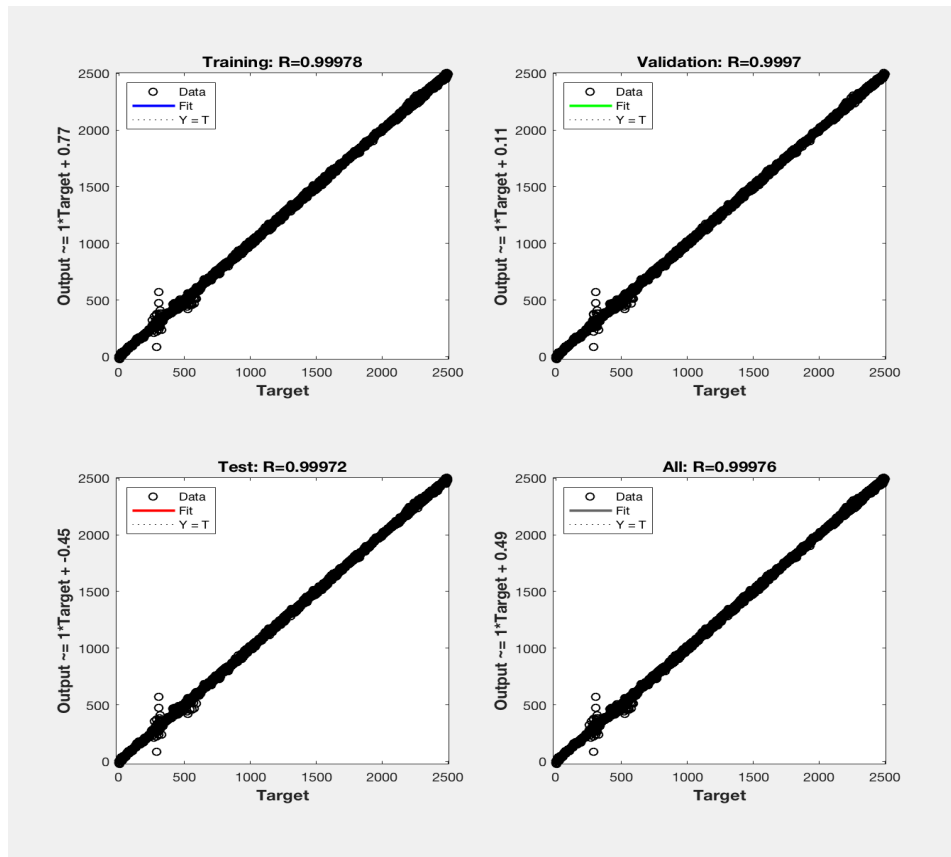
Εικόνα 4. 30 Ιστόγραμμα σφαλμάτων δικτύου με 21440 δείγματα και 10 κόμβους

4.2.16 ΔΕΙΓΜΑ 4, 11 ΚΟΜΒΟΙ

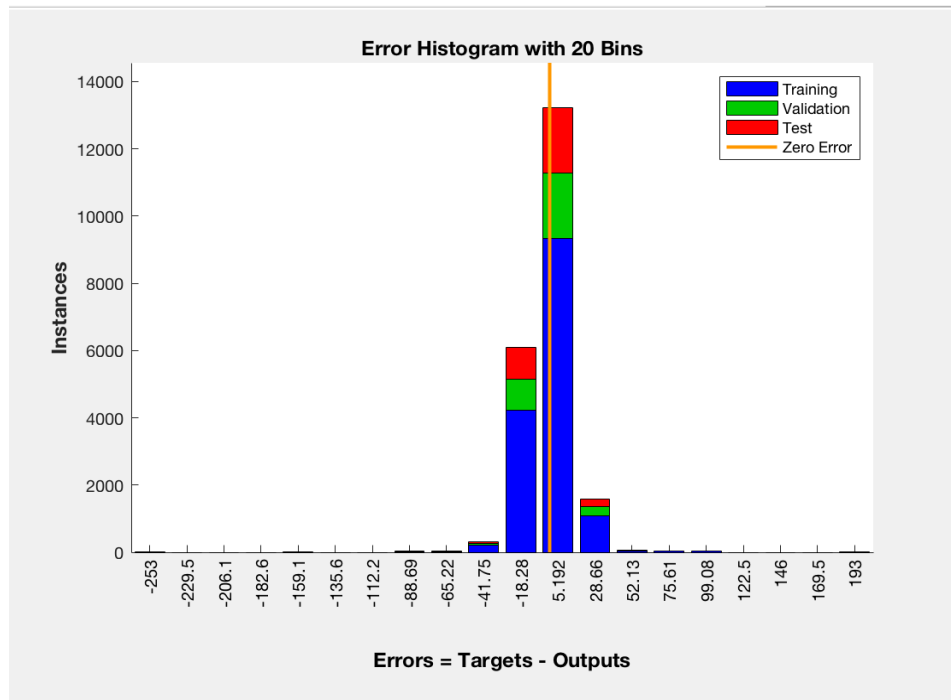
Για το δέκατο έκτο δίκτυο που εκπαιδεύτηκε, χρησιμοποιήθηκε δείγμα με πληθυσμό 21440 και για την γεωμετρία επιλέχθηκαν 11 κόμβοι. Παρατηρούνται ακόμα περισσότερο αυξημένες τιμές μέσου τετραγωνικού σφάλματος σε σχέση με προηγούμενα δίκτυα και τιμές παλινδρόμησης που δεν απέχουν σημαντικά από την μονάδα. (Πίνακας 4.16). Επίσης παρατηρείται κατανομή των αποτελεσμάτων (outputs) που δεν απέχει από την ευθεία $y=x$ στο διάγραμμα αποτελεσμάτων ως προς τις επιθυμητές τιμές (outputs-targets) (Εικόνα 4.31), πράγμα που επιβεβαιώνεται μερικώς και από την μορφή του ιστογράμματος σφαλμάτων (Εικόνα 4.32). Βάσει των παρατηρήσεων αυτών προκύπτει ότι η αύξηση των κόμβων δεν έφερε καλύτερα αποτελέσματα, αντιθέτως αύξησε σημαντικά το μέγεθος του σφάλματος σε σχέση με την προηγούμενη γεωμετρία. Αυτή τη φορά ακολουθεί ακόμα μια σημαντική αύξηση στον αριθμό των κόμβων.

