



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΧΗΜΙΚΩΝ ΜΗΧΑΝΙΚΩΝ

**ΤΟΜΕΑΣ ΙΙ: ΑΝΑΛΥΣΗΣ, ΣΧΕΔΙΑΣΜΟΥ ΚΑΙ ΑΝΑΠΤΥΞΗΣ ΔΙΕΡΓΑΣΙΩΝ
ΚΑΙ ΣΥΣΤΗΜΑΤΩΝ**

ΜΟΝΑΔΑ ΑΥΤΟΜΑΤΗΣ ΡΥΘΜΙΣΗΣ ΚΑΙ ΠΛΗΡΟΦΟΡΙΚΗΣ

**ΑΝΑΠΤΥΞΗ ΝΕΥΡΩΝΙΚΟΥ ΡΥΘΜΙΣΤΗ ΓΙΑ ΤΟΝ ΑΥΤΟΜΑΤΟ
ΕΛΕΓΧΟ ΑΣΤΑΘΩΝ ΜΗ ΓΡΑΜΜΙΚΩΝ ΔΥΝΑΜΙΚΩΝ
ΣΥΣΤΗΜΑΤΩΝ**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Αλεξάνδρα Ι. Κυρίου

Επιβλέπων : Χ. Σαρίμβεης

Αναπληρωτής Καθηγητής

Αθήνα, Ιούλιος 2012

ΠΕΡΙΛΗΨΗ

Στόχος αυτής της διπλωματικής εργασίας ήταν η ανάπτυξη ενός νευρωνικού ρυθμιστή, δηλαδή ενός ρυθμιστή που υλοποιείται με χρήση νευρωνικών δικτύων με σκοπό να επιτυγχάνεται η αυτόματη ρύθμιση μη γραμμικών συστημάτων σε ασταθείς περιοχές λειτουργίας.

Τα νευρωνικά δίκτυα είναι ένας κλάδος της τεχνητής νοημοσύνης που εξελίσσεται διαρκώς τα τελευταία χρόνια. Εξαιτίας της μη γραμμική δομής τους χρησιμοποιούνται εκτενώς για τη μοντελοποίηση μη γραμμικών συστημάτων. Στην συγκεκριμένη εργασία επιλέχθηκε η αρχιτεκτονική νευρωνικών δικτύων ακτινικής συνάρτησης βάσης (Radial Basis Function, RBF), εξαιτίας της απλής δομής και των γρήγορων αλγορίθμων μάθησης που χρησιμοποιούν. Συγκεκριμένα, για την εκπαίδευση των δικτύων χρησιμοποιήθηκε ο αλγόριθμος των ασαφών μέσων, ο οποίος διακρίνεται για την ταχύτητά του και την αξιοπιστία του, έναντι των κλασσικών μεθόδων εκπαίδευσης νευρωνικών δικτύων RBF.

Η προτεινόμενη μεθοδολογία εφαρμόστηκε στην αυτόματη ρύθμιση ενός αντιδραστήρα Continuous Stirred Tank reactor (CSTR) με πολλαπλά σημεία ισορροπίας. Χρησιμοποιώντας την αρχιτεκτονική δομή RBF, ο προτεινόμενος νευρωνικός ρυθμιστής αναπτύχθηκε, με βάση δεδομένα εισόδου- εξόδου που συλλέχθηκαν κατά την προσομοίωση της δυναμικής λειτουργίας του αντιδραστήρα. Ως είσοδοι στο νευρωνικό δίκτυο χρησιμοποιήθηκαν η επιθυμητή τιμή και το τρέχον διάλυμα κατάστασης του συστήματος, ενώ σαν έξοδος προκύπτει η τρέχουσα τιμή της μεταβλητής εκχειρισμού.

Τα αποτελέσματα στην αυτόματη ρύθμιση του αντιδραστήρα CSTR δείχνουν ότι η προτεινόμενη μέθοδος είναι ταχύτερη και αξιόπιστη ενώ πλεονεκτεί έναντι κλασσικών μεθόδων αυτόματης ρύθμισης, όπως οι ρυθμιστές τύπου PID.

Λέξεις-κλειδιά

Νευρωνικά δίκτυα, Συναρτήσεις Ακτινικής Βάσης, Αντιδραστήρας CSTR, Νευρωνικός Ρυθμιστής.

ABSTRACT

The scope of this thesis was the development of a neurocontroller i.e. a controller based on neural networks for the automatic control of nonlinear systems in unstable regions of operation.

Neural Networks are a class of Artificial Intelligence that have experienced continuous evolution during the last decades. Due to their nonlinear structure, they are used extensively in modelling nonlinear systems. In this thesis, the Radial Basis Function (RBF) neural network architecture was chosen, due to their simple structure and the fast learning algorithms. In particular, the training of the RBF networks is performed using the fuzzy means algorithm which guarantees increased accuracy and lower computational times over classical training procedures of RBF Neural Networks.

The proposed methodology was applied for controlling a CSTR reactor with multiple steady states, which was simulated in MATLAB .Using the RBF architecture, the proposed neurocontroller was developed based on input-output data that were collected during simulation of the reactor dynamic behavior. The neural network receives as inputs the set point and the current state variables and produces as output the current value for the manipulated variable.

The results of applying the neurocontroller for controlling the CSTR reactor illustrate that the proposed methodology is fast and accurate, and has superior performance compare to classical control schemes, such as PID controllers.

Keywords

Neural Networks, Radial Basis Function, CSTR Reactor, Neural Control

Ευχαριστίες

Στο σημείο αυτό θα ήθελα να ευχαριστήσω θερμά για τη συμβολή τους τον επιβλέποντα Αναπληρωτή Καθηγητή Χαράλαμπο Σαρίμβει καθώς και τον Επίκουρο Καθηγητή Αλέξανδρο Αλεξανδρίδη από το Τμήμα Ηλεκτρονικής του ΤΕΙ Αθήνας.

Επίσης θα ήθελα να ευχαριστήσω την οικογένεια μου που με στήριξε αμέριστα στη διάρκεια των σπουδών μου.

ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

ΠΕΡΙΛΗΨΗ	2
ABSTRACT	3
ΚΕΦΑΛΑΙΟ Ι	7
ΕΙΣΑΓΩΓΗ	7
1.1 Ευφύης Ρύθμιση.....	7
1.2 Δομή Εργασίας.....	8
ΚΕΦΑΛΑΙΟ ΙΙ.....	9
ΤΕΧΝΗΤΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ	9
2.1 Ιστορική Αναδρομή.....	9
2.2 Βιολογικά Νευρωνικά Δίκτυα.....	11
2.3 Βασικό Μοντέλο Τεχνητού Νευρώνα.....	13
2.4 Αρχιτεκτονικές Νευρωνικών Δικτύων.....	16
2.4.1 Μονοεπίπεδα Νευρωνικά Δίκτυα (Single Layered Perceptron)	16
2.4.2 Πολυεπίπεδα Νευρωνικά Δίκτυα (Multi Layered Perceptron, MLP).....	17
2.4.3 Νευρωνικά Δίκτυα Ακτινικών Συναρτήσεων Βάσης.....	18
2.5 Εκπαίδευση Τεχνητών Νευρωνικών Δικτύων	20
2.5.1 Εκπαίδευση με Επίβλεψη.....	20
2.5.2 Εκπαίδευση με Ενίσχυση.....	21
2.5.3 Εκπαίδευση χωρίς Επίβλεψη.....	22
2.6 Γενίκευση.....	23
2.7 Πλεονεκτήματα και Μειονεκτήματα Τεχνητών Νευρωνικών Δικτύων.....	23
2.8 Εφαρμογές των Τεχνητών Νευρωνικών Δικτύων.....	25
ΚΕΦΑΛΑΙΟ ΙΙΙ.....	26
ΔΙΚΤΥΑ ΑΚΤΙΝΙΚΗΣ ΣΥΝΑΡΤΗΣΗΣ ΒΑΣΗΣ RBF	26
3.1 Ακτινικές Συναρτήσεις Βάσης στην Παρεμβολή.....	26
3.2 Νευρωνικά Δίκτυα Ακτινικών Συναρτήσεων Βάσης RBF	28
3.2.1 Γενική Επισκόπηση των Δικτύων RBF.....	28
3.2.2 Βελτίωση του Αλγορίθμου Εκπαίδευσης.....	31
ΚΕΦΑΛΑΙΟ ΙV	39
ΝΕΥΡΩΝΙΚΟΙ ΡΥΘΜΙΣΤΕΣ.....	39
4.1 Εισαγωγή.....	39

4.2 Ρύθμιση Προβλεπτικού Μοντέλου (Model Predictive Control)	41
4.3 Ρύθμιση με Αντίστροφο Μοντέλο	42
4.3.1 Άμεση Αντίστροφη Ρύθμιση (Direct Inverse Control)	44
4.3.2. Ρύθμιση Εσωτερικού Μοντέλου (Internal Model Control, IMC)	44
ΚΕΦΑΛΑΙΟ V	45
ΑΝΤΙΔΡΑΣΤΗΡΑΣ CSTR	45
5.1 Εισαγωγή.....	45
5.2. ΟΙ Εξισώσεις Μοντελοποίησης	46
5.2.1 Παράμετροι και Μεταβλητές	46
5.2.2 Γενικό Ισοζύγιο Μάζας.....	47
5.2.3 Ισοζύγιο Μάζας του Συστατικού A.....	47
5.2.4 Ισοζύγιο Ενέργειας.....	48
5.2.5 Δυναμικές Εξισώσεις στη Μορφή του Χώρου Κατάστασης	48
5.3. Σημεία Ισορροπίας.....	48
5.4. Δυναμική Συμπεριφορά	50
5.5. Γραμμικοποίηση Δυναμικών Εξισώσεων	54
5.5.1 Ανάλυση Ευστάθειας	54
5.6. Κατανόηση των Πολλαπλών Σημείων Ισορροπίας.....	56
5.6.1 Καμπύλες Παραγωγής και Απομάκρυνσης Θερμότητας	57
5.6.2 Επίδραση των Παραμέτρων Σχεδιασμού	58
5.6.3 Πολλαπλά Σημεία Ισορροπίας	59
5.6.4 Εξέταση της Ευστάθειας (Ανάλυση Μόνιμης Κατάστασης).....	60
ΚΕΦΑΛΑΙΟ VI	62
ΣΧΕΔΙΑΣΜΟΣ ΚΑΙ ΕΦΑΡΜΟΓΗ ΝΕΥΡΩΝΙΚΟΥ ΡΥΘΜΙΣΤΗ	62
6.1 Προσομοίωση του Πειράματος.....	62
6.2 Πρόσθιο Μοντέλο	65
6.3 Αντίστροφο Μοντέλο-Σχεδιασμός Νευρωνικού Ρυθμιστή.....	66
6.4 Εφαρμογή και Αξιολόγηση Νευρωνικού Ρυθμιστή.....	67
ΚΕΦΑΛΑΙΟ VII.....	75
ΣΥΜΠΕΡΑΣΜΑΤΑ	75
ΑΝΑΦΟΡΕΣ	76
ΠΑΡΑΡΤΗΜΑ	78

ΚΕΦΑΛΑΙΟ Ι

ΕΙΣΑΓΩΓΗ

1.1 Ευφυής Ρύθμιση

Οι μεθοδολογίες σχεδιασμού συστημάτων αυτόματης ρύθμισης είναι ιδιαίτερα ανεπτυγμένες και αποτελεσματικές για γραμμικά συστήματα τόσο στο χώρο της συχνότητας όσο και στο χώρο του χρόνου (χώρος κατάστασης). Οι δυο προσεγγίσεις αποτελούν την αποκαλούμενη κλασική θεωρία αυτόματης ρύθμισης. Όμως πολλά συστήματα με ερευνητικό αλλά και βιομηχανικό ενδιαφέρον είναι συστήματα μη γραμμικά ή και χρονικά μεταβαλλόμενα. Ιδιαίτερα στη βιομηχανία, πολλές διεργασίες παρουσιάζουν εγγενείς μη γραμμικότητες και συχνά έχουν μεγάλο εύρος λειτουργίας, που καθιστά ανέφικτη τη γραμμικοποίηση γύρω από ένα συγκεκριμένο σημείο λειτουργίας. Επίσης, αρκετές διεργασίες εμφανίζουν ασυνέχεια, όπως νεκρές ζώνες, κορεσμό και υστέρηση, κάνοντας τις γραμμικές λύσεις ελέγχου αναποτελεσματικές.

Η αυξανόμενη ζήτηση για συστήματα μεγαλύτερης ακρίβειας και ενεργειακής απόδοσης καθώς και τα πεπερασμένα όρια αποτελεσματικότητας που είναι το αποτέλεσμα της εφαρμογής γραμμικών τεχνικών σχεδίασης οδήγησαν την ερευνητική κοινότητα στην ανάπτυξη μη γραμμικών τεχνικών αυτόματης ρύθμισης, κάτι που επιτεύχθηκε τα τελευταία χρόνια με την ανάπτυξη της θεωρίας του μη γραμμικού ελέγχου [13]. Οι τεχνικές όμως αυτές δεν έχουν επιτύχει την προσδοκώμενη απήχηση από τη βιομηχανία, λόγω της απόστασης ανάμεσα στη θεωρία και την εφαρμογή και του υψηλού βαθμού εξειδίκευσης, δηλαδή το περιορισμένο εύρος μη γραμμικών συστημάτων στα οποία βρίσκει εφαρμογή κάθε επιμέρους μεθοδολογία.

Αντίθετα με το συμβατικό μη-γραμμικό έλεγχο, οι λεγόμενες ευφυείς τεχνικές αυτόματης ρύθμισης παρουσιάζουν μεγαλύτερο βαθμό διείσδυσης στη βιομηχανία [17]. Ως «ευφυή συστήματα ρύθμισης» (Intelligent Control) χαρακτηρίζονται εκείνα που βασίζονται σε μεθόδους τεχνητής νοημοσύνης (Artificial Intelligence), δηλαδή όταν προσπαθούν να μιμηθούν σημαντικά χαρακτηριστικά της ανθρώπινης νοημοσύνης, όπως η προσαρμογή και η μάθηση, ο σχεδιασμός υπό μεγάλη αβεβαιότητα και η επεξεργασία μεγάλου όγκου δεδομένων. Η τεχνητή νοημοσύνη περιλαμβάνει τα έμπειρα συστήματα, την ασαφή λογική, τα νευρωνικά δίκτυα, τους εξελικτικούς αλγορίθμους και οποιοδήποτε άλλο μέσο με το οποίο τα δεδομένα αναλύονται, οργανώνονται και μετατρέπονται σε γνώση (RayChaudhuri et al., 1996).

Τα νευρωνικά δίκτυα, με την ικανότητά τους να μαθαίνουν και την έμφυτη μαζική παράλληλη επεξεργασία τους εισαγάγουν πολυάριθμες ευκαιρίες για την ανάπτυξη ανώτερων δομών

ρύθμισης.

1.2 Δομή Εργασίας

Η παρούσα εργασία στοχεύει στην ανάπτυξη μιας νέας μεθόδου άμεσης αντίστροφης ρύθμισης με το σχεδιασμό ενός πρωτότυπου νευρωνικού ρυθμιστή που βασίζεται στην αρχιτεκτονική ακτινικής συνάρτησης βάσης με ιδιαίτερη έμφαση στην αποτελεσματικότητα σε περιοχές μη ευσταθούς λειτουργίας. Η διπλωματική εργασία αποτελείται από 7 κεφάλαια.

Το δεύτερο κεφάλαιο πραγματεύεται το πεδίο των Τεχνητών Νευρωνικών Δικτύων. Αρχικά, η ιστορική αναδρομή δείχνει την εξελικτική πορεία των νευρωνικών δικτύων μέχρι σήμερα. Στη συνέχεια περιγράφονται η δομή και οι αρχές λειτουργίας τους καθώς και οι μέθοδοι με τις οποίες εκπαιδεύονται. Τέλος, αναφέρονται τα πλεονεκτήματα και τα μειονεκτήματα των νευρωνικών δικτύων και μερικές εφαρμογές τους.

Το τρίτο κεφάλαιο είναι αφιερωμένο στα Νευρωνικά δίκτυα Ακτινικής Συνάρτησης Βάσης. Αρχικά, γίνεται περιγραφή των Συναρτήσεων Ακτινικής Βάσης. Στη συνέχεια, ακολουθεί μια γενική επισκόπηση στα Νευρωνικά δίκτυα Ακτινικής Συνάρτησης Βάσης. Στο τέλος του κεφαλαίου γίνεται εκτενής αναφορά στον αλγόριθμο των ασαφών μέσων.

Το τέταρτο κεφάλαιο επικεντρώνεται στη χρήση των Νευρωνικών δικτύων στην αυτόματη ρύθμιση. Αναφέρονται οι τρόποι με τους οποίους τα νευρωνικά δίκτυα μπορούν να εφαρμοστούν στην αυτόματη ρύθμιση και αναλύονται οι κύριες μεθοδολογίες που είναι η μεθοδολογία προβλεπτικού μοντέλου και η μεθοδολογία αντίστροφης ρύθμισης.

Στο πέμπτο κεφάλαιο γίνεται ανάλυση του αντιδραστήρα συνεχούς έργου που θα χρησιμοποιηθεί ως παράδειγμα εφαρμογής της προτεινόμενης μεθοδολογίας. Συγκεκριμένα, αναπτύσσεται το δυναμικό μοντέλο της διεργασίας, υπολογίζονται τα πολλαπλά σημεία ισορροπίας και γίνεται ανάλυση ευστάθειας του συστήματος. Τέλος, γίνεται αναφορά στο φαινόμενο της υστέρησης και καθορίζεται το πώς προκύπτουν τα πολλαπλά σημεία ισορροπίας.

Το έκτο κεφάλαιο περιλαμβάνει το σχεδιασμό και την εφαρμογή του νευρωνικού ρυθμιστή. Αρχικά, παρουσιάζεται το πείραμα που σχεδιάστηκε με σκοπό τη δημιουργία βάσης δεδομένων για την εκπαίδευση του νευρωνικού δικτύου. Ακολουθεί ο σχεδιασμός και η εφαρμογή του νευρωνικού ρυθμιστή και η αξιολόγηση της απόδοσής του. Τέλος γίνεται σύγκριση του νευρωνικού ρυθμιστή με έναν συμβατικό PID ρυθμιστή.

Στο έβδομο κεφάλαιο περιλαμβάνονται τα συμπεράσματα που εξάγονται από την εργασία και αναφέρονται κάποιες προτάσεις για περαιτέρω έρευνα.

Τέλος, παρατίθεται η βιβλιογραφία η οποία χρησιμοποιήθηκε στην εργασία.

ΚΕΦΑΛΑΙΟ II

ΤΕΧΝΗΤΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ

Το Τεχνητό Νευρωνικό Δίκτυο (ΤΝΔ, Artificial Neural Network ANN), το οποίο συχνά καλείται Προσομοιωμένο Νευρωνικό Δίκτυο ή απλά Νευρωνικό δίκτυο, είναι ένα μαθηματικό ή υπολογιστικό μοντέλο βασισμένο στα βιολογικά νευρωνικά δίκτυα. Τα ΤΝΔ προέκυψαν από τη διαπίστωση ότι ο τρόπος λειτουργίας του ανθρώπινου εγκεφάλου είναι εντελώς διαφορετικός από τον τρόπο λειτουργίας ενός τυπικού ηλεκτρονικού υπολογιστή. Ένα ΤΝΔ αποτελείται από διασυνδεδεμένες ομάδες τεχνητών νευρώνων και επεξεργάζεται τις πληροφορίες χρησιμοποιώντας την συνδετική προσέγγιση (connectionist approach) στους υπολογισμούς. Στις περισσότερες περιπτώσεις ένα ΤΝΔ αποτελεί ένα προσαρμόσιμο /δυναμικό σύστημα, το οποίο αλλάζει τη δομή του είτε λόγω εξωτερικών είτε λόγω εσωτερικών πληροφοριών που προκύπτουν κατά τη φάση της εκπαίδευσης.

2.1 Ιστορική Αναδρομή

Η μελέτη του ανθρώπινου εγκεφάλου έχει ξεκινήσει εδώ και χιλιάδες χρόνια. Η σημαντικότερη πρόοδος όμως έχει επιτελεστεί τον τελευταίο αιώνα όταν και με την βοήθεια των ηλεκτρονικών υπολογιστών ο άνθρωπος ξεκίνησε την προσπάθεια προσομοίωσης της ανθρώπινης σκέψης.

Η σύγχρονη εποχή της μελέτης των τεχνητών νευρωνικών δικτύων [30] ξεκινά από την εργασία του νευροφυσιολόγου Warren McCulloch και του μαθηματικού Walter Pitts το 1943. Ο McCulloch μελετούσε για περισσότερα από 20 χρόνια το «γεγονός» που αποτελούσε το εφελτήριο της αίσθησης, της σκέψης κλπ. Μόνο όμως όταν ο McCulloch και ο Pitts συνεργάστηκαν, κατάφεραν να εκδώσουν μια εργασία για την πιθανή λειτουργία των νευρώνων και να κατασκευάσουν ένα πρώιμο τεχνητό νευρωνικό δίκτυο βασισμένο σε απλά ηλεκτρικά κυκλώματα. Σε αυτούς πιστώνεται η θεωρία McCulloch-Pitts για Τυπικά Νευρωνικά Δίκτυα.

Η αμέσως επόμενη μεγάλη εξέλιξη στο πεδίο των Τεχνητών Νευρωνικών Δικτύων ήρθε το 1949 με το βιβλίο "The Organization of Behavior" από τον Donald Hebb. Το βιβλίο αυτό ενίσχυε την θεωρία των McCulloch-Pitts για τους νευρώνες και την λειτουργία τους. Μια σημαντική πρόταση του βιβλίου ήταν η ιδιότητα των νευρώνων να ενδυναμώνονται κάθε φορά που χρησιμοποιούνται. Αυτό όπως θα δούμε έχει ιδιαίτερη εφαρμογή, ειδικά στην εκπαίδευση του δικτύου.

Καθώς οι συμβατικοί υπολογιστές άρχισαν να εξελίσσονται κατά την δεκαετία του 50, η έρευνα για τα νευρωνικά δίκτυα άρχισε να φθίνει. Παρόλα αυτά κάποιοι μεμονωμένοι επιστήμονες συνέχισαν την έρευνα. Ένας από αυτούς ήταν ο Marvin Minsky, που το 1954 τελείωσε την διδακτορική του διατριβή με τίτλο "Theory of Neural-Analog Reinforcement Systems and its Application to the Brain-Model Problem", στην οποία εξέφραζε και τους προβληματισμούς του για την έρευνα στο πεδίο των Τεχνητών Νευρωνικών Δικτύων. Επίσης ήταν και ένας από τους πρώτους που καταπατίστηκαν με την τεχνητή νοημοσύνη (artificial intelligence).

Το 1958 Frank Rosenblatt, ένας νευροβιολόγος, άρχισε να δουλεύει πάνω στο Perceptron. Το Perceptron το πρώτο «πρακτικό» Τεχνητό Νευρωνικό Δίκτυο, ήταν φτιαγμένο από σχετικά πρωτόγονα και παλιά για εκείνη την εποχή υλικά. Το Perceptron βασίστηκε στην έρευνα πάνω στο μάτι της μύγας και το πώς επεξεργάζεται αυτό τα δεδομένα που λαμβάνει. Το συγκεκριμένο είδος Τεχνητού Νευρωνικού Δικτύου αποδείχτηκε, όπως αναφέρεται στο βιβλίο "Perceptrons" των Marvin Minsky και Seymour Papert που εκδόθηκε το 1969, πως είχε περιορισμένες δυνατότητες.

Τα Τεχνητά Νευρωνικά Δίκτυα που εφαρμόστηκαν πρώτα σε πραγματικά προβλήματα ήταν τα μοντέλα ADALINE (ADaptive LINEar Elements) και MADELINE (Multiple ADaptive LINEar Elements). Αυτά δημιουργήθηκαν το 1959-1960 από τους Bernard Widrow και Marcian Hoff. Το μοντέλο ADALINE χρησιμοποιείται για την αφαίρεση του θορύβου από τις τηλεφωνικές γραμμές.

Την περίοδο μεταξύ 1969 και 1981 η προσοχή της κοινής γνώμης στράφηκε έντονα στα νευρωνικά δίκτυα. Οι δυνατότητες τους υπερεκτιμήθηκαν σε μεγάλο βαθμό και δημιουργήθηκε η εντύπωση πως μπορούσαν να κάνουν τα πάντα με αποτέλεσμα την απογοήτευση όταν οι προσδοκίες αυτές δεν είχαν σχέση με τη πραγματικότητα, και την μείωση της χρηματοδότησης για έρευνα στην επιστημονική αυτή περιοχή.

Το 1982 ο John Hopfield παρουσίασε στην επιστημονική κοινότητα το έργο του πάνω στην τεχνητή νοημοσύνη και πρότεινε μια αλλαγή της προσέγγισης της. Συγκεκριμένα ισχυρίστηκε πως η αποστολή της επιστημονικής κοινότητας δεν θα έπρεπε να είναι τόσο η προσομοίωση του ανθρωπίνου εγκεφάλου αυτού καθεαυτού, όσο η χρήση της γενικότερης λειτουργίας του, με απώτερο σκοπό την δημιουργία μηχανών που θα έχουν την δυνατότητα να λύσουν δυναμικά προβλήματα. Έδειξε επίσης πως τέτοια δίκτυα ήταν δυνατόν να υπάρξουν και περιέγραψε το πώς θα λειτουργούσαν. Ο χαρακτήρας του και οι γνώσεις του στο πεδίο της Μαθηματικής Ανάλυσης έπεισαν την επιστημονική κοινότητα πως το ενδιαφέρον και η έρευνα για την τεχνητή νοημοσύνη και τα τεχνητά νευρωνικά δίκτυα έπρεπε να ανανεωθεί. Οι ιδέες του «γέννησαν» ένα νέο τύπο τεχνητού νευρωνικού δικτύου που πλέον ονομάζεται μοντέλο Hopfield.

Σχεδόν ταυτόχρονα με την αλλαγή πλεύσης που πρότεινε ο Hopfield, σε ένα συνέδριο τεχνητής νοημοσύνης στην Ιαπωνία ανακοινώθηκε από την τοπική επιστημονική κοινότητα η επανεκκίνηση των ερευνών γύρω από τα τεχνητά νευρωνικά δίκτυα. Μπροστά στο ενδεχόμενο να μείνουν πίσω στην έρευνα οι Η.Π.Α. επανέφεραν επίσης την χρηματοδότηση στον συγκεκριμένο επιστημονικό κλάδο.

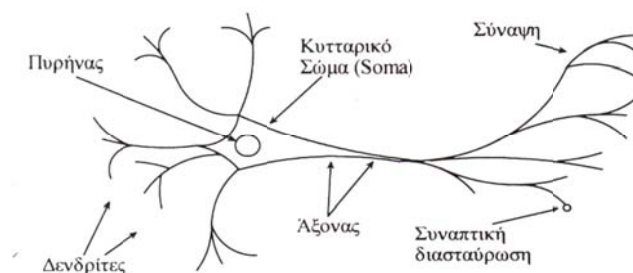
Το 1986 έγινε το πρώτο συνέδριο νευρωνικών δικτύων με την συμμετοχή 1800 συνέδρων. Επίσης οι Rumelhart, Hinton και Williams ανακοίνωσαν μετά από την έρευνα τους γύρω από τον αλγόριθμο αντίστροφης διάδοσης (back propagation algorithm), πως ο αλγόριθμος αυτός ήταν ο αποτελεσματικότερος για την εκπαίδευση δικτύων perceptron πολλών επιπέδων.

Από τις αρχές τις δεκαετίας του 1990 πολλές τεχνολογικές εξελίξεις έχουν επιτελεστεί στο πεδίο των τεχνητών νευρωνικών δικτύων ανά τον κόσμο. Οι εταιρίες του κλάδου της τεχνολογίας αναζητούν τρόπους εφαρμογής των Τεχνητών Νευρωνικών Δικτύων στην ηλεκτρονική. Συγκεκριμένα επενδύουν στην δημιουργία τριών ειδών νευρωνικών τσιπ: αναλογικών, ψηφιακών και οπτικών.

2.2 Βιολογικά Νευρωνικά Δίκτυα

Ο νευρώνας (*neuron*) είναι η θεμελιώδης κυτταρική δομική μονάδα του ανθρώπινου εγκεφάλου. Πρόκειται για ένα απλό στοιχείο επεξεργασίας, το οποίο δέχεται και συνδυάζει σήματα από άλλους νευρώνες μέσα από μονοπάτια εισόδων που καλούνται *δενδρίτες*, όπως φαίνεται στο Σχήμα 2.1.

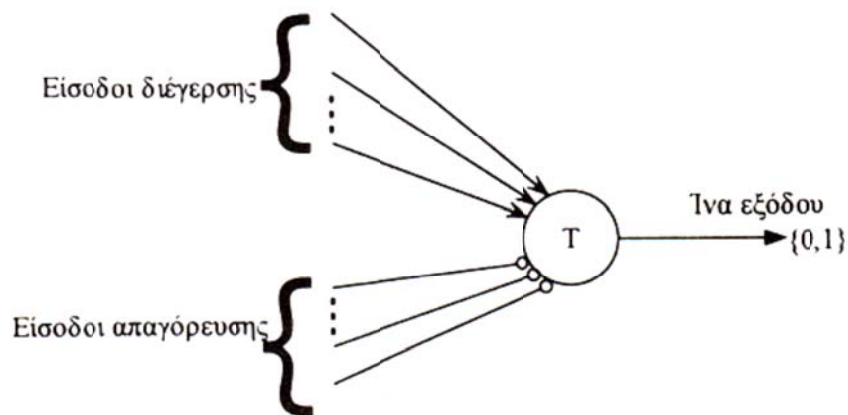
Αν το συνδυασμένο σήμα εισόδου είναι αρκετά ισχυρό προκαλεί ενεργοποίηση του νευρώνα, ο οποίος παράγει με τη σειρά του ένα σήμα εξόδου. Το σήμα αυτό διαδίδεται μέσω του άξονα, ο οποίος συνδέεται με τους δενδρίτες αρκετών άλλων νευρώνων. Κάθε σήμα που έρχεται σε ένα νευρώνα κατά μήκος ενός δενδρίτη περνάει μέσα από μια *σύναψη* ή αλλιώς συναπτικό σύνδεσμο. Αυτός ο σύνδεσμος είναι ένα πολύ μικρό διάκενο που είναι γεμάτο με νευροδιαβιβαστικό υγρό, το οποίο είτε επιταχύνει είτε επιβραδύνει τη ροή των ηλεκτρικών φορτίων. Η βασική δράση των νευρώνων είναι κατ' ουσία χημική και το νευροδιαβιβαστικό υγρό παράγει ηλεκτρικά σήματα που πηγαίνουν στον πυρήνα (στο σώμα) του νευρώνα. Η ρύθμιση της αντίστασης ή της αγωγιμότητας του κενού της σύναψης είναι ιδιαίτερης σημασίας γιατί επηρεάζει τις ικανότητες *μάθησης* και *απομνημόνευσης*. Καθώς οι συναπτικές δυνάμεις του νευρώνα προσαρμόζονται, ο εγκέφαλος μαθαίνει και αποθηκεύει πληροφορίες.



Σχήμα 2.1: Δομή τυπικού βιολογικού νευρώνα

Η επεξεργασία της πληροφορίας (σημάτων) στους εγκεφαλικούς νευρώνες και τους συναπτικούς δεσμούς γίνεται με τη βοήθεια ηλεκτροχημικών διεργασιών που είναι κατά πολύ βραδύτερες από την ηλεκτρονική επεξεργασία στα σύγχρονα ολοκληρωμένα κυκλώματα. Παρόλα αυτά ο ανθρώπινος εγκέφαλος είναι ικανός να φέρνει σε πέρας πολύ σύνθετες εργασίες σε εξαιρετικά σύντομο χρονικό διάστημα. Το πλεονέκτημα του ανθρώπινου εγκεφάλου είναι ότι τα διάφορα υποσυστήματά του (ομάδες νευρώνων) *λειτουργούν παράλληλα* με αποτέλεσμα ξεχωριστές μεταξύ τους λειτουργίες να εκτελούνται στο ίδιο χρονικό διάστημα και στη συνέχεια να γίνεται η σύνδεσή τους.

Η σημαντικότερη προσπάθεια μοντελοποίησης του βιολογικού νευρώνα έγινε από τους McCulloch και Pitts. Το μοντέλο των McCulloch - Pitts δεν είναι ένα πλήρες μοντέλο του νευρικού κυττάρου, αλλά μια απλή παρουσίαση δύο καταστάσεων της λογικής διαδικασίας που λαμβάνει χώρα σ' αυτό. Το βασικό μοντέλο νευρώνα των McCulloch και Pitts έχει τη μορφή του Σχήματος 2.2 .



Σχήμα 2.2: Μοντέλο νευρώνα McCulloch – Pitts.

