



Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Χημικών Μηχανικών

Διπλωματική εργασία

Αλγόριθμος ενισχυτικής μάθησης για ρύθμιση διεργασιών με
κατασκευή νευρωνικών δικτύων ακτινικής βάσης σε πραγματικό
χρόνο

Πετσαγκουράκης Παναγιώτης

Επιβλέπων:

Αν. Καθηγητής Σαρίμβεης Χαράλαμπος

Ιούνιος 2015

Περίληψη

Η ρύθμιση ενός μη γραμμικού συστήματος δεν μπορεί να πραγματοποιηθεί με ευκολία, ειδικά όταν υπάρχει μη γραμμικότητα και στη μεταβλητή εισόδου (non affine). Η ενισχυτική μάθηση δίνει τη δυνατότητα να σχεδιαστεί online ένας ευσταθής ρυθμιστής χωρίς καμία γνώση για το σύστημα. Ο εν λόγω ρυθμιστής έχει δύο οντότητες, το action νευρωνικό δίκτυο και το critic νευρωνικό δίκτυο. Το πρώτο προσεγγίζει το μη γραμμικό κομμάτι του συστήματος και παράγει ένα μέρος του σήματος εισόδου, ενώ από την άλλη το critic εφαρμόζεται για να προσεγγίζει τη συνάρτηση κόστους με στόχο να ανανεώσει τα βάρη του action νευρωνικού δικτύου. Οι περισσότερες εφαρμογές που εμφανίζονται βιβλιογραφικά στην online μάθηση, μέσω νευρωνικών δικτύων λαμβάνουν τυχαίες αρχικές τιμές για το διάνυμα των βαρών που συνδέει τους νευρώνες εισόδου με τους κρυφούς νευρώνες. Το αποτέλεσμα αυτής της διαδικασίας είναι η απόδοση του συστήματος να αλλάζει σε κάθε προσομοίωση. Στην παρούσα εργασία, εφαρμόστηκε μια διαφορετική προσέγγιση, όπου χρησιμοποιήθηκαν νευρωνικά δίκτυα ακτινικής βάσης (RBF neural networks), όπου τα κέντρα των συναρτήσεων ενεργοποίησης, αλλά και ο αριθμός νευρώνων (nodes) πλέον αλλάζουν και αυτά με online κινήσεις. Για το σκοπό αυτό χρησιμοποιήθηκε ο αλγόριθμος online fuzzy means, όπου επιλέγει σε κάθε χρονική στιγμή αν θα αυξηθούν τα κέντρα. Αυτή η προσέγγιση εφαρμόστηκε σε 3 προβλήματα. Το πρώτο αποτελείται από ένα μη γραμμικό non affine μαθηματικό σύστημα, ενώ τα άλλα δυο εξάγονται από το χώρο της χημικής μηχανικής.

Λέξεις Κλειδιά

Νευρωνικά Δίκτυα Ακτινικής Βάσης, Ασαφή Λογική, Ενισχυτική μάθηση, Έξυπνη ρύθμιση

Abstract

Controlling a nonlinear system is not an easy task, especially, when the input variable consists of a non-affine term. Reinforcement learning provides the opportunity to design online a stable controller without any knowledge of the system itself. The proposed controller has two entities, an action neural network and a critic neural network. The action neural network approximates the nonlinear part of the system's dynamics, and generates the control signal. On the other hand, the critic neural network is implemented to approximate the cost function, aiming to update the adaptable weights of the action network. In literature, there are many applications of the control theory applied on online learning. Most of them initialize randomly the weight vector from the input layer to the hidden layer and continuously hold it constant. The result of this is that the performance changes at each simulation. In this project a different approach is implemented, using Radial Basis Function (RBF) networks as artificial neural networks to tackle this problem. The online fuzzy means algorithm is used at the learning procedure, updating not only the parameters (centers of the RBF) but also the number of the nodes of the networks. Consequently, every parameter of the networks is trained online instead of offline. This approach is applied on three examples. The first one is a nonlinear, non-affine discrete system and the other two are taken from chemical engineering application.

Keywords

Radial Basis Neural Networks, Fuzzy Logic, Reinforcement Learning, Intelligent Control

Ευχαριστίες

Πριν ξεκινήσω την Διπλωματική μου ήμουν πεπεισμένος ότι θέλω μετά το τέλος των σπουδών μου να κατευθυνθώ απευθείας στην αγορά εργασίας, μάλιστα έδινα στον εαυτό μου το πολύ άλλον ένα χρόνο για να κάνω ένα μεταπτυχιακό και να αυξήσω τις υπάρχουσες γνώσεις μου. Όμως η τριβή μου με ένα τόσο ενδιαφέρον αντικείμενο, όπως αυτό που ασχολήθηκα κατά την διάρκεια της διπλωματικής με έκανε να καταλάβω, το πόσο μου αρέσει να ασχολούμαι και να εμβαθύνω σε θέματα του κλάδου μας. Για αυτό το λόγο πιστεύω πως αυτή η εργασία δεν μου χάρισε μόνο γνώσεις και εμπειρίες αλλά και θέληση για τη συνέχιση ενός διδακτορικού. Σε αυτό το σημείο θα ήθελα να ευχαριστήσω τον αναπληρωτή καθηγητή Χ.Σαρίμβεη που μου εμπιστεύτηκε ένα τόσο πολύ ενδιαφέρον θέμα, και θέλω πραγματικά να τον ευχαριστήσω που σε όλη διάρκεια της εκπόνησης με συμβούλεψε, με καθοδήγησε σε πολύ καθοριστικά σημεία, αλλά επίσης με βοήθησε στις μεταπτυχιακές μου επιλογές μέχρι και στη προετοιμασία για συνέντευξη. Σε αυτή την εργασία είχε σημαντικό ρόλο και η υποψήφια διδάκτορας Ευαγγελία Χονδροδήμα που την ευχαριστώ για την πολύ σημαντική βοήθεια της, καθώς με βοήθησε να 'ξεκολλήσω' και να συνεχίσω σε πολύ σημαντικά σημεία της εργασίας.

Επίσης, θα ήθελα να ευχαριστήσω τον καθηγητή Α.Μπουντουβή για τις συμβουλές του σχετικά με τις επιλογές μου αλλά και τη βοήθεια στις διαδικασίες για το διδακτορικό που θα ακολουθηθεί.

Θέλω επίσης να ευχαριστήσω όλους τους φίλους μου (συμφοιτητές ή μη) που μου στάθηκαν σε όλη την πενταετή πορεία των σπουδών μου. Ωστόσο ειδικό ευχαριστώ χρειάζεται να δοθεί στους Δημήτρη Σκαρπαλέζο, Αναστασία Νικολακοπούλου, Κώστα Πανόπουλο, Νίκο Ρακοβίτη, Σπύρο Κυριαζή, Δημητρή Λεμπέκο, Κω-

σταντίνο Παναγόπουλο, Αθανασία Ρούσση, Διονύση Μπαλτά, Δημήτρη Ζουράρη, Σοφία Τζαβίδη, επίκουρο καθηγητή Αντώνη Καραντώνη και Γιάννη Αντωνόπουλο.

Τέλος από τα μεγαλύτερα ευχαριστώ στους γονείς μου, στον αδερφό μου Γιάννη που ήταν πάντα εκεί και να με ηρεμεί και συμβουλεύει και φυσικά στην Άννα Μαρία Τσακίρογλου που δεν σταμάτησε λεπτό να είναι δίπλα μου όλα αυτά τα πέντε χρόνια, στηρίζοντας, συμβουλεύοντας και ενθαρρύνοντας με.

Περιεχόμενα

1	Εισαγωγή	1
2	Τεχνητά Νευρωνικά Δίκτυα	3
2.1	Ιστορική αναδρομή	3
2.2	Βιολογική Προέλευση	5
2.3	Το μαθηματικό μοντέλο	7
2.4	Αρχιτεκτονικές νευρωνικών δικτύων	10
2.4.1	Ενός επιπέδου νευρωνικά δίκτυα πρόσθιας τροφοδότησης	10
2.4.2	Πολυεπίπεδα δίκτυα πρόσθιας τροφοδότησης	11
2.4.3	Δίκτυα ακτινικής βάσης RBF	12
2.5	Εκπαίδευση Νευρωνικών Δικτύων	13
2.5.1	Επιβλεπόμενη μάθηση	13
2.5.2	Ενισχυτική Μάθηση	15
2.5.3	Μάθηση χωρίς επίβλεψη	16
2.6	Ιδιότητες	16
2.6.1	Προσέγγιση Συναρτήσεων (Function Approximation)	17
2.7	Ταξινόμηση σχημάτων μάθησης	18
2.7.1	Απευθείας υπολογισμός vs Εκπαίδευσης	18
2.7.2	Μάθηση σε πραγματικό χρόνο vs σε μη πραγματικό χρόνο	18
2.8	Πλεονεκτήματα	19
2.9	Μειονεκτήματα	20
2.10	Εφαρμογές των Τεχνητών Νευρωνικών Δικτύων	20

3	Δυναμικός Προγραμματισμός	23
3.1	Εισαγωγή στον Δυναμικό Προγραμματισμό	23
3.2	Δυναμικός προγραμματισμός και συστήματα διακριτού χρόνου . . .	24
3.3	Αρχή βελτιστοποίηση του Bellman	25
4	Ενισχυτική Μάθηση	27
4.1	Σύστημα ή Περιβάλλον	28
4.2	Βασιζόμενη στο μοντέλο και ελεύθερο μοντέλου	29
4.2.1	Νευρωνικά δίκτυα και ενισχυτική μάθηση	29
4.3	Στόχος η βέλτιστη απόδοση	30
4.4	Εξίσωση του Bellman	30
5	Νευρωνικά Δίκτυα Ακτινικής Βάσης	33
5.1	Δομή Δικτύων Ακτινικής Βάσης	34
6	Ο Αλγόριθμος των Ασαφών Μέσων (Fuzzy Means Algorithm)	37
6.1	Μη συμμετρικός off-line αλγόριθμος	37
6.2	Αλγόριθμος ασαφών μέσων σε πραγματικό χρόνο	41
6.2.1	Αρχικοποίηση του αλγορίθμου	47
6.2.2	Προσδιορισμός του πρώτου κέντρου	47
6.2.3	Έλεγχος νέου κέντρου	48
6.2.4	Επιλογή του νέου κέντρου	49
6.2.5	Ανανέωση του διανύσματος ιστορικού ενεργοποίησης	49
6.2.6	Έλεγχος διαγραφής υπάρχοντος κόμβου	49
6.2.7	Διαγραφή κόμβου	50
6.2.8	Υπολογισμός νέων συναπτικών βαρών με τη μέθοδο των ελαχίστων τετραγώνων.	50
7	Μεθοδολογία Ρύθμισης Μη Γραμμικών Συστημάτων με Ενισχυτική Μάθηση	51
7.1	Σχεδιασμός Ρυθμιστή	52
7.2	Μάθηση Νευρωνικών Δικτύων Σε Πραγματικό Χρόνο- Δίκτυο Κριτής	54

7.3	Μάθηση Νευρωνικών Δικτύων Σε Πραγματικό Χρόνο- Δίκτυο Ενέργειας	56
7.4	Διάγραμμα Ροής	57
7.5	Εφαρμογές	58
7.5.1	Εφαρμογή 1	59
7.5.2	Εφαρμογή 2-Αντιδραστήρας με παράλληλες Αντιδράσεις . .	67
8	Τροποποίηση της συνάρτησης κόστους στην μεθοδολογία ενισχυτικής μάθησης	75
8.1	Τροποποίηση του δικτύου Κριτής	76
8.2	Εφαρμογές	77
8.2.1	Εφαρμογή 1	77
8.2.2	Εφαρμογή 2-Αντιδραστήρας με παράλληλες Αντιδράσεις . .	82
8.2.3	Εφαρμογή 3 - Χημικός αντιδραστήρας	87
9	Συμπεράσματα	95

Κατάλογος σχημάτων

2.1	Βιολογικά Νευρωνικά Δίκτυα	7
2.2	Ένας Νευρώνας	9
2.3	Ενός επιπέδου νευρωνικά δίκτυα	11
2.4	Νευρωνικά Δίκτυα Ακτινικής Βάσης	13
2.5	Μάθηση με επίβλεψη	15
4.1	Σχηματική αναπαράσταση ενισχυτικής μάθησης	28
5.1	Δομή νευρωνικών δικτύων ακτινικής βάσης	35
6.1	Ασαφής διαμέριση του χώρου των μεταβλητών εισόδου σε 5 ασαφή σύνολα για ένα δισδιάστατο πρόβλημα	41
6.2	Επιλογή των κατάλληλων ασαφών συνόλων	42
6.3	Προσδιορισμός του κοντινότερου ασαφή υπόχωρου	42
6.4	Αλγόριθμος Ασαφών μέσων σε πραγματικό χρόνο	46
7.1	Διάγραμμα ροής της μεθοδολογίας ρύθμισης	58
7.2	Η πορεία της μεταβλητής εξόδου και της επιθυμητής τιμής σε συ- νάρτηση με το χρόνο, περίπτωση 1.1.	60
7.3	Η πορεία του σφάλματος σε συνάρτηση με το χρόνο, περίπτωση 1.1.	61
7.4	Η εξέλιξη του αριθμού των κέντρων των δύο δικτύων, περίπτωση 1.1.	61
7.5	Η πορεία της μεταβλητής εξόδου και της επιθυμητής τιμής σε συ- νάρτηση με το χρόνο, περίπτωση 1.2.	63
7.6	Η πορεία του σφάλματος σε συνάρτηση με το χρόνο, περίπτωση 1.2.	63
7.7	Η εξέλιξη του αριθμού των κέντρων των δύο δικτύων, περίπτωση 1.2.	64

7.8	Η πορεία της μεταβλητής εξόδου και της επιθυμητής τιμής σε συνάρτηση με το χρόνο, περίπτωση 1.3	65
7.9	Η πορεία του σφάλματος σε συνάρτηση με το χρόνο, περίπτωση 1.3	66
7.10	Η εξέλιξη του αριθμού των κέντρων των δύο δικτύων, περίπτωση 1.3	66
7.11	Αντιδραστήρας παραγωγής ενδιάμεσου προϊόντος Β	68
7.12	Η πορεία της μεταβλητής εξόδου και της επιθυμητής τιμής σε συνάρτηση με το χρόνο, περίπτωση 2.1	70
7.13	Η πορεία του σφάλματος σε συνάρτηση με το χρόνο, περίπτωση 2.1	70
7.14	Η εξέλιξη του αριθμού των κέντρων των δύο δικτύων, περίπτωση 2.1	71
7.15	Η πορεία της μεταβλητής εξόδου και της επιθυμητής τιμής σε συνάρτηση με το χρόνο, περίπτωση 2.2	72
7.16	Η πορεία του σφάλματος σε συνάρτηση με το χρόνο, περίπτωση 2.2	72
7.17	Η εξέλιξη του αριθμού των κέντρων των δύο δικτύων, περίπτωση 2.2	73
8.1	Η πορεία της μεταβλητής εξόδου και της επιθυμητής τιμής σε συνάρτηση με το χρόνο, περίπτωση 1.2.	78
8.2	Η πορεία του σφάλματος σε συνάρτηση με το χρόνο, περίπτωση 1.2.	79
8.3	Η εξέλιξη του αριθμού των κέντρων των δύο δικτύων, περίπτωση 1.2.	79
8.4	Η πορεία της μεταβλητής εξόδου και της επιθυμητής τιμής σε συνάρτηση με το χρόνο, περίπτωση 1.3.	81
8.5	Η πορεία του σφάλματος σε συνάρτηση με το χρόνο, περίπτωση 1.3.	81
8.6	Η εξέλιξη του αριθμού των κέντρων των δύο δικτύων, περίπτωση 1.3.	82
8.7	Η πορεία της μεταβλητής εξόδου και της επιθυμητής τιμής σε συνάρτηση με το χρόνο, διαφορετικές επιθυμητές τιμές.	84
8.8	Η πορεία του σφάλματος σε συνάρτηση με το χρόνο, περίπτωση 2.1.	84
8.9	Η εξέλιξη του αριθμού των κέντρων των δύο δικτύων, περίπτωση 2.1.	85
8.10	Η πορεία της μεταβλητής εξόδου και της επιθυμητής τιμής σε συνάρτηση με το χρόνο, περίπτωση 2.2.	86
8.11	Η πορεία του σφάλματος σε συνάρτηση με το χρόνο, περίπτωση 2.2.	86
8.12	Η εξέλιξη του αριθμού των κέντρων των δύο δικτύων, περίπτωση 2.2.	87
8.13	Η πορεία της μεταβλητής εξόδου και της επιθυμητής τιμής σε συνάρτηση με το χρόνο, περίπτωση 3.1	91

8.14	Η πορεία του σφάλματος σε συνάρτηση με το χρόνο, περίπτωση 3.1	91
8.15	Η εξέλιξη του αριθμού των κέντρων των δύο δικτύων, περίπτωση 3.1	92
8.16	Η πορεία της μεταβλητής εξόδου και της επιθυμητής τιμής σε συνάρτηση με το χρόνο, περίπτωση 3.2	92
8.17	Η πορεία του σφάλματος σε συνάρτηση με το χρόνο, περίπτωση 3.2	93
8.18	Η εξέλιξη του αριθμού των κέντρων των δύο δικτύων, περίπτωση 3.2	93

Κατάλογος πινάκων

6.1	Αλγόριθμος ασαφών μέσων με μη συμμετρική διαμέριση του χώρου	43
7.1	Τιμές Σχεδιασμού περίπτωση 1.1	60
7.2	Τιμές Σχεδιασμού Περίπτωση 1.2	62
7.3	Τιμές Σχεδιασμού Περίπτωση 1.3	65
7.4	Τιμές Σχεδιασμού Περίπτωση 2.1	69
8.1	Τιμές Σχεδιασμού Περίπτωση 1.2	78
8.2	Τιμές Σχεδιασμού Περίπτωση 1.3	80
8.3	Τιμές Σχεδιασμού Περίπτωση 2.1,2.2	83
8.4	Παράμετροι αντιδραστήρα	88
8.5	Τιμές παραμέτρων δυναμικής συμπεριφοράς αντιδραστήρα	89
8.6	Τιμές Σχεδιασμού Περίπτωση Χημικού αντιδραστήρα με εξώθερμη αντίδραση	89
9.1	Συνοπτικός πίνακας αποτελεσμάτων (SSE)	96

Κεφάλαιο 1

Εισαγωγή

Μία από τις πιο ενδιαφέρουσες τεχνικές μηχανικής μάθησης είναι αυτή της ενισχυτικής μάθησης και ειδικότερα οι μεθοδολογίες «ενέργειας-κριτικής» (actor-critic). Οι μεθοδολογίες αυτές έχουν ξεχωριστές και ανεξάρτητες δομές μνήμης. Η πρώτη χρησιμοποιείται για την επιλογή των ενεργειών και η δεύτερη για την αξιολόγησή τους. Ανάμεσα στις πολλές εφαρμογές της, η ενισχυτική μάθηση έχει χρησιμοποιηθεί για το σχεδιασμό ρυθμιστών για μη γραμμικά συστήματα και διεργασίες.

Τα νευρωνικά δίκτυα είναι ένα σύνολο από ισχυρά μαθηματικά εργαλεία [1], τα οποία προσομοιώνουν τον τρόπο που ο ανθρώπινος εγκέφαλος επεξεργάζεται πληροφορίες και μαθαίνει. Τα νευρωνικά δίκτυα έχουν την ικανότητα να εντοπίζουν και να μαθαίνουν πολύπλοκες και μη γραμμικές συσχετίσεις, χωρίς να απαιτείται εκ των προτέρων καμία γνώση του συστήματος αλλά με βάση μόνο δεδομένα εισόδου-εξόδου.

Η ενισχυτική μάθηση σε συνδυασμό με τα νευρωνικά δίκτυα έχει χρησιμοποιηθεί ευρέως στη ρύθμιση μη γραμμικών συστημάτων [2]. Στη βιβλιογραφία υπάρχουν πολλές μέθοδοι οι οποίες βασίζονται σε νευρωνικά δίκτυα προς τα εμπρός τροφοδότησης [2]. Οι εν λόγω μέθοδοι επιλέγουν τυχαία, πριν ξεκινήσει η αλληλεπίδραση, τον αριθμό των νευρώνων και τις συνάψεις της κρυφής στοιβάδας, οι οποίες δε μεταβάλλονται κατά τη διαδικασία μάθησης. Σε πραγματικό χρόνο ενημερώνονται μόνο οι συνδέσεις ανάμεσα στην κρυφή στοιβάδα και τη στοιβάδα εξόδου.

Στην παρούσα εργασία παρουσιάζεται μία πρωτότυπη μεθοδολογία ρύθμισης «ενέργειας-

κριτικής», στην οποία νευρωνικά δίκτυα ακτινικής συνάρτησης βάσης (RBF neural networks) χρησιμοποιούνται για την κατασκευή και την ενημέρωση σε πραγματικό χρόνο τόσο του δικτύου ενέργειας όσο και του δικτύου κριτικής του συστήματος ρύθμισης. Το πρώτο δίκτυο προσεγγίζει τη μη γραμμική δυναμική του συστήματος και παράγει ένα μέρος της ρυθμιστικής δράσης, ενώ το δεύτερο δίκτυο προσεγγίζει τη συνάρτηση κόστους με στόχο να ανανεώσει τα βάρη του νευρωνικού δικτύου ενέργειας. Η εκπαίδευση των δύο νευρωνικών δικτύων γίνεται με την μέθοδο των ασαφών μέσων (Fuzzy Means – FM) [3], και συγκεκριμένα την online έκδοση του αλγορίθμου [4], που επιτρέπει (με βάση πληροφορίες που λαμβάνονται από το περιβάλλον) την πλήρη κατασκευή και αναπροσαρμογή των δικτύων σε πραγματικό χρόνο, συμπεριλαμβανομένης της ίδιας της δομής των δικτύων, δηλαδή του πλήθους των νευρώνων της κρυφής στοιβάδας, αλλά και των κέντρων των συναρτήσεων ενεργοποίησης. Η προτεινόμενη μεθοδολογία εφαρμόστηκε σε ένα μη γραμμικό μη αφινικό (non-affine) δυναμικό σύστημα και σε δύο παραδείγματα μη γραμμικού αντιδραστήρα, στο ένα πραγματοποιούνται παράλληλες αντιδράσεις ενώ στο άλλο η αντίδραση είναι εξώθερμη.