SAMPLE 4 number of nodes 11	Sample size	MSE (mean squared error)	R(regression)
Training	15008	238.77e-0	9.99776e-1
Validation	3216	314.61e-0	9.99704e-1
testing	3216	297.62e-0	9.99716e-1

Πίνακας 4. 16 Χαρακτηριστικά στοιχεία δικτύου για πληθυσμό 21440 και 11 κόμβους



Εικόνα 4. 31 Γραφική απεικόνιση των αποτελεσμάτων δικτύου με 21440 δείγματα και 11 κόμβους σε σχέση με τους στόχους



Εικόνα 4. 32 Ιστογράμμο σφαλμάτων δικτύου με 21440 δείγματα και 11 κόμβους

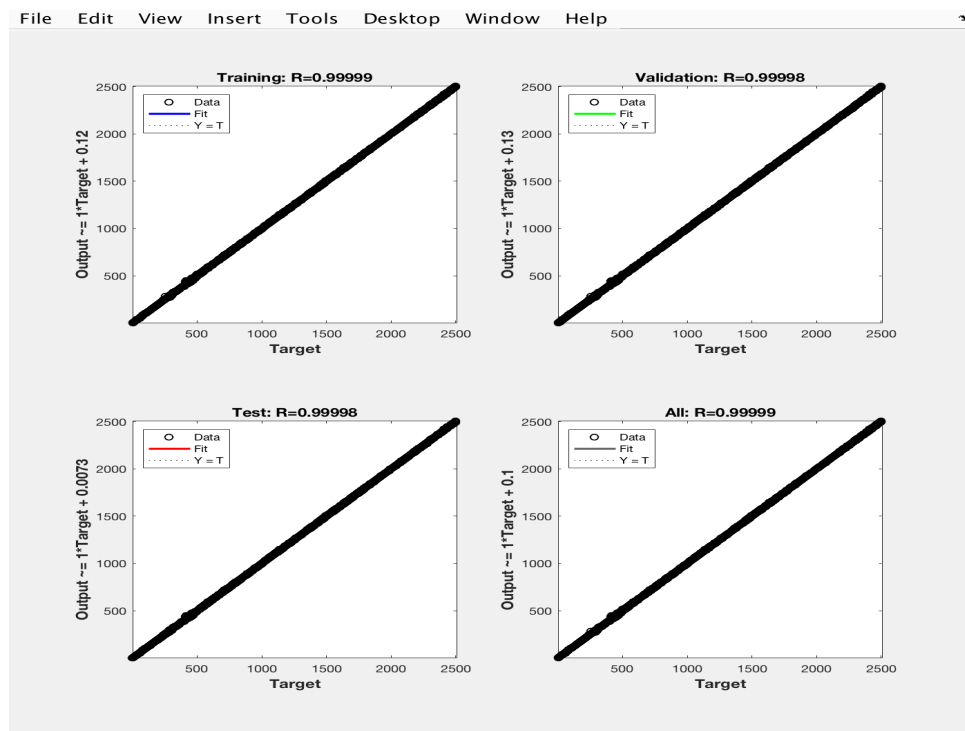
4.2.17 ΔΕΙΓΜΑ 4, 20 ΚΟΜΒΟΙ

Για το δέκατο έβδομο δίκτυο που εκπαιδεύτηκε, χρησιμοποιήθηκε δείγμα με πληθυσμό 21440 και για την γεωμετρία επιλέχθηκαν 20 κόμβοι. Παρατηρούνται οι χαμηλότερες μέχρι στιγμής τιμές μέσου τετραγωνικού σφάλματος σε σχέση με προηγούμενα δίκτυα και τιμές συντελεστή παλινδρόμησης που απέχουν ελάχιστα από την μονάδα. (Πίνακας 4.17). Επίσης παρατηρείται κατανομή των αποτελεσμάτων (outputs) που δεν απέχει σχεδόν καθόλου από την ευθεία $y=x$ στο διάγραμμα αποτελεσμάτων ως προς τις επιθυμητές τιμές (outputs-targets) (Εικόνα 4.33), πράγμα που επιβεβαιώνεται μερικώς και από την μορφή του ιστογράμματος σφαλμάτων (Εικόνα 4.34).