Η ένα εξόδου του κυττάρου συσχετίζει τις πληροφορίες που αφορούν στην κατάσταση του νευρώνα (διέγερση ή όχι) και τελικά γίνεται είσοδος σε ένα άλλο κύτταρο. Επιτρέπεται οποιοσδήποτε αριθμός εισόδων και οι εισόδοι διαιρούνται σε δύο κατηγορίες: στις εισόδους που διεγείρουν το κύτταρο και στις απαγορευτικές εισόδους. Οι εισόδοι διέγερσης που αναπαρίστανται με μια γραμμή, η οποία καταλήγει σε ένα βελάκι, τείνουν στο να διεγείρουν το κύτταρο κατά αυξανόμενο τρόπο μέχρι την τελική διέγερση. Οι απαγορευτικές εισόδοι αναπαρίστανται από μια γραμμή η οποία καταλήγει σε ένα μικρό κύκλο. Οι εισόδοι αυτές είναι ικανές να απαγορεύσουν εντελώς τη διέγερση του κυττάρου. Έστω T η αριθμητική τιμή του κατωφλίου διέγερσης του κυττάρου. Το κύτταρο McCulloch - Pitts θα διεγερθεί όταν το άθροισμα των εισόδων διέγερσης γίνει μεγαλύτερο ή ίσο με την τιμή κατωφλίου και δεν υπάρχουν απαγορευτικές εισόδοι.

Με βάση το βιολογικό ανάλογο, μπορούμε να ορίσουμε ένα Τεχνητό Νευρωνικό Δίκτυο ως εξής:

“Τεχνητό Νευρωνικό Δίκτυο είναι μια αρχιτεκτονική δομή (δίκτυο) αποτελούμενη από ένα πλήθος διασυνδεδεμένων μονάδων επεξεργασίας (τεχνητοί νευρώνες). Κάθε μονάδα επεξεργασίας

χαρακτηρίζεται από εισόδους και εξόδους. Υλοποιεί τοπικά ένα υπολογισμό με βάση τις εισόδους που δέχεται και μεταδίδει το αποτέλεσμα (έξοδος) σε άλλες μονάδες επεξεργασίας με τις οποίες συνδέεται. Οι τιμές των βαρών των συνδέσεων αποτελούν τη γνώση που είναι αποθηκευμένη στο ΤΝΔ και καθορίζουν τη λειτουργικότητά του. Συνήθως ένα ΤΝΔ αναπτύσσει μια συνολική λειτουργικότητα μέσω μιας μορφής εκπαίδευσης (μάθησης)”.

2.3 Βασικό Μοντέλο Τεχνητού Νευρώνα

Ένας τεχνητός νευρώνας είναι μια μονάδα επεξεργασίας πληροφοριών, η οποία είναι θεμελιώδης για τη λειτουργία ενός νευρωνικού δικτύου. Το Σχήμα 2.3 δείχνει το μοντέλο ενός νευρώνα, ο οποίος αποτελεί τη βάση για τον σχεδιασμό των τεχνητών νευρωνικών δικτύων. Εδώ αναγνωρίζονται τα βασικά στοιχεία των νευρωνικών μοντέλων:

Τα δεδομένα εισόδου (Inputs), τα οποία είναι συνήθως ερεθίσματα από το εξωτερικό περιβάλλον αλλά μπορεί να είναι ακόμα και αποτελέσματα ενός άλλου Τεχνητού Νευρωνικού Δικτύου. Μπορεί να είναι πολλών μορφών όπως για παράδειγμα ακέραιοι ή πραγματικοί αριθμοί.

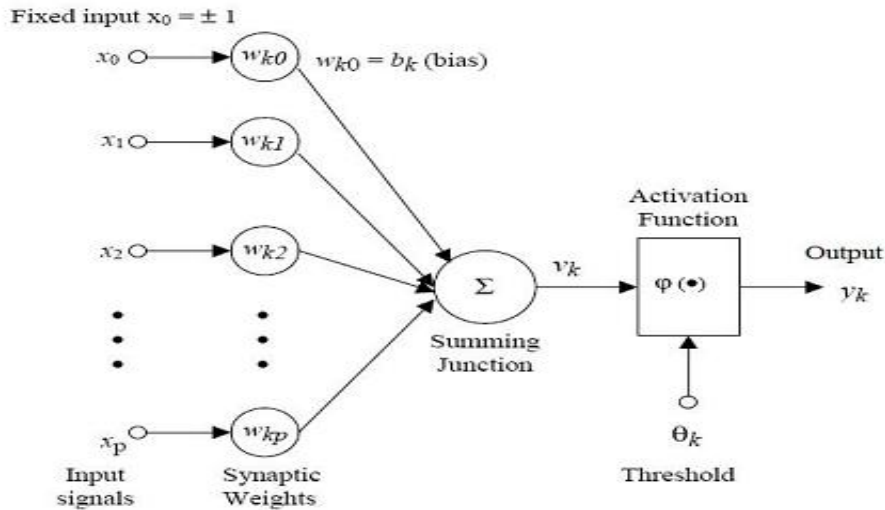
Τα συναπτικά βάρη (Synaptic Weights). Αν η είσοδος του νευρώνα k στη σύναψη j είναι το σήμα x_j , τότε αυτό πολλαπλασιάζεται με το συναπτικό βάρος w_{kj} , όπου ο πρώτος δείκτης αναφέρεται στο νευρώνα και ο δεύτερος στη σύναψη που δέχεται στην είσοδό της το σήμα w_{kj} . Σκοπός της εκπαίδευσης του Τεχνητού Νευρωνικού Δικτύου είναι η διαμόρφωση των σωστών τιμών για τα συναπτικά βάρη έτσι ώστε να έχουμε την επιθυμητή έξοδο.

Ο αθροιστής (Weighted Sum), ο οποίος στην έξοδό του δίνει το άθροισμα των σταθμισμένων εισόδων.

Το κατώφλι (Threshold) θ_k . Πρόκειται για δευτερεύουσα παράμετρο του συστήματος, η οποία συνήθως επιλέγεται με στόχο την καλύτερη ευελιξία του και αφαιρείται από το άθροισμα το οποίο παίρνουμε από την έξοδο του αθροιστή.

Η συνάρτηση ενεργοποίησης (Activation Function). Από αυτή περνά η έξοδος του αθροιστή και δίνει αποτέλεσμα στο διάστημα $[0, 1]$ ή $[-1, 1]$ ανάλογα με τον τύπο της συνάρτησης που επιλέχθηκε. Η συνάρτηση που χρησιμοποιήθηκε για τον πρωτότυπο νευρώνα των McCulloch-Pitts ήταν η βηματική συνάρτηση (unit step function).

Δεδομένο εξόδου (Output) Αποτελεί το τελικό προϊόν του νευρώνα και είναι το αποτέλεσμα της εφαρμογής της συνάρτησης ενεργοποίησης στο σταθμισμένο άθροισμα των δεδομένων εισόδων αφού έχουμε αφαιρέσει την παράμετρο του κατωφλιού. Ο τύπος του δεδομένου εξόδου εξαρτάται άμεσα από την συνάρτηση ενεργοποίησης.



Σχήμα 2.3: Τυπική δομή τεχνητού νευρώνα.

Οι εξισώσεις που περιγράφουν το παραπάνω μοντέλο είναι:

$$u_k = \sum_{j=0}^N w_{kj} \cdot x_j \quad (2.1)$$

$$y_k = \varphi(u_k - \theta_k) \quad (2.2)$$

Όπου x_j , $j=1, \dots, p$ τα σήματα εισόδου, w_{kj} τα συναπτικά βάρη του νευρώνα k , u_k η είσοδος του γραμμικού συνδυαστή, θ_k το κατώφλι, $\varphi(\cdot)$ η συνάρτηση ενεργοποίησης και y_k είναι το σήμα εξόδου του νευρώνα k .

Θεωρώντας και το κατώφλι ως συναπτικό βάρος w_{kj0} της εισόδου $x_0 = -1$, το v_k , δηλαδή η έξοδος του αθροιστή, δίνεται από τη σχέση :

$$v_k = \sum_{j=0}^N w_{kj} \cdot x_j \quad (2.3)$$

ενώ η έξοδος του νευρώνα είναι:

$$y_k = \varphi(v_k) \quad (2.4)$$

Το κατώφλι θ_k μειώνει την είσοδο της συνάρτησης ενεργοποίησης όταν το σήμα είναι θετικό και την αυξάνει όταν είναι αρνητικό.

Όσον αφορά τη **συνάρτηση ενεργοποίησης**, ευθύνεται σε μεγάλο βαθμό για τη συμπεριφορά του νευρώνα. Συνήθως, χρησιμοποιείται η ίδια συνάρτηση ενεργοποίησης για όλους τους νευρώνες σε όλα τα επίπεδα του ΤΝΔ, παρόλο που δεν είναι υποχρεωτικό κάτι τέτοιο. Η συνάρτηση ενεργοποίησης μπορεί να είναι βηματική, γραμμική ή μη γραμμική. Παρακάτω δίνονται οι εξισώσεις

μερικών συναρτήσεων ενεργοποίησης και οι γραφικές παραστάσεις τους στο Σχήμα 2.4

Η βηματική συνάρτηση ενεργοποίησης έχει τη μορφή:

$$\varphi(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases} \quad (2.5)$$

Η γραμμική συνάρτηση ενεργοποίησης μπορεί να είναι η:

$$\varphi(x) = x \quad (2.6\alpha)$$

ή και η

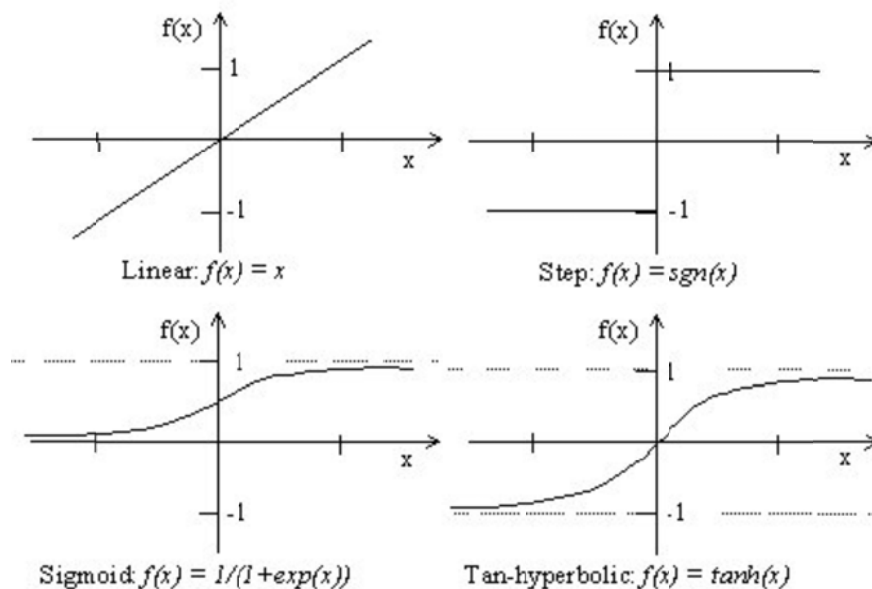
$$\varphi(x) = \begin{cases} 0, & x \leq x_{\min} \\ mx + b, & x_{\max} > x > x_{\min} \\ 1, & x > x_{\max} \end{cases} \quad (2.6\beta)$$

Χαρακτηριστικό παράδειγμα μη γραμμικών συναρτήσεων ενεργοποίησης αποτελεί η οικογένεια των σιγμοειδών συναρτήσεων, όπως είναι η σιγμοειδής συνάρτηση υπερβολικής εφαπτομένης (Hyperbolic Tangent Function), που ορίζεται ως εξής:

$$\varphi(v) = \tanh\left(\frac{v}{2}\right) = \frac{1 - e^{-v}}{1 + e^{-v}} \quad (2.7\alpha)$$

Μια άλλη σιγμοειδής συνάρτηση, είναι η λογιστική σιγμοειδής συνάρτηση, η οποία έχει την εξής μορφή:

$$\varphi(v) = \frac{1}{1 + e^{-av}} \quad (2.7\beta)$$



Σχήμα 2.4 : Τυπικές Συναρτήσεις Ενεργοποίησης

Τα σύγχρονα Τεχνητά Νευρωνικά Δίκτυα αποτελούνται από αρκετούς απλούς νευρώνες,

όπως αυτόν που περιγράφηκε παραπάνω, οι οποίοι συνθέτουν αρχικά ένα στρώμα και στη συνέχεια ένα σύστημα πολλών στρωμάτων.

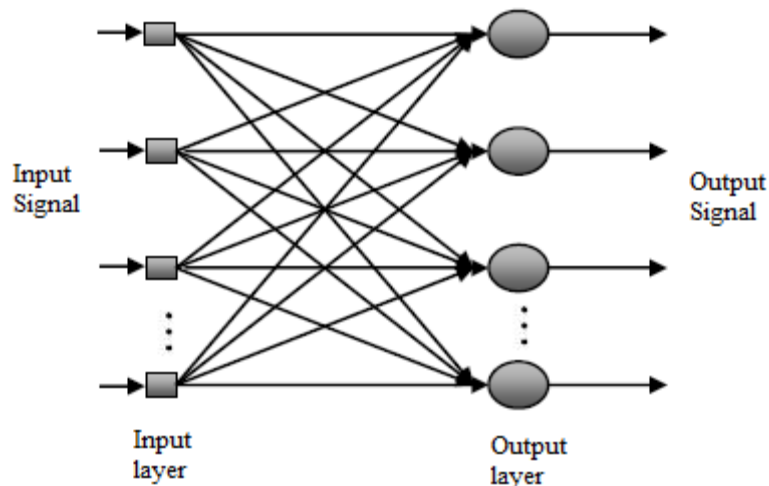
2.4 Αρχιτεκτονικές Νευρωνικών Δικτύων

Η τοπολογική δομή είναι το κύριο χαρακτηριστικό των ΤΝΔ και αναφέρεται στην αρχιτεκτονική στην οποία διευθετούνται και διασυνδέονται πολλαπλοί νευρώνες. Οι δύο βασικές ιδιότητες που καθορίζουν την αρχιτεκτονική ενός ΤΝΔ είναι το *πλήθος των επιπέδων* (layers) και οι *συνδέσεις* μεταξύ των νευρώνων. Το τρίτο χαρακτηριστικό, το οποίο σχετίζεται με τον τρόπο κατά τον οποίο είναι δομημένοι οι νευρώνες, είναι ο *αλγόριθμος μάθησης* που χρησιμοποιείται για την εκπαίδευση του δικτύου.

Η απλούστερη και ευρύτερα χρησιμοποιούμενη κατηγορία ΤΝΔ είναι τα *δίκτυα πρόσθιας τροφοδότησης* (feedforward networks). Το χαρακτηριστικό των δικτύων αυτών είναι ότι δεν υπάρχει ανατροφοδότηση της εξόδου ενός νευρώνα προς τους νευρώνες από τους οποίους επηρεάζεται άμεσα ή έμμεσα. Με άλλα λόγια, οι υπολογισμοί πραγματοποιούνται από την είσοδο με κατεύθυνση προς την έξοδο.

2.4.1 Μονοεπίπεδα Νευρωνικά Δίκτυα (Single Layered Perceptron)

Μολονότι ένας νευρώνας μπορεί να εκτελέσει μερικές απλές λειτουργίες που αφορούν την ανίχνευση απλών προτύπων, η δύναμη της υπολογιστικότητας των νευρώνων εμφανίζεται μόνο όταν αυτά συνδεθούν σε δίκτυο. Το πιο απλό δίκτυο αποτελείται από ένα στρώμα εισόδου το οποίο προβάλλεται στους νευρώνες του επιπέδου εξόδου, αλλά όχι αντίστροφα. Ένα τέτοιο ΤΝΔ είναι αυστηρά τύπου πρόσθιας τροφοδότησης (feedforward) και καλείται ΤΝΔ πρόσθιας τροφοδότησης ενός επιπέδου. Τα νευρώνια του στρώματος εισόδου που φαίνονται στα αριστερά του Σχήματος 2.5 εργάζονται μόνο για το μοίρασμα των τιμών εισόδου, δεν εκτελούν υπολογισμούς και έτσι δεν θεωρούνται ότι αποτελούν στρώμα. Οι υπολογισμοί πραγματοποιούνται στο επίπεδο εξόδου όπου ο κάθε νευρώνας παράγει ένα άθροισμα από τις τιμές εισόδου του δικτύου που έχουν πολλαπλασιαστεί με τα αντίστοιχα βάρη.



Σχήμα 2.5 : Μονοεπίπεδο Νευρωνικό Δίκτυο

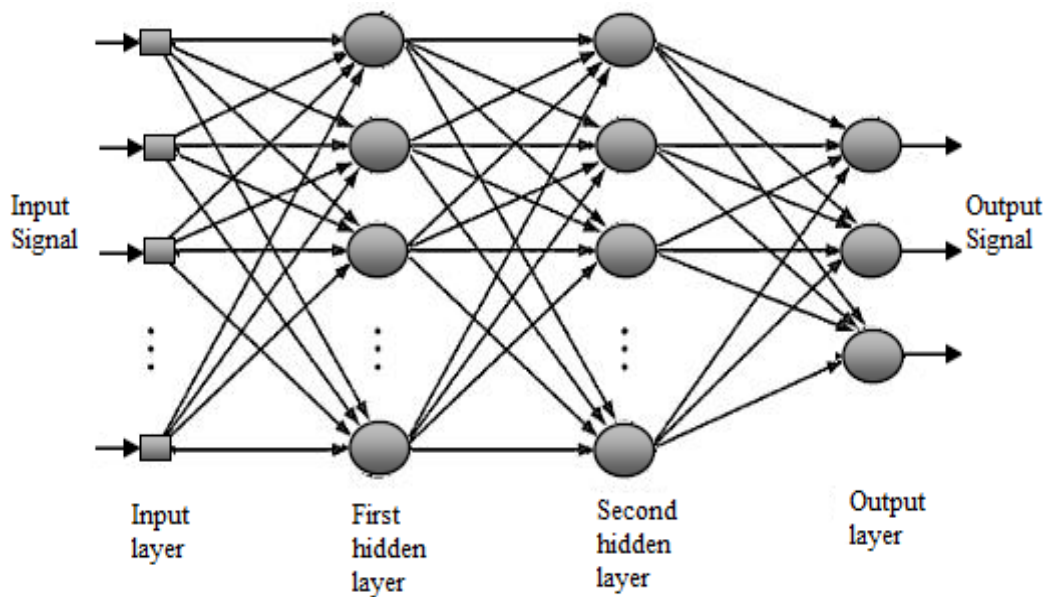
Κάθε μία από τις εισόδους (x_1, x_2, \dots, x_N) συνδέεται με κάθε νευρώνα του στρώματος εξόδου, μέσω της σύνδεσης με το αντίστοιχο βάρος. Δεδομένου ότι όλες οι τιμές εξόδου (y_1, y_2, \dots, y_N) υπολογίζονται από το ίδιο σύνολο τιμών εισόδου, κάθε έξοδος εξαρτάται από τα συναπτικά βάρη .

$$y_k = \varphi\left(\sum_{j=1}^N w_{kj} \cdot x_j\right) \quad (2.8)$$

Η διαδικασία εκπαίδευσης ενός τέτοιου δικτύου τυπικά περιλαμβάνει την προσαρμογή του πίνακα των συναπτικών βαρών, έτσι ώστε το νευρωνικό δίκτυο να μιμείται την απόκριση για μια γνωστή αντιστοίχιση εισόδου-εξόδου. Στην πραγματικότητα πολλές από τις συνδέσεις που φαίνονται μπορεί να μην υπάρχουν, οι συνδέσεις αυτές θα έχουν τιμή για το συναπτικό βάρος ίση με μηδέν.

2.4.2 Πολυεπίπεδα Νευρωνικά Δίκτυα (Multi Layered Perceptron, MLP)

Μεγαλύτερα περισσότερο πολύπλοκα δίκτυα, γενικά, προσφέρουν μεγαλύτερη ικανότητα υπολογισμών. Τα πολυεπίπεδα δίκτυα έχει αποδειχθεί ότι έχουν ικανότητες πέρα από αυτές των μονόστρωματικών δικτύων και στα πρόσφατα χρόνια έχουν αναπτυχθεί αλγόριθμοι για την εκπαίδευσή τους. Τα πολυστρωματικά δίκτυα μπορούν να σχηματιστούν από ομάδες μονόστρωματικών δικτύων, όπου η έξοδος ενός στρώματος αποτελεί την είσοδο του επόμενου στρώματος, όπως δείχνει το Σχήμα 2.6. Τα επίπεδα (στρώματα) που παρεμβάλλονται μεταξύ του στρώματος εισόδου και εξόδου ονομάζονται κρυφά επίπεδα.



Σχήμα 2.6 : Πολυεπίπεδο Νευρωνικό Δίκτυο

Η επιλογή του αριθμού των κρυφών επιπέδων είναι γενικά ένας συμβιβασμός μεταξύ της πολυπλοκότητας του προβλήματος και του υπολογιστικού φόρτου. Τα πολυστρωματικά νευρωνικά δίκτυα εκπαιδεύονται με επιβλεπόμενο τρόπο με ένα δημοφιλή αλγόριθμο γνωστό ως αλγόριθμο οπισθοδιάδοσης σφάλματος (Error Back Propagation, BP). Στη διαδικασία της εκπαίδευσης διακρίνονται δυο φάσεις :

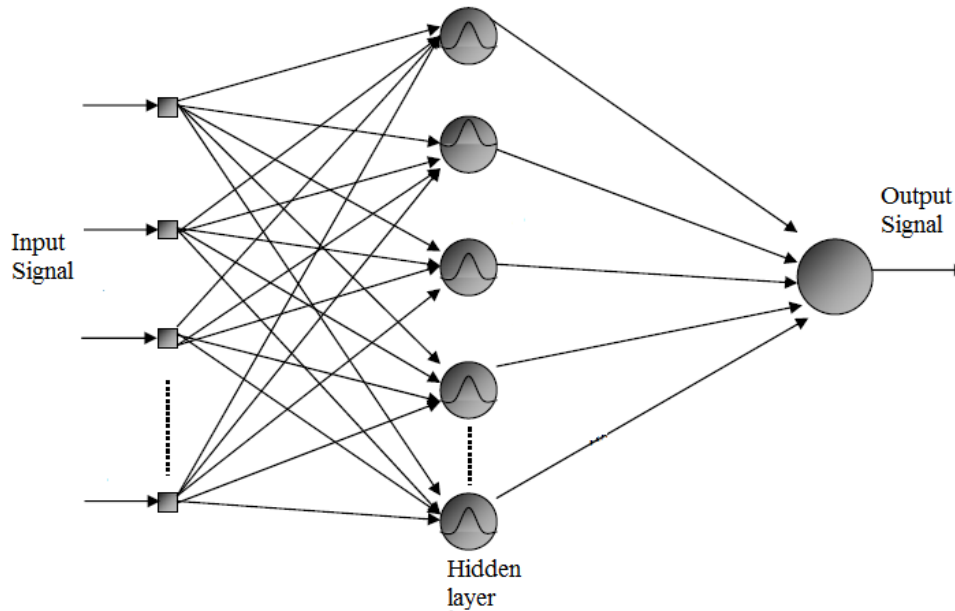
- Εμπρόσθια σάρωση (forward pass) κατά την οποία υπολογίζεται η έξοδος του δικτύου για κάθε είσοδο.
- Ανάστροφη σάρωση (backward pass) κατά την οποία ανανεώνονται τα βάρη του δικτύου. Η πραγματική έξοδος της προηγούμενης φάσης συγκρίνεται με την επιθυμητή, δίνοντας το σφάλμα που αντιστοιχεί σε κάθε κόμβο εξόδου. Με κατάλληλο τρόπο τα σφάλματα αυτά μεταφέρονται προς τα πίσω συμβάλλοντας στην ανανέωση των βαρών.

2.4.3 Νευρωνικά Δίκτυα Ακτινικών Συναρτήσεων Βάσης

Τα ΤΝΔ ακτινικών συναρτήσεων βάσης (Radial Basis Functions, RBF) αποτελούν και αυτά ΤΝΔ πρόσθιας τροφοδότησης, αλλά θεωρείται ότι αποτελούν ξεχωριστή κατηγορία , λόγω των διαφορών τους στη δομή, τη λειτουργία αλλά και τον τρόπο εκπαίδευσης τους. Είναι ένας σχετικά νέος τύπος αρχιτεκτονικής νευρωνικών δικτύων που εισήχθη στη δεκαετία του 1980 ως πρόταση βελτίωσης των ΤΝΔ πρόσθιας τροφοδότησης.

Τα RBF νευρωνικά δίκτυα διαφέρουν από την παραδοσιακή MLP αρχιτεκτονική στην

ενεργοποίηση του κρυφού επιπέδου, η οποία καθορίζεται από την Ευκλείδεια απόσταση μεταξύ του διανύσματος εισόδου και ενός πρότυπου διανύσματος. Η συνάρτηση ενεργοποίησης των νευρώνων του κρυφού επιπέδου επιλέγεται από ένα σύνολο ακτινικών συναρτήσεων βάσης και συνήθως είναι η Γκαουσιανή.



Σχήμα 2.7 : Νευρωνικό Δίκτυο Ακτινικής Συνάρτησης Βάσης

Τα δίκτυα RBF, όπως δείχνει το Σχήμα 2.7 έχουν μόνο ένα κρυφό στρώμα, στο οποίο κάθε νευρώνας ορίζεται από το κέντρο του και την ακτίνα του. Κάθε κρυφός νευρώνας δίνει μια έξοδο η οποία στη συνέχεια σταθμίζεται και περνάει στο στρώμα εξόδου. Οι έξοδοι του δικτύου αποτελούνται από τα αθροίσματα των σταθμισμένων νευρώνων του κρυφού στρώματος.

Δεδομένου ότι οι συναρτήσεις ενεργοποίησης είναι μη γραμμικές, δεν είναι απαραίτητο να υπάρχουν περισσότερα από ένα κρυφά επίπεδα για να προσεγγίσουν οποιαδήποτε συνάρτηση. Η εκπαίδευση των RBF νευρωνικών δικτύων είναι γενικά μια διαδικασία δύο σταδίων.

- (α) Προσδιορισμός των πρότυπων διανυσμάτων (κέντρο και ακτίνα)
- (β) Προσαρμογή των βαρών του στρώματος εξόδου.

Τα δίκτυα RBF θα συζητηθούν αναλυτικά στο Κεφάλαιο III.

Σύγκριση RBF και MLP

Και οι δύο αυτές αρχιτεκτονικές ΤΝΔ έχουν το κοινό χαρακτηριστικό ότι υλοποιούν μη γραμμικές απεικονίσεις από το χώρο των εισόδων στο χώρο των εξόδων. Τόσο τα MLP όσο και τα RBF χαρακτηρίζονται από την ιδιότητα της καθολικής προσέγγισης (global approximation), επομένως

μπορούν να προσεγγίσουν (με οσοδήποτε καλή ακρίβεια) οποιαδήποτε συνάρτηση, αν χρησιμοποιηθεί ο κατάλληλος αριθμός νευρώνων.

Τα δίκτυα RBF έχουν μια σειρά από πλεονεκτήματα σε σχέση με τα MLP. Πρώτον, όπως αναφέρθηκε προηγουμένως, μπορούν να προσεγγίσουν οποιαδήποτε μη γραμμική συνάρτηση, χρησιμοποιώντας ένα μόνο κρυφό επίπεδο, μειώνοντας έτσι κάποιες μεταβλητές απόφασης σχετικά με τον αριθμό των στρωμάτων. Δεύτερον, λειτουργούν δημιουργώντας *τοπικές αναπαραστάσεις*, δηλ. για τον υπολογισμό της εξόδου συμμετέχουν μόνο οι κρυφοί νευρώνες που περιλαμβάνουν την είσοδο στη σφαίρα επιρροής τους. Τέλος, η απλή γραμμική μετατροπή στο στρώμα εξόδου μπορεί να βελτιστοποιηθεί πλήρως με την παραδοσιακή μέθοδο της γραμμικής παλινδρόμησης. Ως εκ τούτου, τα δίκτυα RBF μπορούν να εκπαιδευτούν τάξεις μεγέθους γρηγορότερα σε σχέση με τα δίκτυα MLP.

Από την άλλη πλευρά, οι αλγόριθμοι εκπαίδευσης των RBF είναι επιρρεπείς στο να συγκλίνουν σε μη βέλτιστους συνδυασμούς. Επίσης, τα RBF δεν προεκτείνουν σε περιοχές που δεν έχουν εκπαιδευτεί. Όταν τα δεδομένα εισόδου είναι μακριά από τα ακτινικά κέντρα η έξοδος πλησιάζει το μηδέν, κάτι που οφείλεται στις ακτινικές συναρτήσεις βάσης που χρησιμοποιούνται ως συναρτήσεις ενεργοποίησης. Μια εφαρμογή στην οποία μπορεί να προκύψει το παραπάνω πρόβλημα είναι τα προβλήματα παλινδρόμησης, τα οποία απαιτούν μη περιορισμένη συνάρτηση ενεργοποίησης.

2.5 Εκπαίδευση Τεχνητών Νευρωνικών Δικτύων

Ένας ορισμός για την εκπαίδευση ΤΝΔ που πρότειναν οι Mendel και McClaren είναι ο εξής:

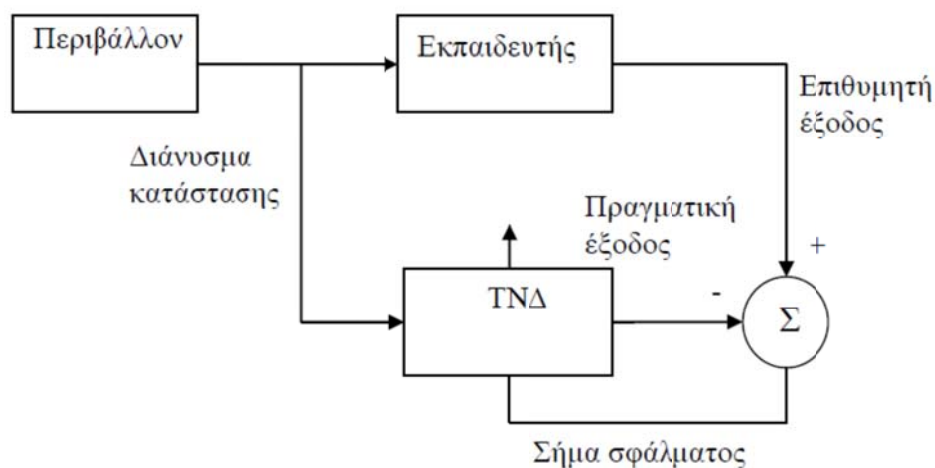
“Η *εκπαίδευση* (training) ενός ΤΝΔ είναι μια διαδικασία από την οποία οι ελεύθερες παράμετροι του νευρωνικού δικτύου προσαρμόζονται μέσω μιας διαδικασίας διέγερσης από το περιβάλλον στο οποίο είναι εμπεδωμένο το δίκτυο. Ο τύπος της εκπαίδευσης καθορίζεται από τον τρόπο με τον οποίο λαμβάνει χώρα η αλλαγή των παραμέτρων.”

2.5.1 Εκπαίδευση με Επίβλεψη

Η δομή της *εκπαίδευσης με επίβλεψη* (supervised learning) έχει τη γενική μορφή του Σχήματος 2.8 και περιλαμβάνει δύο κύριες συνιστώσες, το *δάσκαλο* και το *σύστημα μάθησης*. Ο δάσκαλος έχει γνώση του περιβάλλοντος, και αυτή η γνώση αντιπροσωπεύεται από ένα σύνολο παραδειγμάτων εισόδου-εξόδου. Το περιβάλλον είναι άγνωστο για το νευρωνικό δίκτυο και έτσι κατά τη διαδικασία της εκπαίδευσης ο δάσκαλος παρέχει στο νευρωνικό δίκτυο τις επιθυμητές αποκρίσεις για τα διανύσματα εκπαίδευσης. Για να φτάσει στη βέλτιστη απόκριση το νευρωνικό δίκτυο προσαρμόζονται οι παράμετροι του υπό τη συνδυασμένη επιρροή του διανύσματος εκπαίδευσης και του σήματος σφάλματος. Το σήμα σφάλματος είναι η διαφορά μεταξύ της

επιθυμητής απόκρισης και της πραγματικής απόκρισης του δικτύου. Οι αλλαγές που επιτελούνται στις παραμέτρους του νευρωνικού δικτύου γίνονται επαναληπτικά έως ότου το δίκτυο να προσομοιώνει τη συμπεριφορά του δασκάλου. Η γνώση αυτή του δασκάλου αποθηκεύεται στο νευρωνικό δίκτυο με την μορφή σταθερών συναπτικών βαρών. Όπως φαίνεται στο Σχήμα 2.8 η επιβλεπόμενη μάθηση αποτελείται από ένα σύστημα ανάδρασης κλειστού βρόχου, με το περιβάλλον είναι εκτός του συστήματος αυτού.