Η δομή της παρούσας εργασίας είναι η ακόλουθη. Στο κεφάλαιο 2 γίνεται εισαγωγή στα νευρωνικά δίκτυα, στα είδη τα οποία χωρίζονται στις ιδιότητες τους αλλά και τους διάφορους τρόπους εκπαίδευσης. Στο κεφάλαιο 3 και 4 παρουσιάζονται οι βασικές αρχές του Δυναμικού προγραμματισμού και της ενισχυτικής Μάθησης αντίστοιχα. Στο κεφάλαιο 5 παρουσιάζονται τα νευρωνικά δίκτυα ακτινικής βάσης, τα είδη και οι ιδιότητες τους. Στο κεφάλαιο 6 περιγράφεται αναλυτικά οι μεθοδολογίες Fuzzy Means (FM) [3],[5] και online FM [4]. Στη συνέχεια στο κεφάλαιο 7 αναλύεται η πρώτη προσέγγιση του σχεδιασμού του ρυθμιστή, ενώ στο κεφάλαιο 8 περιγράφεται ο νέος τρόπος ρύθμισης.

Κεφάλαιο 2

Τεχνητά Νευρωνικά Δίκτυα

Τα τεχνητά Νευρωνικά δίκτυα είναι απλοποιημένα μοντέλα του κεντρικού νευρικού συστήματος του ανθρώπου, μιμούνται τη λειτουργία αλλά και τη δομή των βιολογικών νευρωνικών δικτύων. Αποτελούνται από διασυνδεδεμένα υπολογιστικά στοιχεία που έχουν την ικανότητα να ανταποκρίνονται σε ερεθίσματα που δέχονται στην είσοδο τους και να μαθαίνουν να προσαρμόζονται στο περιβάλλον τους. Η τελευταία ικανότητα τους είναι πολύ κρίσιμη για τη χρήση τους σε προβλήματα της μηχανικής. [6]

2.1 Ιστορική αναδρομή

Για πολλά χρόνια γινόντουσαν μελέτες σχετικά με τη δράση των νευρώνων στον ανθρώπινο εγκέφαλο, το 1943 όμως ο νευροφυσιολόγος Warren McCulloch και ο μαθηματικός Walter Pitts δημοσίευσαν μια εργασία όπου σχετικά με το πώς λειτουργούν οι νευρώνες, σκοπεύοντας να περιγράψουν τον τρόπο λειτουργίας του εγκεφάλου, ενώ μοντελοποίησαν ένα πρώιμο και απλό νευρωνικό δίκτυο χρησιμοποιώντας ηλεκτρικά κυκλώματα. Το 1949, ο Donald Hebb δημοσίευσε την εργασία με τίτλο 'The organization of Behavior' όπου τόνιζε το γεγονός ότι οι νευρώνες εξελίσσονται συνεχώς, η ιδέα αυτή ήταν από τις βασικότερες και ουσιώδεις για τον τρόπο όπου ο άνθρωπος μαθαίνει. Καθώς εξελίχθηκαν οι ηλεκτρονικοί υπολογιστές τη δεκαετία του 50', ήταν πλέον πιθανή η προσομοίωση ενός υποθετικού

νευρωνικού δικτύου. Τα πρώτα βήματα έγιναν από τον Rochester από τα ερευνητικά εργαστήρια IBM, ωστόσο οι πρώτες προσπάθειες απέτυχαν. 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. Το 1959, οι Bernard Widrow και Macarian Hoff από το Stanford ανέπτυξαν δύο μοντέλα που ονομάζονταν ‘ADALINE’ και ‘MADALINE’, όπου σημαίνουν ADaptive LINear Elements και Multiple ADaptive LINear Elements αντίστοιχα. Το πρώτο αναπτύχθηκε με σκοπό να αναγνωρίζει δυαδικά πρότυπα και να αφαιρεί τον θόρυβο που δημιουργείται.

Υπήρξε η περίοδος της δυσφήμισης κατά την οποία μετά την έκδοση ενός βιβλίου από τους Minsky και Papert στο οποίο τόνιζαν τους περιορισμούς ενός Αισθητήρα απλού στρώματος μειώθηκε δραματικά η χρηματοδότηση για την έρευνα στα νευρωνικά δίκτυα και προκλήθηκε προκατάληψη απέναντι σε αυτόν τον τομέα και τις δυνατότητες των δικτύων. Παρά την μείωση της χρηματοδότησης, αρκετοί ερευνητές συνέχισαν να ασχολούνται με τη δημιουργία υπολογιστικών μεθόδων που να βασίζονται σε δίκτυα νευρώνων για προβλήματα όπως η αναγνώριση προτύπων. Μεταξύ των επιστημόνων που συνέβαλλαν στην ανάπτυξη του κλάδου ήταν και οι Grossberg και Gail οι οποίοι ανακάλυψαν τα ART(Adaptive Resonance Theory) δίκτυα βασιζόμενοι σε βιολογικά μοντέλα. Παράλληλα οι Anderson και Kohonen δημιούργησαν παρόμοια δίκτυα. Ο Klopff το 1972 ανέπτυξε μία μέθοδο εκμάθησης στα τεχνητά νευρωνικά δίκτυα, βασιζόμενος σε μια βασική αρχή εκμάθησης των νευρώνων την ετερόσταση. Ο Werbos το 1974 δημιούργησε και χρησιμοποίησε τη μέθοδο εκμάθησης back-propagation (ανάδρασης) που πέρασαν αρκετά χρόνια πριν αρχίσει να χρησιμοποιείται ευρέως. Υπήρξαν και αρκετοί άλλοι επιστήμονες οι οποίοι συνέβαλλαν στην ανάπτυξη του κλάδου όπως ήταν οι Amari και Fukushima. [7] Το 1982, ο John Hopfield από το CalTech παρουσίασε ένα άρθρο στη National

Academy of Sciences. Η προσέγγιση του σκοπεύε στη δημιουργία μηχανών που χρησιμοποιούν αμφίδρομες συνδέσεις, ενώ μέχρι τώρα ήταν ενός δρόμου όλες οι συνδέσεις. Η νέα αυτή μέθοδος πρέπει να τονισθεί ότι δεν έλυσε για πρώτη φορά γνωστά άλυτα προβλήματα, αφού όλα τα προβλήματα με τα οποία ασχολείται είχαν λυθεί με άλλες μεθόδους. Μάλιστα οι λύσεις που προτάθηκαν δεν ήταν καν οι καλύτερες, συνήθως ήταν αργές. Εν τούτοις η χρήση των δικτύων αυτών είναι πολύ χρήσιμη για την διορατικότητα που μας δίνει για την φύση του προβλήματος. Την ίδια χρονιά οι Reilly και Cooper χρησιμοποίησαν ένα Υβριδικό δίκτυο με πολλές στοιβάδες όπου κάθε στοιβάδα χρησιμοποιείτο για την επίλυση διαφορετικού προβλήματος. Εκείνη την περίοδο σε ένα συνέδριο τεχνητής νοημοσύνης στην Ιαπωνία σηματοδοτήθηκε η επανεκκίνηση ερευνών γύρω από τον κλάδο των νευρωνικών δικτύων και τότε η Αμερική αντικρίζοντας τον κίνδυνο να μείνει πίσω στον ερευνητικό αυτό τομέα των νευρωνικών δικτύων ξεκίνησε τη χρηματοδότηση στο τομέα αυτό.

Το 1986, τρεις ανεξάρτητες ερευνητικές ομάδες, όπου η μια συμπεριελάμβανε το David Rumelhart, πρώην μέλος του τμήματος ψυχολογίας του πανεπιστημίου του Stanford, προέβλεψαν παρόμοιες ιδέες, τα λεγόμενα οπισθοδιάδοσης νευρωνικά δίκτυα (back propagation neural networks), καθώς καταμερίζεται το σφάλμα της αναγνώρισης προτύπων μέσα από το δίκτυο. Τα υβριδικά δίκτυα χρησιμοποιούσαν δύο στοιβάδες, ενώ αυτά πολλές. Το αποτέλεσμα είναι τα τελευταία να είναι αργά στην μάθηση χρειάζοντας χιλιάδες επαναλήψεις. Σήμερα, τα νευρωνικά δίκτυα έχουν μεγάλη εξέλιξη και επιλύουν επιτυχώς ένα μεγάλο εύρος προβλημάτων, συμπεριλαμβάνοντας την αναζήτηση τάσης σε πολύ μεγάλες ποσότητες δεδομένων. [8]

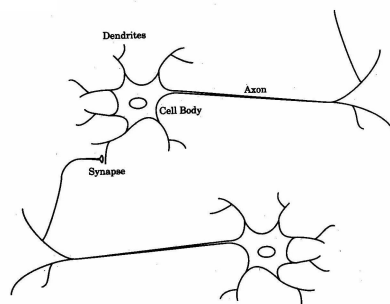
2.2 Βιολογική Προέλευση

Το επιστημονικό πεδίο των Τεχνητών Νευρωνικών Δικτύων (ΤΝΔ) βασίστηκε στο γεγονός ότι ο ανθρώπινος εγκέφαλος εκτελεί υπολογισμούς με εντελώς διαφορετικό τρόπο από το συμβατικό υπολογιστή. Η αναλογία μεταξύ νευροφυσιολογίας του ανθρώπινου εγκεφάλου και των ΤΝΔ είναι πολύ σημαντική για την κατανόησή τους. Ο νευρώνας είναι ένας εξειδικευμένος τύπος κυττάρου που αποτελεί βασική μονάδα των συστημάτων επεξεργασίας πληροφοριών που απαρτίζουν

το νευρωνικό σύστημα του ανθρώπου. [9] Είναι το βασικό δομικό στοιχείο του εγκεφάλου τόσο στον άνθρωπο όσο και στα ζώα. Ο εγκέφαλος ενός νεογέννητου ανθρώπου αποτελείται από περίπου 100 δισεκατομμύρια νευρώνες κάθε ένας από τους οποίους συνδέεται με περίπου 1000 άλλους νευρώνες. Αυτό πραγματοποιείται μέσω του άξονα κάθε νευρώνα στον οποίο καταλήγουν ισάριθμοι δενδρίτες άλλων νευρώνων. [10],[11]

Εκτιμάται ότι ο μέσος ανθρώπινος εγκέφαλος αποτελείται από 109 νευρώνες που συνδέονται μεταξύ τους με διάφορους τρόπους. Οι νευρώνες είναι στοιχεία μνήμης και λογικών κυκλωμάτων. Ανατομικά ο νευρώνας αποτελείται από το σώμα (body) που αποτελεί τον πυρήνα του, τους δενδρίτες (dendrites) μέσω των οποίων λαμβάνει σήματα από γειτονικούς νευρώνες (σημεία εισόδου) και τον άξονα (axon) που είναι η έξοδος του νευρώνα και το μέσο σύνδεσης του με άλλους νευρώνες. Σε κάθε δενδρίτη υπάρχει ένα απειροελάχιστο κενό που ονομάζεται σύναψη (synapse). Οι συνάψεις αποτελούν τα σημεία όπου συνδέονται οι διακλαδώσεις του άξονα ενός νευρώνα με τους δενδρίτες άλλων νευρώνων. Το υλικό τους μεταδίδει ηλεκτρικό σήμα από τον αποστολέα δενδρίτη στο παραλήπτη δενδρίτη. Η πυκνότητα αλλά και οι αποστάσεις του υλικού επηρεάζουν την ευκολία που πραγματοποιείται η ηλεκτρική δραστηριότητα. Το ποσοστό που μεταδίδεται στο παραλήπτη ονομάζεται συναπτικό βάρος. Στο σχήμα 2.1 απεικονίζεται ένας βιολογικός νευρώνας. Ο νευρώνας λειτουργεί σαν δυαδικό στοιχείο καθώς αν το άθροισμα των ηλεκτρικών σημάτων που λαμβάνει είναι μεγαλύτερο από κάποιο κατώφλι τότε αρχίζει να παράγει παλμούς με μεγάλη συχνότητα, όπου τότε αναφέρεται ότι πυροβολεί, ενώ όταν δεν συμβαίνει αυτό το γεγονός ο νευρώνας παράγει πολύ αραιούς παλμούς σε τυχαίες χρονικές στιγμές όποτε θεωρείται ότι παραμένει αδρανής. Οι αριθμοί των μονάδων των νευρώνων και οι συνδέσεις τους στον άνθρωπο ως τάξη μεγέθους είναι πολύ μεγαλύτεροι από τους αριθμούς μονάδων που μπορεί να χειριστεί εύκολα σήμερα ένας υπολογιστής. Αυτό είναι λογικό εξαιτίας της πολυπλοκότητας του ανθρώπινου εγκεφάλου και των διεργασιών που επιτελεί. Τα ΤΝΔ αδυνατούν να πραγματοποιήσουν δραστηριότητες που ακόμα και ο εγκέφαλος ενός μικρού παιδιού κάνει πολύ εύκολα. Από την άλλη τα ΤΝΔ μπορούν να λύσουν προβλήματα που ο ανθρώπινος εγκέφαλος δεν μπορεί να πραγματοποιήσει με ευκολία όπως μαθηματικά προβλή-

ματα, αναγνώριση προτύπων ή πρόβλεψη. Ο κύρια διαφορά μεταξύ των βιολογικών και των τεχνητών είναι η πολυπλοκότητα στις συνάψεις όπου στα βιολογικά είναι πολύ πιο περίπλοκες, ενώ ακόμα εμφανίζονται μεγαλύτερες ταχύτητες στη διάδοση των σημάτων στα τεχνητά νευρωνικά δίκτυα, ωστόσο αυτή η διαφορά στη ταχύτητα δεν αρκεί για να αντισταθμίσει την διαφορά τους στη πολυπλοκότητα. Είναι σημαντικό να τονισθεί ότι ο εγκέφαλος μαθαίνει πολύ πιο εύκολα και γρήγορα από τον πιο γρήγορο ηλεκτρονικό υπολογιστή, επίσης ο ανθρώπινος εγκέφαλος κάνει σύγχρονη και ασύγχρονη ενημέρωση των μονάδων του, άρα πραγματοποιείται η εκπαίδευση με συνεχή τρόπο ενώ τα ΤΝΔ κάνει μόνο σύγχρονη ενημέρωση, δηλαδή σε διακριτό χρόνο. [12],[13]



Σχήμα 2.1: Βιολογικά Νευρωνικά Δίκτυα

2.3 Το μαθηματικό μοντέλο

Ένας νευρώνας είναι μονάδα επεξεργασίας πληροφορίας, η οποία είναι θεμελιώδης για τη λειτουργία ενός νευρωνικού δικτύου. Το σχηματικό διάγραμμα του Σχήματος παρουσιάζει το μοντέλο ενός νευρώνα που αποτελεί τη βάση για τη σχεδίαση μιας μεγάλης οικογένειας νευρωνικών δικτύων. Τα βασικά στοιχεία αυτού του μοντέλου είναι

1. **Τα δεδομένα εισόδου** όπου είναι ερεθίσματα από το εξωτερικό περιβάλλον, τα οποία μπορεί να αποτελούν και αποτελέσματα άλλων νευρωνικών δικτύων.

2. **Ένα σύνολο συνάψεων**, κάθε μια εκ των οποίων χαρακτηρίζεται από το δικό της βάρος ή δύναμη. Συγκεκριμένα, ένα σήμα x_j στην είσοδο της σύναψης j που συνδέεται με τον νευρώνα k πολλαπλασιάζεται επί το συναπτικό βάρος w_{kj} .
3. **Ο αθροιστής (adder)** για την άθροιση των σημάτων εισόδου, σταθμισμένων από τα αντίστοιχα συναπτικά βάρη του νευρώνα, οι λειτουργίες που περιγράφονται εδώ συνιστούν έναν γραμμικό συνδυαστή (linear combiner)
4. **Η συνάρτηση ενεργοποίησης (activation function)** για τον περιορισμό του πλατους του σήματος εξόδου ενός νευρώνα. Τυπικά το κανονικοποιημένο εύρος τιμών της εξόδου ενός νευρώνα είναι γράφεται ως μοναδιαίο κλειστό διάστημα, με τη μορφή $[0,1]$ ή $[-1,1]$
5. **Το κατώφλι (Threshold) θ** . 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.
6. **Δεδομένο εξόδου (Output)** Αποτελεί το τελικό προϊόν του νευρώνα και είναι το αποτέλεσμα της εφαρμογής της συνάρτησης ενεργοποίησης στο σταθμισμένο άθροισμα των δεδομένων εισόδου αφού έχουμε αφαιρέσει την παράμετρο του κατωφλιού. Ο τύπος του δεδομένου εξόδου εξαρτάται άμεσα από την συνάρτηση ενεργοποίησης. [14]

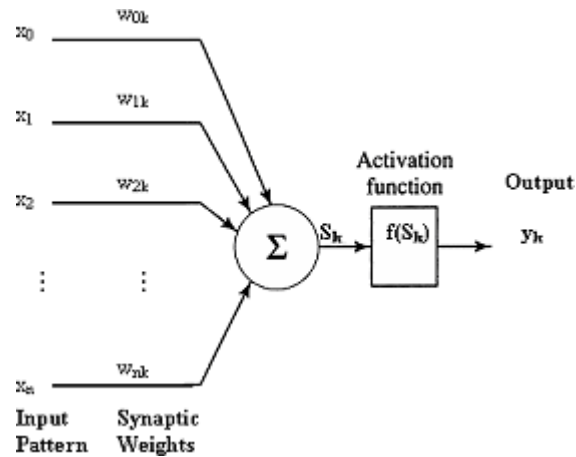
Η μαθηματική σχέση που περιγράφει έναν νευρώνα παρουσιάζεται παρακάτω.

$$v_k = \sum_{j=0}^m (w_{jk}x_j) \quad (2.1)$$

$$y_k = f(v_k) \quad (2.2)$$

Όπου x_j αποτελεί το σήμα εισόδου με $j=1 \dots m$, w_{jk} αποτελούν τα συναπτικά βάρη του νευρώνα k , v_k είσοδος του γραμμικού συνδιαστή, $\theta=w_{k0}$ x_0 το κατώφλι με $x_0=\pm 1$, $f(\cdot)$ η συνάρτηση ενεργοποίησης και y_k είναι το σήμα εξόδου του νευρώνα k , όπως φαίνεται και στο σχήμα 2.2.

Η συνάρτηση ενεργοποίησης αποτελεί κρίσιμο παράγοντα για τη συμπεριφορά του νευρωνικού δικτύου. Η συνάρτηση ενεργοποίησης μπορεί να είναι η ίδια για όλο δίκτυο ή και διαφορετική. Στη συνέχεια παρατίθενται κάποιες βασικές συναρτήσεις ενεργοποίησης.



Σχήμα 2.2: Ένας Νευρώνας

Βηματική. Η συνάρτηση η οποία αναφέρεται και ως *heavyside*, περιγράφεται από την παρακάτω εξίσωση

$$f(u) = \begin{cases} -1, & \text{if } u \leq 0 \\ 1, & \text{if } u > 0 \end{cases} \quad (2.3)$$

Γραμμική συνάρτηση ενεργοποίησης που είναι η

$$f(u) = \begin{cases} 0, & \text{if } u \leq 0 \\ u, & \text{if } 0 < u < 1 \\ 1, & \text{if } u \geq 1 \end{cases} \quad (2.4)$$

Σύνηθες είναι να χρησιμοποιούνται μη γραμμικές συναρτήσεις ενεργοποίησης, μερικές από αυτές είναι η **συγμοειδής (sigmoid)** και η **υπερβολική εφαπτομένη (hyperbolic tangent)**

$$f(u) = \frac{1}{1 + e^{-u}} \quad (2.5)$$

$$f(u) = \frac{1 - e^{-u}}{1 + e^{-u}} \quad (2.6)$$

2.4 Αρχιτεκτονικές νευρωνικών δικτύων

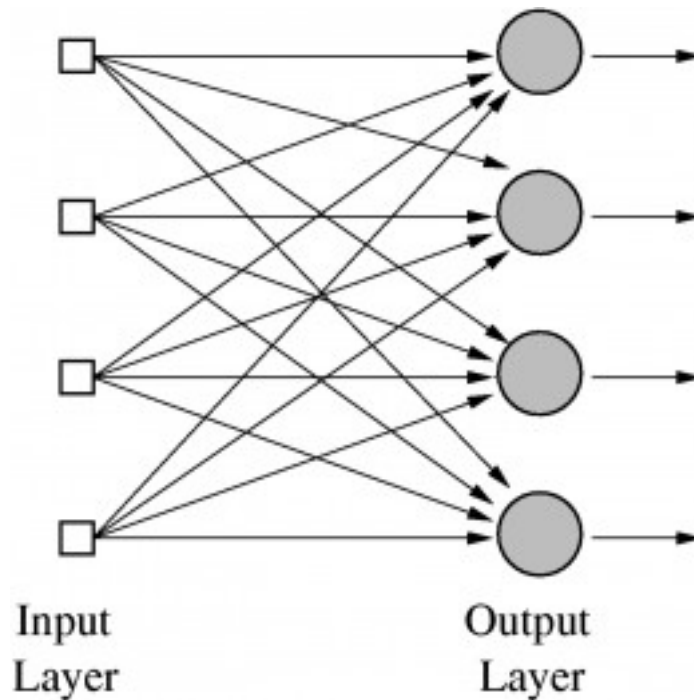
Βασικά στοιχεία που χαρακτηρίζουν τα νευρωνικά δίκτυα είναι ο αριθμός των επιπέδων των νευρώνων (layers) και οι συνδέσεις μεταξύ των νευρώνων. Ένα ακόμα χαρακτηριστικό το οποίο επηρεάζεται η τοπολογική δομή των νευρωνικών δικτύων είναι ο τρόπος εκπαίδευσης. Η απλούστερη και ευρύτερα χρησιμοποιούμενη 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 Ενός επιπέδου νευρωνικά δίκτυα πρόσθιας τροφοδότησης

Σε ένα νευρωνικό δίκτυο, οι νευρώνες οργανώνονται σε μορφή επιπέδων. Η πιο απλή μορφή δικτύων που έχουν σχεδιαστεί αποτελείται ένα επίπεδο εισόδου το οποίο συνδέεται απευθείας με ένα επίπεδο νευρώνων, και ποτέ αντίστροφα. Έτσι αυτά τα δίκτυα ονομάστηκαν πρόσθιας τροφοδότησης (feedforward), η μορφή του οποίου απεικονίζεται στο σχήμα 2.3. Ένα τέτοιο δίκτυο αποκαλείται ενός επιπέδου, καθώς υπάρχει ένα επίπεδο εξόδου από τους νευρώνες. Με το επίπεδο εισόδου να μην προσμετράται σε αυτή τη μέτρηση.