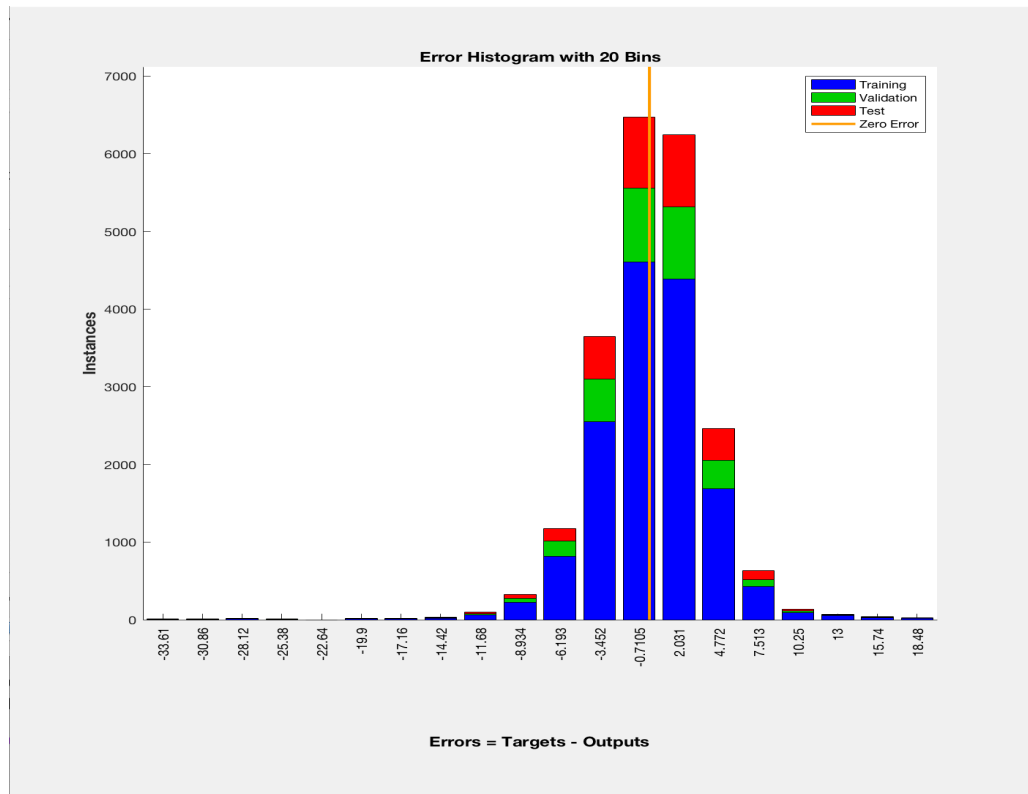
Βάσει των παρατηρήσεων αυτών προκύπτει ότι η αύξηση των κόμβων έφερε καλύτερα αποτελέσματα. Η διαδικασία σταματάει εδώ καθώς προέκυψαν αρκετά αποτελέσματα για να εξαχθεί ένα γενικό συμπέρασμα

SAMPLE 4 number of nodes 20	Sample size	MSE (mean squared error)	R(regression)
Training	15008	15.04e-0	9.99985e-1
Validation	3216	17.19e-0	9.99983e-1
Testing	3216	16.55e-0	9.99984e-1

Πίνακας 4. 17 Χαρακτηριστικά στοιχεία δικτύου για πληθυσμό 21440 και 20 κόμβους



Εικόνα 4. 33 Γραφική απεικόνιση των αποτελεσμάτων δικτύου με 21440 δείγματα και 20 κόμβους σε σχέση με τους στόχους



Εικόνα 4. 34 Ιστόγραμμα σφαλμάτων δικτύου με 21440 δείγματα και 8 κόμβους

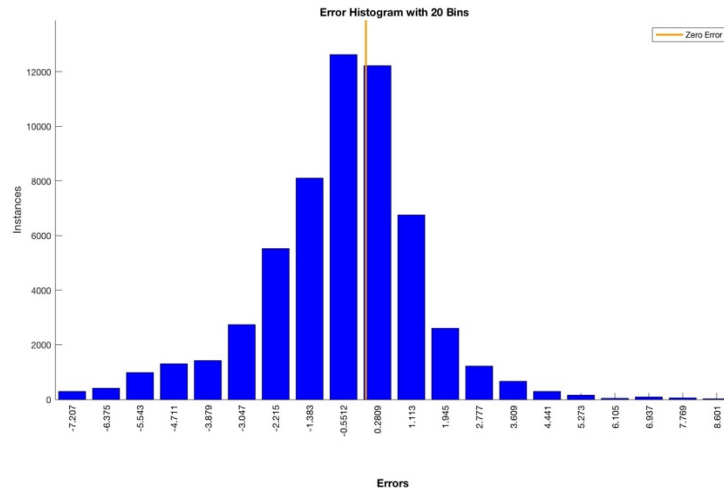
4.3 ΣΥΓΚΡΙΣΗ ΑΠΟΤΕΛΕΣΜΑΤΩΝ ΑΠΟ ΝΕΥΡΩΝΙΚΟ ΔΙΚΤΥΟ ΜΕ ΑΠΟΤΕΛΕΣΜΑΤΑ ΑΠΟ ΕΞΕΙΔΙΚΕΥΜΕΝΟ ΧΡΗΣΤΗ

Σε αυτό το σημείο ακολουθεί μια σύγκριση των πρώτων αφίξεων όπως αυτές επιλέγονται από το νευρωνικό δίκτυο σε αντιδιαστολή με τις θέσεις των πρώτων αφίξεων όπως αυτές επιλέχθηκαν από έναν έμπειρο και εξειδικευμένο χρήστη, Δρ Γ. Κρητικάκη.

4.3.1 ΔΙΑΔΙΚΑΣΙΑ ΣΥΓΚΡΙΣΗΣ

Για τη σύγκριση επιλογών μεταξύ του νευρωνικού και του χρήστη ακολουθήθηκε η εξής διαδικασία:

- Αρχικά επιλέγονται τα δεδομένα που θα επεξεργαστούν και συγκεντρώνονται σε πίνακα αντίστοιχο με αυτόν που χρησιμοποιήθηκε για την εκπαίδευση του δικτύου.
- Χρησιμοποιείται το βέλτιστο νευρωνικό, του οποίου η δημιουργία παρουσιάστηκε παραπάνω και οι πρώτες αφίξεις αποθηκεύονται σε μορφή χρόνων διαδρομής (ms).
- Το επόμενο βήμα είναι η σύγκριση των τιμών του νευρωνικού και του χρήστη, για το σκοπό αυτό δημιουργείται πίνακας που περιέχει τις διαφορές των τιμών αυτών.
- Τελικά με την χρήση των τιμών των διαφορών κατασκευάζεται το ιστόγραμμα σφαλμάτων (Εικόνα 4.35).