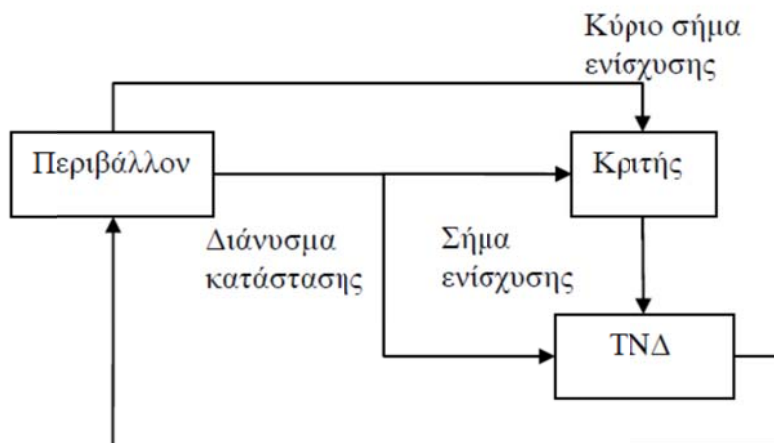
Ως μέτρο απόδοσης χρησιμοποιούμε συχνά το μέσο τετραγωνικό σφάλμα ή το άθροισμα των τετραγώνων των σφαλμάτων. Η *καμπύλη μάθησης* είναι το μέτρο απόδοσης σε συνάρτηση με τον αριθμό των επαναλήψεων των διανυσμάτων εισόδου-εξόδου και μας δίνει μια καλή εικόνα σχετικά με την αποτελεσματικότητα της εκπαίδευσης. Η καμπύλη αυτή πρέπει να είναι φθίνουσα και κατά τη διάρκεια της εκπαίδευσης την πληροφορία αυτή μας την δίνει η κλίση (gradient) της καμπύλης. Με την επιβλεπόμενη μάθηση με τη χρήση κατάλληλου αλγορίθμου που μειώνει την συνάρτηση κόστους μπορεί να προσεγγιστεί μια άγνωστη απεικόνιση εισόδου-εξόδου.



Σχήμα 2.8 : Δομή της Εκπαίδευσης με Επίβλεψη

2.5.2 Εκπαίδευση με Ενίσχυση

Στην *εκπαίδευση με ενίσχυση* (reinforcement learning), η εκπαίδευση μιας απεικόνισης εισόδου-εξόδου εκτελείται μέσω συνεχών αλληλεπιδράσεων με το περιβάλλον, έτσι ώστε να ελαχιστοποιηθεί ένας βαθμωτός δείκτης απόδοσης. Το Σχήμα 2.9 παρουσιάζει ένα σύστημα εκπαίδευσης με ενίσχυση, το οποίο περιλαμβάνει ένα *κριτικό* (critic) που μετατρέπει ένα πρωτεύον σήμα ενίσχυσης λαμβανόμενο από το περιβάλλον σε ένα σήμα ενίσχυσης υψηλότερης ποιότητας που ονομάζεται *ευρετικό σήμα ενίσχυσης*.

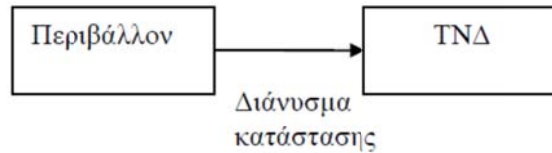


Σχήμα 2.9 : Δομή Εκπαίδευσης με Ενίσχυση

Το σύστημα μαθαίνει κάτω από καθυστερούμενη ενίσχυση, που σημαίνει ότι το σύστημα παρακολουθεί μια χρονική ακολουθία ερεθισμάτων (διανύσματα κατάστασης) τα οποία οδηγούν στην παραγωγή του ευρετικού σήματος ενίσχυσης. Ο στόχος της εκπαίδευσης είναι η *ελαχιστοποίηση μιας συνάρτησης cost-to-go*, που ορίζεται ως η πρόβλεψη του αθροιστικού κόστους των ενεργειών που λαμβάνουν χώρα σε μια ακολουθία από βήματα αντί του απλού άμεσου κόστους. Μπορεί να αποδειχθεί ότι κάποιες ενέργειες που πάρθηκαν νωρίτερα σε αυτή την χρονική ακολουθία βημάτων είναι στην πραγματικότητα οι πιο καθοριστικοί παράγοντες για ολόκληρη την συμπεριφορά του συστήματος. Η λειτουργία της *μηχανής εκπαίδευσης*, που αποτελεί το δεύτερο συστατικό του συστήματος, είναι να ανακαλύψει αυτές τις ενέργειες και να τις τροφοδοτήσει πάλι στο περιβάλλον.

2.5.3 Εκπαίδευση χωρίς Επίβλεψη

Στην εκπαίδευση χωρίς επίβλεψη δεν υπάρχει εξωτερικός δάσκαλος ή κριτικός για να επιβλέπει την διαδικασία εκπαίδευσης, όπως φαίνεται και στο Σχήμα 2.10. Το νευρωνικό δίκτυο μαθαίνει από ένα μέτρο της ποιότητας της αναπαράστασης που καλείται να μάθει και οι ελεύθεροι παράμετροι του δικτύου βελτιστοποιούνται με βάση το μέτρο αυτό. Για να εκτελεστεί εκπαίδευση χωρίς επίβλεψη πρέπει να χρησιμοποιηθεί ένας ανταγωνιστικός κανόνας εκπαίδευσης. Για παράδειγμα, μπορεί να χρησιμοποιηθεί ένα νευρωνικό δίκτυο που αποτελείται από δύο επίπεδα, ένα επίπεδο εισόδου και ένα ανταγωνιστικό επίπεδο. Το επίπεδο εισόδου λαμβάνει τα διαθέσιμα δεδομένα. Το ανταγωνιστικό επίπεδο αποτελείται από νευρώνες που ανταγωνίζονται μεταξύ τους (σύμφωνα με τον κανόνα εκπαίδευσης) για την 'ευκαιρία' να αποκριθούν σε χαρακτηριστικά που περιέχονται στα δεδομένα εισόδου.



Σχήμα 2.10 : Δομή Εκπαίδευσης χωρίς Επίβλεψη

2.6 Γενίκευση

Ένα ΤΝΔ χαρακτηρίζεται από δυνατότητες γενίκευσης όταν η απεικόνιση εισόδου-εξόδου που υπολογίζεται από το δίκτυο είναι σωστή για δεδομένα δοκιμής που δεν χρησιμοποιήθηκαν ποτέ στην εκπαίδευση του δικτύου. Εδώ υποτίθεται ότι τα δεδομένα δοκιμής επιλέγονται από τον ίδιο πληθυσμό που χρησιμοποιείται για την δημιουργία των δεδομένων εκπαίδευσης.

Η διαδικασία της εκπαίδευσης ενός νευρωνικού δικτύου μπορεί να θεωρηθεί ως ένα πρόβλημα *προσαρμογής καμπύλης*. Το ίδιο το δίκτυο μπορεί να θεωρηθεί απλά ως μια μη γραμμική απεικόνιση εισόδου-εξόδου. Μια τέτοια σκοπιά μας επιτρέπει να δούμε τη γενίκευση ως το αποτέλεσμα μιας καλής μη γραμμικής παρεμβολής των δεδομένων εισόδου

Όταν ωστόσο, ένα νευρωνικό δίκτυο μαθαίνει πάρα πολλά παραδείγματα εισόδου-εξόδου, το δίκτυο μπορεί να καταλήξει να απομνημονεύει τα δεδομένα εκπαίδευσης. Μπορεί να το κάνει αυτό βρίσκοντας μια ιδιότητα (λόγω του θορύβου για παράδειγμα) που είναι παρούσα στα δεδομένα εκπαίδευσης, αλλά στην πραγματικότητα, όχι στην υποκείμενη συνάρτηση που πρόκειται να αναπαραχθεί. Ένα τέτοιο φαινόμενο αναφέρεται ως *υπερπροσαρμογή* ή *υπερεκπαίδευση*. Όταν το δίκτυο υπερεκπαιδεύεται, χάνει την ικανότητά του να γενικεύει μεταξύ παρόμοιων προτύπων εισόδου-εξόδου.

Η γενίκευση επηρεάζεται από τρεις παράγοντες :

- (1.) Το μέγεθος του συνόλου εκπαίδευσης και κατά πόσο είναι αντιπροσωπευτικό για το περιβάλλον ενδιαφέροντος.
- (2.) Την αρχιτεκτονική του νευρωνικού δικτύου.
- (3.) Τη φυσική πολυπλοκότητα του προβλήματος.

Σαφώς, δεν έχουμε τον έλεγχο στον τελευταίο παράγοντα και για τους υπόλοιπους συνήθως καταφεύγουμε σε δοκιμές.

2.7 Πλεονεκτήματα και Μειονεκτήματα Τεχνητών Νευρωνικών Δικτύων

Πλεονεκτήματα

Θα αναφέρουμε σε αυτό το σημείο, τα κυριότερα χαρακτηριστικά που κάνουν τα Τεχνητά Νευρωνικά Δίκτυα να ξεχωρίζουν από τις καθιερωμένες υπολογιστικές μεθόδους, αλλά και που συνηγορούν για την εφαρμογή τους σε διάφορα πεδία.

Ευελιξία: Τα Τεχνητά Νευρωνικά Δίκτυα χαρακτηρίζονται από υψηλή προσαρμοστικότητα. Αυτό συμβαίνει γιατί έχουν την δυνατότητα να αλλάζουν τα βάρη τους ανάλογα με τις μεταβολές του περιβάλλοντος τους και άρα ουσιαστικά να επανεκπαιδεύονται. Αυτός είναι ο λόγος για τον οποίο είναι ιδανικά για την επίλυση πολύπλοκων και πολυδιάστατων προβλημάτων. Ένα άλλο χαρακτηριστικό της ευελιξίας τους είναι πως έχουν την δυνατότητα να δέχονται τόσο ψηφιακά όσο και αναλογικά δεδομένα.

Μη γραμμικότητα: Οι νευρώνες, οι οποίοι συνδέονται μεταξύ τους για να σχηματίσουν το Τεχνητό Νευρωνικό Δίκτυο, αποτελούν ουσιαστικά μη γραμμικές μονάδες επεξεργασίας. Έτσι υπάρχει η δυνατότητα μοντελοποίησης και προσομοίωσης φαινομένων των οποίων ο μηχανισμός παραγωγής σημάτων είναι μη γραμμικός.

Αντιστοιχία στους βιολογικούς νευρώνες: Η λειτουργία των Τεχνητών Νευρωνικών Δικτύων βασίζεται στις αρχές που διέπουν την λειτουργία του ανθρώπινου εγκεφάλου. Άρα η περαιτέρω βιολογική έρευνα θα δώσει πιθανότατα την δυνατότητα για βελτιώσεις στον σχεδιασμό και στην αρχιτεκτονική των Τεχνητών Νευρωνικών Δικτύων.

Ανθεκτικότητα: Τα Τεχνητά Νευρωνικά Δίκτυα αποτελούνται, όπως αναφέραμε, από πολλές μονάδες επεξεργασίας που είναι μεταξύ τους συνδεδεμένες. Τα δεδομένα διαχέονται σε όλο το δίκτυο και έτσι δημιουργείται μια «ανοσία» σε κάποια τοπική αστοχία. Αν μέρος των νευρώνων ή των μεταξύ τους συνδέσεων δεν λειτουργήσει, το σύστημα στο σύνολο του δεν καταστρέφεται ενώ η απόδοση του δεν θα φθίνει σημαντικά.

Ευκολία στη χρήση: Ουσιαστικά για να μπορέσει κάποιος να το χρησιμοποιήσει τα ΤΝΔ χρειάζεται γνώσεις για την επιλογή του σωστού δικτύου και την ερμηνεία των αποτελεσμάτων. Οι απαιτήσεις σε επίπεδο γνώσεων είναι αρκετά χαμηλότερες σε σχέση με τις γνώσεις που απαιτούν παραδοσιακές μέθοδοι μοντελοποίησης.

Ταχύτητα υπολογισμών: Λόγω της παράλληλης λειτουργίας όλων των νευρώνων τα Τεχνητά Νευρωνικά Δίκτυα έχουν υψηλές ταχύτητες απόκρισης

Γενικότητα: Τα Τεχνητά Νευρωνικά Δίκτυα μπορούν να αποτυπώσουν και να επεξεργαστούν μια αρκετά μεγάλη ποικιλία προτύπων (patterns) και έτσι μας δίνουν τη δυνατότητα να τα χρησιμοποιήσουμε σε ένα μεγάλο πλήθος εφαρμογών.

Εκπαίδευση: Τα ΤΝΔ δίκτυα εκπαιδεύονται με σκοπό να βελτιώσουν την προσέγγιση τους και να βελτιώσουν τα δεδομένα εξόδου τους.

Μειονεκτήματα

Παρακάτω αναφέρουμε μια σειρά από στοιχεία που αποτελούν τροχοπέδη στην εφαρμογή

τον Τεχνητών Νευρωνικών Δικτύων και στα οποία θα πρέπει να επικεντρωθεί η προσπάθεια για βελτίωση τους.

- Τα Τεχνητά Νευρωνικά Δίκτυα αποτελούν τεχνολογία μαύρου κουτιού (black box) με την έννοια ότι γνωρίζουμε τα δεδομένα εισόδου και εξόδου, αλλά όχι τις διεργασίες που γίνονται ενδιάμεσα.
- Επέκταση του παραπάνω προβλήματος των Τεχνικών Νευρωνικών Δικτύων είναι ότι δεν μας επιτρέπουν την ασφαλή εξαγωγή συμπερασμάτων για την λειτουργία των συστημάτων που προσομοιώνουν.

2.8 Εφαρμογές των Τεχνητών Νευρωνικών Δικτύων

Οι εφαρμογές των νευρωνικών δικτύων έχουν προκύψει τα τελευταία χρόνια και μερικές από αυτές βρίσκονται ήδη στην αγορά ως έτοιμα προϊόντα. Εκτιμάται ότι τα επόμενα χρόνια θα ακολουθήσει ένας μεγαλύτερος αριθμός εφαρμογών, αφού το πεδίο αυτό βρίσκεται ακόμη σε νηπιακή ηλικία. Οι εφαρμογές αυτές περιλαμβάνουν αναγνώριση προτύπων, υπολογισμό συναρτήσεων, βελτιστοποίηση, πρόβλεψη, αυτόματη ρύθμιση.

Μερικές από τις εφαρμογές των τεχνητών νευρωνικών δικτύων σε διάφορους τομείς της επιστήμης και της τεχνολογίας είναι οι ακόλουθες :

Αεροδιαστημική : Αυτόματος πιλότος, έλεγχος αεροσκάφους, ανίχνευση βλαβών

Οικονομία: Εκτίμηση τιμών μετοχών και συναλλάγματος, διαχείριση χαρτοφυλακίων, υπολογισμός κινδύνου σε δάνεια

Βιομηχανία: Αυτόματη ρύθμιση, έλεγχος γραμμής παραγωγής, έλεγχος ποιότητας

Ηλεκτρονική: Τεχνητή όραση, σύνθεση ήχου

Ρομποτική: Έλεγχος κίνησης, όραση

Τηλεπικοινωνίες: Συμπύεση δεδομένων, μετάφραση γλώσσας σε πραγματικό χρόνο

Ιατρική: Διάγνωση βασισμένη σε συμπτώματα, εξαγωγή συμπερασμάτων από ακτινογραφίες, ανάλυση καρδιογραφημάτων και εγκεφαλογραφημάτων

Βιολογία: Βοήθεια στην κατανόηση του εγκεφάλου

ΚΕΦΑΛΑΙΟ ΙΙΙ

ΔΙΚΤΥΑ ΑΚΤΙΝΙΚΗΣ ΣΥΝΑΡΤΗΣΗΣ ΒΑΣΗΣ RBF

Τα Νευρωνικά Δίκτυα Ακτινικών Συναρτήσεων Βάσης δεν προήλθαν από κάποιο βιολογικό μοντέλο αλλά από το πεδίο της προσέγγισης μη γραμμικών συναρτήσεων και προσαρμογής υπερεπιφανειών. Η ονομασία τους οφείλεται στις ακτινικές συναρτήσεις που χρησιμοποιούν οι νευρώνες του κρυφού επιπέδου ως συναρτήσεις ενεργοποίησης.

3.1 Ακτινικές Συναρτήσεις Βάσης στην Παρεμβολή

Θεωρούμε έναν n – διάστατο χώρο εισόδου x με έναν μονοδιάστατο χώρο εξόδου y . Έστω ότι τα δεδομένα που διαθέτουμε περιέχουν N διανύσματα εισόδου \mathbf{x}^n μαζί με τα αντίστοιχα διανύσματα εξόδου \mathbf{y}^n . Ας υποθέσουμε ότι ο στόχος είναι να βρούμε μια συνάρτηση $h(x)$ τέτοια ώστε,

$$h(\mathbf{x}^n) = \mathbf{y}^n \quad n = 1, 2, \dots, N \quad (3.1)$$

Η προσέγγιση της χρήση των ακτινικών συναρτήσεων βάσης στην παρεμβολή (Powell, 1987) εισάγει μια σειρά από N συναρτήσεις βάσης, μια για κάθε δεδομένο, οι οποίες λαμβάνουν τη μορφή $g(\|\mathbf{x} - \mathbf{x}^n\|)$, έτσι ώστε η αντίστοιχη έξοδος να είναι ένας γραμμικός συνδυασμός των συναρτήσεων βάσης, δηλαδή

$$h(x) = \sum w_n g(\|\mathbf{x} - \mathbf{x}^n\|) \quad (3.2)$$

Σύμφωνα με την τεχνική της ακριβούς παρεμβολής του Powell, η παραπάνω εξίσωση μπορεί να γραφεί σε μορφή πίνακα, όπως φαίνεται παρακάτω,

$$G = \begin{pmatrix} \|x^1 - x_1\| & \cdots & \|x^1 - x_n\| \\ \vdots & \ddots & \vdots \\ \|x^{n'} - x_1\| & \cdots & \|x^{n'} - x_n\| \end{pmatrix}, \quad W = \begin{bmatrix} w_1 \\ \vdots \\ w_{n'} \end{bmatrix}, \quad Y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

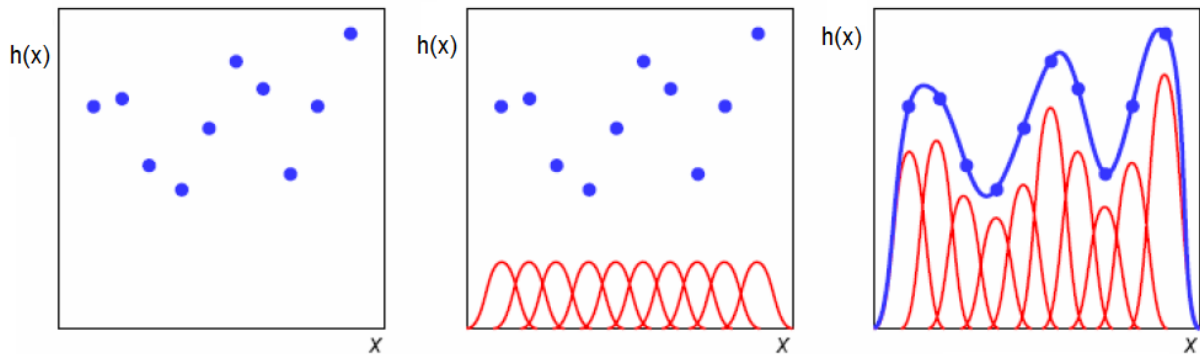
$$G \cdot W = Y \quad (3.3)$$

Εδώ είναι $n = n' = N$ και n' είναι ο αριθμός των κέντρων των ακτινικών συναρτήσεων

Υπό την προϋπόθεση ότι υπάρχει ο αντίστροφος πίνακας G^{-1} , μπορούμε να λύσουμε την εξίσωση (3.3) ως προς W

$$W = G^{-1} \cdot Y \quad (3.4)$$

Επίσης, έχει αποδειχθεί (Michille, 1986) ότι για μεγάλο αριθμό συναρτήσεων $g(\cdot)$ ο πίνακας G είναι αντιστρέψιμος. Έτσι, όταν τα βάρη των RBF ορίζονται από την Εξίσωση (3.4) η συνάρτηση $h(x)$ αντιπροσωπεύει μια συνεχή διαφορίσιμη επιφάνεια που περνά ακριβώς από κάθε σημείο, όπως φαίνεται στο Σχήμα 3.1.



Σχήμα 3.1 : Συνάρτηση παρεμβολής με N σημεία δεδομένων και N συναρτήσεις βάσης (μία ανά κουκκίδα).

Υπάρχουν διάφορες συναρτήσεις ακτινικής βάσης όπως η Γκαουσιανή, η spline RBF λεπτής πλάκας, η λογιστική η τετραγωνική RBF κλπ. Οι πιο συχνά χρησιμοποιούμενες είναι η συνάρτηση Gauss και η συνάρτηση Thin Plate Spline (TPS), οι οποίες αναπαριστώνται στο Σχήμα 3.2.

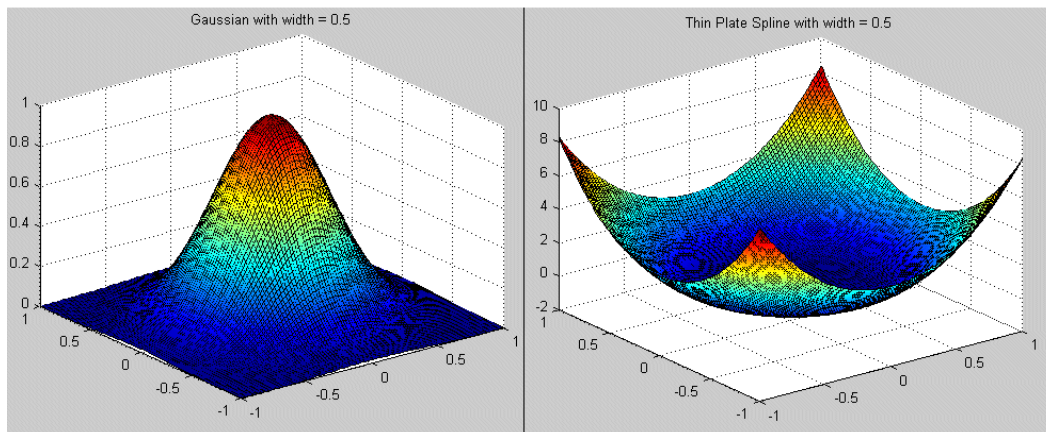
$$g^G_i(r) = \Phi(\|x - c_i\|) = \exp\left[\frac{-\|x - c_i\|^2}{2\sigma_i^2}\right] \quad (3.5)$$

$$g^{TSP}_i(r) = \Phi(\|x - c_i\|) = \|x - c_i\|^2 \cdot \log(\|x - c_i\|) \quad (3.6)$$

Όπου σ είναι η παράμετρος διακύμανσης για μια δεδομένη διάσταση του x και ελέγχει την ομαλότητα της συνάρτησης βάσης.

Για τη συνάρτηση Gauss ισχύει: $\lim_{|r| \rightarrow \infty} g^G(x) = 0$

Αντίθετα για την TPS ισχύει: $\lim_{|r| \rightarrow \infty} g^{TPS}(x) = \infty$



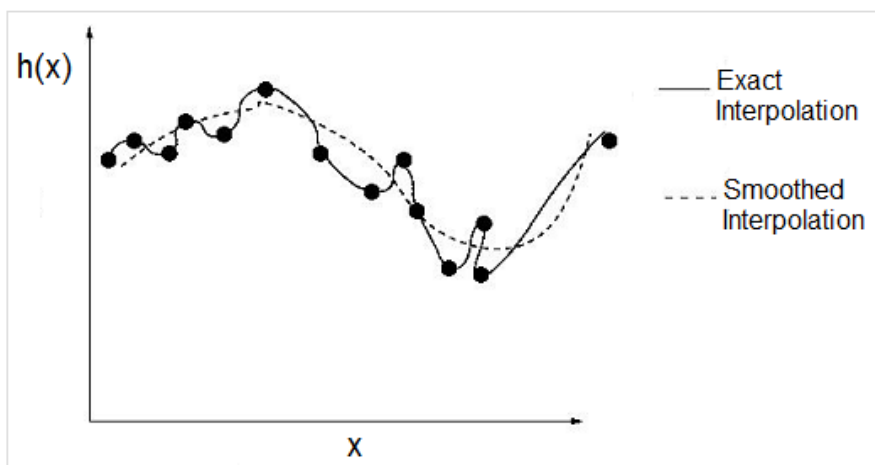
Σχήμα 3.2 : Γραφικές παραστάσεις των καμπυλών Gauss και Thin Plate Spline για συνάρτηση δύο μεταβλητών

3.2 Νευρωνικά Δίκτυα Ακτινικών Συναρτήσεων Βάσης RBF

3.2.1 Γενική Επισκόπηση των Δικτύων RBF

Οι απεικονίσεις των ακτινικών συναρτήσεων βάσης που συζητήθηκαν παραπάνω παρέχουν μια συνάρτηση παρεμβολής που περνά ακριβώς από κάθε σημείο. Ωστόσο, μια τέτοια ακριβής παρεμβολή για δεδομένα με θόρυβο οδηγεί τελικά στην εκμάθηση του θορύβου, κάτι που δεν είναι επιθυμητό. Ένα άλλο σοβαρό εμπόδιο στην τεχνική της ακριβούς παρεμβολής είναι ότι ο αριθμός των συναρτήσεων βάσης πρέπει να είναι ίσος με τον αριθμό των προτύπων στο σετ δεδομένων, το οποίο προσθέτει ιδιαίτερο υπολογιστικό φόρτο στην περίπτωση μεγάλου αριθμού δεδομένων.

Οι Broomhead και Lowe σε δημοσιεύσεις τους πρότειναν μια σειρά από πιθανές τροποποιήσεις στις τεχνικές της ακριβούς παρεμβολής, δίνοντας έτσι το έναυσμα για την ανάπτυξη των νευρωνικών δικτύων ακτινικής συνάρτησης βάσης. Σε αντίθεση με την ακριβή παρεμβολή που περιγράφηκε παραπάνω, η τεχνική αυτή παρέχει μια ομαλή συνάρτηση παρεμβολής (Σχήμα 3.2), στην οποία ο αριθμός των συναρτήσεων βάσης (νευρώνες κρυφού επιπέδου) καθορίζεται από την πολυπλοκότητα του προβλήματος και όχι από τον αριθμό των δεδομένων.



Σχήμα 3.2: Ακριβής και ομαλή συνάρτηση παρεμβολής

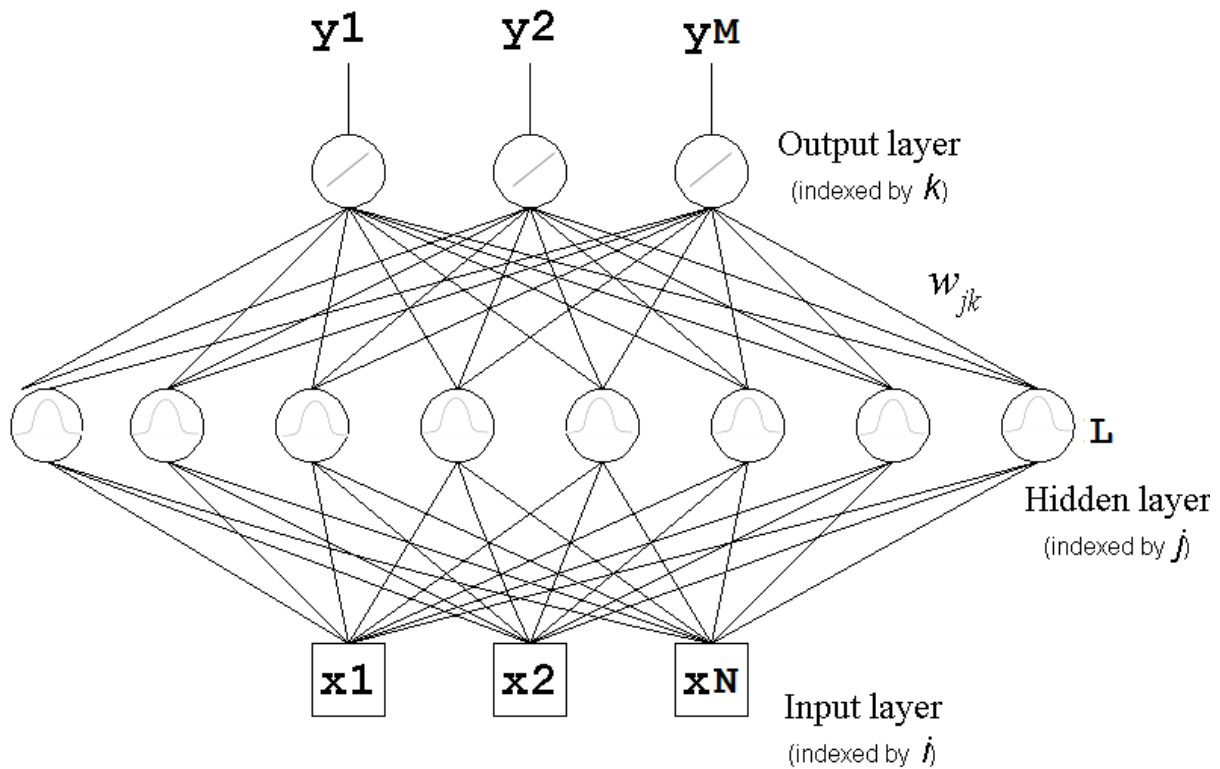
Οι τροποποιήσεις που προτάθηκαν μπορούν να συνοψισθούν ως εξής :

1. Ο αριθμός L των συναρτήσεων βάσης δεν χρειάζεται να είναι ίσος με τον αριθμό N των δεδομένων, και είναι συνήθως πολύ μικρότερος από το N .

2. Για τα κέντρα των συναρτήσεων βάσης δεν υπάρχει πλέον ο περιορισμός να δίνονται από τα διανύσματα εισόδου. Αντίθετα, ο προσδιορισμός των κατάλληλων κέντρων γίνεται μέρος της διαδικασίας εκπαίδευσης.

3. Αντί να έχουμε μια κοινή παράμετρο πλάτους σ , κάθε συνάρτηση βάσης έχει το δικό της πλάτος σ_i η τιμή των οποίων προσδιορίζεται, επίσης κατά τη διάρκεια της εκπαίδευσης.

4. Οι παράμετροι πόλωσης περιλαμβάνονται στο γραμμικό άθροισμα. Αντισταθμίζουν τη διαφορά μεταξύ της μέσης τιμής των συναρτήσεων ενεργοποίησης και την αντίστοιχη μέση τιμή των εξόδων.



Σχήμα 3.3 : Νευρωνικό Δίκτυο RBF

Θεωρούμε το παραπάνω δίκτυο (Σχήμα 3.3) με N νευρώνες εισόδου, L κρυφούς νευρώνες και M νευρώνες στο επίπεδο εξόδου. Κάθε ένας από τους L νευρώνες στον κρυφό επίπεδο εφαρμόζει μια συνάρτηση ενεργοποίησης $g(\cdot)$ η οποία είναι συνάρτηση της Ευκλείδειας απόστασης μεταξύ της εισόδου και ενός N -διάστατου πρότυπου διανύσματος. Κάθε κρυφός νευρώνας με το δικό του πρότυπο διάνυσμα ως παράμετρο δίνει μια έξοδο η οποία στη συνέχεια σταθμίζεται και περνάει στο επίπεδο εξόδου. Οι έξοδοι του δικτύου αποτελούνται από τα αθροίσματα των σταθμισμένων νευρώνων του κρυφού επιπέδου :

$$y_k = \sum_{j=1}^L W_{jk} \cdot g(\|x - c_j\|) \quad (3.7)$$

με την πόλωση να συμπεριλαμβάνεται στο κύριο διάνυσμα των βαρών

Η έντονη διαφορά των RBF νευρωνικών δικτύων με τα MLP έγκειται στον ρόλο των βαρών του πρώτου και του δεύτερου στρώματος και οδηγεί σε μια διαδικασία δύο σταδίων εκπαίδευσης του RBF-ΝΔ. Στο πρώτο στάδιο, μόνο τα δεδομένα εισόδου χρησιμοποιούνται για τον προσδιορισμό των βαρών του πρώτου επιπέδου (παράμετροι των συναρτήσεων βάσης c_j, σ_j) με εκπαίδευση χωρίς επίβλεψη. Τα βάρη του πρώτου επιπέδου στη συνέχεια διατηρούνται σταθερά, ενώ τα βάρη του δεύτερου επιπέδου βελτιστοποιούνται κατά τη δεύτερη φάση της εκπαίδευσης. Ωστόσο, πρέπει να σημειωθεί ότι μετά τις τροποποιήσεις που παρουσιάστηκαν παραπάνω, δεδομένου ότι

χρησιμοποιούνται λιγότερες συναρτήσεις βάσης δεδομένων από τον αριθμό των δεδομένων δεν θα είναι πλέον δυνατό να βρεθεί ένα σύνολο τιμών των βαρών για τις οποίες η συνάρτηση θα περνάει ακριβώς από όλα τα σημεία των δεδομένων.

Για συγκεκριμένα δεδομένα εκπαίδευσης, αρχικά επιλέγονται οι πρότυπες παράμετροι των συναρτήσεων βάσης τα (κέντρα c_j) με το κατάλληλο πλάτος έτσι ώστε να καλύπτουν ολόκληρο τον N -διάστατο χώρο εισόδου που καλύπτεται από τα πρότυπα εισόδου x_n . Αυτό μπορεί να γίνει με παρατήρηση του εύρους και της πυκνότητας των προτύπων εισόδου. Ωστόσο εάν τα εν λόγω κέντρα των συναρτήσεων βάσης χρησιμοποιηθούν για να συμπληρώσουν την υπο-περιοχή όπως αναφέρεται παραπάνω, τότε ο αριθμός των κέντρων θα είναι μια εκθετική συνάρτηση του N . Αυτό το πρόβλημα με τα RBF-ND, γνωστό ως «κατάρρα της μεγάλης διάστασης», είναι πιο έντονο στην περίπτωση που οι κόμβοι εισόδου είναι σε μεγάλο βαθμό ασυσχέτιστοι. Αυτό αυξάνει επίσης τον υπολογιστικό χρόνο και τον αριθμό των προτύπων που απαιτούνται για την εκπαίδευση. Για τη λύση του προβλήματος έχουν προταθεί αλγόριθμοι, όπως ο k-means clustering (Moody & Darken, 1988, Leonard & Kramer, 1991), για να επιλέξουν τις βέλτιστες παραμέτρους του πρώτου επιπέδου βάση της πυκνότητας των δεδομένων.