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_m) υπολογίζονται από το ίδιο σύνολο τιμών εισόδου, κάθε έξοδος εξαρτάται από τα συναπτικά βάρη. 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, έτσι οι συνδέσεις αυτές θα έχουν τιμή για το συναπτικό βάρος ίσο με μηδέν. [9]



Σχήμα 2.3: Ενός επιπέδου νευρωνικά δίκτυα

2.4.2 Πολυεπίπεδα δίκτυα πρόσθιας τροφοδότησης

Αυτή η αρχιτεκτονική νευρωνικών δικτύων αποτελείται από πολλά ενδιάμεσα επίπεδα νευρών (hidden layers) και το επίπεδο εξόδου (output layer). Οι έξοδοι του πρώτου κρυφού επιπέδου είναι είσοδος του δεύτερου κρυφού επιπέδου και η διαδικασία αυτή συνεχίζεται μέχρι να φτάσει το επίπεδο εξόδου. Η επιλογή τόσο του αριθμού των νευρώνων κάθε επιπέδου αλλά και του αριθμού των επιπέδων είναι συμβιβασμός μεταξύ της πολυπλοκότητας του προβλήματος και του υπολογιστικού κόστους.

Η εκπαίδευση ενός δικτύου πολλών στρωμάτων είναι η διαδικασία ρύθμισης των συναπτικών βαρών του έτσι ώστε να ικανοποιείται κάποιο κριτήριο καταλληλότητας.

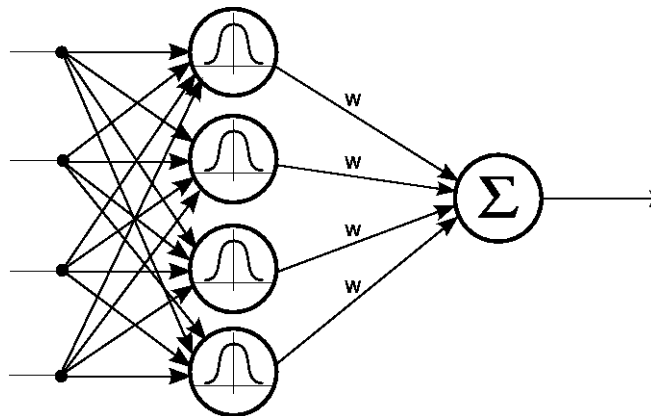
Αυτός άλλωστε είναι και ο στόχος της εκπαίδευσης σε οποιοδήποτε νευρωνικό δίκτυο. Ένας πολύ διαδεδομένος αλγόριθμος εκπαίδευσης των νευρωνικών δικτύων είναι ο Back Propagation, όπου αποτελεί μέθοδος εκπαίδευση με επίβλεψη. Αρχικά υπολογίζεται η έξοδος του νευρωνικού δικτύου και στη συνέχεια γίνεται αντίστροφη σάρωση όπου πραγματοποιείται ανανέωση των συναπτικών βαρών, συγκρίνοντας την έξοδο του δικτύου με την αντίστοιχη επιθυμητή τιμή. [9]

2.4.3 Δίκτυα ακτινικής βάσης RBF

Τα νευρωνικά δίκτυα ακτινικής βάσης είναι δίκτυα δύο στρωμάτων όπου το κρυφό επίπεδο αποτελείται από συναρτήσεις βάσης όπου έχουν σαν είσοδο την ευκλείδεια απόσταση μεταξύ του διανύσματος του σήματος εισόδου και ενός πρότυπου διανύσματος που ονομάζεται κέντρο του κάθε νευρώνα. Οι συναρτήσεις αυτές είναι ακτινικής βάσης, μεταξύ των οποίων είναι η συνάρτηση Gauss και TSP. Στο σχήμα 2.4 παρουσιάζεται η τοπολογία ενός δικτύου RBF. Αυτά τα δίκτυα θα μπορούσαν να αναφερθούν σαν πολυεπίπεδα νευρωνικά δίκτυα, ωστόσο τόσο οι διαφορές στο τρόπο εκπαίδευσης, της δομής και λειτουργίας τα καθιστά ξεχωριστά νευρωνικά δίκτυα. 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.

1. Προσδιορισμός των πρότυπων διανυσμάτων (κέντρο και ακτίνα)
2. Προσαρμογή των βαρών του στρώματος εξόδου.

Το γεγονός ότι μπορεί να υπολογιστούν πρώτα τα χαρακτηριστικού των συναρτήσεων βάσης και ύστερα η προσαρμογή των βαρών, προσθέτει σε αυτά τα δίκτυα πολλά πλεονεκτήματα. [14] Τα δίκτυα αυτού του τύπου χρησιμοποιήθηκαν στη παρούσα εργασία και για αυτό το λόγο θα αναλυθούν περαιτέρω στη συνέχεια.



Σχήμα 2.4: Νευρωνικά Δίκτυα Ακτινικής Βάσης

2.5 Εκπαίδευση Νευρωνικών Δικτύων

Τα νευρωνικά δίκτυα αποτελούν ένα πολύπλοκο πλέγμα μαθηματικών συναρτήσεων με πολλές ελεύθερες μεταβλητές-παράμετροι, έτσι αλληλεπιδρώντας με το περιβάλλον μπορούν να προσαρμοστούν αυτές οι μεταβλητές ώστε να πετύχουν τον εκάστοτε στόχο. Ο τρόπος με τον οποίο αλληλεπιδρούν τα νευρωνικά δίκτυα με το περιβάλλον, καθορίζει και τον τύπο της εκπαίδευσης. Έχει αναπτυχθεί ένας μεγάλος αριθμός αλγορίθμων εκπαίδευσης με τις κυριότερες να είναι η εκπαίδευση με επίβλεψη, χωρίς επίβλεψη και ενίσχυση.

2.5.1 Επιβλεπόμενη μάθηση

Στην επιβλεπόμενη μάθηση δίνεται ένα μεγάλο πλήθος δεδομένων εισόδου-εξόδου στοχεύοντας να γίνει η κατάλληλη αλλαγή των παραμέτρων ώστε να ταιριάζουν οι εξοδοί του δικτύου με τα δεδομένα. Χρησιμοποιείται μια συνάρτηση κόστους που συγκρίνει τις εξόδους των νευρωνικών δικτύων και των δεδομένων, αλλάζοντας τις παραμέτρους του συστήματος με στόχο να ελαχιστοποιείται κάποια συνάρτηση κόστους.

Η πιο συνηθισμένη συνάρτηση κόστους είναι το μέσο τετραγωνικό σφάλμα (means squared error), που προσπαθεί να ελαχιστοποιήσει το μέσο σφάλμα των τετραγώνων

των διαφορών μεταξύ των εξόδων των δικτύων και των επιθυμητών τιμών. Όπως αναφέρθηκε και πριν μια πολύ γνωστή μέθοδος εκπαίδευσης με αυτόν τον τρόπο είναι με τον αλγόριθμο της οπισθοδιάδοσης (Backpropagation), χρησιμοποιώντας τη μέθοδο της κλίσης (gradient descent). Αυτός ο τρόπος εκπαίδευσης χρησιμοποιείται πολύ συχνά στην αναγνώριση προτύπων (pattern recognition-classification) και τη προσέγγιση συναρτήσεων (regression-function approximation). Οι παράμετροι που επηρεάζουν κυρίως αυτόν τον τρόπο μάθησης είναι οι εξής:

1. *Η ποσότητα των διαθέσιμων δεδομένων*

Αν η συνάρτηση που προσπαθεί να προσεγγίσει το δίκτυο είναι γενικά απλή τότε δεν χρειάζεται μεγάλη ποσότητα δεδομένων, ωστόσο αν η αληθινή αυτή συνάρτηση είναι πολύ περίπλοκη (καθώς μπορεί να υπάρχουν πολύ διαφορετικές αλληλεπιδράσεις μεταξύ των σημάτων εισόδου) τότε είναι αναγκαίο να υπάρχει μεγάλος όγκος δεδομένων.

2. *Διάσταση του χώρου εισόδου*

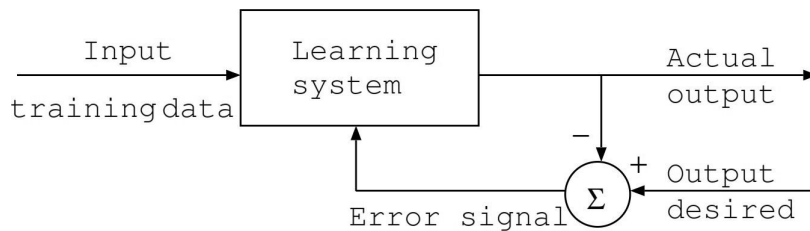
Η διάσταση των μεταβλητών εισόδου δεν θα πρέπει να είναι μεγάλη καθώς, αν είναι τότε πιθανότατα το δίκτυο θα μπερδευτεί στις συσχετίσεις, και έτσι θα πρέπει να ελέγχεται έτσι ώστε ο χώρος να είναι στο κατάλληλο μέγεθος. Έχουν αναπτυχθεί πολλές τεχνικές σχετικά με τη μείωση του χώρου εισόδου. [15]

3. *Θόρυβος στις επιθυμητές τιμές*

Ένας ακόμα πολύ σημαντικός παράγοντας είναι ο θόρυβος που μπορεί να εμφανιστεί στις επιθυμητές τιμές (target values). Αν υπάρχει θόρυβος λόγω ανθρώπινου λάθους ή λάθος του αισθητήρα, δεν θα καταφέρει να ταιριάζει ακριβώς τα αποτελέσματα του με τα επιθυμητά, ενώ αν γίνει προσπάθεια με υπερβολική αύξηση των δεδομένων ή του δικτύου υπάρχει κίνδυνος για υπερπροσαρμογή. [16]

Η δομή αυτού του τύπου εκπαίδευσης φαίνεται στο σχήμα 2.5, που αποτελείται από το δάσκαλο και το περιβάλλον, το περιβάλλον δίνει τις μεταβλητές εισόδου, και ο δάσκαλος τις επιθυμητές τιμές έτσι σε κάθε περίπτωση γίνεται αξιολόγηση

από τον δάσκαλο, αλλάζοντας τις παραμέτρους του δικτύου ώστε να βελτιωθεί η απόδοση.



Σχήμα 2.5: Μάθηση με επίβλεψη

2.5.2 Ενισχυτική Μάθηση

Η ενισχυτική μάθηση είναι ένα πρόβλημα μάθησης συμπεριφοράς. Σε αντίθεση με την επιβλεπόμενη μάθηση όπου χρειάζονται εξαρχής δεδομένα εισόδου εξόδου, η παρούσα μέθοδος χρησιμοποιεί την αλληλεπίδραση ενός κριτή-πράκτορα με το περιβάλλον του. [1] Συγκεκριμένα ο κριτής προσπαθεί να πετύχει έναν συγκεκριμένο στόχο, συνήθως τη βελτιστοποίηση μιας συνάρτησης κόστους, ενώ συνυπάρχουν αβεβαιότητες. [17] Το γεγονός αυτό, συνυπολογίζοντας ότι πολλές φορές είναι πολύ δύσκολο να συγκεντρωθεί ο κατάλληλος αριθμός δεδομένων εισόδου-εξόδου, καθιστά την ενισχυτική μάθηση ιδιαίτερα ελκυστική. Η σύγχρονη προσέγγιση της ενισχυτικής μάθησης βασίζεται στον Δυναμικό προγραμματισμό [1], για την επιλογή της πορείας μιας σειράς ενεργειών. Ο δυναμικός προγραμματισμός είναι μία υπολογιστική μέθοδος η οποία εφαρμόζεται όταν πρόκειται να ληφθεί μία σύνθετη απόφαση η οποία προκύπτει από τη σύνθεση επιμέρους αποφάσεων που αλληλεξαρτώνται. Η αλληλεξάρτηση μπορεί να προκύψει επειδή οι αποφάσεις είτε παρουσιάζουν κάποια χρονική διαδοχή (όπως στην περίπτωση αναζήτησης της συντομότερης διαδρομής), είτε συνδέονται με κοινούς περιορισμούς (όπως στην περίπτωση κατανομής περιορισμένων πόρων μεταξύ ανταγωνιστικών δραστηριοτήτων). Πατέρας

του δυναμικού προγραμματισμού θεωρείται ο Richard Bellman. Η τεχνική του δυναμικού προγραμματισμού εδράζεται σε μια πολύ απλή ιδέα, που είναι γνωστή ως αρχή του βέλτιστου αποτελέσματος κατά Bellman, όπου αναφέρει ότι μια βέλτιστη πολιτική (συνολική επιλογή λειτουργιών ελέγχου) έχει την ιδιότητα ότι, όποια και αν είναι η αρχική κατάσταση και η αρχική απόφαση, οι υπόλοιπες αποφάσεις πρέπει να συνιστούν μια βέλτιστη πολιτική ξεκινώντας από την κατάσταση που προκύπτει μετά από την πρώτη απόφαση. [18]

2.5.3 Μάθηση χωρίς επίβλεψη

Στην εκπαίδευση χωρίς επίβλεψη δεν υπάρχει εξωτερικός δάσκαλος ή κριτικός για να επιβλέπει την διαδικασία εκπαίδευσης. Το νευρωνικό δίκτυο μαθαίνει από ένα μέτρο της ποιότητας της αναπαράστασης που καλείται να μάθει και οι ελεύθερες παράμετροι του δικτύου βελτιστοποιούνται με βάση το μέτρο αυτό. Για να εκτελεστεί εκπαίδευση χωρίς επίβλεψη θα 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, ενώ 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. 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. [14]

2.6 Ιδιότητες

Τα νευρωνικά δίκτυα είναι πολύπλοκες μη γραμμικές συναρτήσεις (complex nonlinear interconnected systems), και μπορούν να χρησιμοποιηθούν ευρέως με πολλές χαρακτηριστικές ιδιότητες. Δύο βασικές ιδιότητες που χαρακτηρίζουν τα νευρωνικά δίκτυα είναι η αναγνώριση προτύπων και η προσέγγιση συναρτήσεων. Οι ιδιότητες αυτές μπορούν να χρησιμοποιηθούν σε ανοιχτού βρόχου (open loop) εφαρμο-

γές, όπου τα ΤΝΔ δεν ρυθμίζουν κάποιο δυναμικό σύστημα με ανατροφοδότηση. Ωστόσο, η ιδιότητα της προσέγγισης συναρτήσεων είναι αυτή που χρησιμοποιείται ευρέως σε συστήματα κλειστού βρόχου.

Από τη μία τα ΤΝΔ είναι αρκετά πολύπλοκα ώστε να μπορούν να έχουν τις επιθυμητές ιδιότητες, αλλά χρειάζεται σε κάθε περίπτωση να υπολογιστούν οι κατάλληλες συνάψεις μεταξύ των κόμβων των νευρωνικών δικτύων. Έτσι έχουν αναπτυχθεί, και συνεχίζονται να αναπτύσσονται αλγόριθμοι που υπολογίζουν εκπαιδύοντας τα συναπτικά βάρη στο ΤΝΔ (tune the nn weights by training the nn), με τέτοιο τρόπο ώστε όταν η εκπαίδευση ολοκληρωθεί το δίκτυο να έχει τις επιθυμητές ιδιότητες.

Μπορούν να κατασκευαστούν νευρωνικά δίκτυα με πολλές κρυφές στοιβάδες, ωστόσο σε πολλές πρακτικές εφαρμογές έχει φανεί ότι και μια κρυφή στοιβάδα είναι αρκετή, ειδικά σε εφαρμογές ρύθμισης. Σε αυτήν την ενότητα θα αναλυθεί μόνο η ιδιότητα της προσέγγισης συναρτήσεων αφού μόνο αυτή χρησιμοποιήθηκε.

2.6.1 Προσέγγιση Συναρτήσεων (Function Approximation)

Απαραίτητη ιδιότητα των ΤΝΔ σε κλειστού βρόχου εφαρμογές ρύθμισης είναι αυτή της καθολικής προσέγγισης (universal function approximation) όταν υπάρχουν παραπάνω από μια κρυφή στοιβάδα. Αυτή η ιδιότητα μελετήθηκε από πολλούς ερευνητές συμπεριλαμβανομένων τους (Cybenko (1989); Hrník, Stinchcombe and White (1989); και Standberg (1991)).

Αποδείχθηκε ότι μια λεία συνάρτηση (smooth function) $f(x)$ μπορεί να προσεγγιστεί με οποιαδήποτε επιθυμητή ακρίβεια σε ένα συγκεκριμένο πεδίο τιμών (compact set). Συγκεκριμένα, έστω $f(x): \mathbb{R}^n \rightarrow \mathbb{R}^m$ μια λεία συνάρτηση, τότε δεδομένου ένα πεδίο τιμών $S \in \mathbb{R}^n$ και μια σταθερή τιμή ε_N , υπάρχει νευρωνικό δίκτυο μία ή παραπάνω κρυφή στοιβάδα τέτοια ώστε ισχύει

$$f(x) = W^T \sigma(\Omega^T x) + \varepsilon \quad (2.7)$$

Όπου $\|\varepsilon\| < \varepsilon_N$ για κάθε τιμή του $x \in S$, για κάποιον αρκετά μεγάλο αριθμό L από κόμβους του ΤΝΔ. Η τιμή ε καλείται ως σφάλμα προσέγγισης των ΤΝΔ (function approximation error) και μειώνεται όσο αυξάνεται ο L . Το γεγονός ότι μπορεί να

αποδειχθεί ότι υπάρχει ΤΝΔ τέτοιο ώστε να προσεγγίζεται μια συνάρτηση δεν δίνει και τον τρόπο όπου θα υπολογιστούν τα συναπτικά βάρη.

Για να λειτουργεί ορθά ένα νευρωνικό δίκτυο πρέπει να επιλεγούν σωστά τα συναπτικά βάρη, όπου οι αλγόριθμοι που πραγματοποιούν αυτούς τους υπολογισμούς μπορούν να ταξινομηθούν σε διάφορες κατηγορίες όπως φαίνεται στην επόμενη ενότητα. [6]

2.7 Ταξινόμηση σχημάτων μάθησης

2.7.1 Απευθείας υπολογισμός vs Εκπαίδευσης

Η διαδικασία της εκπαίδευσης μπορεί να μην πραγματοποιηθεί σε πραγματικό χρόνο (off line), όπου πραγματοποιείται η φάση της εκπαίδευσης πριν πραγματοποιηθεί η φάση της λειτουργίας του ΤΝΔ. Αντίθετα όταν η εκπαίδευση πραγματοποιείται σε πραγματικό χρόνο το νευρωνικό δίκτυο εκπαιδεύεται κατά τη διάρκεια της λειτουργίας του. Η εκπαίδευση σε πραγματικό χρόνο είναι μια πολύ δύσκολη διαδικασία και χρησιμοποιείται ευρέως σε προβλήματα ρύθμισης με ανατροφοδότηση. Τα ΤΝΔ δίκτυα πρέπει από τη μια να σιγουρέψουν ότι η δυναμική του συστήματος παραμένει σταθεροποιημένη και ταυτόχρονα οι συνάψεις τους παραμένουν φραγμένες.

2.7.2 Μάθηση σε πραγματικό χρόνο vs σε μη πραγματικό χρόνο

Η διαδικασία της εκπαίδευσης μπορεί να μην πραγματοποιηθεί σε πραγματικό χρόνο (off line), όπου πραγματοποιείται η φάση της εκπαίδευσης πριν πραγματοποιηθεί η φάση της λειτουργίας του ΤΝΔ. Αντίθετα όταν η εκπαίδευση πραγματοποιείται σε πραγματικό χρόνο το νευρωνικό δίκτυο εκπαιδεύεται κατά τη διάρκεια της λειτουργίας του. Η εκπαίδευση σε πραγματικό χρόνο είναι μια πολύ δύσκολη διαδικασία και χρησιμοποιείται ευρέως σε προβλήματα ρύθμισης με ανατροφοδότηση. Τα ΤΝΔ δίκτυα πρέπει από τη μια να σιγουρέψουν ότι η δυναμική του συστήματος παραμένει σταθεροποιημένη και ταυτόχρονα οι συνάψεις τους παραμένουν

φραγμένες. [6]

2.8 Πλεονεκτήματα

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

Ευελιξία: Τα ΤΝΔ 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 με τους 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.