Εικόνα 4. 35 Ιστόγραμμα σφαλμάτων για την σύγκριση των αποτελεσμάτων δικτύου-χρήστη σε ms

Στο ιστόγραμμα αυτό παρατηρείται μια κανονική κατανομή γύρω από την ευθεία που αντιστοιχεί στο μηδενικό σφάλμα, αυτό σημαίνει ότι δεν υπάρχει συστηματικό σφάλμα. Παρατηρούνται επίσης κάποιες αρκετά αυξημένες τιμές σφάλματος (έως και 8ms), αυτό οφείλεται στον διαφορετικό τρόπο επιλογής των πρώτων αφίξεων μεταξύ του συγγραφέα βάσει του οποίου εκπαιδεύτηκε το νευρωνικό και του εξειδικευμένου χρήστη.

5 ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΜΕΛΛΟΝΤΙΚΕΣ ΠΡΟΤΑΣΕΙΣ

5.1 ΕΙΣΑΓΩΓΗ

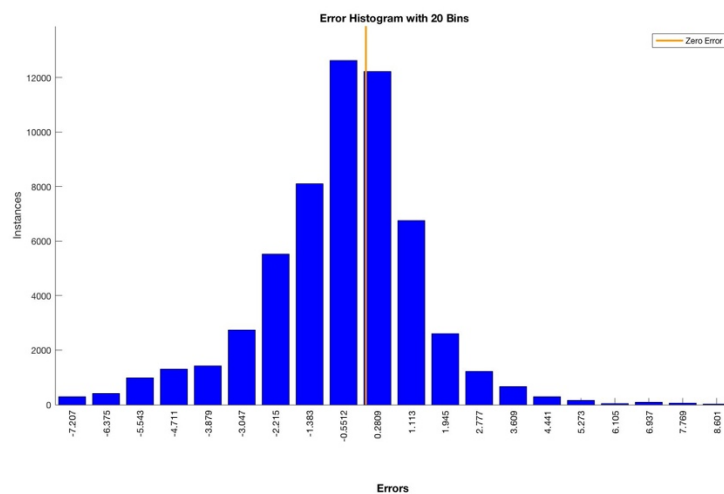
Ο σκοπός της εργασίας αυτής όπως αναφέρθηκε και σε προηγούμενα κεφάλαια ήταν η επιβεβαίωση του κατά πόσον είναι δυνατό να υπάρξει μια αυτόματη διαδικασία επιλογής πρώτων αφίξεων με την χρήση τεχνητών νευρωνικών δικτύων. Τα αποτελέσματα που προέκυψαν όπως θα φανεί και παρακάτω είναι αρκετά ενθαρρυντικά ώστε να δικαιώσουν την ιδέα και να δημιουργήσουν την ανάγκη για εκτενέστερη μελέτη του αντικειμένου στο μέλλον.

5.2 ΣΥΜΠΕΡΑΣΜΑΤΑ

Προέκυψαν αρκετά και ενδιαφέροντα συμπεράσματα. Όσον αφορά το γενικό κομμάτι της εργασίας, αποδείχτηκε ότι η χρήση νευρωνικών δικτύων και εργαλείων τεχνητής νοημοσύνης βρίσκει σημαντική εφαρμογή στα κομμάτια της γεωφυσικής. Τα μοντέλα που αναπτυχθήκαν για την επιλογή των πρώτων αφίξεων δίνουν αξιόλογα αποτελέσματα και επιβεβαιώνουν την ιδέα καθώς και ενθαρρύνουν την περαιτέρω εφαρμογή αυτών και ίσως, την ενσωμάτωση τους σε διαδικασίες επιλογής πρώτων αφίξεων. Όπως φαίνεται και παρακάτω (Εικόνα 5.1) το βέλτιστο δίκτυο που προέκυψε από δοκιμές έχει πολύ καλά στατιστικά χαρακτηριστικά και τα αποτελέσματα του είναι πολύ κοντά σε αυτά που προέρχονται από έναν έμπειρο χρήστη(Εικόνα 5.2)

SAMPLE 4 population of number of nodes 20	Sample size	MSE (mean squared error)	R(regression)
Training	15008	15.04100e-0	9.99985e-1
Validation	3216	17.19121e-0	9.99983e-1
Testing	3216	16.55102e-0	9.99984e-1

Εικόνα 5. 1 Χαρακτηριστικά βέλτιστου δικτύου



Εικόνα 5. 2 Σύγκριση Νευρωνικού δικτύου με εξειδικευμένο χρήστη

Όσον αφορά τα κομμάτια των νευρωνικών δικτύων εξήχθησαν δυο βασικά συμπεράσματα. Το πρώτο είναι ότι το μέγεθος του δείγματος επηρεάζει σημαντικά το μοντέλο καθώς όπως φάνηκε με μεγαλύτερο πλήθος παρατηρήσεων, υπάρχει σημαντική βελτίωση στα αποτελέσματα. Το δεύτερο συμπέρασμα είναι ότι ο αριθμός των κόμβων φαίνεται να επηρεάζει θετικά τα αποτελέσματα όταν αυτός αυξάνεται, αλλά όταν ξεπεραστεί ένα συγκεκριμένο σημείο δεν υπάρχει βελτίωση στα αποτελέσματα.