Μετά τον καθορισμό των βαρών του πρώτου επιπέδου (πρότυπα διανύσματα με c_j κέντρα, και πλάτη σ_j), ακολουθεί ο υπολογισμός των συναπτικών βαρών. Ο υπολογισμός γίνεται με εκπαίδευση με επίβλεψη και η μέθοδος που χρησιμοποιείται συνήθως είναι αυτή των ελαχίστων τετραγώνων, αφού πρόκειται για ένα γραμμικό πρόβλημα. Σε μορφή πινάκων οι τιμές των συναπτικών βαρών προκύπτουν από τη σχέση:

$$\mathbf{w} = (\mathbf{F}^T \mathbf{F})^{-1} \mathbf{F}^T \mathbf{y}_{sp} \quad (3.8)$$

όπου με $\mathbf{F} \in \mathfrak{R}^{K \times L^n}$ συμβολίζεται ο πίνακας εξόδου από της κρυφής στοιβάδας (πίνακας ενεργοποίησης) όπως αυτός ορίστηκε στη Σχέση 4.11, και με $\mathbf{y}_{sp} \in \mathfrak{R}^{K \times 1}$ συμβολίζεται το διάνυσμα των επιθυμητών τιμών εξόδου.

3.2.2 Βελτίωση του Αλγορίθμου Εκπαίδευσης

3.2.2.1 Μειονεκτήματα Συμβατικών Μεθόδων Εκπαίδευσης

Το κρίσιμο σημείο στο σχεδιασμό ενός νευρωνικού δικτύου RBF είναι η επιλογή των κέντρων. Το πρόβλημα αυτό, δηλαδή διαμορφώνεται ως ένα μη πολυωνυμικό πρόβλημα βελτιστοποίησης (non-polynomial, NP) (Garrey & Johnson, 1979). Αυτό σημαίνει ότι δεν έχει βρεθεί αλγόριθμος ο οποίος να εντοπίζει τη βέλτιστη λύση σε πολυωνυμικό χρόνο. Για τη λύση του

προβλήματος έχουν προταθεί στη βιβλιογραφία πολλές προσεγγιστικές μέθοδοι, όπως η k-means clustering (Moody & Darken, 1988, Leonard & Kramer, 1991), τα ορθογώνια ελάχιστα τετράγωνα ή OLS (Chen, Cowan & Grant, 1991) και άλλες.

Οι τεχνικές αυτές, αν και θεωρούνται ως οι πιο δημοφιλείς μέθοδοι εκπαίδευσης των νευρωνικών δικτύων ακτινικής συνάρτησης βάσης, παρουσιάζουν κάποια μειονεκτήματα τα οποία δυσκολεύουν τη χρήση τους σε ορισμένες εφαρμογές. Για παράδειγμα, η πιο κλασική από τις μεθόδους, η k-means clustering, παρουσιάζει τα εξής μειονεκτήματα:

- Καταρχήν ο αριθμός των νευρώνων της κρυφής στοιβάδας, πρέπει να προκαθοριστεί από τον χρήστη. Αυτό σημαίνει ότι πρέπει να ακολουθηθεί μία τακτική δοκιμής και σφάλματος ώστε να επιλεγεί ο βέλτιστος αριθμός των νευρώνων της κρυφής στοιβάδας. Η τεχνική αυτή αυξάνει τον υπολογιστικό χρόνο της διαδικασίας εκπαίδευσης.
- Κατά την επιλογή των κέντρων των κρυφών νευρώνων, απαιτούνται πολλά περάσματα από τα δεδομένα εκπαίδευσης. Η επαναλαμβανόμενη αυτή διαδικασία αυξάνει τον υπολογιστικό φόρτο, ιδιαίτερα αν η βάση των δεδομένων είναι μεγάλη.
- Τέλος, ο αλγόριθμος k-means βασίζεται στην τυχαία αρχική επιλογή των κέντρων. Αυτό σημαίνει ότι για τα ίδια δεδομένα, διαφορετικά τρεξίματα του αλγορίθμου εκπαίδευσης θα μας δώσουν διαφορετικά κέντρα.

Τα παραπάνω μειονεκτήματα καθιστούν ιδιαίτερα δύσκολη τη χρήση αυτών των τεχνικών εκπαίδευσης σε διάφορες εφαρμογές των νευρωνικών δικτύων RBF.

3.2.2.2 Περιγραφή του Αλγορίθμου των Ασαφών Μέσων (fuzzy means)

Μία μέθοδος εκπαίδευσης η οποία δεν χαρακτηρίζεται από τα προβλήματα που αναφέρθηκαν πιο πάνω είναι αλγόριθμος των ασαφών μέσων (Sagimveis και συνεργάτες, 2002). Πιο συγκεκριμένα, ο αλγόριθμος αυτός με βάση μία ασαφή διαμέριση του χώρου των μεταβλητών εισόδου επιλέγει τόσο τη δομή (τον αριθμό των νευρώνων), όσο και τις παραμέτρους (τα κέντρα των νευρώνων) σε ένα μόνο βήμα, κάνοντας μόνο ένα πέρασμα από τα δεδομένα εκπαίδευσης. Ακόμα, για μία συγκεκριμένη ασαφή διαμέριση του χώρου των μεταβλητών εισόδου, ο αλγόριθμος παράγει πάντα το ίδιο σύνολο κέντρων, αφού δεν εμπλέκεται τυχαία επιλογή κέντρων.

Η θεωρία της ασαφούς λογικής (fuzzy logic) χρησιμοποιείται για να περιγράψει συστήματα με μεγάλη αβεβαιότητα. Η βάση όλων των ασαφών συστημάτων είναι η διαμέριση του χώρου των μεταβλητών εισόδου σε ένα αριθμό ασαφών συνόλων. Στην προτεινόμενη μεθοδολογία ο χώρος κάθε μεταβλητής εισόδου x_i ($i = 1, 2, \dots, N$) διαμελίζεται σε c_i τριγωνικά ασαφή σύνολα $A_{i,n}^1, A_{i,n}^2, \dots, A_{i,n}^{c_i}$ με συνάρτηση συμμετοχής της μορφής:

$$\mu_A(x) = \begin{cases} 1 - \frac{|x-a|}{\delta a}, & x \in [a - \delta a, a + \delta a] \\ 0, & \text{διαφορετικα} \end{cases} \quad (3.9)$$

όπου a είναι το κεντρικό στοιχείο στο οποίο η τιμή συμμετοχής γίνεται ίση με τη μονάδα και δa είναι το αντίστοιχο πλάτος, το οποίο επιλέγεται έτσι ώστε οι δύο κορυφές του τριγώνου να βρίσκονται ακριβώς στα κέντρα των δύο γειτονικών ασαφών συνόλων. Προκύπτει ότι κάθε ασαφές σύνολο A μπορεί να περιγραφεί πλήρως από το αντίστοιχο κέντρο και πλάτος του:

$$A_n = \{a, \delta a\} \quad (3.10)$$

Ορίζουμε τη διαμέριση του χώρου κάθε μεταβλητής εισόδου σαν :

$$T_i = \{A_i^1, A_i^2, \dots, A_i^{c_i}\}, \quad 1 \leq i \leq N \quad (3.11)$$

Έπεται ότι η ασαφής διαμέλιση του συνολικού χώρου των μεταβλητών εισόδου \mathbf{X} μπορεί να οριστεί διαμερίζοντας τον σε C ασαφή σύνολα, $\mathbf{A}_n^1, \mathbf{A}_n^2, \dots, \mathbf{A}_n^c$, όπου

$$C = \prod_{i=1}^N c_i \quad (3.12)$$

Κάθε ασαφής υπόχωρος A_n^l $1 \leq l \leq C$ μπορεί να οριστεί από το συνδυασμό N συγκεκριμένων ασαφών συνόλων $A_{1,n}^l \in T_1, A_{2,n}^l \in T_2, \dots, A_{N,n}^l \in T_N$ και μπορεί να αναπαρασταθεί ως:

$$A_n^l = [A_1^l, A_2^l, \dots, A_N^l]^T \quad (3.13)$$

Ορίζουμε τα διανύσματα:

$$\mathbf{a}^l = [a_1^l, a_2^l, \dots, a_N^l]^T, \quad \delta \mathbf{a} = [\delta a_1, \delta a_2, \dots, \delta a_N]^T \quad (3.14)$$

Με τη βοήθεια της (4.17), ο υπόχωρος A^l μπορεί να γραφτεί ως:

$$\mathbf{A}_n^l = \{\mathbf{a}^l, \delta \mathbf{a}\} \quad (3.15)$$

Σημειώνεται ότι το διάνυσμα $\delta \mathbf{a}$ είναι το ίδιο για όλους τους ασαφείς υποχώρους. Η ιδέα της συνάρτησης συμμετοχής μπορεί να επεκταθεί και σε πολλαπλές διαστάσεις εισάγοντας την έννοια της πολυδιάστατης συνάρτησης συμμετοχής $\mu_{A^l}(\mathbf{x}(k))$ του διανύσματος εισόδου $\mathbf{x}(k)$ στο \mathbf{A}^l :

$$\mu_{A^l}(\mathbf{x}(k)) = \begin{cases} 1 - rd^l(\mathbf{x}(k)), & rd^l(\mathbf{x}(k)) \leq 1 \\ 0, & \text{διαφορετικα} \end{cases} \quad (3.16)$$

όπου με $rd^l(\mathbf{x}(k))$ συμβολίζεται η ευκλείδεια σχετική απόσταση μεταξύ του \mathbf{A}^l και του διανύσματος εισόδου $\mathbf{x}(k)$:

$$rd^l(\mathbf{x}(k)) = \frac{\left[\sum_{i=1}^N (a_i^l - x_i(k))^2 \right]^{1/2}}{\left[\sum_{i=1}^N (\delta a_i)^2 \right]^{1/2}} \quad (3.17)$$

Προφανώς, ο ασαφής υπόχωρος που περιγράφει καλύτερα ένα διάνυσμα εισόδου είναι αυτός με τη μικρότερη ευκλείδεια σχετική απόσταση, αφού αυτός ο υπόχωρος αναθέτει στο διάνυσμα αυτό το μεγαλύτερο βαθμό συμμετοχής. Για τον καθορισμό του υποχώρου που αντιστοιχεί σε κάθε διάνυσμα εισόδου, χρησιμοποιούμε την παρακάτω διαδικασία θεωρώντας ότι οι χώροι των μεταβλητών εισόδου έχουν διαμεριστεί σε ένα αριθμό ασαφών συνόλων.

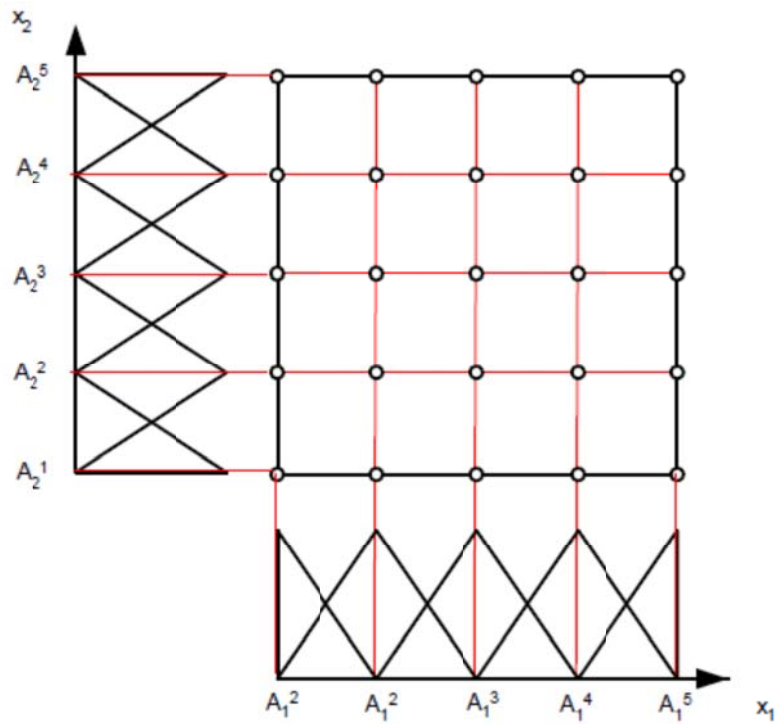
ΑΛΓΟΡΙΘΜΟΣ : Καθορισμός του κοντινότερου ασαφούς υποχώρου για ένα δεδομένο διάνυσμα εισόδου

Έστω ότι έχουμε ένα διάνυσμα μεταβλητών εισόδου $\mathbf{x} \in \mathfrak{R}^N$, και όλοι οι χώροι των μεταβλητών έχουν διαμερισθεί σε ασαφή σύνολα. Για παράδειγμα ας θεωρήσουμε το δισδιάστατο χώρο του Σχήματος 3.5 και ένα διάνυσμα εισόδου $\mathbf{x} = [x_1, x_2]$. Τα ασαφή σύνολα για τις μεταβλητές x_1 και x_2 είναι αντιστοίχως:

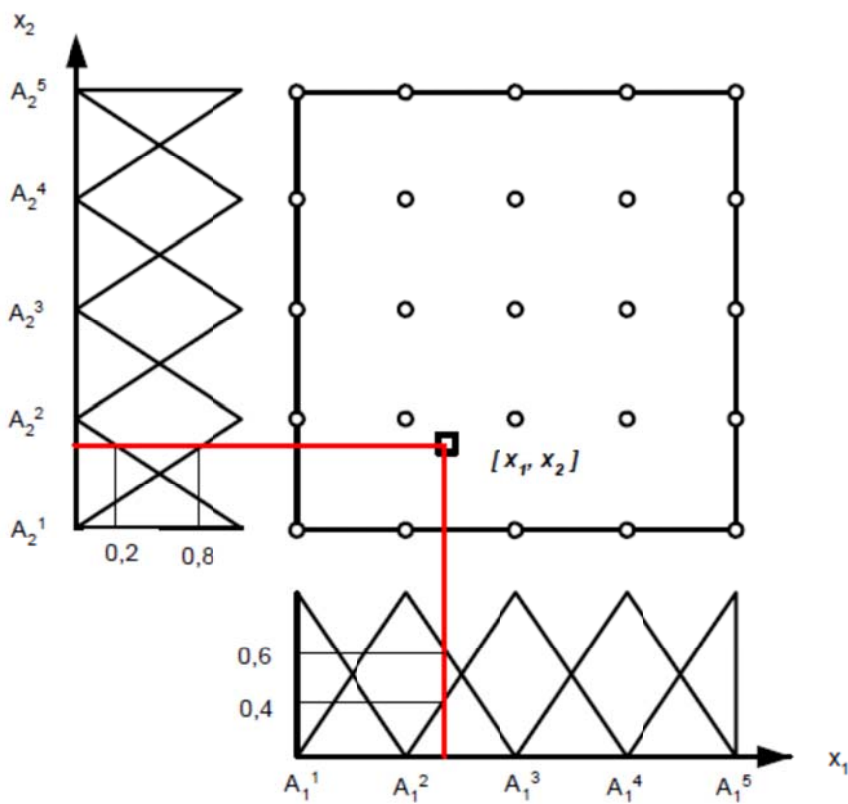
$$T_1 = \{A_{1,n}^1, A_{1,n}^2, A_{1,n}^3, A_{1,n}^4, A_{1,n}^5\}$$

$$T_2 = \{A_{1,n}^1, A_{2,n}^2, A_{3,n}^3, A_{4,n}^4, A_{5,n}^5\}$$

ΒΗΜΑ 1: Καθορίζονται οι βαθμοί συμμετοχής των x_1, x_2, \dots, x_N στα αντίστοιχα ασαφή σύνολα. Για το παράδειγμα έχουμε από το Σχήμα 3.6, ότι ο βαθμός συμμετοχής του x_1 είναι: 0.6 στο $A_{1,n}^2$, 0.4 στο $A_{1,n}^3$ και μηδενικός στα άλλα ασαφή σύνολα. Ομοίως, ο βαθμός συμμετοχής του x_2 είναι: 0.2 στο $A_{2,n}^1$, 0.8 στο $A_{2,n}^2$ και μηδέν στα άλλα ασαφή σύνολα.



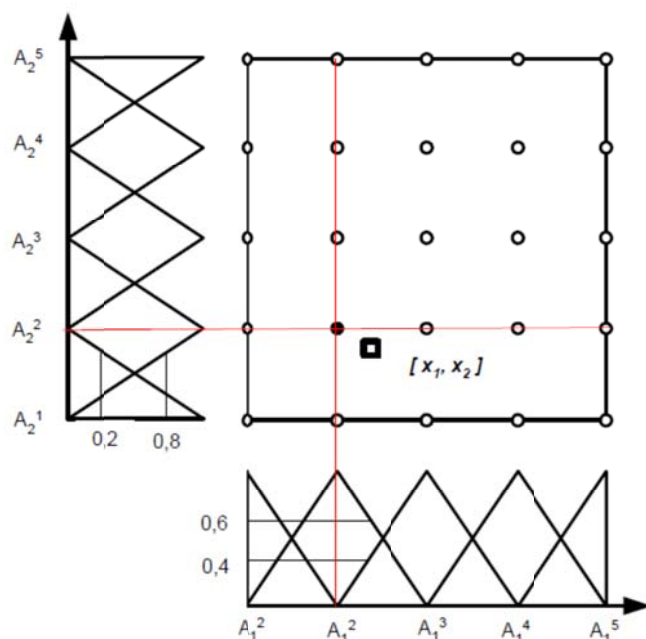
Σχήμα 3.5 : Ασαφής διαμερισμός του χώρου των μεταβλητών εισόδου σε 5 ασαφή σύνολα για ένα δισδιάστατο πρόβλημα



Σχήμα 3.6 : Καθορισμός βαθμού συμμετοχής του διανύσματος $x = [x_1, x_2]$ στο δισδιάστατο χώρο

ΒΗΜΑ 2: Αναθέτουμε στις μεταβλητές x_1, x_2, \dots, x_N τα ασαφή σύνολα με το μεγαλύτερο βαθμό συμμετοχής. Για παράδειγμα στο Σχήμα 4.13, το $A_{1,n}^2$ ανατίθεται στη μεταβλητή x_1 και το $A_{2,n}^2$ ανατίθεται στη μεταβλητή x_2 .

ΒΗΜΑ 3: Ο κοντινότερος ασαφής υπόχωρος $A_n^{j_0}$ για το δεδομένο διάνυσμα εισόδου \mathbf{x} δίνεται ως ο συνδυασμός των ασαφών συνόλων που επιλέχθηκαν στο βήμα 2. Για παράδειγμα στο Σχήμα 3.7 έχουμε: $A_n^{j_0} = [A_{1,n}^2, A_{2,n}^2] = \{[a_1^2, a_2^2], [\delta a_1, \delta a_2]\}$



Σχήμα 3.7 : Ανάθεση του κοντινότερου ασαφούς υπόχωρου σε κάθε μεταβλητή

Με βάση τον **αλγόριθμο 1**, αναπτύσσεται η μεθοδολογία εκπαίδευσης των ασαφών μέσω (fuzzy-means algorithm). Η μεθοδολογία προϋποθέτει την ύπαρξη K^n διανυσμάτων εισόδου $\mathbf{x}^T(k) = [x_1(k), x_2(k), \dots, x_N(k)]$, ($k = 1, 2, \dots, K^n$) και μία ασαφή διαμέλιση του χώρου κάθε μεταβλητής x_i ($i = 1, 2, \dots, N$) σε c_i συμμετρικά τριγωνικά ασαφή σύνολα.

Αλγόριθμος Fuzzy-Means

Πάρε το πρώτο διάνυσμα μεταβλητών εισόδου $x(1)$

Θέσε $L = 1$

Εφάρμοσε τον αλγόριθμο 1 στο διάνυσμα εισόδου $x(1)$:

{

FOR $i = 1$ TO N

$$A_{i,n}^1 = \{a_i^1, \delta a_i\} \leftarrow \max_{1 \leq j \leq c_i} [\mu_{A_i^j}(x_i(1))] \quad (3.18)$$

END

Δημιουργία του πρώτου ασαφούς υπόχωρου

$$\mathbf{A}_n^1 = \{\mathbf{a}^1, \delta \mathbf{a}\} = \{[a_1^1, a_2^1, \dots, a_N^1]^T, [\delta a_1, \delta a_2, \dots, \delta a_N]^T\} \quad (3.19)$$

FOR $k = 2$ TO K^n

$$rd^{l_0}(\mathbf{x}(k)) = \min_{1 \leq l \leq L} [rd^l(\mathbf{x}(k))] \quad (3.20)$$

IF $rd^l(\mathbf{x}(k)) > 1$ THEN

$L = L + 1$

Εφαρμογή του αλγορίθμου 1 στο διάνυσμα εισόδου $\mathbf{x}(k)$:

{

FOR $i = 1$ έως N

$$A_{i,n}^L = \{a_i^L, \delta a_i\} \leftarrow \max_{1 \leq j \leq c_i} [\mu_{A_i^j}(x_i(k))] \quad (3.21)$$

END

Παραγωγή του L^{0v} ασαφούς υπόχωρου

$$\mathbf{A}_n^L = \{\mathbf{a}^L, \delta \mathbf{a}\} = \{[a_1^L, a_2^L, \dots, a_N^L]^T, [\delta a_1, \delta a_2, \dots, \delta a_N]^T\} \quad (3.22)$$

}

ENDIF

END

Μετά το τέλος του αλγορίθμου σχηματίζεται η κρυφή στοιβάδα με ανάθεση κάθε παραγόμενου ασαφούς υπόχωρου σε ένα κρυφό νευρώνα. Το κέντρο κάθε κρυφού νευρώνα είναι ίσο με το κέντρο του αντίστοιχου υπόχωρου.

Ο αλγόριθμος ξεκινάει με το πρώτο δεδομένο $\mathbf{x}(1)$ και υπολογίζει τον πρώτο ασαφή υπόχωρο, κάνοντας χρήση του αλγορίθμου 1: Για $i = 1, 2, \dots, N$, επιλέγει το ασαφές σύνολο που αναθέτει το μεγαλύτερο βαθμό συμμετοχής στο $\mathbf{x}_i(1)$. Αυτή η διαδικασία επιλογής αντιπροσωπεύεται από τη Σχέση 3.16. Ο συνδυασμός των N επιλεγμένων ασαφών συνόλων σχηματίζει τον πρώτο ασαφή υπόχωρο, ο οποίος περιγράφει καλύτερα το διάνυσμα εισόδου $\mathbf{x}(1)$ (Εξίσωση 3.17).

Στη συνέχεια, για κάθε ένα από τα εναπομείναντα διανύσματα εισόδου $\mathbf{x}(k)$ ($k = 1, 2, \dots, K^n$), ο αλγόριθμος καθορίζει τον κοντινότερο υπόχωρο, από τους L που έχουν διαμορφωθεί μέχρι την

εξέταση του εκάστοτε διανύσματος. Με βάση την Εξίσωση 3.18 υπολογίζεται η σχετική ευκλείδεια απόσταση μεταξύ του εκάστοτε διανύσματος εισόδου και κάθε υπόχωρου, $rd^l(\mathbf{x}(k))$. Εάν η ελάχιστη από αυτές τις αποστάσεις $rd^{l_0}(\mathbf{x}(k))$ είναι μεγαλύτερη της μονάδας, τότε το διάνυσμα $\mathbf{x}(k)$ δεν περιγράφεται επαρκώς από κανένα από τους ήδη υπάρχοντες ασαφείς υποχώρους και ο αλγόριθμος παράγει ένα νέο ασαφή υπόχωρο. Εάν η ελάχιστη απόσταση είναι μικρότερη της μονάδας, ένας υπάρχων ασαφής υπόχωρος ανατίθεται σε αυτό το διάνυσμα.

Ουσιαστικά, η μέθοδος των ασαφών μέσων ορίζει ένα πολυδιάστατο πλέγμα στο χώρο εισόδου και επιλέγει κάποιους από τους κόμβους του πλέγματος ως κέντρα. Ωστόσο, δεν εξετάζει όλους τους κόμβους, αλλά μόνο αυτούς που βρίσκονται κοντά στα διανύσματα εισόδου. Για παράδειγμα, σε ένα πρόβλημα όπου υπάρχουν 15 μεταβλητές εισόδου και επιλέγεται ο χώρος κάθε μεταβλητής εισόδου να διαμεριστεί σε πέντε ασαφή σύνολα, δημιουργούνται $5^{15} \approx 3 \times 10^{10}$ ασαφείς υπόχωροι. Ωστόσο, η μέθοδος εξετάζει μόνο ένα πολύ μικρό ποσοστό από αυτούς.

Στη συνέχεια, αφού έχει υπολογιστεί ο αριθμός των κρυφών νευρώνων καθώς και οι θέσεις των κέντρων, ακολουθεί ο υπολογισμός των συναπτικών βαρών. Όπως αναφέρθηκε παραπάνω, ο υπολογισμός γίνεται με εκπαίδευση με επίβλεψη και συνήθως χρησιμοποιείται η μέθοδος των ελαχίστων τετραγώνων, αφού πρόκειται για ένα γραμμικό πρόβλημα.

3.2.2.3 Πλεονεκτήματα του Αλγορίθμου των Ασαφών Μέσων

Τα πλεονεκτήματα που παρουσιάζει η μέθοδος εκπαίδευσης των ασαφών μέσων έναντι των συμβατικών μεθόδων εκπαίδευσης νευρωνικών δικτύων RBF, την καθιστά πολύτιμη για πολλές εφαρμογές. Τα πλεονεκτήματα αυτά είναι:

I. Μειώνεται δραστικά ο υπολογιστικός χρόνος για την εκπαίδευση του δικτύου, αφού ο αλγόριθμος χρειάζεται μόνο ένα πέρασμα από τα δεδομένα εκπαίδευσης. Αντίθετα, οι κλασικές μέθοδοι απαιτούν πολλαπλά περάσματα μέχρι να συγκλίνουν.

II. Η μέθοδος έχει επαναληψιμότητα, αφού δεν εμπλέκεται αρχική τυχαία επιλογή κέντρων όπως συμβαίνει στην k-means clustering.

III. Δεν απαιτείται εκ των προτέρων καθορισμός του αριθμού των κόμβων της κρυφής στοιβάδας, αφού για δεδομένο αριθμό ασαφών συνόλων ο αλγόριθμος υπολογίζει μόνος του τη διάσταση της κρυφής στοιβάδας.

ΚΕΦΑΛΑΙΟ IV

ΝΕΥΡΩΝΙΚΟΙ ΡΥΘΜΙΣΤΕΣ

4.1 Εισαγωγή

Οι συμβατικές μέθοδοι σχεδιασμού συστημάτων ρύθμισης περιλαμβάνουν την κατασκευή ενός μαθηματικού μοντέλου που αναπαριστά τη δυναμική του συστήματος και τη χρήση αναλυτικών τεχνικών για την παραγωγή ενός ρυθμιστικού κανόνα. Η πιο διαδεδομένη μέθοδος ανάλυσης και ελέγχου είναι η γραμμικοποίηση γύρω από κάποιο επιθυμητό σημείο και χρήση ανατροφοδότησης για τη βελτίωση της απόδοσης του συστήματος έναντι των αβεβαιοτήτων του μοντέλου και των εξωτερικών διαταραχών.

Όταν οι σταθεροί ρυθμιστές ανατροφοδότησης δεν είναι επαρκείς, τότε χρησιμοποιούνται οι προσαρμοστικοί ρυθμιστές. Ο σχεδιασμός ενός προσαρμοστικού ρυθμιστή περιλαμβάνει τον σχεδιασμό ρυθμιστών που πληρούν τις προδιαγραφές γύρω από ένα σημείο λειτουργίας, όπου το γραμμικό μοντέλο είναι έγκυρο και τη χρήση ενός προγραμματιστή (scheduler). Ο προσαρμοστικός ρυθμιστής που προκύπτει με αυτόν τον τρόπο μπορεί να επιτύχει τους ρυθμιστικούς στόχους σε όλο το εύρος λειτουργίας.

Σε αρκετές όμως περιπτώσεις απαιτείται να αυξηθεί σημαντικά το εύρος λειτουργίας του συστήματος. Επίσης, χρειάζεται να αντιμετωπιστούν αποτελεσματικά στοιχεία αβεβαιότητας σε πολύπλοκα μοντέλα καθώς και απρόβλεπτες μεταβολές στο σύστημα, στο περιβάλλον και στους ρυθμιστικούς στόχους. Για την επίτευξη των παραπάνω στόχων, η χρήση ευφυών διαδικασιών λήψης αποφάσεων είναι σε θέση να διατηρήσει το επίπεδο απόδοσης σε υψηλά επίπεδα ακόμα και όταν συμβαίνουν δραστικές αλλαγές στις συνθήκες λειτουργίας.

Τα νευρωνικά δίκτυα, με την ικανότητά τους να μαθαίνουν και την έμφυτη μαζική παράλληλη επεξεργασία τους εισαγάγουν πολυάριθμες ευκαιρίες για την ανάπτυξη ανώτερων δομών ρύθμισης για σύνθετα συστήματα. Η δυνατότητα των νευρωνικών δικτύων να προσεγγίζουν μη γραμμικές συναρτήσεις, επιτρέπει τον αποτελεσματικό σχεδιασμό μη γραμμικών ρυθμιστών (Hornik et al., 1989). Επίσης, η ικανότητα τους για παράλληλη επεξεργασία κάνει ελκυστική την χρήση τους σε πολυμεταβλητά συστήματα ρύθμισης. Στον παρακάτω πίνακα συνοψίζονται τα πλεονεκτικά χαρακτηριστικά των νευρωνικών δικτύων που τα καθιστούν εύχρηστα εργαλεία στην αυτόματη ρύθμιση

Πίνακας 4-1 : Χαρακτηριστικά νευρωνικών δικτύων που τα καθιστούν κατάλληλα για εφαρμογή στα συστήματα ρύθμισης

Χαρακτηριστικά νευρωνικών δικτύων
<ul style="list-style-type: none">• Ικανότητα να αντιπροσωπεύουν αυθαίρετες μη-γραμμικές σχέσεις• Δυνατότητα εκμάθησης αβέβαιων συστημάτων μέσω off-line και on-line προσαρμογή βαρών• Ευελιξία να χειρίζονται τις πληροφορίες εισόδου που μετατρέπονται σε εσωτερική αναπαράσταση επιτρέποντας σύντηξη δεδομένων, τόσο σε ποσοτικά όσο και σε ποιοτικά δεδομένα• Η παράλληλη κατανεμημένη αρχιτεκτονική επεξεργασίας επιτρέπει τη γρήγορη επεξεργασία για μεγάλης κλίμακας δυναμικά συστήματα

Τα νευρωνικά δίκτυα έχουν εφαρμοστεί στη χημική βιομηχανία στη μοντελοποίηση διεργασιών, στην αυτόματη ρύθμιση, σε επαγωγικές εκτιμήσεις και εποπτικές εφαρμογές. Η μοντελοποίηση διεργασιών περιλαμβάνει τη μη-γραμμική πρόβλεψη των ρυθμιζόμενων μεταβλητών ή/και τη μοντελοποίηση άγνωστων παραμέτρων του μοντέλου (π.χ., κινητική αντίδρασης) σε συνήθεις διαφορικές εξισώσεις σχηματίζοντας ημιεμπειρικές εξισώσεις. Στη ρύθμιση διεργασιών τα νευρωνικά δίκτυα έχουν χρησιμοποιηθεί ως προγραμματιστές (scheduler) των παραμέτρων PID ρυθμιστών, αντιστοιχίζοντας τον χώρο κατάστασης σε ένα συγκεκριμένο σύνολο παραμέτρων του ρυθμιστή (Narendra & Parthasarathy, 1990). Επίσης, τα νευρωνικά δίκτυα έχουν λειτουργήσει τα ίδια ως ρυθμιστικός νόμος, αντιστοιχίζοντας τον χώρο κατάστασης σε ρυθμιστικές ενέργειες

Σε εφαρμογές ρύθμισης τα νευρωνικά δίκτυα μπορούν να ενσωματωθούν στον βρόχο ελέγχου μέσω άμεσων ή έμμεσων μεθόδων ρύθμισης. Στην άμεση μέθοδο, ένα νευρωνικό δίκτυο εκπαιδεύεται με δεδομένα εισόδου-εξόδου για να παραστήσει την αντίστροφη δυναμική του μοντέλου και στη συνέχεια λειτουργεί σαν ρυθμιστής σε πρόσθια μετάδοση. Στην έμμεση μέθοδο, το νευρωνικό δίκτυο εκπαιδεύεται για να αναπαραστήσει την δυναμική του συστήματος. Για την τρέχουσα κατάσταση και την τρέχουσα ρυθμιστική δράση, το δίκτυο μαθαίνει να προβλέπει την επόμενη κατάσταση του συστήματος. Το μοντέλο αυτό της διεργασίας μπορεί να χρησιμοποιηθεί στη συνέχεια από έναν αλγόριθμο ρύθμισης για τον υπολογισμό της ρυθμιστικής δράσης (Psychogios & Ungar, 1991).