Γενικευμένη Μορφή: 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.

2.9 Μειονεκτήματα

Παρακάτω αναφέρουμε μια σειρά από στοιχεία που αποτελούν τροχοπέδη στην εφαρμογή των Τεχνητών Νευρωνικών Δικτύων και στα οποία θα πρέπει να επικεντρωθεί η προσπάθεια για βελτίωση τους. [19]

1. 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.
2. Επέκταση του παραπάνω προβλήματος των Τεχνικών Νευρωνικών Δικτύων είναι ότι δεν μας επιτρέπουν την ασφαλή εξαγωγή συμπερασμάτων για την λειτουργία των συστημάτων που προσομοιώνουν.
3. Είναι πολύ συχνό το φαινόμενο να μην μπορούν να βρουν το παγκόσμιο άριστο άλλα το τοπικό.

2.10 Εφαρμογές των Τεχνητών Νευρωνικών Δικτύων

Τα νευρωνικά δίκτυα εφαρμόζονται ευρέως τα τελευταία χρόνια, ενώ εκτιμάται ότι θα υπάρξει ακόμα μεγαλύτερη εφαρμογή στα επόμενα χρόνια. 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: [20]

Αεροδιαστημική: Αυτόματος πιλότος, έλεγχος αεροσκάφους, ανίχνευση βλαβών.

Οικονομία: Οι διακυμάνσεις των τιμών των μετοχών και των χρηματιστηριακών δεικτών είναι ακόμα ένα παράδειγμα ενός πολύπλοκου, πολυδιάστατου, αλλά και σε ορισμένες περιπτώσεις εν μέρει ντετερμινιστικού φαινομένου. Τα νευρωνικά δίκτυα χρησιμοποιούνται από πολλούς τεχνικούς αναλυτές, ώστε να κάνουν προβλέψεις σχετικά με τις τιμές των μετοχών, βασιζόμενοι σε ένα μεγάλο αριθμό παραγόντων, όπως δηλαδή, τις προηγούμενες επιδόσεις άλλων αποθεμάτων και διαφόρων οικονομικών δεικτών.

Έλεγχος συστημάτων κινητήρα: Τα νευρωνικά δίκτυα έχουν χρησιμοποιηθεί για την ανάλυση των εισροών που δέχονται οι αισθητήρες ενός κινητήρα. Το νευρωνικό δίκτυο ελέγχει μια ποικιλία παραμέτρων με τις οποίες λειτουργεί ο κινητήρας, προκειμένου να επιτευχθεί ένας συγκεκριμένος στόχος. Για παράδειγμα, το δίκτυο αυτό επιχειρεί την ελαχιστοποίηση της κατανάλωσης των καυσίμων.

Ηλεκτρονική: Τεχνητή όραση, σύνθεση ήχου.

Τηλεπικοινωνίες: Συμπύεση δεδομένων, μετάφραση γλώσσας σε πραγματικό χρόνο

Ιατρική διάγνωση: Ένα ευρύ φάσμα ιατρικά συσχετιζόμενων ενδείξεων, όπως ο συνδυασμός της καρδιακής συχνότητας, τα επίπεδα των διαφόρων ουσιών στο αίμα, ο ρυθμός της αναπνοής μπορούν να παρακολουθηθούν. Η εκδήλωση μιας συγκεκριμένης ιατρικής κατάστασης, γίνεται να συσχετιστεί με ένα πολύπλοκο συνδυασμό μεταβολών σε ένα υποσύνολο μεταβλητών που παρακολουθούνται. Τα νευρωνικά δίκτυα έχουν χρησιμοποιηθεί για την αναγνώριση αυτού του προτύπου πρόβλεψης, ώστε να χορηγηθεί η κατάλληλη θεραπεία.

Κεφάλαιο 3

Δυναμικός Προγραμματισμός

Ο δυναμικός προγραμματισμός είναι μία υπολογιστική μέθοδος η οποία εφαρμόζεται όταν πρόκειται να ληφθεί μία σύνθετη απόφαση η οποία προκύπτει από τη σύνθεση επιμέρους αποφάσεων που αλληλεξαρτώνται. Η αλληλεξάρτηση μπορεί να προκύψει επειδή οι αποφάσεις είτε παρουσιάζουν κάποια χρονική διαδοχή (όπως στην περίπτωση αναζήτησης της συντομότερης διαδρομής), είτε συνδέονται με κοινούς περιορισμούς (όπως στην περίπτωση κατανομής περιορισμένων πόρων μεταξύ ανταγωνιστικών δραστηριοτήτων). Πατέρας του δυναμικού προγραμματισμού θεωρείται ο Richard Bellman. [1]

3.1 Εισαγωγή στον Δυναμικό Προγραμματισμό

Ο Bellman με τον όρο δυναμικός προγραμματισμός ονόμασε τον τρόπο με τον οποίο επιλύονται τα προβλήματα στα οποία πρέπει να ληφθούν μία σειρά αποφάσεων που η κάθε μια από αυτές επηρεάζει τις επόμενες αποφάσεις, και που όλες μαζί δημιουργούν ένα βέλτιστο αποτέλεσμα. Η λέξη ‘προγραμματισμός’ δεν είχε καμία σχέση με τον προγραμματισμό υπολογιστών, αντίθετα προήλθε από τον όρο ‘μαθηματικός προγραμματισμός’, ο οποίος σήμερα είναι γενικά αποδεκτός ως συνώνυμο της βελτιστοποίησης. Τα προβλήματα που επιλύονται με τη μέθοδο του δυναμικού προγραμματισμού παρουσιάζουν τα ακόλουθα χαρακτηριστικά: [18]

1. Οι αποφάσεις λαμβάνονται διαδοχικά
2. Το πρόβλημα μπορεί να διαιρεθεί σε βήματα και κάθε φορά απαιτείται να ληφθεί μια στρατηγική απόφαση
3. Κάθε βήμα έχει ορισμένο αριθμό καταστάσεων που συνδέονται με αυτό
4. Το αποτέλεσμα μιας στρατηγικής απόφασης που λαμβάνεται σε κάθε βήμα είναι να μετατρέπει την παρούσα κατάσταση σε μια κατάσταση που συνδέεται με το επόμενο βήμα
5. Κάθε απόφαση συνδέεται με ένα κέρδος ή κόστος
6. Ο αντικειμενικός στόχος είναι να μεγιστοποιηθεί το συνολικό κέρδος ή να ελαχιστοποιηθεί το κόστος ή γενικότερα να επιτευχθεί το βέλτιστο δυνατό αποτέλεσμα
7. Οι αποφάσεις που παίρνονται σε κάθε βήμα εξαρτώνται μόνον από την κατάσταση του προηγούμενου βήματος και όχι από τον τρόπο με τον οποίο πραγματοποιήθηκε το βήμα αυτό

3.2 Δυναμικός προγραμματισμός και συστήματα διακριτού χρόνου

Τα βασικά χαρακτηριστικά αυτής της προσέγγισης είναι το δυναμικό μοντέλο διακριτού χρόνου

$$x_{k+1} = f(x_k, u_k, w_k) \quad (3.1)$$

Όπου x_k είναι η κατάσταση του συστήματος, u_k η επιβαλλόμενη μεταβλητή εισόδου και w_k η άγνωστη κατά βάση διαταραχή τη χρονική στιγμή k . Ενώ το k κινείται γύρω από τις τιμές 1 μέχρι $N-1$, όπου N ορίζοντας της ρύθμισης. Ενώ ακόμα ένα χαρακτηριστικό του συστήματος είναι η συνάρτηση κόστους η οποία αυξάνει με το χρόνο καθώς προστίθενται νέες τιμές. [18]

Το κόστος είναι προσθετικό με την έννοια ότι κάθε χρονική στιγμή προστίθεται νέα τιμή, με $g_k(x_k, u_k, w_k)$ το κόστος σε κάθε χρονική στιγμή, ενώ $g_N(x_N)$ το κόστος στη τελευταία χρονική στιγμή. Έτσι η συνάρτηση κόστους είναι

$$g_N(x_N) + \sum_{i=0}^{N-1} (g_k(x_i, u_i, w_i)) \quad (3.2)$$

Εξαιτίας του όρου w_k το κόστος περιλαμβάνει αβεβαιότητα και έτσι δεν γίνεται ντετερμινιστική βελτιστοποίηση. Έτσι εφαρμόζεται βελτιστοποίηση του αναμενόμενου κόστους (**expected cost**)

$$\mathbb{E}[g_N(x_N) + \sum_{i=0}^{N-1} (g_i(x_k, u_k, w_k))] \quad (3.3)$$

Όπου το αναμενόμενο κόστος γίνεται σύμφωνα με την κατανομή των τυχαίων μεταβλητών που συμπεριλαμβάνονται στο σύστημα. Η βελτιστοποίηση γίνεται ως προς τις μεταβλητές εισόδου u_0, u_1, \dots, u_{N-1} , λαμβάνοντας υπόψιν τη κατάσταση του συστήματος τη στιγμή k .

3.3 Αρχή βελτιστοποίηση του Bellman

Η επίλυση προβλήματος με δυναμικό προγραμματισμό βασίζεται στην αρχή του Bellman (principle of optimality), σύμφωνα με την οποία: ‘Για οποιαδήποτε αρχική (προηγούμενη) σειρά αποφάσεων και κατάσταση του συστήματος (initial stage and decision), οι υπολειπόμενες αποφάσεις πρέπει να αποτελούν μια βέλτιστη πολιτική για τη κατάσταση του συστήματος που προέκυψε μετά τις αρχικές ή προηγούμενες αποφάσεις’.

Έστω $\pi^* = \mu_0^*, \mu_1^*, \dots, \mu_{N-1}^*$ μια βέλτιστη πολιτική και υποθέτοντας ότι όταν χρησιμοποιείται π^* , είναι διαθέσιμη η κατάσταση του συστήματος x_i , τότε στοχεύεται από τη κατάσταση x_i τη χρονική κατάσταση i να ελαχιστοποιηθεί η συνάρτηση κόστους ‘cost-to-go’ από την χρονική στιγμή i μέχρι τη στιγμή N

$$J_i(x_i) = \mathbb{E}[g_N(x_N) + \sum_{k=i}^{N-1} (g_k(x_k, \mu_k(x_k), w_k))] \quad (3.4)$$

για $i=0,1,\dots,N$. Τότε η ‘περιομμένη’ πολιτική $\pi^*=\mu_i^*, \mu_{i+1}^*, \dots, \mu_{N-1}^*$ είναι βέλτιστη για το εν λόγω πρόβλημα.

Η αρχή αυτή της βελτιστότητας μπορεί να αιτιολογηθεί ως εξής: Εάν η ‘περιομμένη’ πολιτική $\pi^*=\mu_i^*, \mu_{i+1}^*, \dots, \mu_{N-1}^*$ δεν ήταν βέλτιστη τότε η κατάσταση x_i τη χρονική στιγμή i , θα μπορούσε να μειωθεί η συνάρτηση ‘cost-to-go’ αλλάζοντας απλώς τη πολιτική με μια άλλη που θα ήταν βέλτιστη για το υπό πρόβλημα.

Η αρχή της βελτιστότητας βασίζεται στο σκεπτικό του ‘διαίρει και βασίλευε’. Ουσιαστικά, μια βέλτιστη πολιτική για ένα πολύπλοκο πολλαπλών διαστάσεων πρόβλημα ελέγχου κατασκευάζεται ως εξής:

1. Δημιουργείται μια βέλτιστη πολιτική για το ‘τελευταίο υπο-πρόβλημα’ που αφορά μόνο το τελευταίο στάδιο του συστήματος
2. Επεκτείνεται η βέλτιστη πολιτική στο ‘τελικό υπο-πρόβλημα’ που αφορά τα δύο τελευταία στάδια του συστήματος
3. Συνεχίζεται κατά τον ίδιο τρόπο, μέχρι να καλυφθεί ολόκληρο το πρόβλημα

Κεφάλαιο 4

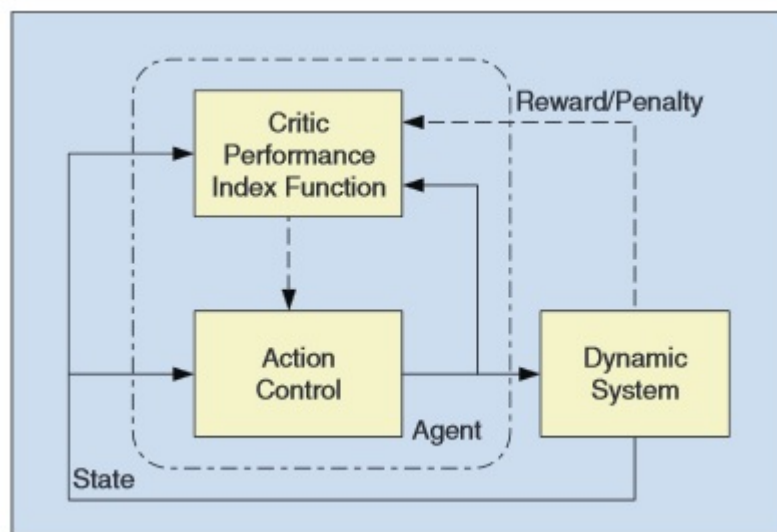
Ενισχυτική Μάθηση

Η ενισχυτική μάθηση βασίζεται, στο γεγονός ότι κάθε ζωντανός οργανισμός όταν αλληλεπιδρά με το περιβάλλον βελτιώνει της δράσεις-ενέργειες του με σκοπό τόσο την ανάπτυξη του όσο και την επιβίωση του. Έχει παρατηρηθεί ότι κάθε κοινωνία, κάθε οργανισμός εφαρμόζοντας ποινές ή επιβραβεύσεις ενεργοποιούν την συμπεριφορά με σκοπό να ανταποκρίνονται στο περιβάλλον.

Η εκμάθηση από την αλληλοεπίδραση με το περιβάλλον ονομάζεται ενισχυτική μάθηση. Όπως έχει αναφερθεί στα προηγούμενα κεφάλαια υπάρχουν διάφοροι τύποι μάθησης (supervised, reinforcement, unsupervised). Η ενισχυτική μάθηση παραπέμπει σε έναν πράκτορα (actor-agent), αλληλεπιδρά με το περιβάλλον και αλλάζει τις ρυθμιστικές ενέργειες, βασιζόμενες στην απόκριση των προηγούμενων πράξεων. Αυτή η διαδικασία που βασίζεται στην πληροφορία που εξάγεται από το περιβάλλον, ονομάζεται 'βασισμένη στην ενέργεια μάθηση' (action-based learning). Με την ενισχυτική μάθηση συνεπάγεται και σχέση αιτίου-αποτελέσματος μεταξύ δράσεων (actions) και της ανταμοιβής-τιμωρίας (reward-punishment).

Η βασική εφαρμογή της ενισχυτικής μάθησης είναι της δομής ενέργειας-κριτικής [17], όπου εφαρμόζεται η ρυθμιστική δράση στο περιβάλλον, και ο κριτής αξιολογεί αυτή τη δράση. Με βάση αυτή την αξιολόγηση η ρυθμιστική δράση αναβαθμίζεται. Άρα υπάρχουν δύο στάδια, αρχικά η αξιολόγηση της δράσης (policy evaluation) από τον κριτή και ύστερα η αναβάθμιση της (policy improvement). Όπως κάθε ζώντας οργανισμός αλληλεπιδρά με σκοπό να βελτιώσει ένα κριτήριο, έτσι και η ενισχυ-

τική μάθηση, ο κριτής αξιολογεί της ρυθμιστικές δράσεις (πολιτικές) βασιζόμενος σε ένα κριτήριο βελτιστοποίησης. [21] Με αυτό τον τρόπο η ενισχυτική μάθηση που σκοπεύει να βελτιστοποιήσει ένα κριτήριο (optimal reinforcement learning) δρα σε κάθε χρονική στιγμή χωρίς να ενδιαφέρεται για την λεπτομερή δυναμική του συστήματος αλλά μονάχα για την απόκριση του, παρατηρώντας τη συμπεριφορά του. Μια σχηματική αναπαράσταση της ενισχυτικής μάθησης παρουσιάζεται στο σχήμα 4.1



Σχήμα 4.1: Σχηματική αναπαράσταση ενισχυτικής μάθησης

4.1 Σύστημα ή Περιβάλλον

Στην ενισχυτική μάθηση ο φορέας των ενεργειών (actor) είναι πράκτορας (agent) που παράγει τη ρυθμιστική πολιτική. Ο φορέας μαθηματικά περιγράφεται από την εξής εξίσωση (4.1)

$$k = h(x_k) \quad (4.1)$$

Λαμβάνει x_k την κατάσταση του συστήματος σαν είσοδο και την ρυθμιστική δράση σαν έξοδο. Οτιδήποτε βρίσκεται εκτός από το φορέα θεωρείται περιβάλλον. Έτσι το σύστημα (3.1) λαμβάνεται σαν μέρος του περιβάλλοντος, ενώ συχνά η δυναμική του συστήματος δεν λαμβάνεται καθόλου υπόψη, δηλαδή δεν χρησιμοποιείται

συγκεκριμένο μοντέλο για το σχεδιασμό.

4.2 Βασιζόμενη στο μοντέλο και ελεύθερο μοντέλου

Οι αλγόριθμοι της ενισχυτικής μάθησης μπορούν να χωριστούν σε δύο κατηγορίες: Σε αυτές που βασίζονται στο μοντέλο (model-based) και έτσι χρησιμοποιείται μια προσέγγιση της δυναμικής του συστήματος και αυτές που δεν χρειάζονται κανένα μοντέλο για το σύστημα (model-free). Δεν μπορεί να διευκρινιστεί αν κάποια από τις δύο είναι καλύτερη αφού εξαρτάται από τον τύπο του προβλήματος. Τα κύρια πλεονεκτήματα από τη πρώτη κατηγορία (model-based) είναι ότι το πραγματικό μοντέλο συμπληρώνεται από το προσομοιωμένο (simulated), ενώ η γνώση των καταστάσεων αρκούν για να υπολογιστεί η βέλτιστη ρυθμιστική δράση. Από την άλλη, τα μοντέλα είναι πολύ σύνθετα ή μπορεί να μην είναι ακριβές.

4.2.1 Νευρωνικά δίκτυα και ενισχυτική μάθηση

Σχεδόν όλοι οι αλγόριθμοι ενισχυτικής μάθησης περιέχουν την εκτίμηση-προσέγγιση μιας συνάρτησης αξίας-κόστους (value function) που δείχνει πόσο καλή είναι η παρούσα κατάσταση ή πόσο καλή ήταν η ρυθμιστική δράση. Η αρχική βασική ιδέα ήταν να κατασκευαστεί αυτή η συνάρτηση περιέχοντας έναν πίνακα (lookup table) που να προσδιορίζει για κάθε πιθανή κατάσταση την τιμή της, αλλά προφανώς αυτή η προσέγγιση δεν είναι εφικτή για μεγάλα προβλήματα.

Για αυτό το λόγο συνδυάστηκαν τα νευρωνικά δίκτυα με την ενισχυτική μάθηση, καθώς έχουν την ικανότητα να προσεγγίζουν οποιαδήποτε συνάρτηση. Χαρακτηριστικό είναι το παράδειγμα όπου έγινε προσπάθεια να φτιαχτεί ένας 'παίχτης' τάβλι φτάνοντας το ανθρώπινο επίπεδο μετά από μήνες παιχνιδιού με τον εαυτό του. Στο τάβλι οι πιθανές θέσεις είναι 10^{20} , άρα προφανώς οι τιμές της συνάρτησης αξίας δεν μπορούν να αποθηκευτούν σε έναν πίνακα. [22]

4.3 Στόχος η βέλτιστη απόδοση

Με την έννοια του στόχου της βέλτιστης απόδοσης, ποσοτικοποιούμε την απόδοση μέσω της συνάρτησης κόστους (εξ: 4.2)

$$V_h(k) = \sum_{i=k}^{\infty} \gamma^{i-k} r(x_i, u_i) \quad (4.2)$$

με $0 < \gamma < 1$ τον εκ πτωτικό παράγοντα. Αυτή η συνάρτηση κόστους χρησιμοποιείται σαν 'cost to go' συνάρτηση όπως περιγράφηκε στο κεφάλαιο 2, αλλά για άπειρου ορίζοντα πρόβλημα. Η έννοια της παραμέτρου γ είναι ότι δίνεται βάση στις κοντινές χρονικές στιγμές παρά της μακρινές. Η συνάρτηση $r(x_i, u_i)$ ονομάζεται συνάρτηση χρησιμότητας (*utility function*) και αξιολογεί κάθε χρονική στιγμή το κόστος. Η επιλογή της γίνεται σύμφωνα με τον εκάστοτε σχεδιασμό, ενώ συχνά επιλέγεται μια τετραγωνική μορφή της

$$r(x_i, u_i) = x_i^T Q x_i + u_i^T R u_i \quad (4.3)$$

Ο αντικειμενικός σκοπός της θεωρίας βέλτιστου ελέγχου είναι να επιλέγει την πολική αυτή (ρυθμιστικές δράσεις) τέτοιες ώστε να ελαχιστοποιείται το εκάστοτε κόστος, άρα μπορεί να ληφθεί το εξής

$$V^*(k) = \min_{h(\cdot)} \sum_{i=k}^{\infty} \gamma^{i-k} r(x_i, h(x_i)) \quad (4.4)$$

Όπου ονομάζεται βέλτιστο κόστος ή αξία.