5.3 ΠΡΟΒΛΗΜΑΤΑ

Τα δύο σημαντικά προβλήματα που εντοπίστηκαν είναι ότι ένα νευρωνικό δίκτυο εκπαιδεύεται με βάση δεδομένα που προέκυψαν από κάποιο ανθρώπινο χρήστη, πράγμα που σημαίνει ότι το δίκτυο είναι πιθανό να υποπέσει στα ίδια λάθη που θα έκανε και ο χρήστης βάσει του οποίου εκπαιδεύτηκε. Το δεύτερο πρόβλημα που προκύπτει είναι η περιορισμένη βάση δεδομένων που υπάρχει αυτήν τη στιγμή όσον αφορά την επιλογή πρώτων αφίξεων, μια διαδικασία τεχνικά δύσκολη και αρκετά χρονοβόρα.

5.4 ΜΕΛΛΟΝΤΙΚΕΣ ΠΡΟΤΑΣΕΙΣ

Γενικά, προτείνεται η δοκιμαστική εφαρμογή των τεχνητών νευρωνικών δικτύων και σε άλλα προβλήματα στο αντικείμενο της γεωφυσικής καθώς είναι μια τεχνολογία που βοηθάει σημαντικά στην εξοικονόμηση του χρόνου που χρειάζεται για την επίλυση

κάποιων από αυτών, στην κατηγορία αυτή εμπίπτουν όλα τα προβλήματα που σχετίζονται με αναγνώριση μοτίβων (pattern recognition), για παράδειγμα στην ανάλυση ταχυτήτων.

Ειδικά, προτείνεται η ενσωμάτωση της αυτόματης πρόβλεψης με χρήση τεχνητών νευρωνικών δικτύων σε γεωφυσικές διασκοπήσεις, σαν μια πρώτη εικόνα και σε συνδυασμό με τον ανθρώπινο παράγοντα δημιουργείται μια μέθοδος που εξοικονομεί σημαντικά χρόνο.

Επίσης προτείνεται η δημιουργία βάσης δεδομένων από επιλογή πρώτων αφίξεων, η οποία θα συμβάλει στην καλύτερη εκπαίδευση των μελλοντικών δικτύων.

Τέλος με την χρήση απλών δομών τεχνητών νευρωνικών δικτύων ανοίγει ο δρόμος για την δημιουργία δικτύων βαθιάς μάθησης (deep neural networks) για την πλήρη αυτοματοποίηση της διαδικασίας επεξεργασίας σεισμικών δεδομένων.

6 ΚΩΔΙΚΑΣ MATLAB ΓΙΑ ΤΗΝ ΔΗΜΙΟΥΡΓΙΑ ΝΕΥΡΩΝΙΚΟΥ ΔΙΚΤΥΟΥ

```
function [Y,Xf,Af] = nnet(X,~,~)
%
%
% [Y] = myNeuralNetworkFunction(X,~,~) takes these arguments:
%
% X = 1xTS cell, 1 inputs over TS timesteps
% Each X{1,ts} = 6400xQ matrix, input #1 at timestep ts.
%
% and returns:
% Y = 1xTS cell of 1 outputs over TS timesteps.
% Each Y{1,ts} = 1xQ matrix, output #1 at timestep ts.
%
% where Q is number of samples (or series) and TS is the number of timesteps.

%#ok<*RPMT0>

% ===== NEURAL NETWORK CONSTANTS =====

% Input 1
x1_step1.xoffset = ....
x1_step1.gain = ....
x1_step1.ymin = -1;

% Layer 2
b2 = 0.50352337330744523225;
LW2_1 = ...
% Output 1
y1_step1.ymin = -1;
y1_step1.gain = 0.000800320128051221;
y1_step1.xoffset = 1;

% ===== SIMULATION =====

% Format Input Arguments
isCellX = iscell(X);
```

```

if ~isCellX
    X = {X};
end

% Dimensions
TS = size(X,2); % timesteps
if ~isempty(X)
    Q = size(X{1},2); % samples/series
else
    Q = 0;
end

% Allocate Outputs
Y = cell(1,TS);

% Time loop
for ts=1:TS

    % Input 1
    Xp1 = mapminmax_apply(X{1,ts},x1_step1);

    % Layer 1
    a1 = tansig_apply(repmat(b1,1,Q) + IW1_1*Xp1);

    % Layer 2
    a2 = repmat(b2,1,Q) + LW2_1*a1;

    % Output 1
    Y{1,ts} = mapminmax_reverse(a2,y1_step1);
end

% Final Delay States
Xf = cell(1,0);
Af = cell(2,0);

% Format Output Arguments
if ~isCellX
    Y = cell2mat(Y);
end
end

% ===== MODULE FUNCTIONS =====

% Map Minimum and Maximum Input Processing Function
function y = mapminmax_apply(x,settings)
y = bsxfun(@minus,x,settings.xoffset);
y = bsxfun(@times,y,settings.gain);
y = bsxfun(@plus,y,settings.ymin);
end

% Sigmoid Symmetric Transfer Function
function a = tansig_apply(n,~)
a = 2 ./ (1 + exp(-2*n)) - 1;
end

% Map Minimum and Maximum Output Reverse-Processing Function
function x = mapminmax_reverse(y,settings)
x = bsxfun(@minus,y,settings.ymin);

```



```

x = bsxfun(@rdivide,x,settings.gain);
x = bsxfun(@plus,x,settings.xoffset);
end

```

7 Παράρτημα I-Αλγόριθμοι: Steepest descent, Newton, Gauss-Newton

7.1 Αλγόριθμος Steepest descent

Ο αλγόριθμος αυτός είναι ένας πρώτης τάξης αλγόριθμος. Χρησιμοποιεί την πρώτη παράγωγο της συνάρτησης του ολικού σφάλματος για να βρει την ελάχιστη τιμή στον χώρο λαθών. Φυσιολογικά, η κλίση g ορίζεται ως η πρώτη παράγωγος της συνάρτησης ολικού σφάλματος:

$$g = \frac{\partial E(x, w)}{\partial w} = \left[\frac{\partial E}{\partial w_1} \quad \frac{\partial E}{\partial w_2} \quad \dots \quad \frac{\partial E}{\partial w_N} \right]^T \quad (7.1)$$

Με τον ορισμό της κλίσης g στην 7.1 ο κανόνας για την μέθοδο Steepest descent μπορεί να γραφτεί ως:

$$w_{k+1} = w_k - \alpha g_k \quad (7.2)$$

Όπου α είναι η σταθερά εκπαίδευσης.