Η βιβλιογραφία σε θέματα αυτόματης ρύθμισης παρουσιάζει μια σειρά από μεθοδολογίες που στηρίζονται στα νευρωνικά δίκτυα. Ο Hunt et al. (1992) βασίστηκε σε αυτές τις δομές που έχουν άμεση εξάρτηση στο μοντέλο του συστήματος και τις κατέταξε σε ρύθμιση με επίβλεψη, άμεση αντίστροφη ρύθμιση, ρύθμιση μοντέλου αναφοράς, ρύθμιση εσωτερικού μοντέλου, ρύθμιση

προβλεπτικού μοντέλου και προγραμματισμός κερδών. Από αυτές, τα νευρωνικά δίκτυα έχουν βρει σημαντική εφαρμογή στη ρύθμιση προβλεπτικού μοντέλου και στην άμεση αντίστροφη ρύθμιση.

4.2 Ρύθμιση Προβλεπτικού Μοντέλου (Model Predictive Control)

Η πιο εκτεταμένη χρήση των νευρωνικών δικτύων σε εφαρμογές ρύθμισης έχει γίνει στην ρύθμιση προβλεπτικού μοντέλου (Model Predictive Control, MPC). Η μεθοδολογία MPC περιλαμβάνει την χρήση ενός μοντέλου για την πρόβλεψη των μεταβλητών εξόδου της διεργασίας, σε ένα πεπερασμένο μελλοντικό ορίζοντα (*ορίζοντα πρόβλεψης*). Το μοντέλο αυτό χρησιμοποιείται για την διαμόρφωση ενός προβλήματος βελτιστοποίησης που ελαχιστοποιεί μία κατάλληλα επιλεγμένη αντικειμενική συνάρτηση. Μεταβλητές απόφασης αυτού του προβλήματος είναι οι μελλοντικές τιμές των μεταβλητών εκ χειρισμού, σε ένα εν γένει μικρότερο μελλοντικό χρονικό ορίζοντα, τον *ορίζοντα ρύθμισης*. Εφόσον η βέλτιστη αλληλουχία μελλοντικών ρυθμιστικών κινήσεων προσδιοριστεί, μόνο η πρώτη τιμή εφαρμόζεται τελικά στο σύστημα. Το πρόβλημα βελτιστοποίησης διαμορφώνεται και επιλύεται εκ νέου την επόμενη χρονική στιγμή, αξιοποιώντας όλες τις νεώτερες πληροφορίες για την διεργασία.

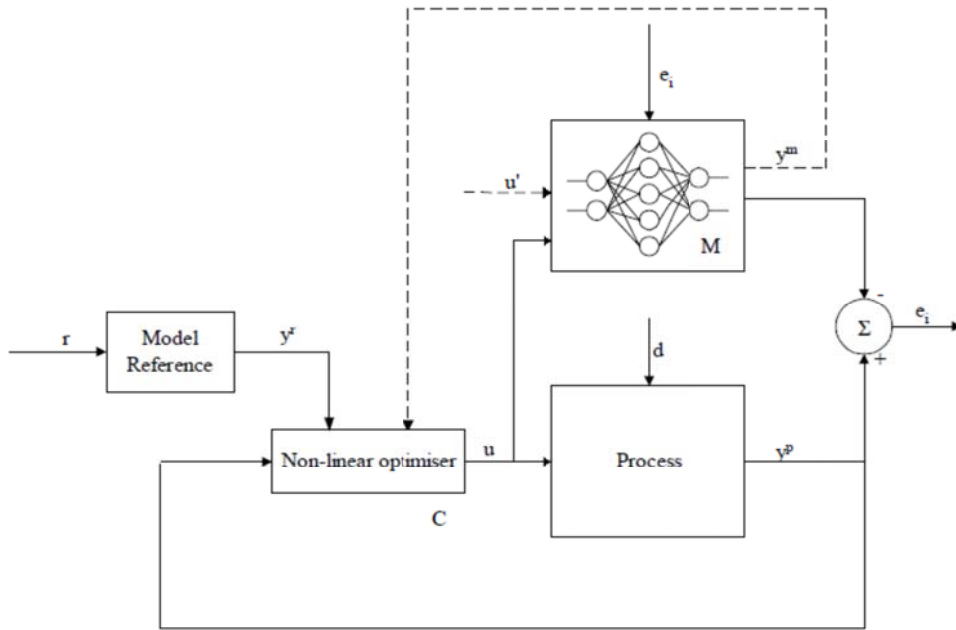
Η εφαρμογή της μεθοδολογίας MPC σε συνδυασμό με τα νευρωνικά δίκτυα, περιλαμβάνει τη χρήση ενός μοντέλου νευρωνικών δικτύων που προβλέπει τις μελλοντικές αποκρίσεις του συστήματος στο καθορισμένο χρονικό ορίζοντα. Οι προβλέψεις του νευρωνικού δικτύου στη συνέχεια περνάνε σε μια ρουτίνα βελτιστοποίησης, η οποία προσπαθεί να ελαχιστοποιήσει ένα κριτήριο απόδοσης στον υπολογισμό του κατάλληλου ρυθμιστικού σήματος. Το ρυθμιστικό σήμα μπορεί να επιλεγεί έτσι ώστε να ελαχιστοποιηθεί το τετραγωνικό κριτήριο απόδοσης

$$J = \sum_{j=N_1}^{N_2} [y^r(t+j) - y^m(t+j)]^2 + \sum_{j=1}^{N_2} \lambda_j [u'(t+j-1) - u'(t+j-2)]^2 \quad (4.1)$$

που υπόκειται στους περιορισμούς του δυναμικού μοντέλου.

Οι σταθερές N_1 και N_2 καθορίζουν τους ορίζοντες στους οποίους θεωρούνται το σφάλμα παρακολούθησης και ελέγχου, το λ είναι το ρυθμιστικό βάρος και οι υπόλοιπες παράμετροι φαίνονται στο Σχήμα 4.1.

Μια άλλη εναλλακτική λύση, είναι να εκπαιδευτεί και ένα δεύτερο νευρωνικό δίκτυο το οποίο θα μιμείται τη δράση της ρουτίνας βελτιστοποίησης. Κατά συνέπεια, ο νευρωνικός ρυθμιστής θα παράγει την ίδια έξοδο ρύθμισης για μια δεδομένη είσοδο του συστήματος (Hunt et al, 1992).



Σχήμα 4.1: Ρύθμιση προβλεπτικού μοντέλου με χρήση νευρωνικών δικτύων

Η εφαρμογή της MPC, λόγω της πολυπλοκότητας της, συνήθως δικαιολογείται μόνο για διεργασίες που δεν μπορούν να ρυθμιστούν επαρκώς με τη χρήση λιγότερο πολύπλοκων αλγορίθμων. Στις διεργασίες αυτές περιλαμβάνονται η παραγωγή χημικών προϊόντων υψηλής καθαρότητας, οι χημικοί αντιδραστήρες με πολλαπλά σημεία ισορροπίας, η υπερκρίσιμη εκχύλιση, η αζεοτροπική απόσταξη και οι διεργασίες που υπόκεινται σε πολλούς σχεδιαστικούς και λειτουργικούς περιορισμούς .

4.3 Ρύθμιση με Αντίστροφο Μοντέλο

Μια απλούστερη εναλλακτική λύση της MPC είναι οι μεθοδολογίες της αντίστροφης ρύθμισης (Inverse Control). Η αντίστροφη ρύθμιση στηρίζεται στη χρήση της αντίστροφης δυναμικής του συστήματος στη θέση του ρυθμιστή. Η βασική ιδέα είναι ότι εάν η διεργασία περιγράφεται από μια εξίσωση της μορφής:

$$y(t+1) = f(y(t), \dots, y(t-n+1), u(t), \dots, u(t-m)) \quad (4.2)$$

ένα νευρωνικό δίκτυο μπορεί να εκπαιδευτεί ως το αντίστροφο μοντέλο της διεργασίας:

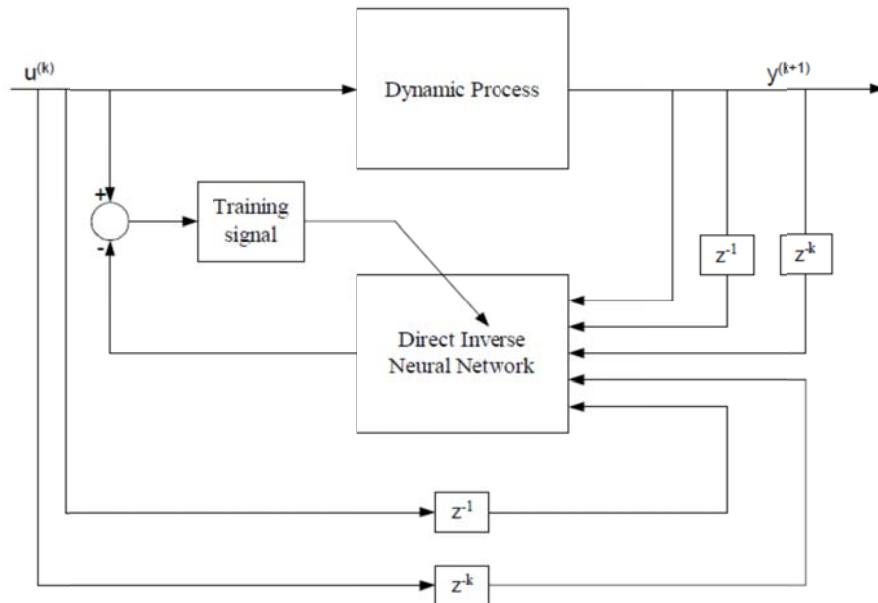
$$\hat{u}(t) = \hat{f}^{-1}(y(t+1), y(t), \dots, y(t-n+1), u(t-1), u(t-m)) \quad (4.3)$$

Το αντίστροφο μοντέλο στη συνέχεια χρησιμοποιείται ως ρυθμιστής για την διεργασία, τοποθετώντας την επιθυμητή τιμή $r(t+1)$, στη θέση της εξόδου $y(t+1)$. Υπάρχουν αρκετές αναφορές που χρησιμοποιούν αυτή την ιδέα, Psaltis et al. (1988), Hunt & Sbarbaro (1991).

Στην γενικευμένη εκπαίδευση (Σχήμα 4.2) το δίκτυο εκπαιδεύεται για να ελαχιστοποιήσει το ακόλουθο κριτήριο (το θ καθορίζει τα βάρη του δικτύου):

$$J_1(\theta) = \sum_{t=1}^N (u(t) - \hat{u}(t))^2 \quad (4.4)$$

Για την εκπαίδευση του νευρωνικού δικτύου προσομοιώνεται ένα πείραμα και μια σειρά από δεδομένα εισόδου-εξόδου αποθηκεύονται.



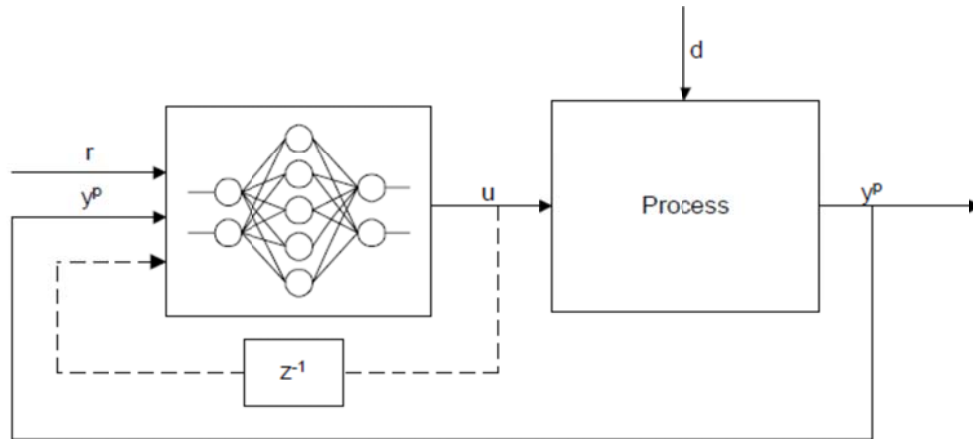
Σχήμα 4.2 : Γενικευμένη εκπαίδευση

Ο χώρος εισόδου, δηλαδή τα διανύσματα των προηγούμενων μετρήσεων των μεταβλητών του συστήματος, έχει μεγάλη επίδραση στην απόδοση του ρυθμιστή. Ωστόσο δεν υπάρχει κάποια συστηματική ή θεωρητική προσέγγιση για την σωστή επιλογή του και έτσι η επιλογή του γίνεται με δοκιμή και σφάλμα. Αν και προτάθηκαν ως κατευθυντήριες γραμμές για την επιλογή του κατάλληλου αριθμού των προηγούμενων εισόδων-εξόδων οι τεχνικές αξιολόγησης, η τελική επιλογή καθορίζεται από το σύστημα που μελετάται (ad hoc) (Hussain and Kershenbaum, 2000). Ο αριθμός των καθυστερημένων εισόδων στο νευρωνικό δίκτυο συχνά εκτιμάται, από την τάξη του συστήματος και νεκρό χρόνο της διεργασίας.

Όταν εκπαιδευτεί το αντίστροφο μοντέλο υπάρχουν διάφοροι τρόποι για να χρησιμοποιηθεί στην ρύθμιση. Εδώ θα αναφερθούμε στην άμεση αντίστροφη ρύθμιση και στην ρύθμιση εσωτερικού μοντέλου.

4.3.1 Άμεση Αντίστροφη Ρύθμιση (Direct Inverse Control)

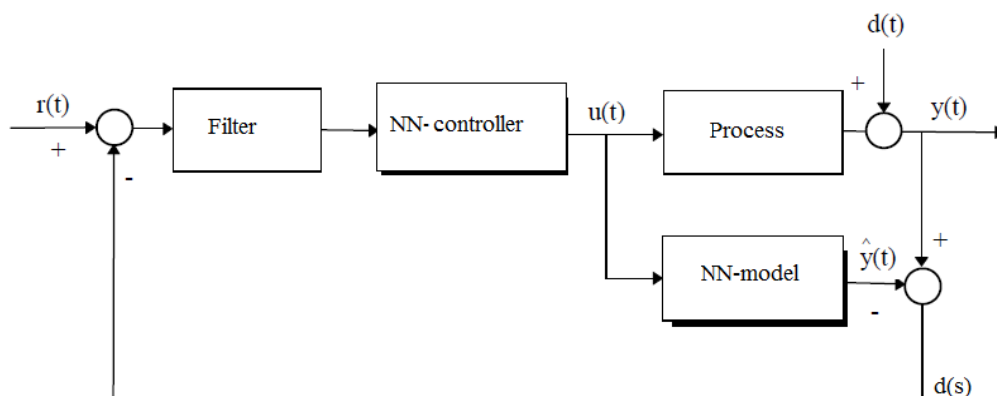
Η απλούστερη προσέγγιση για το σχεδιασμό ενός συστήματος ρύθμισης είναι ο σχεδιασμός ενός συστήματος που βασίζεται στην αντίστροφη δυναμική της διεργασίας, η οποία δρα ως ρυθμιστής παράγοντας την ρυθμιστική δράση (Σχήμα 4.3).



Σχήμα 4.3 : Άμεση αντίστροφη ρύθμιση

4.3.2. Ρύθμιση Εσωτερικού Μοντέλου (Internal Model Control, IMC)

Στο Σχήμα 4.4 φαίνεται το σύστημα IMC, όπου ένα νευρωνικό δίκτυο τοποθετείται παράλληλα με το πραγματικό σύστημα, ο ρυθμιστής είναι το αντίστροφο μοντέλο της διεργασίας και το φίλτρο κάνει το σύστημα σθεναρό στις αναντιστοιχίες διαδικασίας-μοντέλου. Με το IMC σύστημα στόχος είναι να εξαλειφθούν οι άγνωστες διαταραχές που επηρεάζουν το σύστημα. Για το σκοπό αυτό, προσδιορίζεται η διαφορά μεταξύ της εξόδου της διεργασίας και του μοντέλου, $d(s)$, και η διαφορά αυτή ανατροφοδοτείται στο σύστημα. Αν το ΤΝΔ μοντέλο είναι μια καλή προσέγγιση της διεργασίας, τότε η $d(s)$ είναι ίση με τις άγνωστες διαταραχές. Το σήμα $d(s)$ είναι η πληροφορία που λείπει από το ΤΝΔ μοντέλο και μπορεί να χρησιμοποιηθεί για τη βελτίωση του συστήματος ρύθμισης.



Σχήμα 4.4: Ρύθμιση εσωτερικού μοντέλου

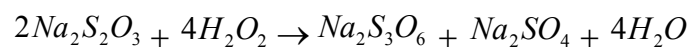
ΚΕΦΑΛΑΙΟ V

ΑΝΤΙΔΡΑΣΤΗΡΑΣ CSTR

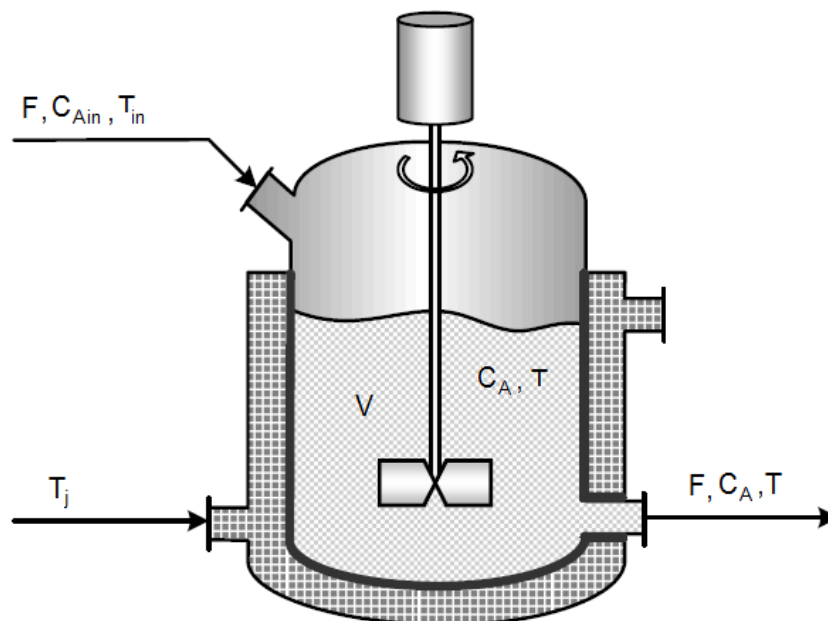
5.1 Εισαγωγή

Οι αντιδραστήρες είναι σημαντικές μονάδες στις χημικές βιομηχανίες. Οι χημικές αντιδράσεις είναι εξώθερμες ή ενδόθερμες και ως εκ τούτου απαιτείται να απομακρύνεται ή να προστίθεται ενέργεια στον αντιδραστήρα. Οι εξώθερμες αντιδράσεις είναι πιο ενδιαφέρουσες για μελέτη λόγω των δυνητικών προβλημάτων ασφαλείας (ταχεία αύξηση της θερμοκρασίας, που μερικές φορές ονομάζεται συμπεριφορά "ανάφλεξης"). Ενδιαφέρον έχουν επίσης οι περιπτώσεις πολλαπλών σημείων ισορροπίας (για την ίδια τιμή της μεταβλητής εισόδου μπορεί να υπάρχουν αρκετές δυνατές τιμές της μεταβλητής εξόδου σε μόνιμη κατάσταση).

Θεωρούμε έναν αντιδραστήρα συνεχούς έργου πλήρους ανάμιξης (CSTR) στον οποίον διεξάγεται η μη αντιστρεπτή εξώθερμη αντίδραση δεύτερης τάξης



Από την ανάλυση που ακολουθεί θα φανεί η πολύ ενδιαφέρουσα συμπεριφορά που μπορεί να προκύψει σε ένα τέτοιο απλό σύστημα.



Σχήμα 5.1: Αντιδραστήρας συνεχούς έργου πλήρους ανάδευσης (CSTR) με μανδύα ψύξης

Στο Σχήμα 5.1 φαίνεται ένα ρεύμα να τροφοδοτείται συνεχώς στον αντιδραστήρα και ένα

άλλο να απομακρύνεται. Δεδομένου ότι στον αντιδραστήρα γίνεται τέλεια ανάμειξη, το ρεύμα εξόδου έχει την ίδια συγκέντρωση και θερμοκρασία με την κύρια μάζα που περιέχεται στο εσωτερικό του αντιδραστήρα. Ένας μανδύας περιβάλλει τον αντιδραστήρα και σε αυτόν γίνεται η κυκλοφορία του ψυκτικού υγρού. Η ενέργεια μεταφέρεται από τα τοιχώματα του αντιδραστήρα στον μανδύα και μ' αυτόν τον τρόπο απομακρύνεται η θερμότητα που παράγεται από την αντίδραση.

5.2. ΟΙ ΕΞΙΣΩΣΕΙΣ ΜΟΝΤΕΛΟΠΟΙΗΣΗΣ

Για λόγους απλότητας θεωρούμε ότι η θερμοκρασία του ψυκτικού μπορεί άμεσα να μεταβληθεί, έτσι ώστε να μην απαιτείται το ενεργειακό ισοζύγιο γύρω από τον μανδύα. Κάνουμε επίσης τις ακόλουθες υποθέσεις:

- Τέλεια ανάδευση
- Σταθερός όγκος
- Σταθερές τιμές των παραμέτρων

5.2.1 Παράμετροι και Μεταβλητές

Οι παράμετροι και μεταβλητές που θα εμφανίζονται στις εξισώσεις μοντελοποίησης αναφέρονται στον παρακάτω πίνακα για λόγους ευκολίας.

Πίνακας 5-1 : Παράμετροι και μεταβλητές του αντιδραστήρα

UA	Ολικός συντελεστής μεταφοράς θερμότητας
C_A	Συγκέντρωση του $Na_2S_2O_3$ στον αντιδραστήρα
C_{Ain}	Συγκέντρωση του $Na_2S_2O_3$ στην τροφοδοσία
c_p	Ειδική θερμότητα
F	Ογκομετρική παροχή
k_0	Προεκθετικός παράγοντας
R	Σταθερά ιδανικών αερίων
r	Ρυθμός της αντίδρασης
T	Θερμοκρασία αντιδραστήρα

T_{in}	Θερμοκρασία τροφοδοσίας
T_j	Θερμοκρασία ψυκτικού
T_{ref}	Θερμοκρασία αναφοράς
V	Όγκος αντιδραστήρα
E	Ενέργεια ενεργοποίησης
ΔH	Ενέργεια αντίδρασης
ρ	Πυκνότητα

5.2.2 Γενικό Ισοζύγιο Μάζας

Ο ρυθμός συσσώρευσης του υλικού στο εσωτερικό του αντιδραστήρα είναι ίσος με την ποσότητα που εισέρχεται μείον την ποσότητα που εξέρχεται.

$$\frac{dV\rho}{dt} = F_{in}\rho_{in} - F_{out}\rho \quad (5.1)$$

Υποθέτοντας μια σταθερή ποσότητα του υλικού στο εσωτερικό του αντιδραστήρα ($\frac{dV\rho}{dt} = 0$),

βρίσκουμε ότι

$$F_{out}\rho = F_{in}\rho_{in} \quad (5.2)$$

Αν υποθέσουμε επίσης ότι η πυκνότητα παραμένει σταθερή, τότε

$$F_{out} = F_{in} = F$$

Και

$$\frac{dV}{dt} = 0$$

5.2.3 Ισοζύγιο Μάζας του Συστατικού Α

Το ισοζύγιο για το συστατικό $Na_2S_2O_3$ είναι

$$\frac{dVC_A}{dt} = FC_{Ain} - FC_A - rV \quad (5.3)$$

όπου r είναι ο ρυθμός της αντίδρασης ανά μονάδα όγκου

5.2.4 Ισοζύγιο Ενέργειας

Το ισοζύγιο ενέργειας είναι

$$\frac{d(V\rho c_p(T - T_{ref}))}{dt} = F\rho c_p(T_{in} - T_{ref}) - F\rho c_p(T - T_{ref}) + (-\Delta H)Vr - UA(T - T_j) \quad (5.4)$$

όπου T_{ref} αντιπροσωπεύει μια αυθαίρετη θερμοκρασία αναφοράς για την ενθαλπία.

5.2.5 Δυναμικές Εξισώσεις στη Μορφή του Χώρου Κατάστασης

Μπορούμε να γράψουμε τις Εξισώσεις (5.3) και (5.4) στην παρακάτω μορφή του χώρου κατάστασης (καθώς $dV / dt = 0$)

$$f_1(C_A, T) = \frac{dC_A}{dt} = \frac{F}{V}(C_{Ain} - C_A) - r \quad (5.5)$$

$$f_2(C_A, T) = \frac{dT}{dt} = \frac{F}{V}(T_{in} - T) + \left(\frac{-\Delta H}{\rho c_p} \right) r - \frac{UA}{V\rho c_p}(T - T_j) \quad (5.6)$$

όπου έχουμε υποθέσει ότι ο όγκος είναι σταθερός. Ο ρυθμός της αντίδρασης ανά μονάδα όγκου (έκφραση Arrhenius) είναι

$$r = 2k_o \exp\left(\frac{-E}{RT}\right)C_A^2 \quad (5.7)$$

και είναι δεύτερης τάξης για τη συγκεκριμένη αντίδραση.

5.3. Σημεία Ισορροπίας

Οι λύσεις του συστήματος για μόνιμη κατάσταση βρίσκονται για $dC_A / dt = 0$ και $dT / dt = 0$ δηλαδή

$$f_1(C_A, T) = 0 = \frac{F}{V}(C_{Ain} - C_A) - 2k_o \exp\left(\frac{-E}{RT}\right)C_A^2 \quad (5.8)$$

$$f_2(C_A, T) = 0 = \frac{F}{V}(T_{in} - T) + 2\left(\frac{-\Delta H}{\rho c_p}\right)k_o \exp\left(\frac{-E}{RT}\right)C_A^2 - \frac{UA}{V\rho c_p}(T - T_j) \quad (5.9)$$

Για την επίλυση των δύο εξισώσεων, όλες οι παράμετροι και οι μεταβλητές, με εξαίρεση δύο (C_A και T) πρέπει να είναι γνωστές. Αν δώσουμε αριθμητικές τιμές σε όλες τις παραμέτρους και τις μεταβλητές μπορούμε να χρησιμοποιήσουμε τη μέθοδο Newton για την εύρεση των σημείων ισορροπίας C_A και T .

Παράμετρος	Τιμή
V (L)	100
UA (J/s·K)	20000
k_0 (L/s ·mol)	6.85E+11
ρ (g/L)	1000
c_p (J/g ·K)	4.2
ΔH (J/mol)	596619
E (J/mol)	76534.704
R (J/mol ·K)	8.314
F (mol/s)	20
T_{in} (K)	275
T_{js} (K)	250

Λύσεις με χρήση της εντολής `fsolve` και της function `cstrsteady.m`

Το m-file με τις εξισώσεις για μόνιμη κατάσταση είναι το `cstrsteady.m` και υπάρχει στο Παράρτημα. Η εντολή για να τρέξει αυτό το αρχείο είναι

```
x = fsolve('cstrsteady',x0);
```

όπου x_0 είναι ένα διάνυσμα των αρχικών εκτιμήσεων και x είναι η λύση. Θα δούμε στη συνέχεια ότι διαφορετικές αρχικές εκτιμήσεις για τη συγκέντρωση και τη θερμοκρασία οδηγούν σε διαφορετικές λύσεις.

Κατά την επιλογή των αρχικών εκτιμήσεων σε μια αριθμητική μέθοδο, είναι σημαντικό να χρησιμοποιηθούν τιμές από το δυνατό εύρος των λύσεων. Για παράδειγμα, δεδομένου ότι η συγκέντρωση του A στην είσοδο είναι 1 mol/L και στην αντίδραση το A καταναλώνεται, η πιθανή περιοχή για τη συγκέντρωση του A είναι $0 < C_A < 1$. Επίσης, το κατώτερο όριο για τη θερμοκρασία είναι ανάμεσα στο $250-275 \text{ K}$ και θα προκύψει εάν δεν υπάρξει καμία αντίδραση, αφού οι θερμοκρασίες της τροφοδοσίας και του ψυκτικού είναι 275 K και 250 K αντίστοιχα. Επίσης, πρέπει να υπάρχει συσχέτιση μεταξύ της συγκέντρωσης και της θερμοκρασίας. Αν η συγκέντρωση του A είναι υψηλή, αυτό σημαίνει ότι η αντίδραση δεν έχει προχωρήσει οπότε μικρή ποσότητα ενέργειας έχει απελευθερωθεί και ως εκ τούτου η θερμοκρασία δεν θα είναι πολύ διαφορετική από τις θερμοκρασίες της τροφοδοσίας και του ψυκτικού

Εκτίμηση 1 : Υψηλή συγκέντρωση - χαμηλή θερμοκρασία. Εδώ θεωρούμε μια αρχική υπόθεση $C_A = 0.9$ και $T = 260$.

```
x = fsolve('cstrsteady',[0.9;260]);
x =
```

0.9872
271.6639

έτσι η λύση για την μόνιμη κατάσταση με την εκτίμηση 1 είναι $\begin{bmatrix} C_{As} \\ T_s \end{bmatrix} = \begin{bmatrix} 0.9872 \\ 271.6639 \end{bmatrix}$

Εκτίμηση 2 : Ενδιάμεση συγκέντρωση και θερμοκρασία

```
x = fsolve('cstrsyeady',[0.6;300])  
x =  
0.6661  
308.4989
```

οπότε η λύση για την μόνιμη κατάσταση με την εκτίμηση 2 είναι $\begin{bmatrix} C_{As} \\ T_s \end{bmatrix} = \begin{bmatrix} 0.6661 \\ 308.4989 \end{bmatrix}$

Εκτίμηση 3 : Χαμηλή συγκέντρωση και υψηλή θερμοκρασία.

```
x = fsolve('cstrsteady',[0.1;360])  
x =  
0.0755  
376.2697
```

άρα η λύση για την μόνιμη κατάσταση με την εκτίμηση 3 είναι $\begin{bmatrix} C_{As} \\ T_s \end{bmatrix} = \begin{bmatrix} 0.0755 \\ 376.2697 \end{bmatrix}$

Τα αποτελέσματα συνοψίζονται στον παρακάτω Πίνακα

Πίνακας 5-3 : Εκτιμήσεις και λύσεις με χρήση της fsolve

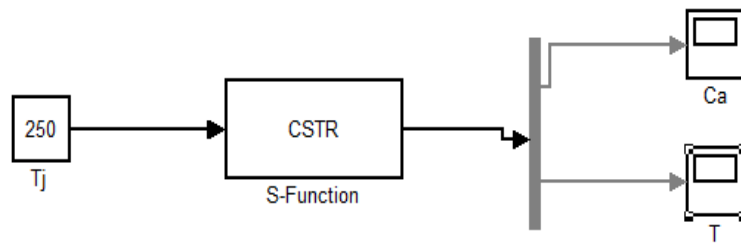
Εκτίμηση & Λύση	Εκτίμηση 1	Εκτίμηση 2	Εκτίμηση 3
$x_0(1)$ CA εκτίμηση	0.9	0.6	0.1
$x_0(2)$ T εκτίμηση	260	300	360
$x_0(1)$ CA λύση	0.987	0.666	0.076
$x_0(2)$ T λύση	271.664	308.499	376.270

Άλλες αρχικές εκτιμήσεις δεν οδηγούν σε άλλες λύσεις, έτσι συμπεραίνουμε ότι υπάρχουν τρεις δυνατές λύσεις.

5.4. Δυναμική Συμπεριφορά

Στην ενότητα αυτή θα μελετήσουμε τη δυναμική συμπεριφορά του αντιδραστήρα. Για τον

λόγο αυτό προσομοιώνουμε το σύστημα στο SIMULINKόπως φαίνεται στο Σχήμα 5.2

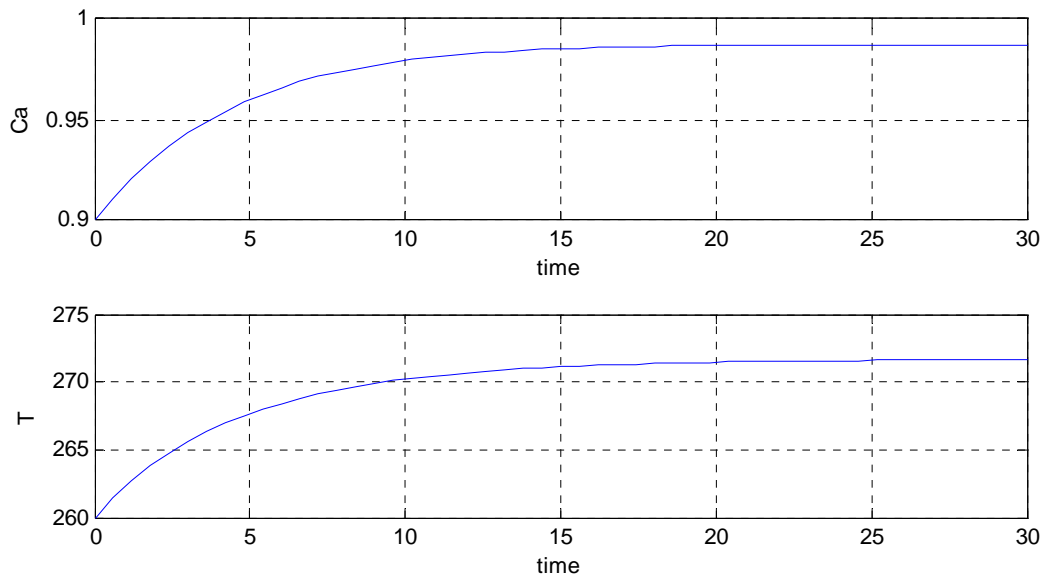


Σχήμα 5.2 : Προσομοίωση του αντιδραστήρα στο Simulink

Όπου το function CSTR.m περιέχει τις παραμέτρους του αντιδραστήρα και τις δυναμικές εξισώσεις που το περιγράφουν

Αρχική κατάσταση 1

Εδώ χρησιμοποιούμε ως αρχικές συνθήκες τις συνθήκες που είναι κοντά στο σημείο ισορροπίας με υψηλή συγκέντρωση A. Η αρχική κατάσταση είναι το διάνυσμα $[C_A, T] = [0.9, 260]$. Οι καμπύλες που απεικονίζονται στο διάγραμμα δείχνουν ότι οι μεταβλητές κατάστασης συγκλίνουν στο σημείο ισορροπίας υψηλής συγκέντρωσης (πάνω σημείο ισορροπίας).