Αντίστοιχα η βέλτιστη ρυθμιστική δράση δίνεται ως

$$h^*(x_k) = \arg \min_{h(\cdot)} \sum_{i=k}^{\infty} \gamma^{i-k} r(x_i, h(x_i)) \quad (4.5)$$

Αυτή η διαδικασία είναι γενικά είναι πολύ δύσκολη ειδικά για μη γραμμικά συστήματα και για αυτό το σκοπό χρησιμοποιούνται τα νευρωνικά δίκτυα. [23]

4.4 Εξίσωση του Bellman

Η εξίσωση (4.2) μπορεί να γραφεί ως εξής

$$V_h(k) = r(x_k, u_k) + \gamma \sum_{i=k+1}^{\infty} \gamma^{i-(k+1)} r(x_i, u_i) \quad (4.6)$$

παρατηρώντας ότι το δεύτερο σκέλος είναι η συνάρτηση κόστους στο $k+1$ βήμα άρα

$$V_h(k) = r(x_k, u_k) + \gamma V_h(k+1) \quad (4.7)$$

Η εξίσωση (4.7) είναι γνωστή ως εξίσωση Bellman. Η αξιολόγηση της αξίας κάθε επιλεγμένης πολιτικής είναι το πρώτο κρίσιμο στοιχείο στην ανάπτυξη μιας τεχνικής ενισχυτικής μάθησης.

Η διακριτού χρόνου Χαμιλτονιανή (*DT Hamiltonian*) μπορεί να ορισθεί ως εξής

$$H(x_k, h(x_k), \Delta V_k) = r(x_k, h(x_k)) + \gamma V_h(x_{k+1}) - V_h(x_k) \quad (4.8)$$

Όπου $\Delta V_k = \gamma V_h(x_{k+1}) - V_h(x_k)$, η εξίσωση (4.8) αποτελεί κρίσιμη για το στάδιο του σχεδιασμού της παρούσας εργασίας. Η εξίσωση Bellman απαιτεί η Χαμιλτονιανή να είναι ίση με 0 για την εκάστοτε πολιτική.

Πλέον η βέλτιστη συνάρτηση κόστους-αξίας (4.4) μπορεί να γραφεί ως εξής, σύμφωνα με την εξίσωση Bellman.

$$V^*(k) = \min_{h(\cdot)} (r(x_k, h(x_k)) + V_h(x_{k+1})) \quad (4.9)$$

Επίσης χρησιμοποιώντας την αρχή βελτιστότητας του Bellman (κεφάλαιο 2) προκύπτει η εξίσωση (4.10)

$$V^*(k) = \min_{h(\cdot)} (r(x_k, h(x_k)) + V^*(k+1)) \quad (4.10)$$

Είναι φανερό ότι για να υπολογιστεί η συνάρτηση αξίας για μια πολιτική τη στιγμή k πρέπει να είναι γνωστή και η τιμή της την χρονική στιγμή k , αυτό μπορεί να επιτευχθεί μέσω του δυναμικού προγραμματισμού και γενικώς *off-line* διαδικασιών όπου υπάρχει πλήρης γνώση του συστήματος. [1]

Κεφάλαιο 5

Νευρωνικά Δίκτυα Ακτινικής Βάσης

Τα δίκτυα ακτινικών συναρτήσεων βάσης (radial basis function networks), τα οποία ονομάζουμε για συντομία RBF δίκτυα έχουν την ρίζα τους στην μαθηματική θεωρία της προσέγγισης συναρτήσεων. Στην ουσία υλοποιούν μία συνάρτηση παρεμβολής (interpolation function), η οποία προσεγγίζει την τιμή μίας συνάρτησης σε κάποιο σημείο ως το μέσο όρο των τιμών της συνάρτησης σε κοντινά σημεία.

Τα RBF δίκτυα είναι νευρωνικά δίκτυα πρόσθιας τροφοδότησης αποτελούμενα συνολικά από δύο επίπεδα (two layer network). Το πρώτο επίπεδο (εσωτερικό επίπεδο-hidden layer), αποτελείται από RBF νευρώνες. Κάθε RBF νευρώνας ορίζει μία περιοχή στον χώρο εισόδου. Αν η είσοδος ανήκει σε μία από αυτές τις περιοχές, τότε η έξοδος του RBF νευρώνα που την αντιπροσωπεύει θα είναι μονάδα, διαφορετικά θα είναι μηδέν. Οι έξοδοι του πρώτου επιπέδου οδηγούνται ως εισοδοί στο δεύτερο επίπεδο (επίπεδο εξόδου-output layer), που αποτελείται από απλούς νευρώνες. Κάθε απλός νευρώνας ορίζει και μία διαφορετική κατάσταση στον χώρο εξόδου. Συνήθως θεωρούμε ότι η έξοδος που ενεργοποιείται από την είσοδο του δικτύου είναι αυτή με την μέγιστη τιμή στον αντίστοιχο νευρώνα εξόδου. Κάθε RBF νευρώνας έχει ισάριθμες εισόδους με τα στοιχεία των διανυσμάτων εισόδου, ενώ κάθε νευρώνας του επιπέδου εξόδου έχει ισάριθμες εισόδους με τους RBF νευρώνες.

Κατά την διάρκεια της εκπαίδευσης ενός RBF δικτύου, οι παράμετροι που πρέπει να καθοριστούν είναι οι παράμετροι των νευρώνων από τους οποίους αποτελείται το δίκτυο. Κάτι που δεν είναι φανερό με την πρώτη ματιά, είναι το γεγονός ότι κατά την εκπαίδευση μπορεί να αλλάξει και η δομή του RBF νευρωνικού δικτύου. Αυτό θα γίνει εάν αυξηθεί ο αριθμός των ορισμένων περιοχών, με την προσθήκη RBF νευρώνων στο εσωτερικό επίπεδο του δικτύου και την δημιουργία ανάλογων συνδέσεων με το επίπεδο εξόδου. [24]

5.1 Δομή Δικτύων Ακτινικής Βάσης

Υπάρχουν πολλά είδη ακτινικών συναρτήσεων βάσης, με τις κυριότερες να είναι η Γκαουσιανή, η spline λεπτής πλάκας (TPS), που φαίνονται στις παρακάτω εξισώσεις 5.1,5.2

$$g_i^G(r) = \Phi(\|x - c_i\|) = \exp\left[-\frac{\|x - c_i\|^2}{2\sigma_i}\right] \quad (5.1)$$

$$g_i^{TSP}(r) = \Phi(\|x - c_i\|) = \|x - c_i\|^2 \log(\|x - c_i\|) \quad (5.2)$$

Όπου σ είναι η παράμετρος διακύμανσης και ελέγχει την ομαλότητα της συνάρτησης βάσης, τα RBF είναι δίκτυα 2 στοιβάδων με την εξωτερική να αποτελεί τον γραμμικό συνδυασμό των εξόδων των συναρτήσεων ακτινικής βάσης.

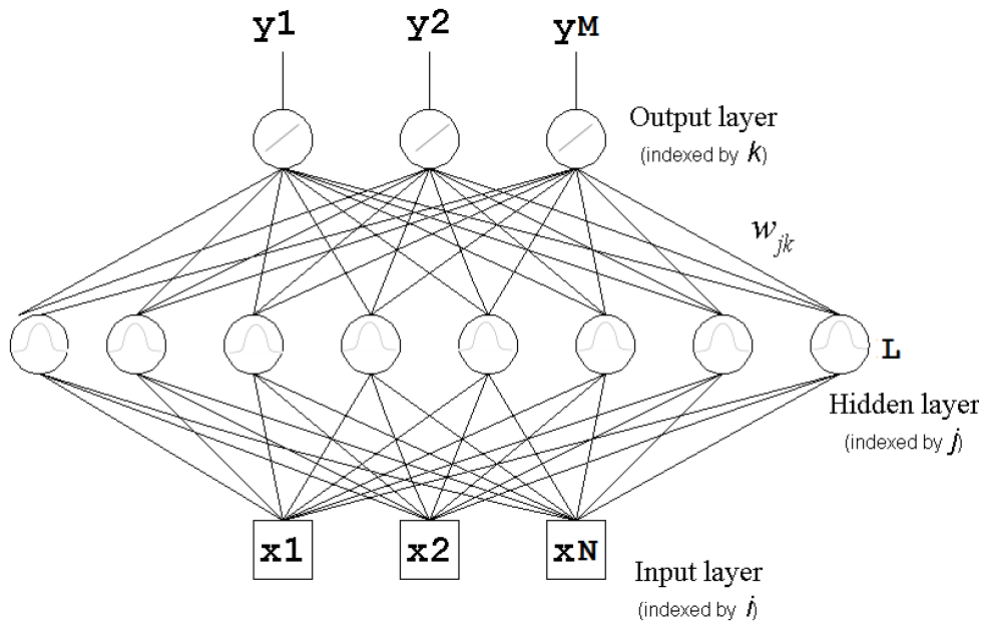
Η δομή των RBF φαίνεται στο σχήμα 5.1

Έτσι οι έξοδοι δίνονται από την εξίσωση 34

$$y_j = \sum_{i=1}^N w_{ij} \varphi(\|x - c_i\|) \quad (5.3)$$

Όπου w_{ij} αποτελούν τα συναπτικά βάρη.

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



Σχήμα 5.1: Δομή νευρωνικών δικτύων ακτινικής βάσης

βάρη του πρώτου επιπέδου στη συνέχεια διατηρούνται σταθερά, ενώ τα βάρη του δεύτερου επιπέδου βελτιστοποιούνται κατά τη δεύτερη φάση της εκπαίδευσης. [14]

Για συγκεκριμένα δεδομένα εκπαίδευσης, αρχικά επιλέγονται οι πρότυπες παράμετροι των συναρτήσεων βάσης με το κατάλληλο πλάτος έτσι ώστε να καλύπτουν ολόκληρο τον N -διάστατο χώρο εισόδου που καλύπτεται από τα πρότυπα εισόδου. 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-ΝΔ, γνωστό ως ‘κατάρτα της μεγάλης διάστασης’, και 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 [19], για να επιλέξουν τις βέλτιστες παραμέτρους του πρώτου επιπέδου βάση της πυκνότητας των δεδομένων. Ωστόσο, η τεχνική αυτή έχει πολλά σημαντικά μειονεκτήματα:

1. Ο αριθμός των νευρώνων της κρυφής στοιβάδας πρέπει να καθοριστεί από πριν, γεγονός που σημαίνει ότι πρέπει να πραγματοποιηθεί μια διαδικασία δοκιμής και σφάλματος ώστε να βρεθεί ο βέλτιστος αριθμός.
 2. Για να ξεκινήσει ο αλγόριθμος πρέπει να γίνει μια αρχική αρχικοποίηση των κέντρων, άρα κάθε φορά που θα επαναλαμβάνεται ο αλγόριθμος θα δίνει και διαφορετικά αποτελέσματα, σχετικά με την επιλογή των κέντρων.
 3. Ο αλγόριθμος αυτός είναι επαναληπτικός με αποτέλεσμα να αυξάνεται ιδιαίτερα το υπολογιστικό κόστος αν υπάρχει μεγάλος όγκος δεδομένων.
- Όταν υπολογιστούν διανυσματικά κέντρα και πλάτη, γίνεται ο καθορισμός των συναπτικών βαρών. Για το σκοπό αυτό χρησιμοποιούνται η μέθοδος των ελαχίστων τετραγώνων καθώς πλέον πρόκειται για γραμμικό πρόβλημα.

Κεφάλαιο 6

Ο Αλγόριθμος των Ασαφών Μέσων (Fuzzy Means Algorithm)

Ο αλγόριθμος των ασαφών προτάθηκε (Sarimveis et al [3]) σαν μια μέθοδο για την επιλογή των κρυφών κέντρων των νευρωνικών δικτύων RBF, που δεν έχουν τα προβλήματα που αναφέρθηκαν παραπάνω. [25] Σε σύγκριση με παραδοσιακές μεθοδολογίες, ο αλγόριθμος των ασαφών μέσων έχει την ικανότητα να καθορίζει αυτόματα το μέγεθος του δικτύου, δηλαδή τον αριθμό των κέντρων του RBF, ενώ αποδεικνύεται ότι είναι πολύ πιο γρήγορος αφού τα δεδομένα εισόδου περνάνε μόνο μια φορά. Η βασική ιδέα είναι ο διαμερισμός του χώρου εισόδου σε έναν αριθμό ασαφών υποσυνόλων.

6.1 Μη συμμετρικός off-line αλγόριθμος

Θεωρώντας την N διάστασης κανονικοποιημένη μεταβλητή εισόδου x_i , για $i=1...N$ του δικτύου όπου μπορεί να διαμερισθεί σε c_i τριγωνικά ασαφή σύνολα:

$$T_i = \{A_{i1}, A_{i2}, \dots, A_{iN}\}, \quad 1 \leq i \leq N \quad (6.1)$$

Υποθέτοντας ότι a_{ij} είναι το κέντρο του ασαφούς συνόλου A_{ij} , και δa_{ij} είναι το αντίστοιχο μισό πλάτος, τότε η συνάρτηση συνυσοφοράς $\mu_{A_{ij}}(x_i(k))$ της μεταβλητής εισόδου x_i στο ασαφές σύνολο A_{ij} ορίζεται ως εξής:

$$\mu_{A_{ij}}(x_i(k)) = \begin{cases} 1 - \frac{|x_i(k) - a_{ij}|}{\delta a_{ij}}, & \text{if } x_i(k) \in [a_{ij} - \delta a_{ij}, a_{ij} + \delta a_{ij}] \\ 0, & \text{otherwise} \end{cases} \quad (6.2)$$

Έτσι, τα ασαφή σύνολα A_{ij} μπορούν να περιγραφούν πλήρως από τα αντίστοιχα κέντρα και πλάτη.

$$A_{ij} = \{a_{ij}, \delta a_{ij}\} \quad (6.3)$$

Θεωρείται ότι σε μια διάσταση i τα πλάτη είναι ίδια σε όλα τα ασαφή σύνολα. Γι αυτό το λόγο στη συνέχεια απαλείφεται ο δείκτης j . Έπεται ότι η ασαφής διαμέλιση του συνολικού χώρου των μεταβλητών εισόδου x μπορεί να οριστεί διαμερίζοντας τον σε C ασαφή σύνολα, όπου

$$C = \prod_{i=1}^N c_i \quad (6.4)$$

Κάθε ασαφές υποσύνολο \mathbf{A}^l ($1 \leq l \leq C$) μπορεί να ορισθεί ως ο συνδυασμός των N ασαφών συνόλων $A^l_{1j_1} \in T_1, A^l_{2j_2} \in T_2, \dots, A^l_{Nj_N} \in T_N$ ($1 \leq j_1 \leq c_1, 1 \leq j_2 \leq c_2, \dots, 1 \leq j_N \leq c_N$), και παρουσιάζεται ως

$$\mathbf{A}^l = [A^l_{1j_1}, A^l_{2j_2}, \dots, A^l_{Nj_N}] = \{[a^l_{1j_1}, a^l_{2j_2}, \dots, a^l_{Nj_N}], [\delta a^l_{1j_1}, \delta a^l_{2j_2}, \dots, \delta a^l_{Nj_N}]\} \quad (6.5)$$

Μια πιο συμπαγής μορφή για να παρουσιαστούν τα ασαφή υποσύνολα \mathbf{A}^l είναι η εξής

$$\mathbf{A}^l = \{\mathbf{a}^l, \delta \mathbf{a}^l\} \quad (6.6)$$

Όπου $\mathbf{a}^l = [a^l_{1j_1}, a^l_{2j_2}, \dots, a^l_{Nj_N}]$ είναι το διάνυσμα του κέντρου και $\delta \mathbf{a}^l = [\delta a^l_{1j_1}, \delta a^l_{2j_2}, \dots, \delta a^l_{Nj_N}]$ είναι το διάνυσμα των πλατών του ασαφής υποχώρου.

Η ιδέα της συνάρτησης συνεισφοράς μπορεί να επεκταθεί σε πολλές διαστάσεις, εισάγοντας την πολυδιάστατη συνάρτηση συνεισφοράς μ της μεταβλητής εισόδου \mathbf{x} στο υποσύνολο \mathbf{A}^l , όπου ορίζεται ως εξής

$$\mu_{\mathbf{A}^l}(\mathbf{x}(k)) = \begin{cases} 1 - rd^l(\mathbf{x}(k)), & \text{if } rd^l(\mathbf{x}(k)) \leq 1 \\ 0, & \text{otherwise} \end{cases} \quad (6.7)$$

Όπου $rd^l(\mathbf{x}(k))$ είναι η ευκλείδεια σχετική απόσταση μεταξύ του συνόλου \mathbf{A}^l και της μεταβλητής εισόδου \mathbf{x}

$$rd^l(\mathbf{x}(k)) = \sqrt{\sum_{i=1}^N \left(\frac{(a_{ij_i}^l - x_i(k))^2}{N\delta\alpha^2} \right)} \quad (6.8)$$

Ο ορισμός της ευκλείδειας απόστασης είναι δυνατόν να χρησιμοποιηθεί μόνο για συμμετρικό διαμερισμό. Έτσι πλέον χρησιμοποιείται η συνάρτηση απόσταση d_r^l , που ορίζει μια υπερεπιφάνεια στο χώρο εισόδου. Αυτή η υπερσφαίρα οριοθετεί το διάλυμα της εισόδου μεταξύ της μη μηδενικής και της μηδενικής συνάρτησης συνεισφοράς στον ασαφή υπόχωρο \mathbf{A}^l . Πρακτικά χρησιμοποιείται μια N-διάστατη υπερέλλειψη για κάθε ασαφή υπόχωρο: [5]

$$\sqrt{\sum_{i=1}^N \left(\frac{(a_{ij_i}^l - x_i(k))^2}{b_i^2} \right)} = 1 \quad (6.9)$$

Όπου b_i είναι το μισό του αντίστοιχου άξονα της έλλειψης για κάθε διάσταση $i=1, \dots, N$. Οι παράμετροι b_i χρειάζεται να ορισθούν, σε αντίθεση με την υπερσφαίρα που είναι ίσες με 1. Η πρώτη προϋπόθεση είναι ότι η επιφάνεια της υπερέλλειψης χρειάζεται να περνά από κάθε κορυφή των υπερ-τριγώνων που ορίζονται από τον ασαφή υπόχωρο, π.χ. σε ένα δισδιάστατο παράδειγμα, πρέπει να διαπερνά 4 κορυφές των τριγώνων. Ωστόσο αυτή η συνθήκη δεν είναι αρκετή για να ορισθεί μια μοναδική υπερέλλειψη, αφού υπάρχουν άπειρες υπερέλλειψεις που να έχουν την παραπάνω ιδιότητα. Από όλες αυτές επιλέγεται αυτή όπου τα b_i είναι αναλογικά με τα πλάτη των τριγώνων

$$\frac{\delta\alpha_1}{b_1} = \frac{\delta\alpha_2}{b_2} = \dots = \frac{\delta\alpha_N}{b_N} \quad (6.10)$$

Έτσι η εξ.(6.8) μετατρέπεται ως εξής

$$\sqrt{\sum_{i=1}^N \left(\frac{(a_{ij_i}^l - x_i(k))^2}{N\delta a_i^2} \right)} = 1 \quad (6.11)$$

Βασιζόμενοι στην εξ (6.11) πλέον ορίζεται η συνάρτηση διαμερισμού από την εξής συνάρτηση απόστασης.

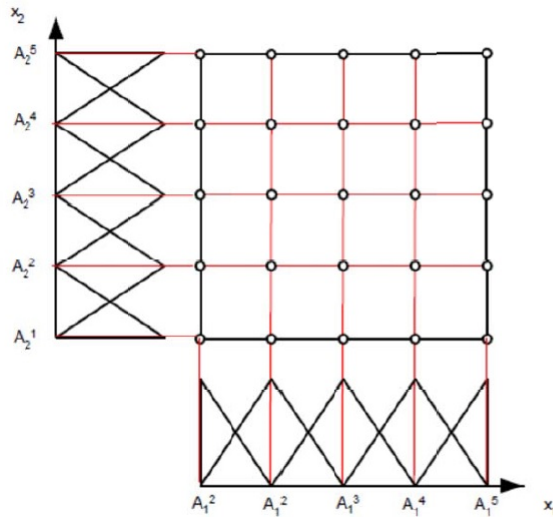
$$d_r^l(\mathbf{x}(k)) = \sqrt{\sum_{i=1}^N \left(\frac{(a_{ij_i}^l - x_i(k))^2}{N\delta a_i^2} \right)} \quad (6.12)$$

Η αντικατάσταση της σφαιρικής σχετικής απόστασης με ελλειψοειδή δημιουργεί ένα μη συμμετρικό διαμερισμό που είναι πιο ευέλικτος, καθώς δίνει τη δυνατότητα να χωρισθεί κάθε διάσταση με διαφορετικό τρόπο. Έτσι, τα αποτελέσματα των RBF νευρωνικών δικτύων καλύπτουν πιο ευρύ φάσμα του χώρου εισόδου, παράγοντας δίκτυα με μεγαλύτερη ακρίβεια και λιγότερη πολυπλοκότητα.

Έτσι αυτή η συνάρτηση συνεισφοράς μπορεί να χρησιμοποιηθεί για να προσδιορίσει το κοντινότερο ασαφές υποσύνολο για ένα συγκεκριμένο δεδομένο εισόδου \mathbf{x} . Προφανώς ο ασαφής υπόχωρος που περιγράφει καλύτερα ένα συγκεκριμένο δεδομένο εισόδου είναι αυτός με τη μεγαλύτερη συνάρτηση συνεισφοράς, και αντίστοιχα τη μικρότερη απόσταση.