Η μέθοδος εκπαίδευσης Steepest descent παρουσιάζει ασυμπτωτική κάλυψη. Γύρω από την λύση όλα τα στοιχεία του διανύσματος της κλίσης θα είναι πολύ μικρά και θα υπάρχει πολύ μικρή αλλαγή στα βάρη.

7.2 Μέθοδος Newton

Η μέθοδος Newton υποθέτει ότι όλα τα στοιχεία της κλίσης g_1, g_2, \dots, g_N είναι συναρτήσεις των βαρών και όλα τα βάρη είναι γραμμικά ανεξάρτητα:

$$\begin{cases} g_1 = F_1(w_1, w_2, \dots, w_N) \\ g_2 = F_2(w_1, w_2, \dots, w_N) \\ \dots \\ g_N = F_N(w_1, w_2, \dots, w_N) \end{cases} \quad (7.3)$$

Όπου F_1, F_2, \dots, F_N είναι μη-γραμμικές σχέσεις μεταξύ των βαρών και των σχετιζόμενων στοιχείων των βαρών. Η (7.3) αναπτύσσεται ως σειρά Taylor και στην συνέχεια λαμβάνεται η προσέγγισή πρώτης τάξης:

$$\begin{cases} g_1 \approx g_{1,0} + \frac{\partial g_1}{\partial w_1} \Delta w_1 + \frac{\partial g_1}{\partial w_2} \Delta w_2 + \dots + \frac{\partial g_1}{\partial w_N} \Delta w_N \\ g_2 \approx g_{2,0} + \frac{\partial g_2}{\partial w_1} \Delta w_1 + \frac{\partial g_2}{\partial w_2} \Delta w_2 + \dots + \frac{\partial g_2}{\partial w_N} \Delta w_N \\ \dots \\ g_N \approx g_{N,0} + \frac{\partial g_N}{\partial w_1} \Delta w_1 + \frac{\partial g_N}{\partial w_2} \Delta w_2 + \dots + \frac{\partial g_N}{\partial w_N} \Delta w_N \end{cases} \quad (7.4)$$

Συνδυάζοντας τον ορισμό της κλίσης από την (7.1) προκύπτει ότι:

$$\frac{\partial g_i}{\partial w_j} = \frac{\partial \left(\frac{\partial E}{\partial w_j} \right)}{\partial w_j} = \frac{\partial^2 E}{\partial w_i \partial w_j} \quad (7.5)$$

Αντικαθιστώντας την (7.5) στην (7.4):

$$\begin{cases} g_1 \approx g_{1,0} + \frac{\partial^2 E}{\partial w_1^2} \Delta w_1 + \frac{\partial^2 E}{\partial w_1 \partial w_2} \Delta w_2 + \dots + \frac{\partial^2 E}{\partial w_1 \partial w_N} \Delta w_N \\ g_2 \approx g_{2,0} + \frac{\partial^2 E}{\partial w_2 \partial w_1} \Delta w_1 + \frac{\partial^2 E}{\partial w_2^2} \Delta w_2 + \dots + \frac{\partial^2 E}{\partial w_2 \partial w_N} \Delta w_N \\ \dots \\ g_N \approx g_{N,0} + \frac{\partial^2 E}{\partial w_N \partial w_1} \Delta w_1 + \frac{\partial^2 E}{\partial w_N \partial w_2} \Delta w_2 + \dots + \frac{\partial^2 E}{\partial w_N^2} \Delta w_N \end{cases} \quad (7.6)$$

Συγκριτικά με τον αλγόριθμο steepest descent πρέπει να υπολογιστούν οι δεύτερες παράγωγοι της συνάρτησης ολικού σφάλματος για κάθε στοιχείο του διανύσματος κλίσης. Για να προκύψει το ελάχιστο της συνάρτησης ολικού σφάλματος E , κάθε στοιχείο του διανύσματος κλίσης πρέπει να είναι 0. Έτσι τα αριστερά μέλη της εξίσωσης (7.6) είναι όλα 0 και:

$$\begin{cases} 0 \approx g_{1,0} + \frac{\partial^2 E}{\partial w_1^2} \Delta w_1 + \frac{\partial^2 E}{\partial w_1 \partial w_2} \Delta w_2 + \dots + \frac{\partial^2 E}{\partial w_1 \partial w_N} \Delta w_N \\ 0 \approx g_{2,0} + \frac{\partial^2 E}{\partial w_2 \partial w_1} \Delta w_1 + \frac{\partial^2 E}{\partial w_2^2} \Delta w_2 + \dots + \frac{\partial^2 E}{\partial w_2 \partial w_N} \Delta w_N \\ \dots \\ 0 \approx g_{N,0} + \frac{\partial^2 E}{\partial w_N \partial w_1} \Delta w_1 + \frac{\partial^2 E}{\partial w_N \partial w_2} \Delta w_2 + \dots + \frac{\partial^2 E}{\partial w_N^2} \Delta w_N \end{cases} \quad (7.7)$$