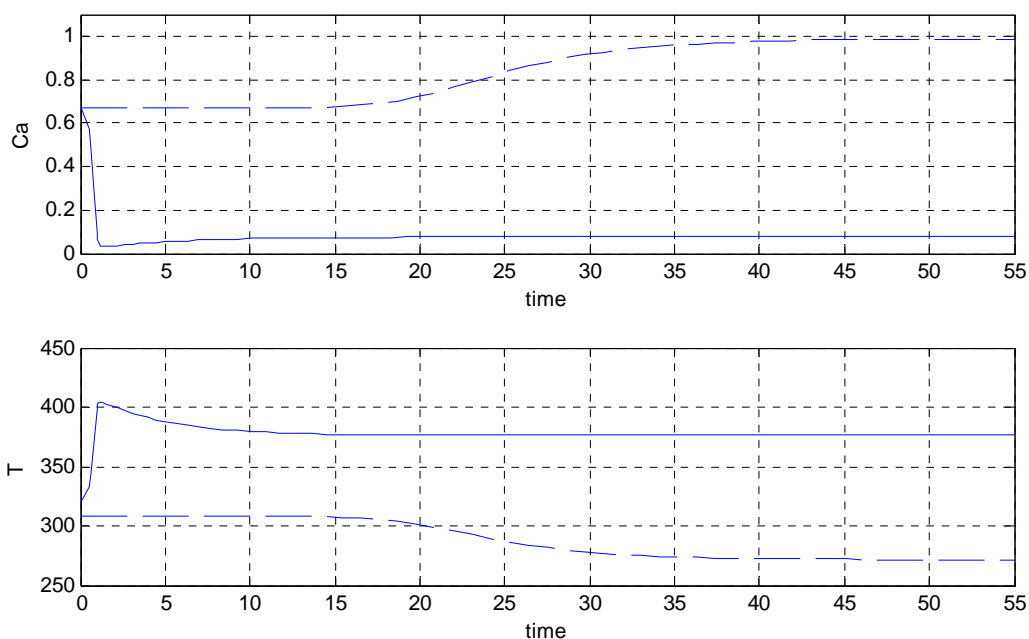
Σχήμα 5.3. Αποκρίσεις των μεταβλητών κατάστασης για αρχικές συνθήκες $x_0 = [0.9, 260]$

Αρχική κατάσταση 2

Εδώ χρησιμοποιούμε τις αρχικές συνθήκες που είναι κοντά στο σημείο ισορροπίας με ενδιάμεση συγκέντρωση. Το αρχικό διάνυσμα για την συνεχή καμπύλη στο Σχήμα 5.4 είναι

$[C_A, T] = [0.6661, 320]$, η οποία συγκλίνει στο σημείο ισορροπίας χαμηλής συγκέντρωσης (κάτω σημείο ισορροπίας). Το αρχικό διάνυσμα για την διακεκομμένη καμπύλη στο Σχήμα 5.4 είναι $[C_A, T] = [0.6661, 308.4989]$, η οποία συγκλίνει στο σημείο ισορροπίας υψηλής συγκέντρωσης.

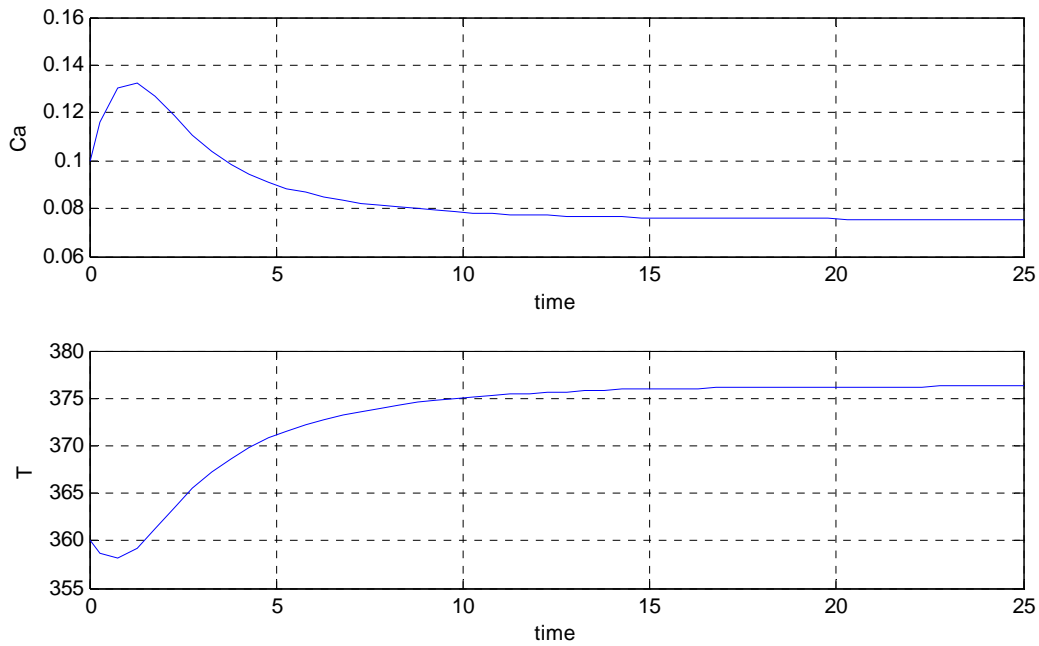
Αν πραγματοποιηθούν πολλές προσομοιώσεις με αρχικές συνθήκες κοντά στο ενδιάμεσο σημείο ισορροπίας, διαπιστώνουμε ότι η συγκέντρωση συγκλίνει είτε στην χαμηλή είτε στην υψηλή συγκέντρωση, αλλά ποτέ στο σημείο ισορροπίας με ενδιάμεση συγκέντρωση (μεσαίο σημείο ισορροπίας). Αυτό δείχνει ότι το σημείο ισορροπίας με ενδιάμεση συγκέντρωση είναι ασταθές. Αυτό θα φανεί πιο καθαρά στην ανάλυση ευστάθειας που θα ακολουθήσει.



Σχήμα 5.4 : Αποκρίσεις των μεταβλητών κατάστασης για αρχική κατάσταση $x_0 = [0.6661, 320]$ (συνεχής γραμμή) και $x_0 = [0.6661, 308.4989]$ (διακεκομμένη γραμμή).

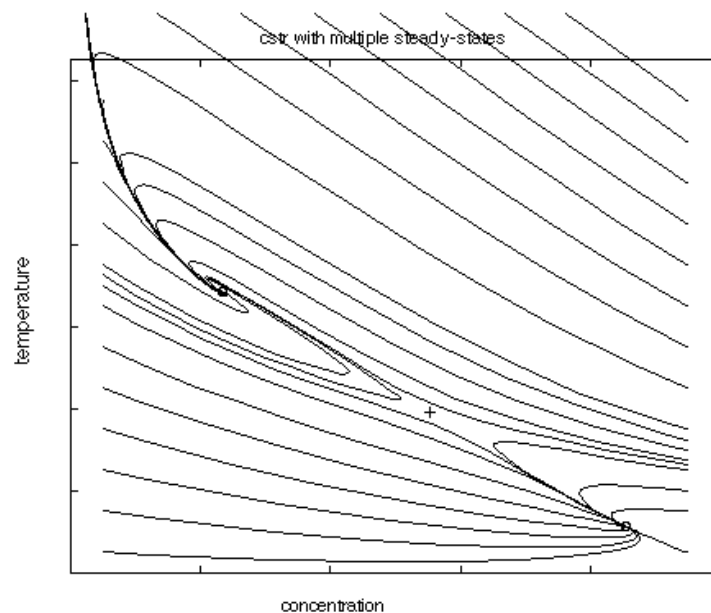
Αρχική Κατάσταση 3

Εδώ χρησιμοποιούμε τις αρχικές συνθήκες που είναι κοντά στο σημείο ισορροπίας με χαμηλή συγκέντρωση. Η αρχική κατάσταση είναι το διάνυσμα $[C_A, T] = [0.1, 260]$. Οι καμπύλες που απεικονίζονται στο Σχήμα 5.5 δείχνουν ότι οι μεταβλητές κατάστασης συγκλίνουν στο σημείο ισορροπίας με χαμηλή συγκέντρωση.



Σχήμα 5.5.: Αποκρίσεις των μεταβλητών κατάστασης για αρχική κατάσταση $x_0 = [0.1 \ 360]$

Σύμφωνα με τα προηγούμενα πραγματοποιήθηκαν προσομοιώσεις για διαφορετικές αρχικές συνθήκες, οι οποίες παρουσιάζονται σε διαφορετικά γραφήματα συναρτήσεως του χρόνου. Οι λύσεις των προσομοιώσεων για μεγάλο αριθμό αρχικών συνθηκών μπορούν να αναπαρασταθούν στο ίδιο διάγραμμα το οποίο ονομάζεται διάγραμμα επίπεδου φάσης (Σχήμα 5.6).



Σχήμα 5.6 : Παράδειγμα διαγράμματος επίπεδου φάσης για αντίδραση πρώτης τάξης, τα ευσταθή σημεία ισορροπίας σημειώνονται με 0 και το ασταθές σημείο με + .

Για να λειτουργήσει ο αντιδραστήρας στην ασταθή ενδιάμεση συγκέντρωση πρέπει να χρησιμοποιηθεί ρύθμιση με ανατροφοδότηση.

5.5. Γραμμικοποίηση Δυναμικών Εξισώσεων

Η ευστάθεια των μη γραμμικών εξισώσεων μπορεί να καθοριστεί και με το προσδιορισμό της μορφής του χώρου κατάστασης

$$\dot{x} = Ax + Bu \quad (5.10)$$

και τον καθορισμό των ιδιοτιμών του πίνακα A .

Οι μη γραμμικές δυναμικές Εξισώσεις (5.5) και (5.6) είναι

$$f_1(C_A, T) = \frac{dC_A}{dt} = \frac{F}{V} C_A - 2k C_A^2 + \frac{F}{V} C_{Ain} \quad (5.11)$$

$$f_2(C_A, T) = \frac{dT}{dt} = 2 \left(\frac{-\Delta H}{\rho c_p} \right) k C_A^2 - \frac{F}{V} T - \frac{UA}{V \rho c_p} T + \frac{UA}{V \rho c_p} T_j + \frac{F}{V} T_{in} \quad (5.12)$$

Οι μεταβλητές κατάστασης και οι μεταβλητές εισόδου μπορούν να οριστούν ως μεταβλητές απόκλισης

$$x = \begin{bmatrix} C_A - C_{As} \\ T - T_s \end{bmatrix}$$

$$u = \begin{bmatrix} T_j - T_{js} \\ C_{Ain} - C_{Ains} \\ T_{in} - T_{ins} \end{bmatrix}$$

5.5.1 Ανάλυση Ευστάθειας

Εκτελώντας την γραμμικοποίηση, έχουμε τα ακόλουθα στοιχεία για τον πίνακα A

$$A_{11} = \frac{\partial f_1}{\partial x_1} = \frac{\partial f_1}{\partial C_A} = -\frac{F}{V} - 4k_s C_{As}$$

$$A_{12} = \frac{\partial f_1}{\partial x_2} = \frac{\partial f_1}{\partial T} = -2C_{As}^2 k_s \quad (5.13)$$

$$A_{21} = \frac{\partial f_2}{\partial x_1} = \frac{\partial f_2}{\partial C_A} = 4 \frac{(-\Delta H)}{\rho c_p} k_s C_{As}$$

$$A_{22} = \frac{\partial f_2}{\partial x_2} = \frac{\partial f_2}{\partial T} = -\frac{F}{V} - \frac{UA}{V \rho c_p} + 2 \frac{(-\Delta H)}{\rho c_p} C_{As}^2 k_s$$

όπου έχουμε ορίσει τις ακόλουθες παραμέτρους για καλύτερη αναπαράσταση

$$k_s = k_o \exp\left(\frac{-E}{RT_s}\right)$$

$$k_s' = \frac{\partial k_s}{\partial T} = k_o \exp\left(\frac{-E}{RT_s}\right) \left(\frac{E}{RT_s^2}\right)$$

Από την ανάλυση που παρουσιάστηκε παραπάνω, ο πίνακας του χώρου κατάστασης A είναι

$$A = \begin{bmatrix} -\frac{F}{V} - 4k_s C_{As} & -2C_{As}^2 k_s' \\ 4\frac{(-\Delta H)}{\rho c_p} k_s C_{As} & -\frac{F}{V} - \frac{UA}{V\rho c_p} + 2\frac{(-\Delta H)}{\rho c_p} C_{As}^2 k_s' \end{bmatrix} \quad (5.14)$$

Τα χαρακτηριστικά ευστάθειας, καθορίζονται από τις ιδιοτιμές του πίνακα A, οι οποίες λαμβάνονται από την επίλυση της εξίσωσης $\det(\lambda I - A) = 0$ ή μέσω της εντολής $\text{eig}(A)$. Χρησιμοποιούμε το function `cstr_amat.m` για να βρούμε τους πίνακες A για τα σημεία ισορροπίας και τις ιδιοτιμές τους.

Χαμηλό σημείο ισορροπίας

Η συγκέντρωση και η θερμοκρασία είναι 0.9872 mol/l και 271.6639 K, αντίστοιχα.

```
>[amat,lambda] = cstr_amat(0.9872,271.6639)
amat =
1.0e+006 *
    -0.0000    -1.7429
     0.0000    -0.0000

lambda =
1.0e+003 *
    -0.0002 + 1.1344i
    -0.0002 - 1.1344i
```

Και οι δύο ιδιοτιμές έχουν αρνητικό πραγματικό μέρος, υποδηλώνοντας ότι το σημείο ισορροπίας είναι ευσταθές. Το αποτέλεσμα αυτό είναι σύμφωνο με τις παρατηρήσεις στο Σχήμα 5.3

Μεσαίο σημείο ισορροπίας

Η συγκέντρωση και η θερμοκρασία είναι 0.6661 και 308.4989 αντίστοιχα

```
»[amat,lambda] = cstr_amat(0.6661,308.4989)
amat =
1.0e+007 *
    -0.0000    -5.8496
     0.0000     0.0000

lambda =
1.0e+004 *
    0.0000 + 4.0815i
    0.0000 - 4.0815i
```

Οι δυο ιδιοτιμές έχουν θετικό πραγματικό μέρος, δείχνοντας ότι το σημείο ισορροπίας είναι ασταθές. Αυτό είναι σύμφωνο με τις απαντήσεις που παρουσιάζονται στο Σχήμα 5.4.

Υψηλό σημείο ισορροπίας

```
»[amat,lambda] = cstr_amat(0.0755,376.2697)
amat =
1.0e+008 *
    -0.0000    -2.4130
     0.0000     0.0000

lambda =
1.0e+005 *
    -0.0000 + 4.1001i
    -0.0000 - 4.1001i
```

Το πραγματικό μέρος της κάθε ιδιοτιμής είναι αρνητικό, υποδηλώνοντας ότι το σημείο είναι ευσταθές. Και πάλι, αυτό είναι συμβατό με τις απαντήσεις στο Σχήμα 5.5

5.6. Κατανόηση των Πολλαπλών Σημείων Ισορροπίας

Ο στόχος αυτής της ενότητας είναι να καθοριστεί πώς προκύπτουν τα πολλαπλά σημεία ισορροπίας και για λόγους απλοστευσης θα γίνει ανάλυση για αντίδραση πρώτης τάξης.

5.6.1 Καμπύλες Παραγωγής και Απομάκρυνσης Θερμότητας

Στην ενότητα 5.3 χρησιμοποιήσαμε αριθμητική μέθοδο για την εύρεση των σημείων ισορροπίας, με την επίλυση 2 εξισώσεων με 2 αγνώστους. Οι εξισώσεις αυτές είναι εύκολο να μειωθούν σε μια ενιαία εξίσωση με έναν άγνωστο.

Επίλυση ως προς τη συγκέντρωση

Το ισοζύγιο μάζας για μόνιμες συνθήκες είναι ($dC_A / dt = 0$)

$$\frac{F}{V}(C_{Ains} - C_{As}) = k_o \exp\left(\frac{-E}{RT_s}\right) C_{As} \quad (5.15)$$

Μπορούμε να αναδιατάξουμε αυτή την εξίσωση για να βρούμε τη συγκέντρωση στη μόνιμη κατάσταση για κάθε δεδομένη θερμοκρασία του αντιδραστήρα T_s ,

$$C_{As} = \frac{\frac{F}{V} C_{Ains}}{\frac{F}{V} = k_o \exp\left(\frac{-E}{RT_s}\right)} \quad (5.16)$$

Επίλυση ως προς την Θερμοκρασία

Το ισοζύγιο ενέργειας για μόνιμες συνθήκες ($dT / dt = 0$) είναι

$$\frac{F}{V}(T_s - T_{ins}) + \frac{UA}{V\rho c_p}(T_s - T_{js}) = \left(\frac{-\Delta H}{\rho c_p}\right) k_o \exp\left(\frac{-E}{RT_s}\right) C_{As} \quad (5.17)$$

Οι όροι στην Εξίσωση (5.17) σχετίζονται με την ενέργεια που παράγεται και απομακρύνεται. Αν πολλαπλασιάσουμε την Εξίσωση (5.17), με $V\rho c_p$ διαπιστώνουμε ότι

$$F\rho c_p(T_s - T_{ins}) + UA(T_s - T_{js}) = -\Delta H V k_o \exp\left(\frac{-E}{RT_s}\right) C_{As} \quad (5.18)$$

$$Q_{rem} = Q_{gen}$$

Η ενέργεια που απομακρύνεται με τη ροή θερμότητας είναι ίση με τη θερμότητα που παράγεται από την αντίδραση. Η εξίσωση της Q_{rem} :

$$Q_{rem} = [-UAT_{js} - F\rho c_p T_{ins}] + [UA + F\rho c_p] T_s \quad (5.19)$$

είναι μια γραμμή, όπου η ανεξάρτητη μεταβλητή είναι η θερμοκρασία του αντιδραστήρα T_s . Η κλίση της γραμμής είναι $[UA + F\rho c_p]$ και ο σταθερός της όρος είναι $[-UAT_{js} - F\rho c_p T_{ins}]$. Οι αλλαγές στη θερμοκρασία του ψυκτικού ή της τροφοδοσίας μετατοπίζουν τον σταθερό όρο, αλλά όχι την κλίση. Αλλαγές στο UA ή το F επηρεάζουν την κλίση αλλά και τον σταθερό όρο.

Ο όρος Q_{gen} γίνεται:

$$Q_{gen} = (-\Delta H)Vk_o \exp\left(\frac{-E}{RT_s}\right)C_{As} \quad (5.20)$$

Αντικαθιστώντας την (5.16) στην (5.20), διαπιστώνουμε ότι

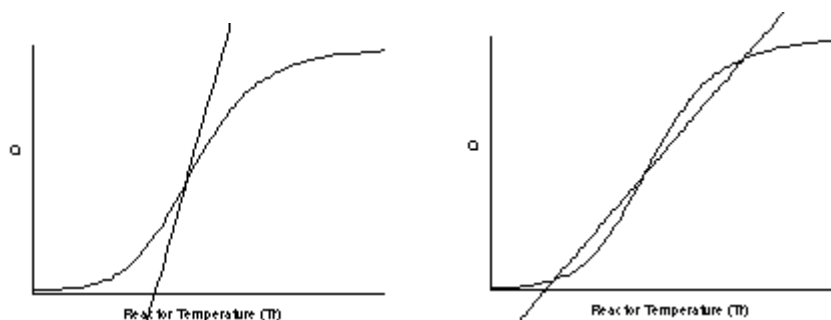
$$Q_{gen} = -\Delta HV \frac{k_o \exp\left(\frac{-\Delta E}{RT_s}\right)\left(\frac{F}{V}C_{Afs}\right)}{\frac{F}{V} + k_o \exp\left(\frac{-\Delta E}{RT_s}\right)} \quad (5.21)$$

Η Εξίσωση (5.21) έχει ένα χαρακτηριστικό S σχήμα όταν αναπαριστάται η Q_{gen} ως συνάρτηση της θερμοκρασίας του αντιδραστήρα. Από την Εξίσωση (5.18) βλέπουμε ότι τα σημεία ισορροπίας αντιστοιχούν στα σημεία τομής των καμπυλών Q_{rem} και Q_{gen} .

5.6.2 Επίδραση των Παραμέτρων Σχεδιασμού

Στο Σχήμα 5.7 φαίνονται τα διάφορα πιθανά σημεία τομής των καμπυλών παραγωγής και απομάκρυνσης της θερμότητας. Αν η κλίση της καμπύλης απομάκρυνσης της θερμότητας είναι μεγαλύτερη από τη μέγιστη κλίση της καμπύλης παραγωγής θερμότητας, υπάρχει μόνο ένα πιθανό σημείο τομής (Σχήμα 5.7α). Καθώς η θερμοκρασία του ψυκτικού ή της τροφοδοσίας αλλάζει, η καμπύλη απομάκρυνσης θερμότητας μετατοπίζεται αριστερά ή δεξιά, έτσι ώστε το σημείο τομής μπορεί να είναι σε υψηλή ή χαμηλή θερμοκρασία, ανάλογα με την τιμή της θερμοκρασίας του ψυκτικού ή της τροφοδοσίας.

Φαίνεται ότι όσο η κλίση της καμπύλης απομάκρυνσης θερμότητας είναι μικρότερη από τη μέγιστη κλίση της καμπύλης παραγωγής θερμότητας, θα υπάρχει πάντα η δυνατότητα ύπαρξης τριών σημείων τομής (Σχήμα 5.7β) με την κατάλληλη ρύθμιση της θερμοκρασίας του ψυκτικού ή της τροφοδοσίας.

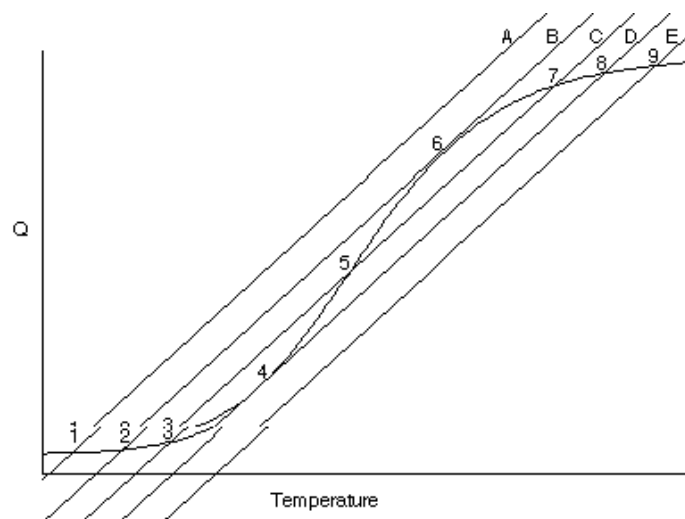


a. b.

Σχήμα 5.7 : Πιθανά σημεία τομής των καμπυλών παραγωγής και απομάκρυνσης της θερμότητας.

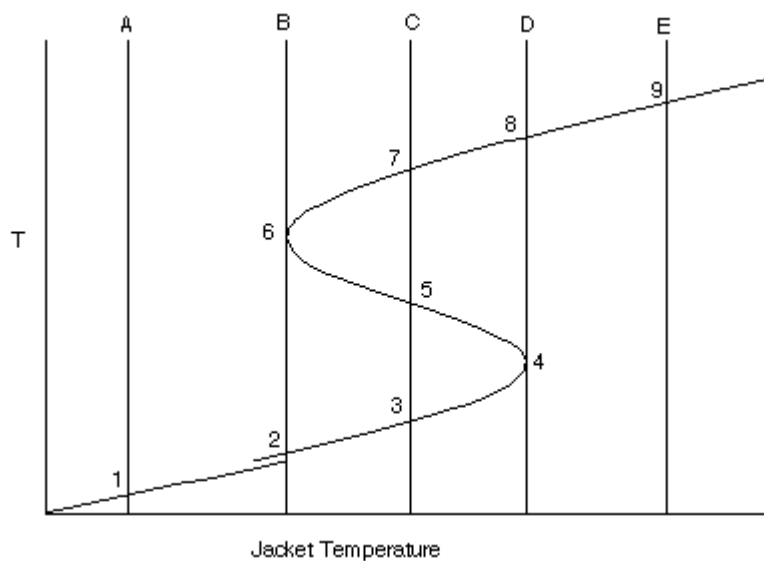
5.6.3 Πολλαπλά Σημεία Ισορροπίας

Στο Σχήμα 5.8 υπάρχουν αρκετές δυνατές γραμμικές καμπύλες απομάκρυνσης θερμότητας με την καμπύλη παραγωγή θερμότητας σχήματος S . Η καμπύλη A τέμνει την καμπύλη παραγωγής θερμότητας σε χαμηλή θερμοκρασία, η καμπύλη B τέμνει σε χαμηλή θερμοκρασία και είναι εφαπτομένη σε υψηλή θερμοκρασία, η καμπύλη C τέμνει σε χαμηλή, μέση και υψηλή θερμοκρασία, η καμπύλη D είναι εφαπτομένη σε χαμηλή θερμοκρασία και τέμνει σε υψηλή θερμοκρασία και η καμπύλη E έχει μόνο ένα σημείο τομής σε υψηλή θερμοκρασία. Οι καμπύλες A, B, C, D και E έχουν τις ίδιες παραμέτρους εκτός από τη θερμοκρασία του ψυκτικού η οποία αυξάνεται καθώς κινούμαστε από την καμπύλη A έως E (από την Εξίσωση (5.19) φαίνεται ότι με την αλλαγή της θερμοκρασίας του ψυκτικού αλλάζει ο σταθερός όρος, αλλά όχι η κλίση της καμπύλης απομάκρυνσης θερμότητας).



Σχήμα 5.8 : Η παραγόμενη ενέργεια και η ενέργεια που απομακρύνεται ως συνάρτηση της θερμοκρασίας του αντιδραστήρα.

Μπορούμε να χρησιμοποιήσουμε το Σχήμα 5.8 για την κατασκευή του διαγράμματος μόνιμης κατάστασης εισόδου -εξόδου όπως φαίνεται στο Σχήμα 5.9, όπου η θερμοκρασία του ψυκτικού είναι η είσοδος και η θερμοκρασία του αντιδραστήρα είναι η έξοδος. Στο Σχήμα 5.9 παρουσιάζεται το φαινόμενο της *υστέρησης*.



Σχήμα 5.9 : Θερμοκρασία αντιδραστήρα ως συνάρτηση της θερμοκρασίας του ψυκτικού.

Ο όρος *υστέρηση* χρησιμοποιείται για να δείξει ότι η συμπεριφορά είναι διαφορετική, ανάλογα με την κατεύθυνση που κινούνται οι είσοδοι. Για παράδειγμα, αν ξεκινήσουμε με χαμηλή θερμοκρασία του ψυκτικού ο αντιδραστήρας λειτουργεί σε χαμηλή θερμοκρασία (σημείο 1). Καθώς αυξάνεται η θερμοκρασία του ψυκτικού αυξάνεται και η θερμοκρασία του αντιδραστήρα (σημεία 2 και 3) μέχρι το χαμηλό όριο της θερμοκρασίας (σημείο 4). Αν η θερμοκρασία του ψυκτικού αυξηθεί περαιτέρω, η θερμοκρασία του αντιδραστήρα κάνει ένα άλμα (*αναφλέγεται*) σε υψηλή θερμοκρασία (σημείο 8). Περαιτέρω αύξηση της θερμοκρασίας του ψυκτικού έχει ως αποτέλεσμα μικρή αύξηση της θερμοκρασίας του αντιδραστήρα.

Όταν κάποιος ξεκινάει από υψηλή θερμοκρασία ψυκτικού (σημείο 9) υπάρχει μόνο μια υψηλή θερμοκρασία του αντιδραστήρα, η οποία μειώνεται καθώς μειώνεται η θερμοκρασία του ψυκτικού (σημεία 8 και 7). Καθώς κινούμαστε ελαφρώς χαμηλότερα από το υψηλό όριο θερμοκρασίας (σημείο 6), η θερμοκρασία του αντιδραστήρα πέφτει απότομα (συμπεριφορά επίσης γνωστή ως *εξαφάνιση* (extinction)) σε χαμηλή θερμοκρασία (σημείο 2). Περαιτέρω μείωση της θερμοκρασίας του ψυκτικού οδηγεί σε μικρή μείωση της θερμοκρασίας του αντιδραστήρα.

Η συμπεριφορά υστέρησης που συζητήθηκε παραπάνω είναι επίσης γνωστή ως η συμπεριφορά *ανάφλεξης-εξαφάνισης*. Η περιοχή μεταξύ των σημείων 4 και 6 φαίνεται να είναι *ασταθής*, επειδή ο αντιδραστήρας δε φαίνεται να λειτουργεί στην περιοχή αυτή.

5.6.4 Εξέταση της Ευστάθειας (Ανάλυση Μόνιμης Κατάστασης)

Θεωρούμε το σύστημα που περιγράφεται από τη καμπύλη απομάκρυνσης ενέργειας C στο Σχήμα 5.8. Στη μόνιμη κατάσταση το ισοζύγιο ενέργειας ικανοποιείται για τα σημεία λειτουργίας 3, 5

και 7τα οποία είναι πιθανά σημεία ισορροπίας.

Θα προσπαθήσουμε να δώσουμε κάποιο φυσικό νόημα στην ευστάθεια του κάθε δυνατού σημείου λειτουργίας.

Χαμηλό σημείο ισορροπίας (Σημείο Λειτουργίας 3)

Ας θεωρήσουμε μια διαταραχή από το σημείο λειτουργίας που είναι ψυχρότερο από το σημείο 3, έστω $T_3 - dT$. Σε αυτό το σημείο παράγεται περισσότερη θερμότητα από ό, τι μπορεί να αφαιρεθεί από τον αντιδραστήρα, με αποτέλεσμα η θερμοκρασία στον αντιδραστήρα να αρχίζει να ανεβαίνει με τελική μετάβαση στην T_3 . Αν αρχίσουμε με $T_3 + dT$, περισσότερη ενέργεια αφαιρείται από ό, τι παράγεται, έτσι ώστε η θερμοκρασία να αρχίζει να μειώνεται με τελική κατάληξη την T_3 . Το σημείο τομής στην χαμηλή θερμοκρασία T_3 μπορεί να είναι ένα ευσταθές σημείο λειτουργίας. Για να διασφαλιστεί πλήρως ότι το σημείο ισορροπίας είναι ευσταθές είναι απαραίτητο να εξεταστούν οι ιδιοτιμές του συστήματος (Ενότητα 5.5).

Μεσαίο σημείο ισορροπίας (σημείο λειτουργίας 5)

Αν αρχίσουμε από το σημείο $T_5 - dT$, στο σημείο αυτό παράγεται λιγότερη ενέργεια από αυτήν που αφαιρείται, προκαλώντας μείωση της T μέχρι την τελική θερμοκρασία T_5 . Αν αρχίσουμε με $T_5 + dT$, περισσότερη ενέργεια παράγεται από ό, τι αφαιρείται, προκαλώντας αύξηση της θερμοκρασίας μέχρι να φτάσει στην T_5 . Το μεσαίο σημείο τομής T_5 είναι ένα ασταθές σημείο λειτουργίας. Εδώ, η ανάλυση μόνιμης κατάστασης είναι αρκετή για να διαπιστωθεί ότι το σημείο 5 είναι ένα ασταθές σημείο λειτουργίας.