Αυτή η διαδικασία παρουσιάζεται στα σχήματα 6.1-6.3, όπου υπάρχει ένας δι-διάστατος χώρος εισόδου. Κάθε διάσταση χωρίζεται σε 5 τριγωνικά ασαφή σύνολα, έτσι δημιουργούνται 25 ασαφή υποσύνολα στο χώρο εισόδου. Τα κέντρα αυτών των υποσυνόλων ορίζουν ένα 2 δισδιάστατο πλέγμα όπως φαίνεται στο σχήμα 6.1. Δεδομένου της είσοδου $\mathbf{x}=[x_1, x_2]$, υπολογίζεται η συνάρτηση συνεισφοράς. Για παράδειγμα το x_1 έχει συνάρτηση συνεισφοράς 0.6 στο A_{12} , 0.4 στο A_{13} , και 0 σε όλα τα άλλα ασαφή σύνολα, ενώ το x_2 συνεισφέρει 0.55 στο A_{22} , 0.45 στο A_{21} και μηδέν σε όλα τα άλλα ασαφή σύνολα. Έτσι τα ασαφή σύνολα που περιγράφουν καλύτερα το δεδομένο εισόδου (έχει τη μεγαλύτερη συνάρτηση συνεισφοράς) είναι το A_{12} για το x_1 και A_{22} για το x_2 . Άρα, ο κοντινότερος ασαφής υπόχωρος \mathbf{A}^{l_0} (σχήμα 6.3) της μεταβλητής εισόδου \mathbf{x} δίνεται από τον συνδυασμό των προηγούμενων επιλεγμένων ασαφών συνόλων \mathbf{A}^{l_0} . Χρησιμοποιώντας το κριτήριο

της ελάχιστης απόστασης, μεταξύ των C δυνατών επιλογών, μπορεί εύκολα να προκύψει το ίδιο αποτέλεσμα.

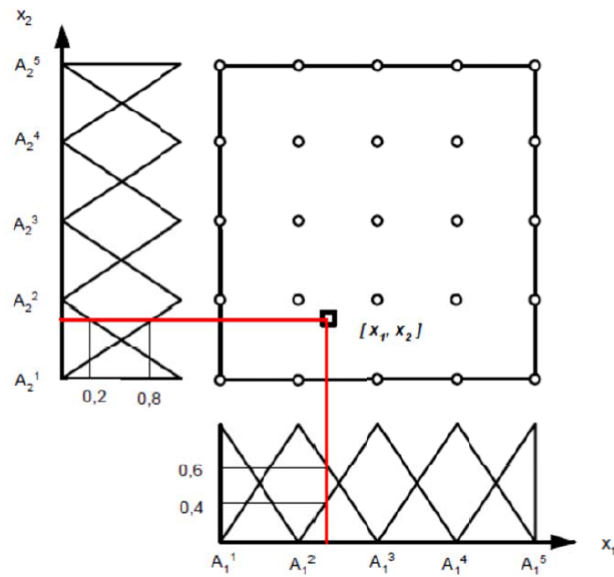


Σχήμα 6.1: Ασαφής διαμέριση του χώρου των μεταβλητών εισόδου σε 5 ασαφή σύνολα για ένα διδιάστατο πρόβλημα

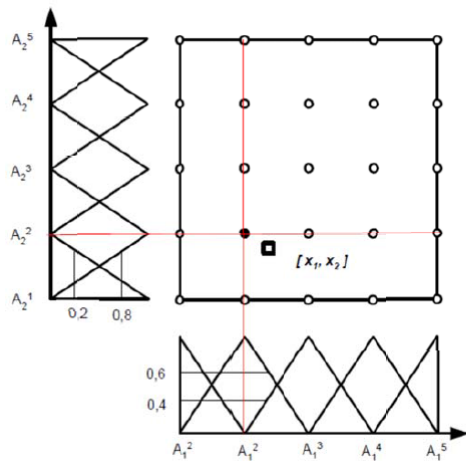
Πλέον μπορεί να παρουσιαστεί ο μη επαναληπτικός αλγόριθμος που απαιτεί μόνο ένα πέρασμα των δεδομένων εισόδου, παράγοντας τα κέντρα πολύ γρήγορα ακόμα και για μεγάλο χώρο εισόδου. Στον αλγόριθμο 1 παρουσιάζεται η συνολική διαδικασία των μη συμμετρικών ασαφών μέσων. Όπου σε κάθε χρονική στιγμή k υπολογίζεται η σχετική απόσταση μεταξύ του νέου δεδομένου εισόδου και των επιλεγμένων κέντρων, υπολογίζοντας την ελάχιστη απόσταση. Αν αυτή η απόσταση είναι μεγαλύτερη από 1 τότε προστίθεται νέο κέντρο σύμφωνα με το παραπάνω παράδειγμα.

6.2 Αλγόριθμος ασαφών μέσων σε πραγματικό χρόνο

Οι Alexandridis et.al. [4] παρουσίασαν μία προσαρμοστική έκδοση της μεθοδολογίας FM, τον αλγόριθμο Online FM. Ο αλγόριθμος Online FM, δίνει την δυνατότητα στο νευρωνικό δίκτυο RBF να προσαρμόζει σε πραγματικό χρόνο τις παραμέ-



Σχήμα 6.2: Επιλογή των κατάλληλων ασαφών συνόλων



Σχήμα 6.3: Προσδιορισμός του κοντινότερου ασαφή υπόχωρου

τρους του. Ο αλγόριθμος ξεκινά με μηδέν κόμβους στην κρυφή στρώση και χτίζει σταδιακά το μοντέλο χρησιμοποιώντας τα διαθέσιμα νέα δεδομένα. Προφανώς, οι προβλέψεις του μοντέλου στα αρχικά βήματα δεν είναι ακριβείς, αλλά η ικανότητα πρόβλεψης του δικτύου βελτιώνεται σταδιακά, καθώς είναι διαθέσιμα όλο και περισ-

Πίνακας 6.1: Αλγόριθμος ασαφών μέσων με μη συμμετρική διαμέριση του χώρου

Algorithm1

Είσοδος: $\{ X_{train}, Y_{train} \}$
 $C=[c_1, c_2, \dots, c_N]$
Έξοδος: L : Αριθμός κέντρων
G: Επιλεγμένα κέντρα

- 1 $k \leftarrow 1$ λήψη του πρώτου δεδομένου
- 2 $L \leftarrow 1$
- 3 **For** $i=1:N$
- 4 Υπολόγισε το ασαφές σύνολο με τη μέγιστη συνάρτηση συνεισφοράς σε κάθε διάσταση i
 $A_i^1 = \{a_i^1, \delta a_i\} \leftarrow \max_{1 \leq j \leq N} \mu_{A_{ij}}(x_i(1))$
- 5 **End For**
- 6 Δημιούργησε το πρώτο κέντρο $G^1 = [a_1^1, a_2^1, \dots, a_N^1]$
- 7 **For** $k=2:\text{size}(X_{train})$
- 8 Αν για το δεδομένο k δεν μπορεί να ανατεθεί κανένα κέντρο, δηλαδή $\min_{1 \leq l \leq L} [d_r^l(x(k))] > 1$
- 9 $L \leftarrow L+1$
- 10 **For** $i=1: N$
- 11 Υπολόγισε το ασαφές σύνολο με τη μέγιστη συνάρτηση συνεισφοράς για κάθε διάσταση i
 $A_i^L = \{a_i^L, \delta a_i\} \leftarrow \max_{1 \leq j \leq N} \mu_{A_{ij}}(x_i(k))$
- 12 **End For**
- 13 Δημιούργησε το L κέντρο $G^L = [a_1^L, a_2^L, \dots, a_N^L]$
- 14 **End If**
- 15 **End For**

σότερα δεδομένα. Ο αλγόριθμος Online FM βασίζεται στην ασαφή διαμέριση του χώρου εισόδου, όπως και ο αρχικός αλγόριθμος FM. Η μέθοδος λαμβάνει υπόψιν όλα τα κέντρα των ασαφών υπόχωρων ως υποψήφια για την συγκρότηση της κρυφής

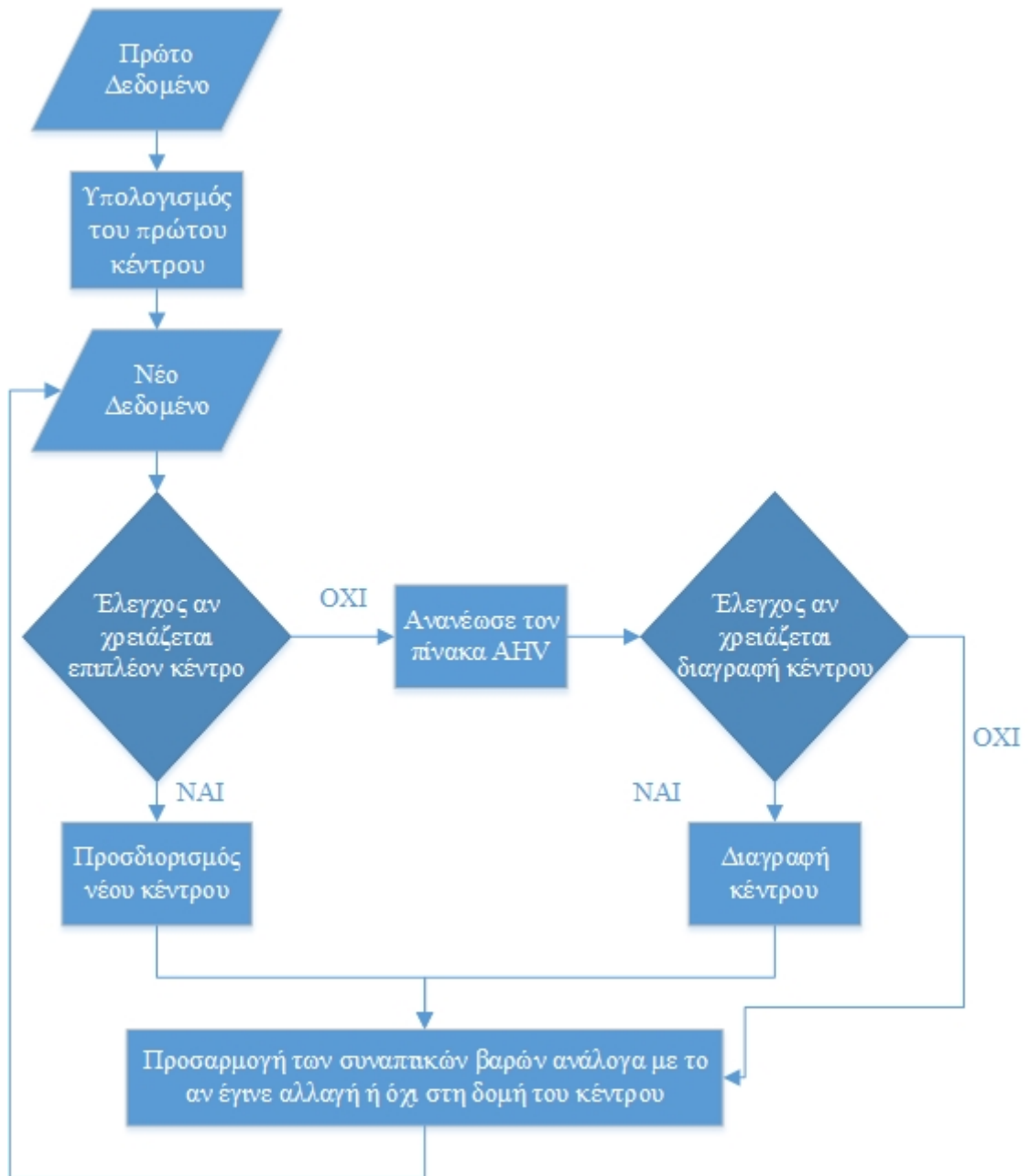
στοιβάδας του δικτύου. Ωστόσο, μεταξύ όλων των υποψηφίων κέντρων, ο αλγόριθμος με δυναμικό τρόπο εντοπίζει και επιλέγει μόνο το υποσύνολο των ασαφών υπόχωρων που βρίσκονται κοντά στα δεδομένα εισόδου. Σε κάθε χρονική στιγμή, ο αριθμός των επιλεγμένων ασαφών υπόχωρων είναι ίσος με τον αριθμό των νευρώνων της κρυφής στοιβάδας και τα κέντρα των επιλεγμένων υπόχωρων συμπίπτουν με τα κέντρα των κρυφών νευρώνων. Ως εκ τούτου, ένα πλήρες εκπαιδευμένο μοντέλο RBF είναι διαθέσιμο σε οποιαδήποτε χρονική στιγμή. Το μοντέλο RBF, αρχικά χρησιμοποιείται για την πρόβλεψη της μελλοντικής συμπεριφοράς των μεταβλητών εξόδου και στη συνέχεια ενημερώνεται βάσει του προτεινόμενου αλγορίθμου. Ο αλγόριθμος Online FM δίνει την δυνατότητα στο δίκτυο RBF να εξελίσσεται με βάση δύο επίπεδα προσαρμογής, συγκεκριμένα:

1. Προσαρμογή των συνανπτικών βαρών μεταξύ των νευρώνων της κρυφής στοιβάδας και της στοιβάδας εξόδου.
2. Προσαρμογή της δομής της κρυφής στοιβάδας με βάση την ασαφή διαμέριση του χώρου εισόδου.

Λόγω της τοπικής προσέγγισης που χρησιμοποιούν τα δίκτυα RBF, το πρώτο επίπεδο προσαρμογής μπορεί να μην είναι επαρκές, όταν ένα νέο δεδομένο δεν καλύπτεται επαρκώς από τα υπάρχοντα κέντρα. Προκειμένου να αντιμετωπιστεί αυτή η περίπτωση, ο αλγόριθμος χρησιμοποιεί το δεύτερο επίπεδο προσαρμογής, όπου προστίθενται νέοι κρυφοί νευρώνες, οι οποίοι περιγράφουν τα δεδομένα που βρίσκονται εκτός της περιοχής που καλύπτεται από τα υπάρχοντα κέντρα. Δεδομένου ότι η συνεχής προσθήκη κρυφών νευρώνων μπορεί να οδηγήσει σε μεγάλες αλλαγές στη δομή του δικτύου και σε αύξηση της υπολογιστικής πολυπλοκότητας, ο αλγόριθμος διαγράφει τους νευρώνες της κρυφής στοιβάδας που καθίστανται περιττοί.

Μόλις εισαχθεί το πρώτο δεδομένο εισόδου στο μοντέλο, ο αλγόριθμος καθορίζει τον ασαφή υπόχωρο που είναι πιο κοντά σε αυτό το δεδομένο βάσει της σχετικής απόστασης [11]. Το κέντρο αυτού του υπόχωρου γίνεται το κέντρο του πρώτου κρυφού νευρώνα. Μόλις προσδιοριστεί ο πρώτος κρυφός κόμβος, ο αλγόριθμος θέτει σε εφαρμογή δύο δυναμικούς πίνακες, οι οποίοι χρησιμοποιούνται για την αποθήκευση σημαντικών πληροφοριών και συγκεκριμένα τους πίνακες τοποθεσίας

κέντρων (Center Location Matrix - CLM) και το διάνυσμα ιστορικού ενεργοποίησης (Activation History Vector - AHV). Ο πίνακας CLM περιέχει τα κέντρα των νευρώνων της κρυφής στοιβάδας σε κάθε χρονική στιγμή και οι διάστασεις του είναι $L \times N$. Το μέγεθος του πίνακα AHV είναι ίσο με τον αριθμό των επιλεγμένων κέντρων του νευρωνικού δικτύου L και περιέχει την τελευταία χρονική στιγμή που ο κάθε ασαφής υπόχωρος ενεργοποιήθηκε από κάποιο δεδομένο για τελευταία φορά. Όταν ένα νέο δεδομένο εισόδου γίνεται διαθέσιμο, ο αλγόριθμος αρχικά ελέγχει αν καλύπτεται από ένα ήδη επιλεγμένο ασαφή υπόχωρο. Στην αντίθετη περίπτωση, προστίθεται ένας νέος νευρώνας στην κρυφή στοιβάδα. Αυτό επιτυγχάνεται επιλέγοντας τον ασαφή υπόχωρο, ο οποίος βάσει της σχετικής απόστασης είναι πιο κοντά στο διάνυσμα εισόδου και τοποθετώντας το κέντρο του νέου κρυφού κόμβου, στο κέντρο του επιλεγμένου υπόχωρου. Στην περίπτωση αυτή, το νέο κέντρο προστίθεται στον πίνακα CLM και ο πίνακας AHV ανανεώνεται. Αν ο αλγόριθμος αποφασίσει ότι δεν χρειάζεται νέος κρυφός κόμβος, τότε ελέγχει εάν ένας υπάρχων κρυφός νευρώνας δεν έχει ανατεθεί πρόσφατα σε ένα διάνυσμα εισόδου. Αν αυτό ισχύει, ο κρυφός κόμβος διαγράφεται και ο αντίστοιχος ασαφής υπόχωρος αφαιρείται από τους επιλεγμένους υπόχωρους. Με τον τρόπο αυτό, ο αλγόριθμος διατηρεί έναν αριθμό κρυφών νευρώνων, οι οποίοι είναι επαρκείς για να περιγράψουν το σύστημα, αλλά και η δομή του δικτύου διατηρείται εντός ενός λογικού μεγέθους. Σε περίπτωση που προστεθεί ή διαγραφεί ένα κέντρο, τα συναπτικά βάρη μεταξύ της κρυφής στοιβάδας και της στοιβάδας εξόδου θα πρέπει να υπολογιστούν εκ νέου. Αυτό επιτυγχάνεται με τη χρήση ενός κινούμενου χρονικού παραθύρου, όπου αποθηκεύεται ένας αριθμός από παλαιά δεδομένα εισόδου-εξόδου. Τα συναπτικά βάρη υπολογίζονται με γραμμική παλινδρόμηση των εξόδων της κρυφής στοιβάδας στις πραγματικές εξόδους του συστήματος. [4] Ο αλγόριθμός αυτός παρουσιάζεται στο σχήμα 6.4



Σχήμα 6.4: Αλγόριθμος Ασαφών μέσων σε πραγματικό χρόνο

6.2.1 Αρχικοποίηση του αλγορίθμου

Στην αρχή του αλγορίθμου κάθε διάσταση της μεταβλητής εισόδου i , $i=1..N$, πρέπει να διαμεριστεί σε c_i ασαφείς χώρους, ενώ επίσης πρέπει να ορισθούν μερικές λειτουργικές παράμετροι:

- Ο αριθμός των N_d συνεχόμενων χρονικών βημάτων όπου ένα κέντρο δεν επιλέχθηκε από κάποια μεταβλητή εισόδου, πριν απομακρυνθεί από τη κρυφή στοιβάδα του δικτύου.
- Ο αριθμός N_s όπου είναι ο αριθμός των παλιών δεδομένων που χρησιμοποιούνται για τους μετέπειτα υπολογισμούς.

Επίσης ο αριθμός των κέντρων στην αρχή L είναι ίσος με μηδέν.

6.2.2 Προσδιορισμός του πρώτου κέντρου

Όταν ληφθεί το πρώτο δεδομένο εισόδου, δημιουργείται το πρώτο κέντρο, άρα προσδιορίζεται ο ασαφής υπόχωρος που είναι πιο κοντά σε αυτό το δεδομένο $\mathbf{A}^1 = [A_{1j_1}^1, A_{2j_2}^1, \dots, A_{Nj_N}^1]$, σύμφωνα με το κριτήριο της ελάχιστης απόστασης ή της μέγιστης συνεισφοράς

$$j_i = \arg \min_{1 \leq j \leq c_i} \mu_{A_{ij}}(x_i(1)) \quad (6.13)$$

Ο ασαφής υπόχωρος είναι ο πρώτος που συλλέγεται και ο πρώτος κρυφός κόμβος ο οποίος δημιουργείται, τοποθετώντας το κέντρο στο κέντρο του υποχώρου. Αφού έχει ορισθεί ο πρώτος υπόχωρος, ο αλγόριθμος ενεργοποιεί τους δύο πίνακες που αναφέρθηκαν παραπάνω CLM (\mathbf{G}) και AHV (\mathbf{h}). Αφού δημιουργηθεί ο πρώτος κρυφός κόμβος, η διάσταση του \mathbf{G} είναι $1 \times N$ και ο \mathbf{h} περιέχει ένα στοιχείο. Συγκεκριμένα ο πρώτος κρυφός κόμβος αποτελεί την πρώτη γραμμή του πίνακα \mathbf{G} :

$$\mathbf{G}^1 \leftarrow \mathbf{a}^1 \quad (6.14)$$

Και η τιμή του πρώτου στοιχείου h είναι ίση με 1:

$$\mathbf{h}^1 \leftarrow 1 \quad (6.15)$$

δηλαδή το πρώτο δεδομένο είναι το τελευταίο σημείο που ανατέθηκε στο ασαφή υπόχωρο. Τέλος σε αυτό το πρώτο βήμα ο αλγόριθμος υπολογίζει τα διανύσματα \mathbf{z} και \mathbf{w} , όπου θα υπολογίζονται σε κάθε χρονική στιγμή. Το διάνυσμα \mathbf{z} είναι διάστασης $L \times 1$ και περιέχει τις αποκρίσεις των κρυφών στοιβάδων, ενώ το διάνυσμα \mathbf{w} (τις ίδιας διάστασης με το \mathbf{z}) αποθηκεύει τις συνάψεις μεταξύ της κρυφής στοιβάδας και της στοιβάδας εξόδου. Προφανώς όταν το πρώτο δεδομένο είναι διαθέσιμο τα παραπάνω διανύσματα έχουν ένα στοιχείο.

$$\mathbf{z}(1) = f(\|\mathbf{x}(1) - \mathbf{G}^1\|_2) \quad (6.16)$$

$$\mathbf{w}(1) = \frac{\mathbf{y}(1)}{\mathbf{z}(1)} \quad (6.17)$$

Όπου f είναι η συνάρτηση ακτινικής βάσης. Έτσι τελειώνουν οι υπολογισμοί που γίνονται στην αρχή του αλγορίθμου. Στη συνέχεια παρατίθενται τα επόμενα βήματα.