Συνδυάζοντας τις εξισώσεις (7.1) και (7.7) προκύπτει:

$$\left\{ \begin{array}{l} -\frac{\partial E}{\partial w_1} = -g_{1,0} \approx \frac{\partial^2 E}{\partial w_1^2} \Delta w_1 + \frac{\partial^2 E}{\partial w_1 \partial w_2} \Delta w_2 + \dots + \frac{\partial^2 E}{\partial w_1 \partial w_N} \Delta w_N \\ -\frac{\partial E}{\partial w_2} = -g_{2,0} \approx \frac{\partial^2 E}{\partial w_2 \partial w_1} \Delta w_1 + \frac{\partial^2 E}{\partial w_2^2} \Delta w_2 + \dots + \frac{\partial^2 E}{\partial w_2 \partial w_N} \Delta w_N \\ \dots \\ -\frac{\partial E}{\partial w_N} = -g_{N,0} \approx \frac{\partial^2 E}{\partial w_N \partial w_1} \Delta w_1 + \frac{\partial^2 E}{\partial w_N \partial w_2} \Delta w_2 + \dots + \frac{\partial^2 E}{\partial w_N^2} \Delta w_N \end{array} \right. \quad (7.8)$$

Υπάρχουν N εξισώσεις για N αγνώστους, έτσι ώστε να μπορούν να υπολογιστούν όλα τα Δw_i . Με τις λύσεις ο χώρος των βαρών μπορεί να ενημερώνεται επαναληπτικά. Η εξίσωση (7.8) γράφεται με μορφή πινάκων:

$$\begin{bmatrix} -g_1 \\ -g_2 \\ \dots \\ -g_N \end{bmatrix} = \begin{bmatrix} -\frac{\partial E}{\partial w_1} \\ -\frac{\partial E}{\partial w_2} \\ \dots \\ -\frac{\partial E}{\partial w_N} \end{bmatrix} = \begin{bmatrix} \frac{\partial^2 E}{\partial w_1^2} & \frac{\partial^2 E}{\partial w_1 \partial w_2} & \dots & \frac{\partial^2 E}{\partial w_1 \partial w_N} \\ \frac{\partial^2 E}{\partial w_2 \partial w_1} & \frac{\partial^2 E}{\partial w_2^2} & \dots & \frac{\partial^2 E}{\partial w_2 \partial w_N} \\ \dots & \dots & \dots & \dots \\ \frac{\partial^2 E}{\partial w_N \partial w_1} & \frac{\partial^2 E}{\partial w_N \partial w_2} & \dots & \frac{\partial^2 E}{\partial w_N^2} \end{bmatrix} \times \begin{bmatrix} \Delta w_1 \\ \Delta w_2 \\ \dots \\ \Delta w_N \end{bmatrix} \quad (7.9)$$

Όπου ο τετραγωνικός πίνακας είναι ο Εσσιανός:

$$H = \begin{bmatrix} \frac{\partial^2 E}{\partial w_1^2} & \frac{\partial^2 E}{\partial w_1 \partial w_2} & \dots & \frac{\partial^2 E}{\partial w_1 \partial w_N} \\ \frac{\partial^2 E}{\partial w_2 \partial w_1} & \frac{\partial^2 E}{\partial w_2^2} & \dots & \frac{\partial^2 E}{\partial w_2 \partial w_N} \\ \dots & \dots & \dots & \dots \\ \frac{\partial^2 E}{\partial w_N \partial w_1} & \frac{\partial^2 E}{\partial w_N \partial w_2} & \dots & \frac{\partial^2 E}{\partial w_N^2} \end{bmatrix} \quad (7.10)$$

Συνδυάζοντας τις (7.1), (7.10) και (7.9):

$$-g = H \Delta w \quad (7.11)$$

και:

$$\Delta w = -H^{-1} g \quad (7.12)$$

Έτσι ο ενημερωμένος κανόνας για την μέθοδο είναι:

$$w_{k+1} = w_k - H_k^{-1} g_k \quad (7.13)$$

Σαν οι δεύτερες παράγωγοι της συνάρτησης ολικού σφάλματος. Ο Εσσιανός πίνακας δίνει μια σωστή εκτίμηση για την αλλαγή του διανύσματος κλίσης.

7.3 Αλγόριθμος Gauss-Newton

Εάν εφαρμοσθεί η μέθοδος Newton για τον υπολογισμό των βαρών, για να υπολογιστεί ο Εσσιανός πίνακας H , θα πρέπει να υπολογιστούν οι δεύτερες παράγωγοι της συνάρτησης ολικού σφάλματος κάτι που θα είναι πολύ πολύπλοκο. Για την απλοποίηση της υπολογιστικής διαδικασίας εισάγεται ο Ιακωβιανός πίνακας J ως:

$$J = \begin{bmatrix} \frac{\partial e_{1,1}}{\partial w_1} & \frac{\partial e_{1,1}}{\partial w_2} & \dots & \frac{\partial e_{1,1}}{\partial w_N} \\ \frac{\partial e_{1,2}}{\partial w_1} & \frac{\partial e_{1,2}}{\partial w_2} & \dots & \frac{\partial e_{1,2}}{\partial w_N} \\ \dots & \dots & \dots & \dots \\ \frac{\partial e_{1,M}}{\partial w_1} & \frac{\partial e_{1,M}}{\partial w_2} & \dots & \frac{\partial e_{1,M}}{\partial w_N} \\ \dots & \dots & \dots & \dots \\ \frac{\partial e_{p,1}}{\partial w_1} & \frac{\partial e_{p,1}}{\partial w_2} & \dots & \frac{\partial e_{p,1}}{\partial w_N} \\ \frac{\partial e_{p,2}}{\partial w_1} & \frac{\partial e_{p,2}}{\partial w_2} & \dots & \frac{\partial e_{p,2}}{\partial w_N} \\ \dots & \dots & \dots & \dots \\ \frac{\partial e_{p,M}}{\partial w_1} & \frac{\partial e_{p,M}}{\partial w_2} & \dots & \frac{\partial e_{p,M}}{\partial w_N} \end{bmatrix} \quad (7.14)$$