Υψηλό σημείο ισορροπίας (σημείο λειτουργίας 7)

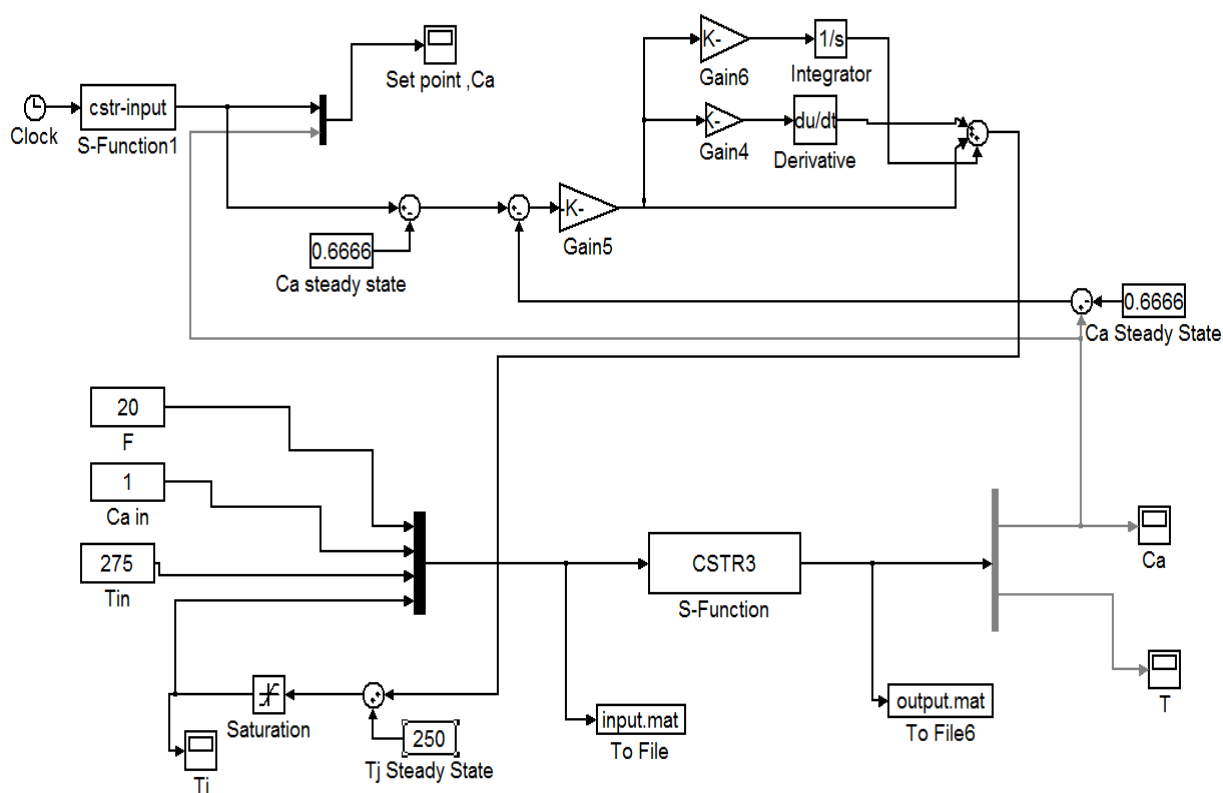
Όμοια για την T_7 όπως στην T_3 βρίσκουμε ότι το σημείο ισορροπίας στην υψηλή θερμοκρασία T_7 μπορεί να είναι ένα ευσταθές σημείο λειτουργίας. Και πάλι, η ευστάθεια του υψηλού σημείου ισορροπίας μπορεί να καθοριστεί μόνο από την εξέταση των ιδιοτιμών του συστήματος.

ΚΕΦΑΛΑΙΟ VI

ΣΧΕΔΙΑΣΜΟΣ ΚΑΙ ΕΦΑΡΜΟΓΗ ΝΕΥΡΩΝΙΚΟΥ ΡΥΘΜΙΣΤΗ

6.1 Προσομοίωση του Πειράματος

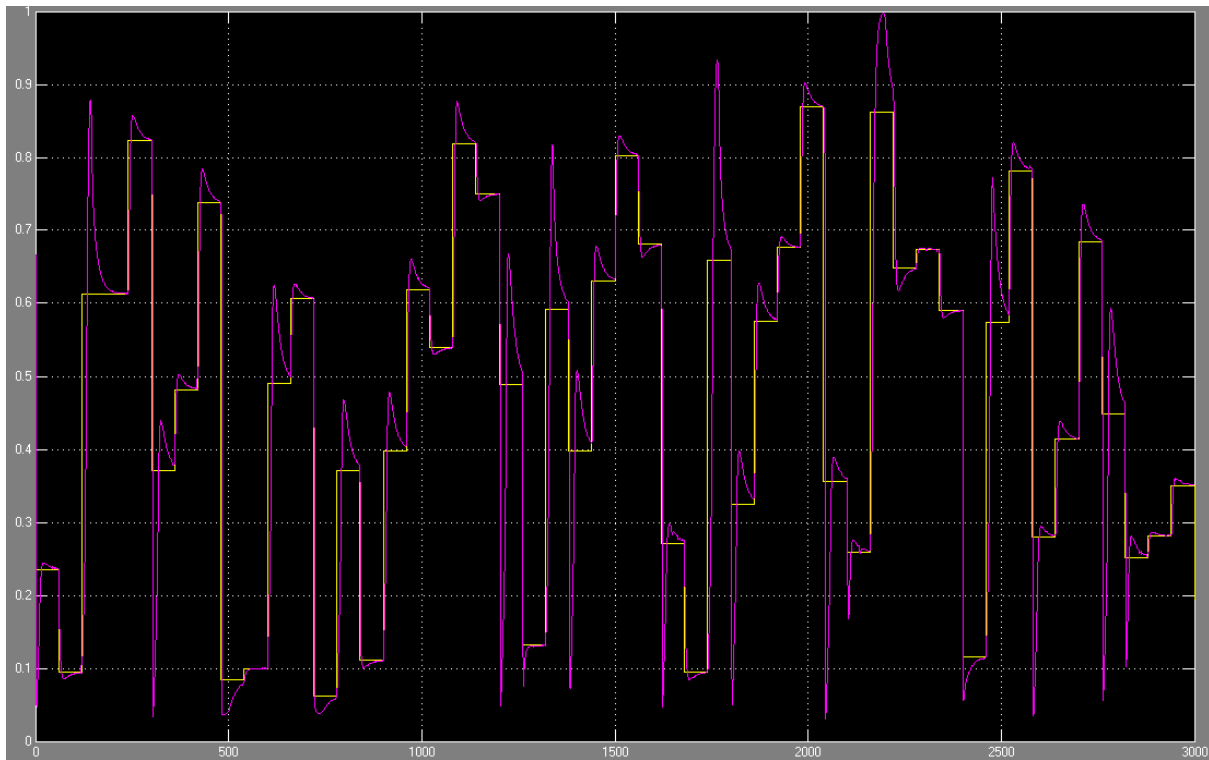
Τα νευρωνικά δίκτυα πραγματοποιούν προβλέψεις βασισμένες στην εκπαίδευσή τους και δεν χρησιμοποιούν μεθόδους επίλυσης της ροής. Επομένως, η ακρίβεια στις προβλέψεις τους εξαρτάται κυρίως από τα δείγματα με τα οποία εκπαιδεύονται. Για τον λόγο αυτό είναι σημαντικό να σχεδιαστεί το κατάλληλο πείραμα, έτσι ώστε τα δεδομένα που θα συλλεχθούν να είναι 'πλούσια' σε πληροφορία για τη δυναμική του συστήματος. Με αυτόν τον τρόπο, παράλληλα με τις ακριβείς προβλέψεις του δικτύου εξασφαλίζονται και ταχείς χρόνοι εκπαίδευσης.



Σχήμα 6.1: Προσομοίωση πειράματος για τη δημιουργία της βάσης δεδομένων

Το πείραμα έχει ως στόχο τη διέγερση του συστήματος ώστε να πυροδοτηθούν έντονες δυναμικές μεταβολές που να καλύπτουν όλο το πεδίο λειτουργίας του αντιδραστήρα. Λόγω του ότι η διεργασία που καλούμαστε να ρυθμίσουμε περιλαμβάνει ένα ασταθές σημείο ισορροπίας, η διέγερση του συστήματος λαμβάνει χώρα σε κλειστό βρόχο χρησιμοποιώντας ρυθμιστή τύπου Proportional-

Integral-Derivative (PID). Το σύστημα προσομοιώνεται στο Simulink όπως φαίνεται στο Σχήμα 6.1. Ο PID σχεδιάζεται συντηρητικά διότι πρωταρχικός του στόχος είναι να σταθεροποιήσει την διεργασία. Η βαθμονόμηση του (tuning) έγινε με δοκιμή και σφάλμα και οι παράμετροι του είναι $K_s = -550$, $K_p = 0.2$, $K_d = 5$. Για τις αλλαγές στην επιθυμητή τιμή (set point), αρχικά χρησιμοποιήθηκε μια γεννήτρια τυχαίων αριθμών με περίοδο εφαρμογής 60 sec.



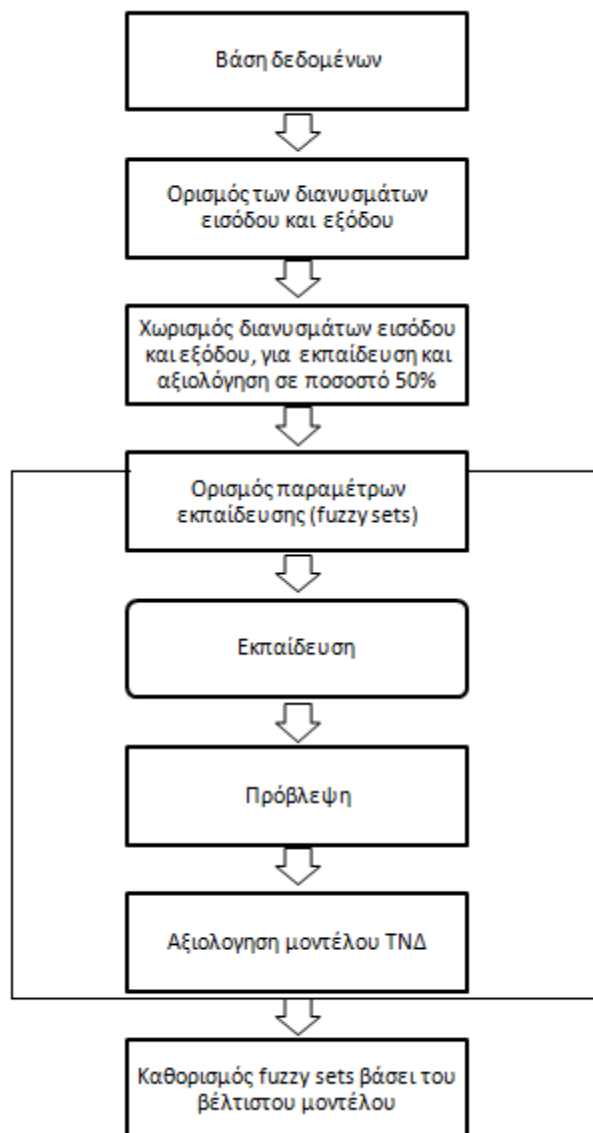
Σχήμα 6.2 : Αποκριση του αντιδραστήρα για αλλαγή στην επιθυμητή τιμή set point ανα 60 sec

Όμως όπως φαίνεται στο Σχήμα 6.2, ο αντιδραστήρας αποκρίνεται πιο γρήγορα στα χαμηλά set points απ'ότι στα υψηλά. Το αποτέλεσμα είναι ότι στα χαμηλά set points συλλέγουμε πολλά δεδομένα σε ισορροπία, που όμως δεν περιέχουν καμιά πληροφορία για τη δυναμική του συστήματος. Για το λόγο αυτό δημιουργήσαμε το function `cstr_input.m` με το οποίο οι αλλαγές στο set point δεν γίνονται με την ίδια ενιαία περίοδο εφαρμογής, αλλά ανάλογα με την τιμή του set point. Συγκεκριμένα, η περίοδος εφαρμογής είναι 20 αν το set point είναι μικρότερο του 0.4 και 30 διαφορετικά

Τα δεδομένα που συλλέγονται από το παραπάνω πείραμα είναι :

- οι μεταβλητές κατάστασης $C_A(k), T(k)$
- η μεταβλητή εκ χειρισμού $T_j(k)$

Μετά την συλλογή των δεδομένων ακολουθεί η εκπαίδευση και αξιολόγηση του νευρωνικού δικτύου σε μια διαδικασία που παρουσιάζεται σε μορφή ψευδο-αλγορίθμου στο Σχήμα 6.3



Σχήμα 6.3 : Διαδικασία που ακολουθήθηκε για την εκπαίδευση και την αξιολόγηση των νευρωνικών δικτύων

Για τον παραπάνω αλγόριθμο δίνονται οι εξής επεξηγήσεις και λεπτομέρειες :

1. Πριν γίνει ο διαχωρισμός των δεδομένων για εκπαίδευση και πρόβλεψη γίνεται αναδιάταξη των δεδομένων έτσι ώστε τα δεδομένα δοκιμής να ανήκουν στον ίδιο πληθυσμό που χρησιμοποιείται για τη δημιουργία των δεδομένων εκπαίδευσης
2. Ο αριθμός των fuzzy sets είναι η μοναδική μεταβλητή απόφασης για την εκπαίδευση του νευρωνικού δικτύου RBF
3. Εκπαιδεύουμε το νευρωνικό δίκτυο με το function RBFtraining.m το οποίο παίρνει σαν

είσοδο τα δεδομένα εκπαίδευσης (που περιλαμβάνουν διανύσματα εισόδου και εξόδου) και τον αριθμό των fuzzy sets και επιστρέφει τις τιμές των παραμέτρων του νευρωνικού δικτύου.

4. Τρέχουμε το function RBFprediction.m το οποίο πραγματοποιεί τις προβλέψεις του νευρωνικού δικτύου. Οι είσοδοι του είναι οι παράμετροι του νευρωνικού δικτύου που έχει ήδη εκπαιδευτεί και τα διανύσματα εισόδου από τα δεδομένα αξιολόγησης.
5. Αξιολογούνται οι προβλέψεις του νευρωνικού δικτύου με κριτήριο τον συντελεστή συσχέτισης

$$R^2 = 1 - \frac{\sum_{i=1}^n (w_o - w_p)^2}{\sum_{i=1}^n (w_o - w_o')^2}$$

Όπου w_o είναι τα διανύσματα εξόδου από τα δεδομένα αξιολόγησης, w_o' η μέση τιμή, n είναι ο συνολικός αριθμός παρατηρήσεων και w_p είναι οι προβλέψεις του νευρωνικού δικτύου

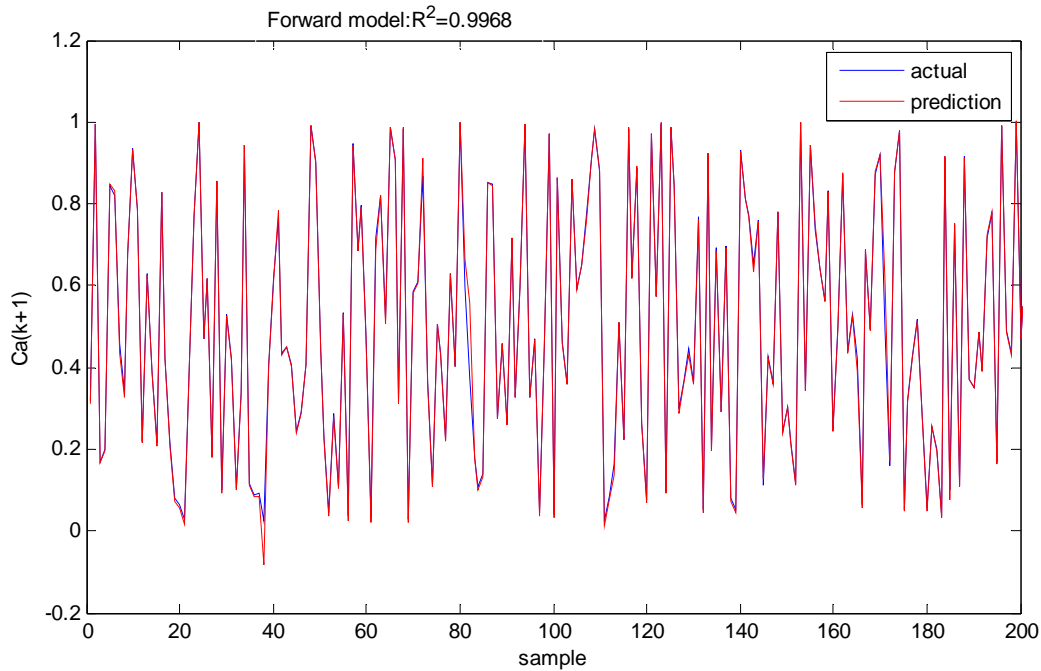
6.2 Πρόσθιο Μοντέλο

Η εκπαίδευση ενός νευρωνικού δικτύου για να αναπαραστήσει την πρόσθια δυναμική του συστήματος αναφέρεται ως πρόσθια μοντελοποίηση και τα μοντέλα που προκύπτουν από αυτή τη διαδικασία ως πρόσθια μοντέλα. Το νευρωνικό δίκτυο τοποθετείται παράλληλα με το σύστημα και το σφάλμα μεταξύ της εξόδου του συστήματος και της πρόβλεψης του δικτύου (δηλαδή το σφάλμα πρόβλεψης) χρησιμοποιείται ως σήμα για την εκπαίδευση του νευρωνικού δικτύου. Η έξοδος του δικτύου είναι η κατά ένα χρονικό βήμα μπροστά πρόβλεψη της συγκέντρωσης, ενώ οι είσοδοι αποτελούνται από τις τρέχουσες τιμές της συγκέντρωσης και της θερμοκρασίας του αντιδραστήρα καθώς και της θερμοκρασίας του ψυκτικού.

$$C_A(k+1) = f(C_A(k), T(k), T_j(k)) \quad (6.1)$$

όπου το f αντιπροσωπεύει το μη γραμμικό νευρωνικό δίκτυο

Στο σύστημα που μελετάμε, το νευρωνικό δίκτυο εκπαιδεύτηκε με 20.000 δεδομένα και έγιναν δοκιμές για τον αριθμό των fuzzy sets στο διάστημα 4 έως 50



Σχήμα 6.4 : Αξιολόγηση του πρόσθιου μοντέλου

Το βέλτιστο μοντέλο νευρωνικού δικτύου προέκυψε για fuzzy sets =50 και είναι αυτό που αναπαριστάται στο Σχήμα 6.1 με $R^2=0,9968$. Γενικά από τις δοκιμές στα fuzzy sets διαπιστώθηκε ότι όταν αυξάνεται ο αριθμός τους στο διάστημα 4 έως 50 βελτιώνεται η απόδοση του νευρωνικού δικτύου.

6.3 Αντίστροφο Μοντέλο-Σχεδιασμός Νευρωνικού Ρυθμιστή

Αντίστροφα μοντέλα είναι τα νευρωνικά δίκτυα που αναπαριστούν την αντίστροφη δυναμική του συστήματος κατά την ολοκλήρωση της εκπαίδευσης τους. Η διαδικασία εκπαίδευσης στην περίπτωση αυτή ονομάζεται αντίστροφη μοντελοποίηση.

Η αναπαράσταση του δικτύου είναι:

$$T_j(k) = f^{-1}(T(k), C_A(k), C_A(k+1)) \quad (6.2)$$

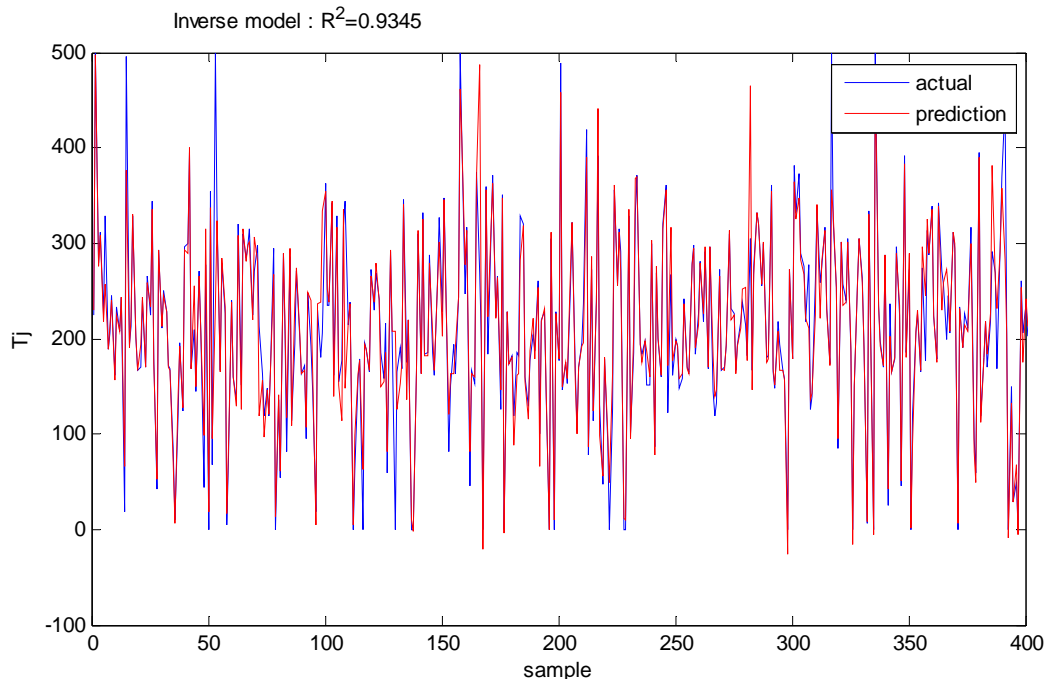
Δηλαδή το μοντέλο προβλέπει την τιμή του T_j που οδηγεί το σύστημα από μια αρχική κατάσταση σε μια τελική κατάσταση μέσα σε ένα χρονικό βήμα και είναι ουσιαστικά ο νευρωνικός ρυθμιστής που θα χρησιμοποιηθεί για τη ρύθμιση της διεργασίας.

Γενίκευση της Εξίσωσης (6.2) αποτελεί η

$$T_j(k) = f^{-1}(T(k), C_A(k), \dots, C_A(k+n)) \quad (6.3)$$

Στην περίπτωση αυτή λοιπόν το νευρωνικό δίκτυο έχει ως στόχο την προσέγγιση της παραπάνω συνάρτησης, χρησιμοποιώντας ως εισόδους την τρέχουσα και μελλοντικές τιμές της συγκέντρωσης και την τρέχουσα τιμή της θερμοκρασίας και λαμβάνοντας στην έξοδο την αντίστοιχη

τρέχουσα τιμή της θερμοκρασίας του ψυκτικού. Όταν το νευρωνικό δίκτυο χρησιμοποιείται ως νευρωνικός ρυθμιστής, η μοναδική αλλά ουσιαστική διαφοροποίηση είναι ότι στις θέσεις των μελλοντικών τιμών της συγκέντρωσης τοποθετούνται οι επιθυμητές τιμές της συγκέντρωσης στους επόμενους χρόνους δειγματοληψίας.



Σχήμα 6.5 : Αξιολόγηση του αντίστροφου μοντέλου

Το παραπάνω διάγραμμα αναπαριστά το αντίστροφο μοντέλο το οποίο εκπαιδεύτηκε με τα ίδια δεδομένα με το πρόσθιο μοντέλο και για fuzzy sets = 50. Φαίνεται ότι οι προβλέψεις του νευρωνικού δικτύου δεν είναι τόσο ακριβείς ωστόσο, για την εφαρμογή του στην ρύθμιση, όπως θα φανεί στη συνέχεια, η ακρίβεια που επιτυγχάνεται κρίνεται επαρκής.

6.4 Εφαρμογή και Αξιολόγηση Νευρωνικού Ρυθμιστή

Για την προσομοίωση του συστήματος κλειστού βρόχου που ρυθμίζεται με το νευρωνικό ρυθμιστή, χρησιμοποιούμε την ακόλουθη διαδικασία, ξεκινώντας από διακριτό χρόνο μηδέν και προχωρώντας στο χρόνο με βήμα ίσο t_s που είναι η περίοδος ρυθμιστικών δράσεων του ρυθμιστή. Για κάθε διακριτή χρονική στιγμή:

- Αρχικά τρέχουμε ένα function το οποίο υλοποιεί τον ρυθμιστή και δίνει ουσιαστικά την πρόβλεψη του νευρωνικού που έχει ήδη εκπαιδευτεί. Δηλαδή παίρνει σαν εισόδους την τρέχουσα κατάσταση του αντιδραστήρα και την επιθυμητή τιμή της συγκέντρωσης του

ρεύματος εξόδου και επιστρέφει την τιμή του T_j . Προφανώς στην αρχή, η προηγούμενη κατάσταση είναι το σημείο ισορροπίας από το οποίο ξεκινάει ο αντιδραστήρας.

- Στη συνέχεια προσομοιώσουμε τον αντιδραστήρα στο Matlab με χρήση επιλυτών διαφορικών εξισώσεων, όπως π.χ. ο επιλυτής ode45. Η σύνταξη είναι η ακόλουθη :

```
[t,x]=ode45(@CSTRfun,[0 ts],x0,[],Tj);
```

Όπου t_s το διάστημα εφαρμογής της απόφασης του ρυθμιστή, x_0 το διάνυσμα με την αρχική κατάσταση του αντιδραστήρα (σε κάθε χρονικό βήμα η αρχική κατάσταση θα είναι ίση με την τελική κατάσταση του προηγούμενου βήματος), και T_j είναι η θερμοκρασία του ψυκτικού που μόλις έχει προβλέψει ο νευρωνικός ελεγκτής.

- Στη συνέχεια από την ακολουθία τιμών του διανύσματος κατάστασης που μας επιστρέφει το *ode45*, κρατάμε την τελευταία τιμή που είναι ουσιαστικά η κατάσταση του συστήματος στο τέλος της προσομοίωσης αυτού του χρονικού βήματος και θα χρησιμοποιηθεί σαν αρχική κατάσταση στο επόμενο βήμα.

Το m-file που υλοποιεί το νευρωνικό ρυθμιστή παρατίθεται στο Παράρτημα.

Για να διαπιστωθεί αν ο RBF νευρωνικός ρυθμιστής έχει τη δυνατότητα να ρυθμίσει τον αντιδραστήρα σε όλη την περιοχή λειτουργίας του ,συμπεριλαμβανομένης της ασταθούς περιοχής, τον εφαρμόζουμε σε ένα πρόβλημα παρακολούθησης επιθυμητής τιμής (set point tracking). Στο πρόβλημα αυτό μετράται η απόδοση του νευρωνικού ρυθμιστή και επιχειρείται η μεγιστοποίηση της.

Δοκιμές έδειξαν πως οι παράμετροι που επηρεάζουν τη συμπεριφορά του νευρωνικού ελεγκτή είναι :

- Η περίοδος δειγματοληψίας δεδομένων εκπαίδευσης
- Η περίοδος ρυθμιστικών δράσεων του ελεγκτή
- Το πλήθος των μελλοντικών επιθυμητών τιμών που περιλαμβάνονται στο διάνυσμα εισόδου

Για την αξιολόγηση του νευρωνικού ελεγκτή χρησιμοποιήθηκαν ως δείκτες απόδοσης

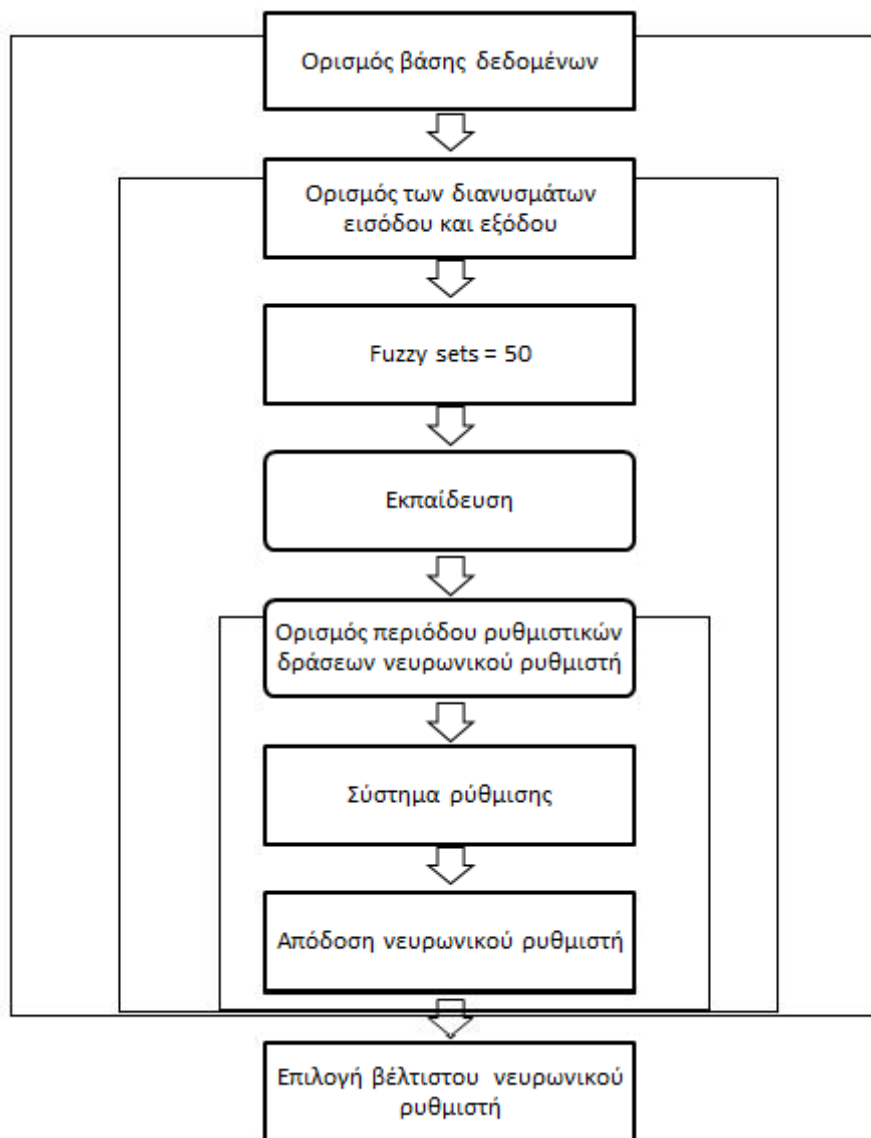
1.Μέσο Απόλυτο Σφάλμα (Mean Absolute Error, MAE)

$$MAE = \frac{1}{n} |Ca - setpoint|$$

2.Μέσο Τετραγωνικό Σφάλμα (Mean Square Error, MSE)

$$MSE = \frac{1}{n} (Ca - setpoint)^2$$

Ακολουθούμε την παρακάτω διαδικασία για να δούμε την επίδραση των παραμέτρων στη συμπεριφορά του νευρωνικού ρυθμιστή.



Σχήμα 6.6 : Διαδικασία που ακολουθήθηκε στην εφαρμογή και αξιολόγηση του νευρωνικού ρυθμιστή

Η ευαισθησία του προτεινόμενου ρυθμιστή σε σχέση με τις τρεις παραμέτρους που αναφέρθηκαν προηγούμενα παρουσιάζεται στους πίνακες 6-1,6-2,6-3.

Πίνακας 6-1 : Επίδραση της περιόδου δειγματοληψίας των δεδομένων εκπαίδευσης στην απόδοση του ρυθμιστή για μία μελλοντική επιθυμητή τιμή

Περίοδος δειγματοληψίας δεδομένων	0/5	1	1.5	2	2.5	3
MAE	0.0754	0.0426	0.0384	0,0442	0.067	0.0456

MSE	0.026	0.0157	0.0149	0,0172	0.0305	0.0158
-----	-------	--------	--------	--------	--------	--------

Η επιλογή μικρότερης περιόδου δειγματοληψίας δεδομένων οδηγεί στην λήψη περισσότερης πληροφορίας από το σύστημα με αποτέλεσμα τη βελτίωση της απόδοσης του ελεγκτή. Ωστόσο στην περίπτωση πολύ μικρής περιόδου δειγματοληψίας, η απόδοση του ρυθμιστή μπορεί να μειωθεί.

Πίνακας 6-2 : Επίδραση της περιόδου των ρυθμιστικών δράσεων στην απόδοση του ρυθμιστή για περίοδο δειγματοληψίας 1.5 και 1 μελλοντική επιθυμητή τιμή

Περίοδος ρυθμιστικών δράσεων	0.1	0.2	0.5	0.8	1
MAE	0.0363	0.0361	0.0471	0.054	0.0522
MSE	0.0145	0.0142	0.0198	0.0198	0.0193

Συνήθως, όσο μικρότερη είναι η περίοδος των ρυθμιστικών δράσεων του ελεγκτή, τόσο καλύτερη συμπεριφορά έχει ο ελεγκτής. Για το συγκεκριμένο πρόβλημα υπάρχει μια βέλτιστη τιμή 0.2. Αν η περίοδος γίνει ακόμη μικρότερη τότε παρατηρείται μείωση στην απόδοση του ρυθμιστικού συστήματος. Το γεγονός αυτό μπορεί να αποδοθεί στη δυναμική του συστήματος, αφού καθώς η περίοδος των ρυθμιστικών δράσεων του ελεγκτή μειώνεται, είναι πιθανόν να μην υπάρχει αρκετός χρόνος ώστε να μεταβεί το σύστημα από την τρέχουσα κατάστασή του σε αυτή που αντιστοιχεί στην καινούργια επιθυμητή τιμή, χρησιμοποιώντας τιμές της μεταβλητής εκ χειρισμού που βρίσκονται στα εντός των ορίων εκπαίδευσης του νευρωνικού δικτύου. Έτσι ο νευρωνικός ρυθμιστής αναγκάζεται να κάνει μια πρόβλεψη για τη μεταβλητή εκ χειρισμού που είναι εκτός των ορίων εκπαίδευσης, γεγονός που επιδρά αρνητικά στους δείκτες απόδοσης.