6.2.3 Έλεγχος νέου κέντρου

Στη χρονική στιγμή k ελέγχεται αν χρειάζεται να προστεθεί νέο κέντρο. Ο αλγόριθμος υπολογίζει τις αποστάσεις μεταξύ του $\mathbf{x}(k)$ και των υπάρχων υποχώρων και υπολογίζει το κέντρο που ελαχιστοποιεί αυτή την απόσταση.

$$d_r^{l_0}(\mathbf{x}(k)) = \min_{1 \leq l \leq L} d_r^l(\mathbf{x}(k)) \quad (6.18)$$

$$l_0 = \arg \max_{1 \leq l \leq L} d_r^l(\mathbf{x}(k)) \quad (6.19)$$

Αν αυτή απόσταση είναι μεγαλύτερη από 1 (όπως και στην offline έκδοση), τότε κανένα από τα κέντρα δεν δύναται να περιγράψει το καινούριο δεδομένο και νέο κέντρο χρειάζεται να προστεθεί ανανεώνοντας το διάνυσμα \mathbf{h} .

6.2.4 Επιλογή του νέου κέντρου

Όταν προστεθεί νέος κόμβος, ο αλγόριθμος αυξάνει τα κέντρα κατά 1 και καθορίζεται ο ασαφής υπόχωρος $\mathbf{A}^l = [A_{1j_1}^L, A_{2j_2}^L, \dots, A_{Nj_N}^L]$ όπου είναι κοντύτερα στον νέο δεδομένο $\mathbf{x}(k)$, χρησιμοποιώντας το κριτήριο της ελάχιστης απόστασης

$$j_i = \arg \max_{1 \leq j \leq c_i} \mu_{A_{ij}}(x_i(1)) \quad (6.20)$$

Το κέντρο ανατέθηκε στη κρυφή στοιβίδα του δικτύου και η διάσταση του πίνακα \mathbf{G} (CLM) αυξάνεται ώστε να περιέχεται και ο νέος κόμβος

$$\mathbf{G}^L \leftarrow \alpha^L \quad (6.21)$$

Το διάνυσμα \mathbf{h} αυξάνεται και αυτό λαμβάνοντας την τιμή k , καθώς το k δεδομένο είναι το τελευταίο δεδομένο που ανατέθηκε

$$\mathbf{h}^L \leftarrow k \quad (6.22)$$

6.2.5 Ανανέωση του διανύσματος ιστορικού ενεργοποίησης

Όταν ο αλγόριθμος αποφασίζει να μην τοποθετήσει νέο κόμβο τότε το διάνυσμα ιστορικού ανανεώνεται ως εξής $\mathbf{h}^l = k$. Δηλαδή το εισερχόμενο δεδομένο k είναι το τελευταίο δεδομένο που ανατέθηκε στο l_0 ασαφή υπόχωρο και υπολογίζεται από την εξ. 6.19

6.2.6 Έλεγχος διαγραφής υπάρχοντος κόμβου

Σε αυτό το βήμα ο αλγόριθμος βρίσκει το στοιχείο του διανύσματος \mathbf{AHV} με τη μικρότερη τιμή και άρα τον υπόχωρο που έκανε τον περισσότερο χρόνο να του ανατεθεί νέο δεδομένο εισόδου.

$$l_r = \arg \max_{1 \leq l \leq L} \mathbf{h}^l(k) \quad (6.23)$$

Αν ισχύει η ανισότητα $\mathbf{h}^{l_r} \leq k - N_d$, τότε κανένα δεδομένο εισόδου δεν έχει τοποθετηθεί στον υπόχωρο l_r για N_d συνεχόμενα βήματα. Σε αυτή τη περίπτωση αυτός ο κόμβος δεν χρειάζεται και πρέπει να διαγραφεί.

6.2.7 Διαγραφή κόμβου

Αν ο αλγόριθμος αποφασίσει να διαγράψει τον κόμβο l_r , τότε διαγράφεται η αντίστοιχη γραμμή του πίνακα CLM (G^{l_r}) και το στοιχείο του διανύσματος ιστορικού ενεργοποίησης \mathbf{h}^{l_r} .

6.2.8 Υπολογισμός νέων συναπτικών βαρών με τη μέθοδο των ελαχίστων τετραγώνων.

Όταν ο αλγόριθμος προσθέτει ή αφαιρεί κάποιον κόμβο τότε οι νέες συνδέσεις πρέπει να υπολογιστούν ξανά. Για να συμβεί αυτό αποθηκεύονται παλιά εισερχόμενα δεδομένα στον πίνακα $R(k)$ που φαίνεται στην εξίσωση .

$$R(i - k + N_s + 1, j) = f(\|\mathbf{x}(i) - \mathbf{G}^j\|_2) \quad (6.24)$$

όπου $i = k - N_s, \dots, k$ και $j = 1, \dots, L$ Έτσι υπολογίζονται οι νέες συνδέσεις σύμφωνα με τις εξ. (6.25)(6.26).

$$\mathbf{w}(k) = (R^T(k)R(k))^{-1}R^T(k)\mathbf{s}(k) \quad (6.25)$$

$$\mathbf{s}(k) = [y(k - N_s), y(k - N_s + 1), \dots, y(k)]^T \quad (6.26)$$

Όταν αριθμός των κόμβων παραμένει σταθερός τότε υπάρχουν πολλές τεχνικές για την ανανέωση των συνδέσεων όπως η μέθοδος RLS και η gradient decent, ανάλογα με την εκάστοτε περίπτωση. Στην παρούσα εργασία για το πρόβλημα ρύθμισης που θα αναλυθεί στη συνέχεια χρησιμοποιήθηκε η μέθοδος gradient descent.

Κεφάλαιο 7

Μεθοδολογία Ρύθμισης Μη Γραμμικών Συστημάτων με Ενισχυτική Μάθηση

Η προτεινόμενη μεθοδολογία αποτελεί επέκταση της μεθοδολογίας που παρουσιάστηκε στην εργασία [26] με την αντικατάσταση των δικτύων MLP από RBF. Αυτή η αντικατάσταση σε συνδυασμό με τον αλγόριθμο εκπαίδευσης των ασαφών μέσων επιτρέπει την ανάπτυξη των νευρωνικών δικτύων από απόλυτα μηδενική βάση, ενώ η δομή μεταβάλλεται όσο εξελίσσεται η προσομοίωση. Στην εργασία αναφοράς [26] η δομή των δικτύων επιλέγεται τυχαία και παραμένει σταθερή. Το ίδιο συμβαίνει με όλες τις παραμέτρους των δικτύων, εκτός από τις παραμέτρους που συσχετίζουν την τελευταία κρυφή στοιβάδα με τη στοιβάδα εξόδου. Είναι εμφανές ότι η μεθοδολογία αυτή παρουσιάζει ευαισθησία σε σχέση με την αρχική τυχαία επιλογή.

Η μεθοδολογία ρύθμισης μη γραμμικών συστημάτων με ενισχυτική μάθηση βασίζεται σε δύο νευρωνικά δίκτυα που αλληλοεπιδρούν μεταξύ τους με τρόπο «ενέργειας-κριτικής». Το πρώτο δίκτυο, επονομαζόμενο και ως ενέργειας (actor), προσεγγίζει το μη γραμμικό κομμάτι του νόμου ελέγχου, παράγοντας έτσι ένα μέρος της ρυθμιστικής δράσης. Το δεύτερο δίκτυο, το οποίο ονομάζεται κριτής (critic), προσεγγίζει τη συνάρτηση κόστους με σκοπό να ανανεώνει τα βάρη του δικτύου ενέργειας. Το σύστημα σε κάθε χρονική στιγμή k , αναπροσαρμόζει τα νευρωνικά δίκτυα, σύμφωνα

με τα δεδομένα που έχει λάβει τις προηγούμενες χρονικές στιγμές και τις τιμές της μεταβλητής εξόδου που λαμβάνει την παρούσα χρονική στιγμή. Το σύστημα το οποίο θα πρέπει να ρυθμιστεί είναι διακριτό, μη γραμμικό και περιγράφεται από τις ακόλουθες εξισώσεις:

$$\begin{aligned} \mathbf{x}(k+1) &= \mathbf{F}(\mathbf{x}(k), u(k)) + \mathbf{d}(k) \\ y(k) &= x_1(k) \end{aligned} \quad (7.1)$$

Όπου $\mathbf{x}(k) = [x_1(k), x_2(k), \dots, x_n(k)]^T \in \mathbb{R}^n$ οι μεταβλητές κατάστασης, $u(k) \in \mathbb{R}^n$ η μεταβλητή εισόδου, $\mathbf{d}(k) \in \mathbb{R}^n$ η τυχαία διαταραχή που περιορίζεται από μια θετική τιμή d_M με $\|d_i(k)\| < d_M$, $\mathbf{F}(\mathbf{x}(k), u(k)) = [F_1(\mathbf{x}(k), u(k)), F_2(\mathbf{x}(k), u(k)), \dots, F_n(\mathbf{x}(k), u(k))] \in \mathbb{R}^n$ μη γραμμική συνάρτηση, με $F_i(0,0) = 0, \forall i$ $y(k)$ η μεταβλητή εξόδου του συστήματος.

Τη χρονική στιγμή k οι επιθυμητές τιμές της μεταβλητής εξόδου για την παρούσα αλλά και για της επόμενες χρονικές στιγμές συμβολίζονται με $y_d(k+i)$ και το σφάλμα ανάμεσα στην επιθυμητή τιμή και τη μεταβλητή εξόδου είναι

$$e_i(k) = y_d(k+i) - y(k+i) \quad (7.2)$$

Στην περίπτωση όπου ισχύει $F_i(\mathbf{x}(k), u(k)) = x_{(i+1)}(k)$ για κάθε τιμή του i από 0 έως $n-1$ τότε ισχύει ότι

$$\begin{aligned} y_d(k+i) - y(k+i) &= y_d(k+i) - x_1(k+i) = \\ y_d(k+i) - x_2(k+i-1) &= y_d(k+i) - x_i(k+i-(i-1)) = \\ &= y_d(k+i) - x_{(i+1)}(k) \end{aligned} \quad (7.3)$$

7.1 Σχεδιασμός Ρυθμιστή

Η σχέση που περιγράφει τη συμπεριφορά της μεταβλητής εξόδου μπορεί να περιγραφεί από μια εξίσωση της μορφής:

$$y(m+k) = F_m = au(k) + f(\mathbf{x}(k), u(k)) + d(k) \quad (7.4)$$

Η μετατροπή της εξίσωσης σε αυτή τη μορφή είναι απαραίτητη για τη συνέχεια του σχεδιασμού, ενώ το m λαμβάνει την τιμή n όταν ισχύει η περίπτωση 7.3, ενώ

διαφορετικά η τιμή είναι 1. Έτσι, πλέον, μπορεί να ορισθεί το σήμα εισόδου ως:

$$u(k) = \frac{1}{\alpha}(u_s(k) - u_a(k)) \quad (7.5)$$

με $u_s(k)$ συμβολίζεται ο ρυθμιστής ανατροφοδότησης που έχει ως στόχο να κάνει ευσταθή τη γραμμικοποιημένη μορφή του σφάλματος, ενώ με $u_a(k)$ συμβολίζεται ο προσαρμοστικός ρυθμιστής που σκοπεύει να προσεγγίσει την συνάρτηση $f(\mathbf{x}, u)$. Σύμφωνα με τα παραπάνω, η εξίσωση 7.4 με βάση την εξίσωση 7.5 μπορεί να γραφεί ως εξής:

$$y(m+k) = u_s(k) - u_a(k) + f(\mathbf{x}(k), u(k)) + d(k) \quad (7.6)$$

Άρα, επειδή το u_a προσεγγίζει την f ισχύει:

$$u_a(k) = \hat{f}(\mathbf{x}(k), u(k)) \quad (7.7)$$

Ενώ το $u_s(k)$ ορίζεται ως εξής

$$u_s(k) = y_d(k+m) + \boldsymbol{\lambda}^\top \mathbf{e}(k) \quad (7.8)$$

Με $\mathbf{e}(k)=[e_0(k), \dots, e_{m-1}(k)]^\top$. Οι τιμές του $\boldsymbol{\lambda}$ επιλέγονται έτσι ώστε να είναι το σύστημα ευσταθές και άρα, βρίσκονται εντός του μοναδιαίου κύκλου ($|z^m + \lambda_1 z^{m-1} + \dots + \lambda_m| = 0$). Πλέον μπορεί προκύψει εύκολα [26]:

$$\mathbf{e}(k+1) = \mathbf{A} \cdot \mathbf{e}(k) + \mathbf{B} \cdot (\hat{f}(\mathbf{x}(k), u(k)) - f(\mathbf{x}(k), u(k)) - \mathbf{d}(k)) \quad (7.9)$$

$$\text{με } \mathbf{A} = \begin{pmatrix} 0 & 1 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ -\lambda_m & -\lambda_{m-1} & \dots & -\lambda_1 \end{pmatrix} \text{ και } \mathbf{B} = \begin{pmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ 1 \end{pmatrix}$$

Είναι προφανές ότι αν υπάρχει u_a τέτοιο ώστε $\hat{f}(\mathbf{x}(k), u(k)) - f(\mathbf{x}(k), u(k)) = 0$ αγνοώντας τη διαταραχή, τότε το σύστημα μετατρέπεται σε γραμμικό, και μάλιστα ευσταθές αφού οι τιμές του $\boldsymbol{\lambda}$ έχουν επιλεγεί ώστε ο πίνακας \mathbf{A} να είναι ευσταθής. Χρησιμοποιώντας το θεώρημα των πεπλεγμένων συναρτήσεων μπορεί να προκύψει το συμπέρασμα ότι υπάρχει u_a τέτοιο ώστε να ικανοποιούνται τα παραπάνω. [26]

7.2 Μάθηση Νευρωνικών Δικτύων Σε Πραγματικό Χρόνο- Δίκτυο Κριτής

Σε αυτή την ενότητα αρχικά ορίζεται η συνάρτηση κόστους J [12], όπως περιγράφεται από την παρακάτω εξίσωση:

$$J(k) = \tau^{N_p} r(\mathbf{e}(k+1)) + \tau^{N_p-1} r(\mathbf{e}(k+2)) + \dots \tau^{k+1} r(\mathbf{e}(k+N_p)) \quad (7.10)$$

Στην οποία r αποτελεί μια βοηθητική συνάρτηση που αναδεικνύει αν η επίδοση είναι καλή ή όχι και N_p ο ορίζοντας ρύθμισης. Η συνάρτηση αυτή ορίζεται ως:

$$r(k) = \begin{cases} 0 & \text{if } f\boldsymbol{\lambda}^T \cdot \mathbf{e}(k) < \varepsilon \\ 1 & \text{if } f\boldsymbol{\lambda}^T \cdot \mathbf{e}(k) > \varepsilon \end{cases} \quad (7.11)$$

Όπου $0 < \tau < 1$ και ε αποτελούν μεταβλητές σχεδιασμού. Από την εξίσωση 7.10 προκύπτει η εξίσωση *Bellman*:

$$J(k) = \tau J(k-1) - \tau^{N_p-1} r(\mathbf{e}(k)) \quad (7.12)$$

Έτσι, το σφάλμα πρόβλεψης προσδιορίζεται ως εξής:

$$e_c(k) = \hat{J}(k) - (\tau \hat{J}(k-1) - \tau^{N_p-1} r(\mathbf{e}(k))) \quad (7.13)$$

Όπου $\hat{J}(k)$ η έξοδος του δικτύου κριτή. Στην παρούσα εργασία εφαρμόστηκαν νευρωνικά δίκτυα RBF και συγκεκριμένα χρησιμοποιήθηκε ως συνάρτηση ενεργοποίησης η συνάρτηση Gauss, άρα το δίκτυο κριτής περιγράφεται από την παρακάτω εξίσωση:

$$\hat{J}(k) = \hat{\mathbf{w}}_c^T(k) g_c(k) \quad (7.14)$$

Όπου:

$$g_c(\mathbf{x}(k)) = [g_{c1}(\mathbf{x}(k)), g_{c2}(\mathbf{x}(k)), \dots, g_{cL_c}(\mathbf{x}(k))] \\ g_{ci}(\mathbf{x}(k)) = \exp\left(-\frac{\|\mathbf{x}(k) - \mathbf{c}_{ci}(k)\|^2}{\sigma_{ci}(k)}\right), \quad i=1, \dots, L_c(k),$$

$L_c(k)$ είναι το πλήθος των νευρώνων της κρυφής στοιβάδας τη χρονική στιγμή k και $\mathbf{c}_{ci}(k)$, $\sigma_{ci}(k)$ αποτελούν το κέντρο και το πλάτος του νευρώνα i , αντίστοιχα.

Στόχος της εκπαίδευσης του δικτύου είναι η ελαχιστοποίηση του σφάλματος:

$$E_c(k) = \frac{1}{2} e_c^T(k) e_c(k) \quad (7.15)$$

Η ελαχιστοποίηση γίνεται μέσω της μεθόδου *gradient decent* όπου:

$$\begin{aligned} \hat{w}_c(k+1) &= \hat{w}_c(k) - l_c (\partial E_c(k)) / (\partial \hat{w}_c(k)) \\ &= \hat{w}_c(k) - l_c \frac{\partial E_c(k)}{\partial e_c(k)} \frac{\partial e_c(k)}{\partial \hat{J}(k)} \frac{\partial \hat{J}(k)}{\partial \hat{w}_c(k)} \end{aligned} \quad (7.16)$$

με $0 < l_c \leq 1$ να είναι ο ρυθμός μάθησης. Με βάση την παραπάνω εξίσωση προκύπτει ότι:

$$\hat{w}_c(k+1) = \hat{w}_c(k) - l_c g_c(\mathbf{x}(k)) [\hat{J}(k) - (\tau \hat{J}(k-1) - \tau^{N_p-1} r(\mathbf{e}(k)))] \quad (7.17)$$

Η παραπάνω διαδικασία μάθησης χρησιμοποιείται όταν η δομή του νευρωνικού δικτύου (δηλαδή το πλήθος των νευρώνων της εσωτερικής στοιβάδας) παραμένει σταθερή, ωστόσο η προτεινόμενη μεθοδολογία επιτρέπει τη μεταβολή της δομής, σύμφωνα με τη μέθοδο που παρουσιάζεται στην ενότητα 5.2. Θα πρέπει να τονισθεί ότι η αρχικοποίηση της μεθοδολογίας θεωρεί ότι δεν υπάρχει καμία γνώση για το σύστημα, επομένως ξεκινάει με μηδέν νευρώνες στην κρυφή στοιβάδα και η δομή του δικτύου κατασκευάζεται σταδιακά με τη συλλογή πληροφοριών από το σύστημα κατά τη δυναμική λειτουργία του.

Σε περίπτωση που υπάρξει αλλαγή της δομής, η διαδικασία επικαιροποίησης των βαρών διαφοροποιείται και ακολουθούνται δύο διαφορετικές τακτικές ανάλογα με το αν το πρόβλημα είναι ‘παρακολούθησης τροχιάς’ (tracking trajectory) ή ‘βηματικής αλλαγής της επιθυμητής τιμής’: Στην πρώτη περίπτωση, τα βάρη αναπροσαρμόζονται στα ιστορικά δεδομένα με τη μέθοδο της γραμμικής παλινδρόμησης, όπως αυτή παρουσιάζεται στην ενότητα 5.2, και στη συνέχεια αν έχει διαγραφεί νευρώνας της κρυφής στοιβάδας, εφαρμόζεται επιπλέον η μέθοδος gradient decent. Στη δεύτερη περίπτωση, η εκπαίδευση συνεχίζεται με τη μέθοδο gradient decent, ενώ όταν προστίθεται νευρώνας, αυτός αποκτά αρχικά βάρος σύνδεσης με τη στοιβάδα εξόδου, που αντιστοιχεί στο αμέσως προηγούμενο κέντρο που προστέθηκε στο δίκτυο. Στη μεθοδολογία που παρουσιάζεται, αποτελούν παραμέτρους σχεδιασμού, η

χρονική στιγμή N_d , όπου από εκεί και πέρα επιτρέπεται η δυνατότητα αφαίρεσης νευρώνων της κρυφής στοιβάδας, η χρονική στιγμή N_{max} στην οποία σταματούν οι μεταβολές στη δομή των δικτύων και το πλήθος των ιστορικών δεδομένων N_s που αποθηκεύονται στη περίπτωση που εφαρμόζονται ελάχιστα τετράγωνα. Επίσης χρειάζεται να επιλεχθεί και ο αριθμός των ασαφών υποχώρων, s , που ορίζονται σε κάθε μεταβλητή εισόδου.