Ολοκληρώνοντας την εξίσωση 12.1 και την εξίσωση του αθροιστικού τετραγωνικού σφάλματος:

$$E(\mathbf{x}, \mathbf{w}) = \frac{1}{2} \sum_{p=1}^P \sum_{m=1}^M e_{p,m}^2 \quad (7.15)$$

το διάνυσμα κλίσης υπολογίζεται ως:

$$g_i = \frac{\partial E}{\partial w_i} = \frac{\partial \left(\frac{1}{2} \sum_{p=1}^P \sum_{m=1}^M e_{p,m}^2 \right)}{\partial w_i} = \sum_{p=1}^P \sum_{m=1}^M \left(\frac{\partial e_{p,m}}{\partial w_i} e_{p,m} \right) \quad (7.16)$$

Συνδυάζοντας την (7.14) και την (7.16) η σχέση μεταξύ Ιακωβιανού πίνακα J και του διανύσματος κλίσης \mathbf{g} θα είναι:

$$\mathbf{g} = J\mathbf{e} \quad (7.17)$$

Όπου το διάνυσμα σφάλματος \mathbf{e} έχει την μορφή

$$\mathbf{e} = \begin{bmatrix} e_{1,1} \\ e_{1,2} \\ \dots \\ e_{1,M} \\ \dots \\ e_{p,1} \\ e_{p,2} \\ \dots \\ e_{p,M} \end{bmatrix} \quad (7.18)$$

Εισάγοντας την εξίσωση (7.1) στην (7.10), το στοιχείο στην ι-οστή σειρά και j-οστή στήλη του Εσσιανού πίνακα μπορεί να υπολογιστεί ως:

$$h_{i,j} = \frac{\partial^2 E}{\partial w_i \partial w_j} = \frac{\partial^2 \left(\frac{1}{2} \sum_{p=1}^P \sum_{m=1}^M e_{p,m}^2 \right)}{\partial w_i \partial w_j} = \sum_{p=1}^P \sum_{m=1}^M \frac{\partial e_{p,m}}{\partial w_i} \frac{\partial e_{p,m}}{\partial w_j} + S_{i,j} \quad (7.19)$$

Όπου το $S_{i,j}$ ισούται με:

$$S_{i,j} = \sum_{p=1}^P \sum_{m=1}^M \frac{\partial^2 e_{p,m}}{\partial w_i \partial w_j} e_{p,m} \quad (7.20)$$

Καθώς η βασική υπόθεση της μεθόδου Newton είναι ότι το $S_{i,j}$ είναι κοντά στο 0, η σχέση μεταξύ του Εσσιανού και του Ιακωβιανού πίνακα μπορεί να γραφτεί ως:

$$H \approx J^T J \quad (7.21)$$

Συνδυάζοντας τις σχέσεις (7.13), (7.17) και (7.21) ο κανόνας για την Gauss-Newton γράφεται ως:

$$w_{k+1} = w_k - (J_k^T J_k)^{-1} J_k e_k \quad (7.22)$$

8 ΒΙΒΛΙΟΓΡΑΦΙΑ

Βαφείδης, Α. 2011. *Εφαρμοσμένη Γεωφυσική Ι - Σεισμικές μέθοδοι, Εργαστήριο Εφαρμοσμένης Γεωφυσικής.* Χανιά : Σχολή Μηχανικών Ορυκτών Πόρων, 2011.

Τριχάκης, Ιωάννης Χ. 2011. *Ανάπτυξη Μοντέλων Υπογείων Υδάτων με Χρήση Εργαλείων Υπολογιστικής Νοημοσύνης και Ποσοτικοποίησης της Αβεβαιότητας.* Χανιά : Πολυτεχνείο Κρήτης, 2011.

Παπαδόπουλος, Δ. Τ. 2010. *Εισαγωγή στη Γεωφυσική.* 2010.

Beroza, Weiqiang Zhu and Gregory C. 2018. *PhaseNet: A Deep-Neural-Network Based Seismic Arrival Time Picking Method.* s.l. : SEG, 2018.

Chun, Tsai Kuo και Hu, Wenyi. *First break picking with deep semi-supervised learning neural network.* s.l. : SEG.

Dix. 2015. Dix Durban Kerry formula. [Ηλεκτρονικό] 2015.
http://wiki.seg.org/wiki/Dictionary:Dix_formula.

Hao Yu, Bogdan M. Wilamowski. 2011. *Training Industrial Electronics Handbook, vol. 5 Intelligent Systems, 2nd Edition.* s.l. : CRC Press, 2011.

Inderwiesen, Tien-when Lo and Philip L. *Fundamentals Of Seismic Tomography.* s.l. : Society Of Exploration Geophysics.

Lo, Tien when and Philip Inderwiesen. *Fundamentals Of Seismic Tomography.* s.l. : Society Of Exploration Geophysicists.

Morgan, N.A. 1970. Wavelet maps: A new analysis tool for reflection seismograms. 1970, σσ. 447-460.

Offshore Operations Subgroup. 2011. *Environmental Management of Geophysical Exploration.* s.l. : Offshore Operations Subgroup of the Operations & Environment Task Group, 2011.

Russel Stuart, Peter Novig. 2010. *Artificial Intelligence, a modern approach .* New Jersey : Pearson Education , 2010.

Shen, Yang, και συν. *Seismic trace editing by applying machine learning.* Beijing : SEG.

Sheriff, R. E. και Geldart, L. P. 1995. *Exploration seismology.* s.l. : Cambridge University Press, 1995.