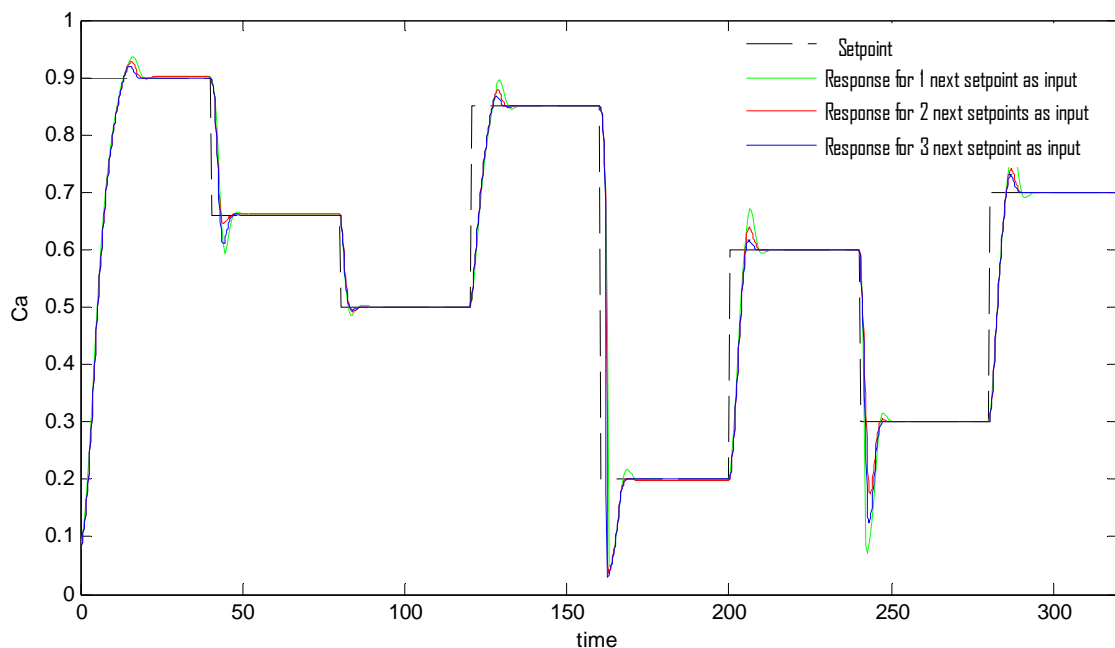
Στη συνέχεια διερευνήθηκε η επίδραση του πλήθους των επιθυμητών τιμών που τροφοδοτούνται σαν είσοδοι στο νευρωνικό ρυθμιστή, στην απόδοσή του. Παρατηρήθηκε ότι η παράμετρος αυτή αλληλεπιδρά με τη περίοδο των ρυθμιστικών δράσεων του ρυθμιστή, οπότε μεταβάλλουμε ταυτόχρονα και τις δυο παραμέτρους.

Πίνακας 6-3: Επίδραση του πλήθους των μελλοντικών επιθυμητών τιμών στην απόδοση του ρυθμιστή για περίοδο δειγματοληψίας δεδομένων 1.5

Αριθμός μελλοντικών επιθυμητών τιμών	1	2	3	4	5
Περίοδος ρυθμιστικών δράσεων του ρυθμιστή	0.5	0.2	0.2	0.1	0.1

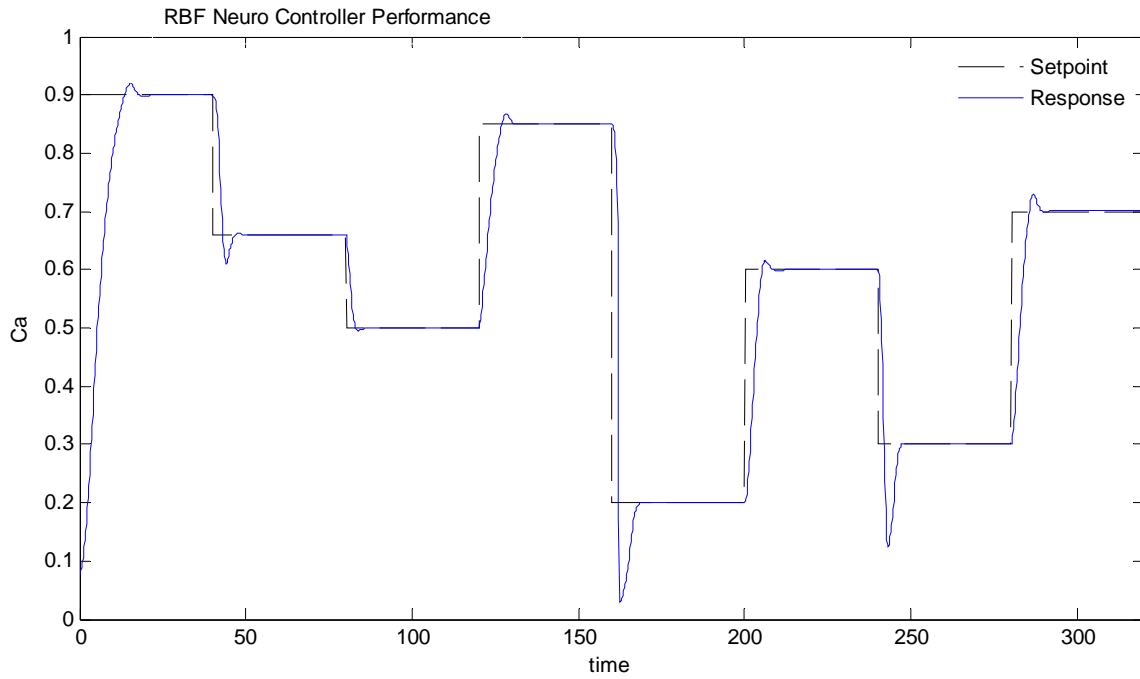
MAE	0.0374	0.0361	0.0358	0.0365	0.0361
MSE	0.0139	0.0141	0.0141	0.0145	0.0181

Παρατηρούμε ότι η απόδοση του νευρωνικού ρυθμιστή δεν παρουσιάζει μεγάλη ευαισθησία ως προς τη συγκεκριμένη παράμετρο. Η καλύτερη δυνατή απόδοση παρατηρείται όταν τροφοδοτούνται ως είσοδοι τρεις μελλοντικές επιθυμητές τιμές. Οι αποκρίσεις του συστήματος όταν ως είσοδοι χρησιμοποιούνται μια, δύο ή τρεις μελλοντικές επιθυμητές τιμές φαίνονται στο Σχήμα 6.8..



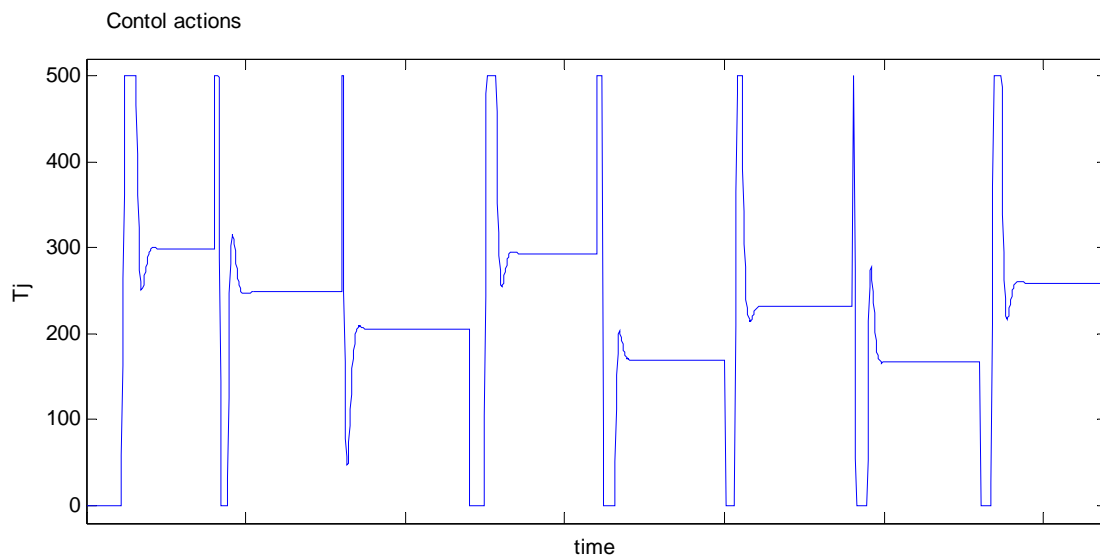
Σχήμα 6.8 : Απόκριση συστήματος σε παρακολούθηση επιθυμητής τιμής για διαφορετικό αριθμό μελλοντικών επιθυμητών τιμών που τροφοδοτούνται σαν είσοδοι στο νευρωνικό ρυθμιστή

Στα παρακάτω διαγράμματα παρουσιάζεται η βέλτιστη δυνατή συμπεριφορά του νευρωνικού ρυθμιστή, που προέκυψε για τις ακόλουθες αριστοποιημένες τιμές παραμέτρων: 1.5 για την περίοδο δειγματοληψίας δεδομένων, 0.2 για τη περίοδο μεταξύ διαδοχικών ρυθμιστικών δράσεων του ρυθμιστή και για πλήθος μελλοντικών επιθυμητών τιμών ίσο με 3.



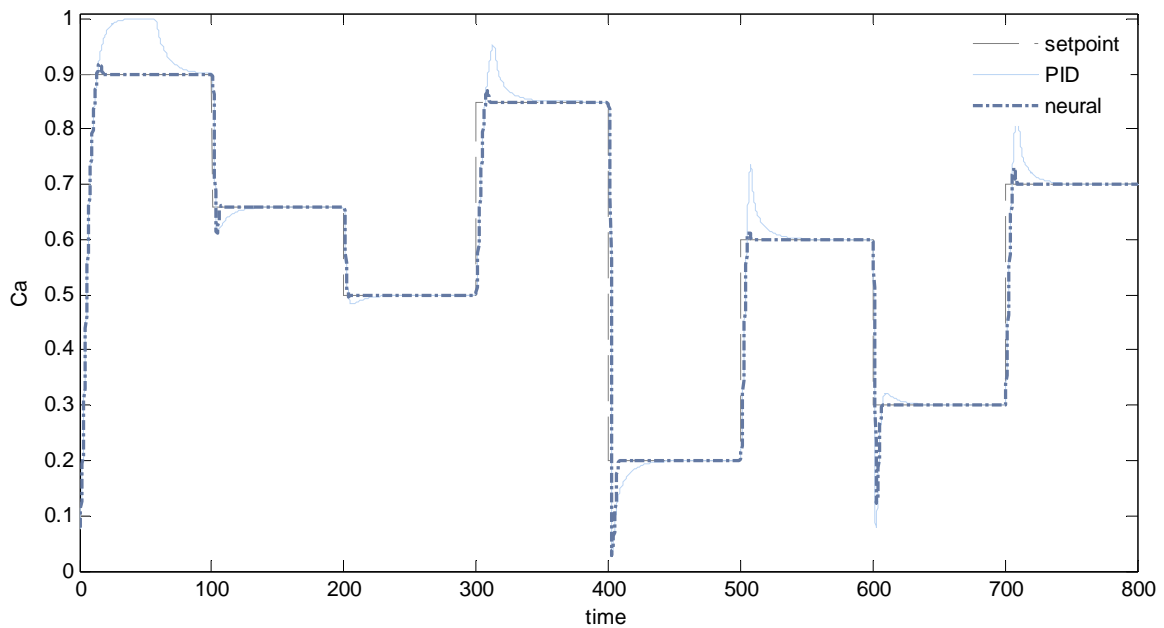
Σχήμα 6.9 : Παρακολούθηση επιθυμητής τιμής από το βέλτιστο νευρωνικό ρυθμιστή

Από το Σχήμα 6.9 γίνεται φανερό ότι ο προτεινόμενος RBF νευρωνικός ρυθμιστής οδηγεί τον αντιδραστήρα στις επιθυμητές τιμές συγκέντρωσης σε όλη τη περιοχή λειτουργίας του. Επίσης, παρατηρούνται μικρές αποκλίσεις μεταξύ της πραγματικής και της επιθυμητής τιμής της συγκέντρωσης.



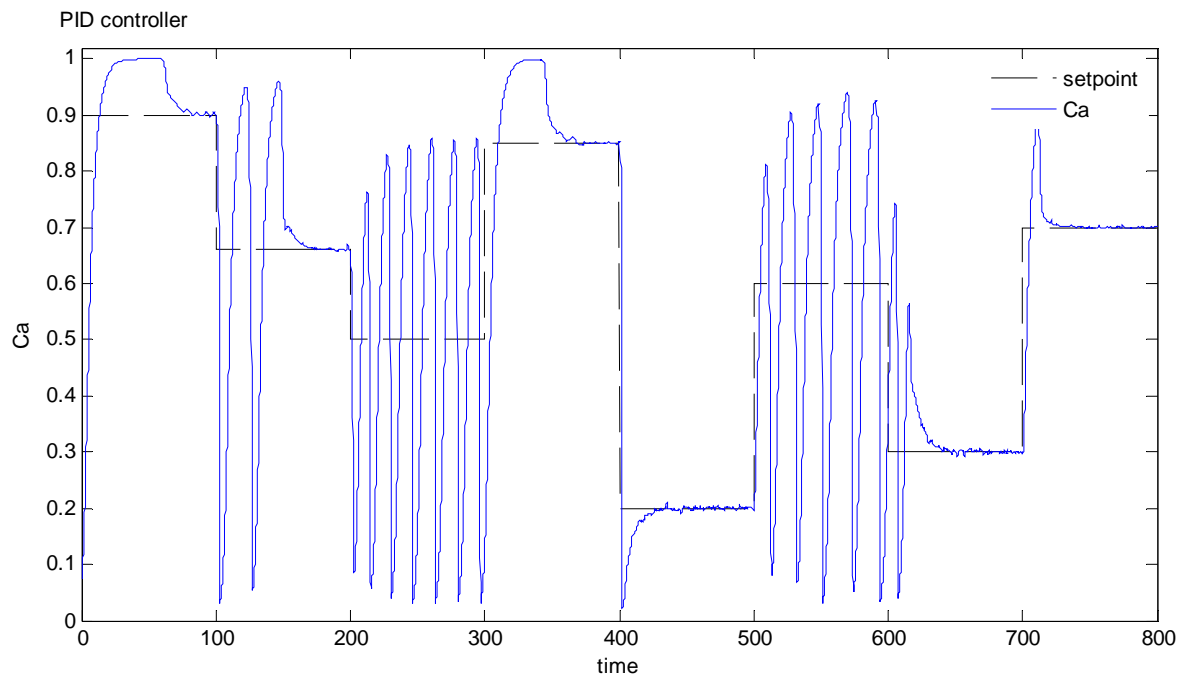
Σχήμα 6.10 : Ρυθμιστικές δράσεις σε συνάρτηση με το χρόνο

Για την περαιτέρω αξιολόγηση του προτεινόμενου συστήματος, έγινε προσομοίωση του συστήματος κλειστού βρόχου με χρήση PID ρυθμιστή βαθμονομημένου αυτόματα από το Simulink με παραμέτρους $P:-10087,4$, $I:-555,95$, $D:-26864$. Τα αποτελέσματα που παρουσιάζονται στο Σχήμα 6.9 καθιστούν εμφανές ότι ο προτεινόμενος νευρωνικός ρυθμιστής υπερτερεί έναντι του PID τόσο σε ταχύτητα όσο και ως προς την υπέρβαση της επιθυμητής τιμής σε κάθε περίπτωση όπου δίνονται σε αυτή βηματικές επιβολές.



Σχήμα 6.11 : Σύγκριση νευρωνικού ρυθμιστή και PID ρυθμιστή

Αν προσπαθήσουμε να βελτιώσουμε την ταχύτητα του συστήματος κλειστού βρόχου με διαφορετική βαθμονόμηση, με παραμέτρους $P:-60646,1$, $I:-7414,8$, $D:-55494,3$ το σύστημα παρουσιάζει ταλαντώσεις στην περιοχή γύρω από το ασταθές σημείο ισορροπίας όπως φαίνεται στο Σχήμα 6.12



Σχήμα 6.12 :Ασταθής συμπεριφορά συστήματος στη προσπάθεια αύξησης της ταχύτητας απόκρισης του PID

ΚΕΦΑΛΑΙΟ VII

ΣΥΜΠΕΡΑΣΜΑΤΑ

Σε αυτήν την εργασία παρουσιάζεται μια νέα μεθοδολογία για τη ρύθμιση μη γραμμικών συστημάτων. Η προτεινόμενη μεθοδολογία ρύθμισης βασίζεται στην προσέγγιση της αντίστροφης δυναμικής του συστήματος με ένα νευρωνικό δίκτυο RBF, το οποίο λαμβάνει ως εισόδους τις μεταβλητές κατάστασης και την επιθυμητή τιμή και παράγει ως έξοδο την κατάλληλη τιμή για τη μεταβλητή εκ χειρισμού. Η εκπαίδευση του RBF νευρωνικού δικτύου πραγματοποιήθηκε με τον αλγόριθμο των ασαφών μέσων. Ο προκύπτων ρυθμιστής εφαρμόστηκε στον έλεγχο ενός αντιδραστήρα CSTR με πολλαπλά σημεία ισορροπίας σε μια μελέτη περίπτωσης που αφορά την παρακολούθηση επιθυμητών τιμών. Τα αποτελέσματα έδειξαν υψηλή απόδοση του προτεινόμενου ρυθμιστή ακόμη και γύρω από ασταθές σημείο ισορροπίας, η οποία στη συνέχεια βελτιώθηκε περαιτέρω αριστοποιώντας τις παραμέτρους που επηρεάζουν τη λειτουργία του. Τέλος, έγινε σύγκριση του προτεινόμενου νευρωνικού ρυθμιστή με ένα συμβατικό PID ρυθμιστή και τα αποτελέσματα κατέδειξαν την ανώτερη απόδοση του προτεινόμενου συστήματος.

Με την ολοκλήρωση της εργασίας αναδείχθηκαν τα παρακάτω σημεία που χρήζουν περαιτέρω διερεύνησης :

- Επέκταση της μεθοδολογίας για την αντιμετώπιση του προβλήματος της απόρριψης διαταραχών (για τη συγκεκριμένη εφαρμογή στον αντιδραστήρα CSTR διαταραχές αποτελούν η συγκέντρωση εισόδου του αντιδρώντος, η θερμοκρασία και η παροχή της τροφοδοσίας).
- Η πειραματική επιβεβαίωση του προτεινόμενου νευρωνικού ελεγκτή σε πειραματική εργαστηριακή διάταξη

ΑΝΑΦΟΡΕΣ

1. Antsaklis J, “*Intelligent Control*”, Encyclopedia of Electrical and Electronics Engineering, Vol. 10, pp. 493-503, John Wiley & Sons, Inc., 1999.
2. Anuradha B., Prabhaker Reddy G*, N.Murthy N., “*Direct Inverse Neural Network Control of A Continuous Stirred Tank Reactor (CSTR)*”, IMECS, Hong Kong, 2009
3. Broomhead, D.S., Lowe, D., 1988, “*Multivariable functional interpolation and adaptive networks*”, Complex Systems, 2, 321–355
4. Chen, S., Cowan, C. F. N., “*Orthogonal Least Squares Learning Algorithm for Radial Basis Neural Networks*”, IEEE Transactions on Neural Networks, vol. 2, No2, 1991
5. Conradie A., “*A Neurocontrol Paradigm for Intelligent Process Control using Evolutionary Reinforcement Learning*”, Ph.d Thesis, Department of Chemical Engineering University of Stellenbosch, South Africa, 2004
6. Dayal S., Taylor A., MacGregor F., “*The design of experiments, training and implementation of nonlinear controllers based on neural network*”, Can J Chem Engng ,72:1067-1079, 1994
7. Garey, R. , Johnson S., “*Computers and Intractability: A Guide to the Theory of NP Completeness*”, Freeman, San Francisco,1979
8. Gumma S., “*A Radial Basis Function Neuro Controller for Permanent Magnet Stepper Motor*”, Master Thesis, Cleveland State University USA, 2004
9. Gupta M. and Rao H., “*Neural-Control Systems theory and Applications*”, Newyork press, 1994
10. Haykin S., “*Neural Networks : A Comprehensive foundation*”, Prentice Hall International,1999
11. Hornik, K., M. Stinchcombe and H. White, “*Multilayer Feedforward Networks are Universal Approximators*”, Neural Networks, 2, 359-366, 1989
12. Hunt, J., D. Sbarbaro R. Zbikowski and P.J. Gawthrop, “*Neural Networks for Control Systems A Survey*”, Automatica, 28(6), 1083-1112, 1992
13. Hussain A., and S. Kershenbaum, “*Implementation of an Inverse-Model-based Control Strategy using Neural Networks on a Partially Simulated Exothermic Reactor*”, Chemical Engineering Research and Design, **78**(A2), 299-311, 2000
14. Isidori A., “*Nonlinear Control Systems*”, Springer, 1999
15. Κυπαρισσίδης Κ., “*Σχεδιασμός των Χημικών Αντιδραστήρων*”, Τμήμα Χημικών Μηχανικών ΑΠΘ,Θεσσαλονίκη, 2008
16. Leonard, J.A., Kramer, M.A., “*Radial Basis Function Networks for Classifying Process Faults*”, IEEE Control Systems, 11, 31-38, 1991
17. Λιβιέρης Ι., “*Αποτίμηση Μεθόδων Εκπαίδευσης Τεχνητών Νευρωνικών Δικτύων και Εφαρμογές*”, Μεταπτυχιακή Διπλωματική Εργασία, Σχολή Μαθηματικών, Πάτρα,2008
18. Moody J., C. Darken, “*Fast Learning in Networks of Locally-tuned Processing Units*”, Neural

- Computation, 2, pp. 281-294, 1989
19. Narendra, S., and K. Parthasarathy, “*Identification and Control of Dynamic Systems using Neural Networks*”, IEEE Transactions on Neural Networks, **1**(1), 1990
 20. Ninos K., Giannakakis C.,Kompogiannis I., Stavrakas I., Alexandridis A., “*Nonlinear Control of a DC-Motor based on Radial Basis Function Neural Networks*”, INISTA ,2011
 21. Norgaard Magnus, “*Neural Network Design Toolkit*”, Technical University of Denmark,1997
 22. Omatu S.,Khalid M. and Rubiyah Y., “*Neuro-control and its Applications*”,Advances in Industrial Control , Springer
 23. Powel D., “*Radial Basis Functions for Multivariable Interpolation: A Review*”,in Algorithms for approximation, J.C. Mason M.G Cox, eds., pp. 143-167 Oxford, 1987
 24. Psaltis D., Sideris A. and Yamamura A.A., “*A Multi-layered Neural Network Controller*”, IEEE Intl. Conf on Neural Networks, VOI.3,pp. 1926-31,SanFransico,CA,1993
 25. Psychogios D.C. & Ungar L.H., “*Direct and Indirect Model based Control Using Artificial Neural Networks*”. Industrial & Engineering Chemistry Research, 30, 2564 – 2573,1991
 26. RayChaudhuri, T., C. Hamey, D. Bell, “*From Conventional Control to Autonomous Intelligent Methods*”, IEEE Control Systems Magazine, 16(5), 78-84
 27. Sarimveis H, Alexandridis A, Tsekouras G, Bafas G., “*A Fast and Efficient Algorithm for Training Radial Basis Function Neural Networks Based on a Fuzzy Partition of the Input Space*”, Industrial and Engineering Chemical Research 41, 751-59,2002
 28. Stephanopoulos G, “*Chemical Process Control: An Introduction to Theory and Practice*”, Department of Chemical Engineering MIT, Prentice Hall International
 29. <http://library.thinkquest.org/C007395/tqweb/history.html>
 30. http://www.rpi.edu/dept/chemeng/WWW/faculty/bequette/education/links_mods/cstr/
 31. <http://www.learnartificialneuralnetworks.com/#Intro>
 32. <http://dspace.lib.ntua.gr>

ΠΑΡΑΡΤΗΜΑ

M.file,function,"cstrsteady"

```
function f = cstrsteady (x)
```

```
V = 100 ;
```

```
UA = 20000 ;
```

```
P = 1000 ;
```

```
Cp = 4.2 ;
```

```
dH= 596619 ;
```

```
k0 = 6.85*10^11 ;
```

```
E = 76534.704 ;
```

```
R = 8.314 ;
```

```
F = 20 ;
```

```
Tin = 275 ;
```

```
Tj = 250 ;
```

```
Cain = 1 ;
```

```
f = [F/V*(Cain-x(1))-2*k0*exp(-E/(R*x(2)))*x(1)^2 ;
```

```
      F/V*(Tin-x(2))+2*dH/(P*Cp)*k0*exp(-E/(R*x(2)))*x(1)^2-
```

```
      UA/(V*P*Cp)*(x(2)-Tj)] ;
```

M.file,"amat.m"

```
A11= -F/V-4*k0*exp(-E/(R*T))*ca
```

```
A12=-2*k0*exp(-E/(R*T))*ca^2*(E/R*T^2)
```

```
A21=4*dh/(p*cp)*k0*exp(-E/(R*T))*ca
```

```
A22=-F/V-UA/(V*p*cp)+2*dh/(p*cp)*k0*exp(-E/(R*T))*(E/(R*T^2))*ca^2
```

```
A=[A11 A12;A21 A22]
```

```
lamda=eig(A)
```

M.file,function,"CSTR3"

```

function [sys,x0,str,Tvector]=CSTR3(t,x,u,flag)

switch flag,

case 0,
    [sys,x0,str,Tvector]=mdlInitializeSizes;

case 1,
    sys=mdlDerivatives(t,x,u);

case 3,
    sys=mdlOutputs(t,x,u);

case {2,4},
    sys=[];

case 9,
    mdlTerminate

otherwise
    error(['Unhandled flag = ',num2str(flag)]);

end

function [sys,x0,str,Tvector]=mdlInitializeSizes

sizes=simsizes;
sizes.NumContStates=2;
sizes.NumDiscStates=0;
sizes.NumOutputs=2;
sizes.NumInputs=-1;
sizes.DirFeedthrough=0;
sizes.NumSampleTimes=1;

```

```

sys=simsizes(sizes);

%x0=[0.9827 268.6]; %Upper steady state for T=250

%x0=[0.666 308.489]; %Middle steady state for T=250

% x0=[0.07545 376.3]; %Lower steady state for T=250

x0=[0.9 200];

str=[];
Tvector=[0 0];

function sys=mdlDerivatives(t,x,u);

V=100;
k0=6.85*10^11;
ER=76534.704/8.314;
DH=-596619;
cp=4.2;
rho=1000;
UA=20000;

sys(1)=u(1)/V*(u(2)-x(1))-2*k0*exp(-ER/x(2))*x(1)^2;
sys(2)=u(1)/V*(u(3)-x(2))+2*(-DH)/rho/cp*k0*exp(-ER/x(2))*x(1)^2-
UA/V/rho/cp*(x(2)-u(4));

function sys=mdlOutputs(t,x,u)

sys(1)=x(1);
sys(2)=x(2);

function mdlTerminate
disp('The simulation has terminated')

```

M.file,function,"cstr_input"

```
function [ spnew ] = cstr_input(u)
tnow=u(1);

% assignin('caller','sp',sp);
% assignin('caller','tprev',tprev);

global sp tprev
if tnow-tprev > 20

    if sp>0.4

        if tnow-tprev > 30
            spnew=rand*0.95;
            assignin('base','sp',spnew);
            assignin('base','tprev',tnow);
            return
        end

    else
        spnew=rand*0.95;
        assignin('base','sp',spnew);
        assignin('base','tprev',tnow);
        return
    end

end

spnew=sp;
```

```
% tprevnew=tprev;
```

```
end
```

M.file,function,"RBFtraining"

```
function [cc]=RBFtraining(x,y,noofsets)
```

```
ntr=size(x,1); %Number of training examples
```

```
niv=size(x,2); %Number of input variables
```

```
nov=1;%size(y,2);
```

```
 %[xsc,xm,xs,ysc,ym,ys]=prestd(x',y');
```

```
[xsc,xmx,xmn,ysc,ymx,ymn]=premmx(x',y');
```

```
 %[xsc,pcam]=prepca(xsc,0.001);
```

```
xsc=xsc';
```

```
ysc=ysc';
```

```
[C,deltaa,cc]=EllipseFCS(xsc,noofsets*ones(1,niv));
```

```
 %[C,deltaa,cc]=FuzzyCenterSelection(xsc,noofsets);
```

```
p=cc-1; %number of nearest neighbors
```

```
cc;
```

```
%-----
```

```
for i=1:cc
```

```

    for j=1:cc
        DC(i,j)=norm(C(i,:)-C(j,:));
    end

end

for i=1:cc

    [srt1,srt2]=sort(DC(i,:));
    for j=1:p
        nneb(i,j)=srt2(j+1);
    end

end

for i=1:cc

    es(i)=0;
    for j=1:p
        es(i)=es(i)+norm(C(i,:)-C(nneb(i,j),:))^2;
    end

end

es=sqrt(1/p*es);

for i=1:cc

    for j=1:ntr
%         tfa(j,i)=exp(-(norm(C(i,:)-xsc(j,:))^2)/(es(i)^2));
        tfa(j,i)=(norm(C(i,:)-xsc(j,:))^2)*log(norm(C(i,:)-
xsc(j,:))+1);
    end
end

```

```

    end

end

tfa(:,cc+1)=1;

w=inv(tfa'*tfa)*tfa'*yssc;

save netinfo C w es xmx xmn ymx ymn %Neural network parameters

```

M.file,function,"RBFprediction"

```

function [yp]=RBFvalidation(xv)
load netinfo

ntr=size(xv,1) ; %Number of training examples
niv=size(xv,2) ; %Number of input variables
nov=1;%size(yv,2)
cc=size(C,1);

[xvssc]=tramnmx(xv',xmx,xmn);
%[xvssc]=trastd(xv',xm,xs);

%xvssc=trapca(xvssc,pcam);
xvssc=xvssc';

clear tfa
for i=1:cc

    for j=1:ntr
        % tfa(j,i)=exp(-(norm(C(i,:)-xvssc(j,:))^2)/(es(i)^2));
        tfa(j,i)=(norm(C(i,:)-xvssc(j,:))^2)*log(norm(C(i,:)-
xvssc(j,:))+1);
    end
end

```

```

end
clear ypsc yp
    tfa(:,cc+1)=1;
ypsc=tfa*w;

yp=postmmx(ypsc',ymx,ymn);
%yp=poststd(ypsc',ym,ys);

yp=yp';

```

M.file,function,"validation"

```

clear
load New15
x1=Tj ;
x2=CaT ;
L=size(x1,1) ;

xall=[x2(1:(L-1),:) x2(2:L,1)];
yall=x1(1:L-1,;) ;
clear x1 x2
N=L-1

rand('seed',2)
perm=randperm(N);
xall=xall(perm,:);
yall=yall(perm,:);

x=xall(1:(N-2)/2,:);
y=yall(1:(N-2)/2,:);

xv=xall(N/2:end-1,:);
y2=yall(N/2:end-1,:);
%end

```

```
clear xall yall
```

```
tmean=mean(y2);
```

```
for noofsets=[40:45] ;
```

```
    [cc]=RBFtraining(x,y,noofsets);
```

```
    [yp]=RBFvalidation(xv);
```

```
    ypall(:,noofsets)=yp;
```

```
    MARE=1/((N-1)/2)*sum(abs(y2-ypall(:,noofsets))./y2)*100;
```

```
    MAREall(noofsets)= MARE
```

```
    R2(noofsets)= 1-(sum((y2-ypall(:,noofsets)).^2)./sum((y2-tmean).^2))
```

```
end
```

M.file,function,"CSTRfunc"

```
function dydt=CSTRfunc(t,x,uin1)
```

```
V=100;
```

```
k0=6.85*10^11;
```

```
ER=76534.704/8.314;
```

```
DH=-596619;
```

```
cp=4.2;
```

```
rho=1000;
```

```
UA=20000;
```

```
F = 20 ;
```

```
sys(1)=F/V*(1-x(1))-2*k0*exp(-ER/x(2))*x(1)^2;
```

```
sys(2)=F/V*(275-x(2))+2*(-DH)/rho/cp*k0*exp(-ER/x(2))*x(1)^2-
```

```
UA/V/rho/cp*(x(2)-uin1);
```

```
dydt=sys';
```

```
ttot=320 ;
ts=0.2;
M=ttot/ts ;
setpoint=[0.9 0.66 0.5 0.85 0.2 0.6 0.3 0.7] ;
x0=[0.0755 376.2697] ;
tsetpoint=ttot/8 ;
N=tsetpoint/ts ;

for j=[1:8]
    for i=[1:N]
        K((j-1)*N+i,:)=((j-1)*N+i)*ts ;
        H((j-1)*N+i,:)=setpoint(j) ;
        xv=[x0 setpoint(j) setpoint(j) setpoint(j)] ;
        yp=RBFvalidation(xv) ;
        if yp>500
            yp=500 ;
        end
        if yp<0
            yp=0;
        end
        Tj=yp ;
        B((j-1)*N+i,:)=Tj ;
        [t,x]=ode45(@CSTRfunc,[0 ts],x0,[],Tj) ;
        x0=x(end,:) ;
        Z((j-1)*N+i,:)=x0 ;
        A(i,:)= x0 ;
    end
    x0=A(end,:) ;
    D(j,:)=x0 ;
end
```

```
SAE=sum(abs(Z(:,1)-H))  
MAE=SAE/M  
SSE= sum((Z(:,1)-H).^2)  
MSE=SSE/M
```