7.3 Μάθηση Νευρωνικών Δικτύων Σε Πραγματικό Χρόνο- Δίκτυο Ενέργειας

Θεωρώντας ότι το σφάλμα του δικτύου είναι $e_a(\mathbf{k}) = \hat{J}(k) - J_d(k) + \hat{f}(\mathbf{x}(k), u(k)) - f(\mathbf{x}(k), u(k))$, με $J_d(k) = 0$, και επειδή το νευρωνικό δίκτυο έχει ως έξοδο τη προσέγγιση της f ισχύει ότι:

$$\hat{f}(k) = \hat{w}_a^T(k) g_a(\mathbf{z}(k)) \quad (7.18)$$

$$\begin{aligned} \text{Όπου } \mathbf{z}(k) &= [\mathbf{x}(k), u_s(k)]^T \\ g_a(\mathbf{z}(k)) &= [g_{a1}(\mathbf{z}(k)), g_{a2}(\mathbf{z}(k)), \dots, g_{aL_a(k)}(\mathbf{z}(k))] \\ g_{ai}(\mathbf{z}(k)) &= \exp\left(-\frac{\|\mathbf{z}(k) - \mathbf{c}_{ai}(k)\|^2}{\sigma_{ai}(k)}\right), \quad i=1, \dots, L_a(k), \end{aligned}$$

$L_a(k)$ είναι το πλήθος των νευρώνων της κρυφής στοιβάδας τη χρονική στιγμή k και $\mathbf{c}_{ai}(k)$, $\sigma_{ai}(k)$ αποτελούν το κέντρο και το πλάτος του νευρώνα i , αντίστοιχα. Στόχος είναι και πάλι η ελαχιστοποίηση του σφάλματος:

$$E_a(k) = \frac{1}{2} e_a^T(k) e_a(k) \quad (7.19)$$

Άρα θεωρώντας $\tilde{f} = \hat{f} - f$, έχουμε:

$$\begin{aligned} \hat{w}_\alpha(k+1) &= \hat{w}_\alpha(k) - l_\alpha \frac{\partial E_\alpha(k)}{\partial w_\alpha(k)} \\ &= \hat{w}_\alpha(k) - l_\alpha \frac{\partial E_\alpha(k)}{\partial e_\alpha(k)} \frac{\partial e_\alpha(k)}{\partial \tilde{f}(k)} \frac{\partial \tilde{f}(k)}{\partial \tilde{w}_\alpha(k)} \end{aligned} \quad (7.20)$$

με $0 < l_\alpha \leq 1$ να αποτελεί τον ρυθμό μάθησης. Άρα:

$$\hat{w}_\alpha(k+1) = \hat{w}_\alpha(k) - l_\alpha g_\alpha(\mathbf{z}(k)) [\hat{w}_c^T(k) g_c(\mathbf{x}(k)) + \tilde{f}] \quad (7.21)$$

με $\tilde{f} = \mathbf{B}^T (\mathbf{e}(k+1) - \mathbf{A} \mathbf{e}(k) + \mathbf{d})$. Ωστόσο η διαταραχή \mathbf{d} κατά βάση είναι άγνωστη και θεωρείται 0 για τη διαδικασία της μάθησης. Άρα:

$$\hat{w}_\alpha(k+1) = \hat{w}_\alpha(k) - l_\alpha g_\alpha(\mathbf{z}(k)) [\hat{w}_c^T(k) g_c(\mathbf{x}(k)) + \mathbf{B}^T (\mathbf{e}(k+1) - \mathbf{A} \mathbf{e}(k))] \quad (7.22)$$

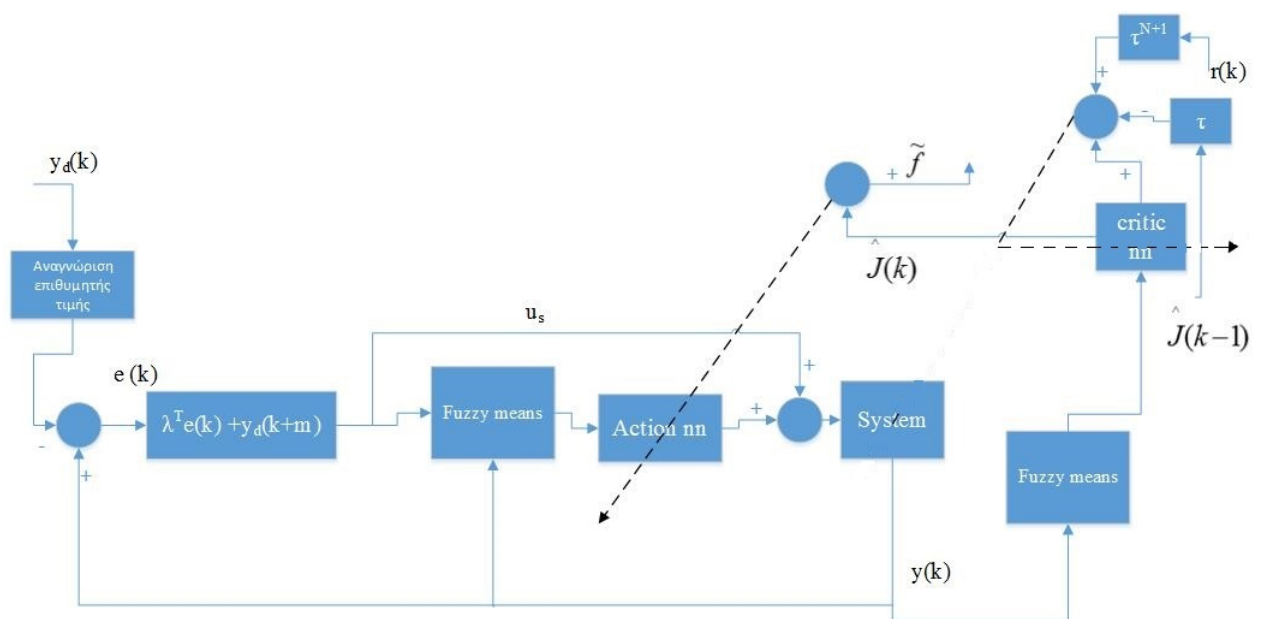
ενώ η γενική διαδικασία μάθησης όταν δηλαδή δεν παραμένουν τα κέντρα σταθερά είναι ίδια με αυτή του προηγούμενου δικτύου.

7.4 Διάγραμμα Ροής

Είναι προφανές ότι ο αλγόριθμος θα πρέπει να αλλάζει ανάλογα με το αν οι επιθυμητές τιμές έχουν σταθερή τιμή και αν ακολουθούν κάποια συγκεκριμένη πορεία (tracking trajectory), για αυτό το λόγο προστέθηκε μια επιπλέον μεταβλητή που αναγνωρίζει την τάση που έχουν οι επιθυμητές τιμές και εφαρμόζει ανάλογα τον αντίστοιχο τρόπο ανανέωσης των συναπτικών βαρών, σύμφωνα με την ενότητα 7.2. Στην περίπτωση όπου οι επιθυμητές τιμές ακολουθούν συγκεκριμένη πορεία, τότε τα βάρη αναπροσαρμόζονται στα ιστορικά δεδομένα με τη μέθοδο της γραμμικής παλινδρόμησης, και αν έχει πραγματοποιηθεί διαγραφή νευρώνων τότε εφαρμόζεται επιπλέον η μέθοδος gradient decent. Στην άλλη περίπτωση, η εκπαίδευση συνεχίζεται με τη μέθοδο gradient decent, και όταν προστεθεί νευρώνας, αυτός λαμβάνει το αρχικό συναπτικό βάρος με τη στοιβάδα εξόδου, που αντιστοιχεί στο προηγούμενο κέντρο που προστέθηκε.

Έτσι γίνεται χρήση της μεταβλητής *slope* που μπορεί και αναγνωρίζει το είδος επιθυμητών, καθώς αφαιρεί την τωρινή επιθυμητή τιμή με την προηγούμενη, και αν δεν είναι μηδέν εφαρμόζεται η περίπτωση όπου το σύστημα χρειάζεται να ακολουθήσει κάποια τροχιά.

Το συνολικό διάγραμμα ροής του αλγορίθμου παρουσιάζεται στο σχήμα 7.1. Είναι φανερό ότι πριν γίνει η είσοδος στα ΤΝΔ εφαρμόζεται και ο αλγόριθμος των ασαφών μέσων, ώστε να γίνει αν χρειαστεί οποιαδήποτε αλλαγή στη δομή τους.



Σχήμα 7.1: Διάγραμμα ροής της μεθοδολογίας ρύθμισης

7.5 Εφαρμογές

Η προτεινόμενη μεθοδολογία εφαρμόστηκε με επιτυχία σε δύο περιπτώσεις. Σκοπός ήταν να μελετηθεί ο αλγόριθμος τόσο στις δυο διαφορετικές αλλαγές των επιθυμητών τιμών αλλά και στο πως ανταποκρίνονται σε διαφορετικά προβλήματα.

7.5.1 Εφαρμογή 1

Το πρώτο πρόβλημα που μελετήθηκε από τους X. Yang et al. [26] αποτελείται από ένα σύστημα που βρίσκεται σε κανονική μορφή (Brunovsky [6]) και περιγράφονται από τις παρακάτω εξισώσεις:

$$\begin{aligned}x_1(k+1) &= x_2(k) \\x_2(k+1) &= \frac{x_1(k)x_2(k)(x_1(k)+2.5)}{1+x_1(k)^2+x_2(k)^2} + u(k) + 0.1u(k)^3 + d(k) \\y(k) &= x_1(k)\end{aligned}\quad (7.23)$$

Στο σύστημα αυτό $u(k)$ αποτελεί την μεταβλητή εκ χειρισμού και $d(k)=0.1 \cos(0.001k)$ η εξωτερική διαταραχή. Σε αυτή τη περίπτωση ο στόχος είναι να ακολουθηθεί η τροχιά:

$$y_d(k) = 0.6 \sin(\pi k / 265) \quad (7.24)$$

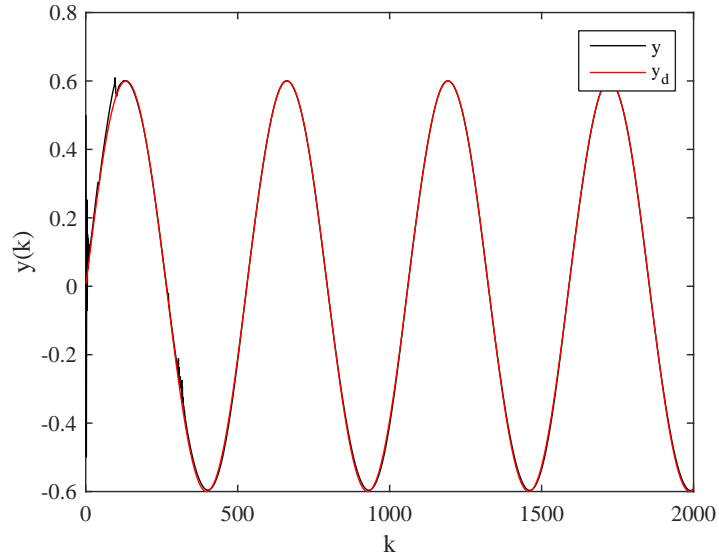
Περίπτωση 1.1

Επιλέχθηκε η μέθοδος εκπαίδευσης που βασίζεται στα συμμετρικά ασαφή σύνολα και οι τιμές των παραμέτρων σχεδιασμού φαίνονται στον πίνακα 7.1

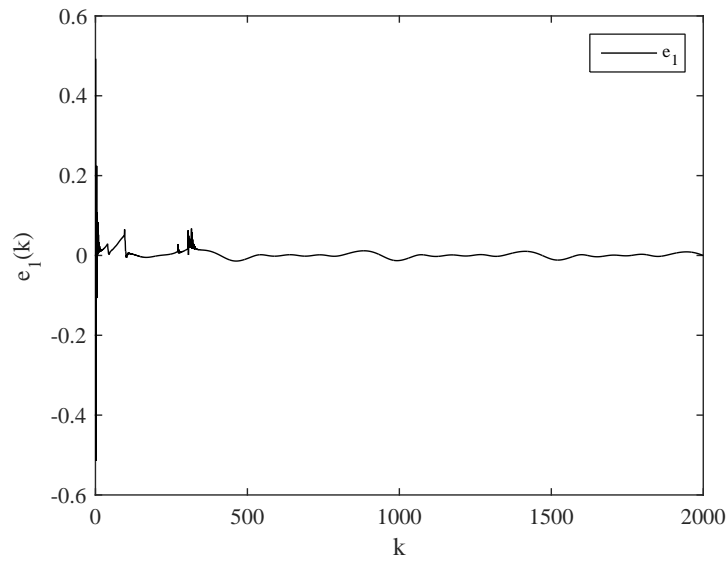
Σημειώνεται ότι και για τα δύο νευρωνικά δίκτυα χρησιμοποιήθηκαν οι ίδιες τιμές των παραμέτρων, με το σύστημα να ξεκινά από την αρχική κατάσταση $x_1(0)=0.5$, $x_2(0)=-0.5$. Στο σχήμα 7.2 παρουσιάζεται η απόκριση του συστήματος, ενώ στο σχήμα 7.3 παρουσιάζεται το σφάλμα ανάμεσα στην απόκριση της μεταβλητής εξόδου και την επιθυμητή τροχιά. Είναι προφανές ότι το σύστημα ακολουθεί με μεγάλη επιτυχία την τροχιά, και το σφάλμα είναι σημαντικό μόνο στα πρώτα βήματα της διαδικασίας. Επίσης, στο σχήμα 7.4 παρουσιάζεται η εξέλιξη του πλήθους των κέντρων των δύο δικτύων, και είναι φανερό ότι δεν χρειάζονται παραπάνω από 7 στο δίκτυο actor ενώ στο δίκτυο critic δεν ξεπερνάει ποτέ τα 8 κέντρα για να υπάρξει ικανοποιητική απόκριση.

Πίνακας 7.1: Τιμές Σχεδιασμού περίπτωση 1.1

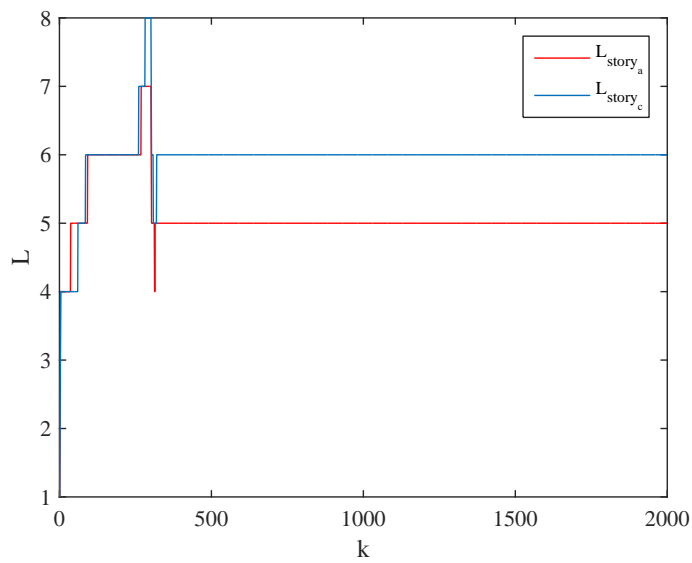
Τιμές Σχεδιασμού περίπτωση 1.1	
λ_1	1
λ_2	0.25
N_d	300
N_s	300
N_{max}	350
s	10
a	1
l_c	0.1
l_α	0.1



Σχήμα 7.2: Η πορεία της μεταβλητής εξόδου και της επιθυμητής τιμής σε συνάρτηση με το χρόνο, περίπτωση 1.1.



Σχήμα 7.3: Η πορεία του σφάλματος σε συνάρτηση με το χρόνο, περίπτωση 1.1.



Σχήμα 7.4: Η εξέλιξη του αριθμού των κέντρων των δύο δικτύων, περίπτωση 1.1.

Σε αυτή τη περίπτωση το σφάλμα (Sum Squared Error, SSE) υπολογίστηκε

0.897, το οποίο θα φανεί ότι είναι μεγαλύτερο σε σχέση με την επόμενη περίπτωση όπου έγινε χρήση του αλγορίθμου εκπαίδευσης που βασίζεται στη μη συμμετρική διαμέριση του χώρου των μεταβλητών εισόδου.

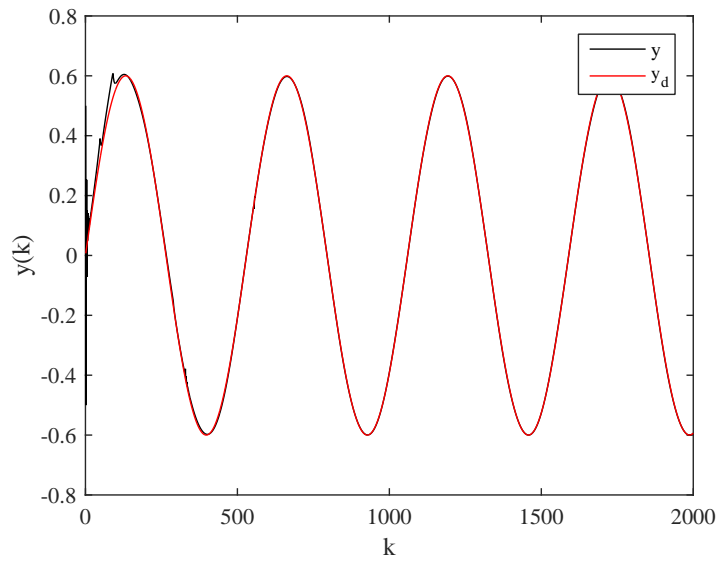
Περίπτωση 1.2

Στη συνέχεια εφαρμόστηκε ο αλγόριθμος εκπαίδευσης που βασίζεται στη μη συμμετρική διαμέριση του χώρου των μεταβλητών εισόδου, συγκεκριμένα οι τιμές σχεδιασμού φαίνονται στον πίνακα 7.2

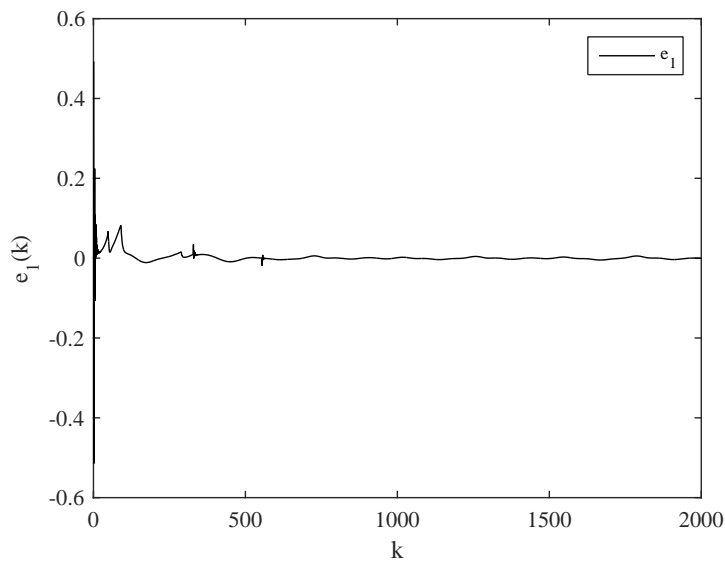
Πίνακας 7.2: Τιμές Σχεδιασμού Περίπτωση 1.2

Τιμές Σχεδιασμού Περίπτωση 1.2	
λ_1	1
λ_2	0.25
N_d	550
N_s	350
N_{max}	2000
s_α	[11 9]
s_c	[13 12 10]
a	1
l_c	0.1
l_α	0.1

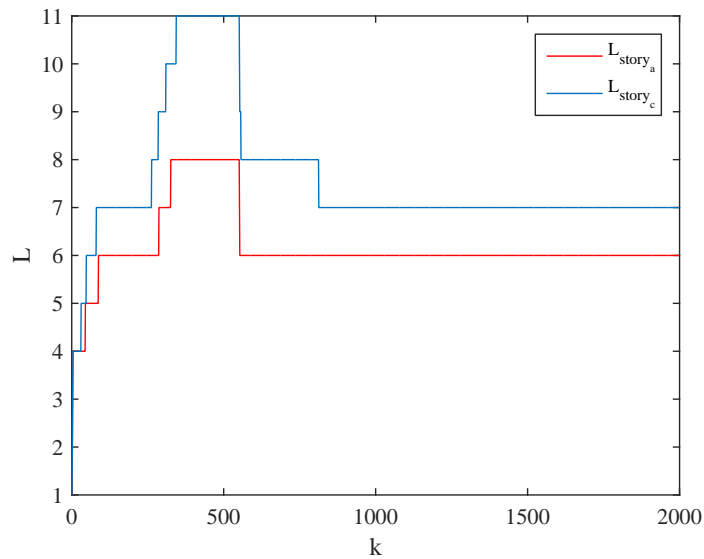
Σε αυτή την περίπτωση οι αλλαγές στη δομή του δικτύου πραγματοποιήθηκαν μέχρι το τέλος της προσομοίωσης, με τα αποτελέσματα να φαίνονται στα σχήματα 7.5-7.7. Το σφάλμα του αθροίσματος των τετραγώνων των σφαλμάτων (Sum Squared Error, SSE) υπολογίστηκε 0.883, ενώ και στα δύο δίκτυα δημιουργήθηκαν παραπάνω κέντρα. Παρατηρούμε ότι ο εναλλακτικός τρόπος εκπαίδευσης του δικτύου οδήγησε στη απομείωση του σφάλματος.



Σχήμα 7.5: Η πορεία της μεταβλητής εξόδου και της επιθυμητής τιμής σε συνάρτηση με το χρόνο, περίπτωση 1.2.



Σχήμα 7.6: Η πορεία του σφάλματος σε συνάρτηση με το χρόνο, περίπτωση 1.2.



Σχήμα 7.7: Η εξέλιξη του αριθμού των κέντρων των δύο δικτύων, περίπτωση 1.2.

Περίπτωση 1.3

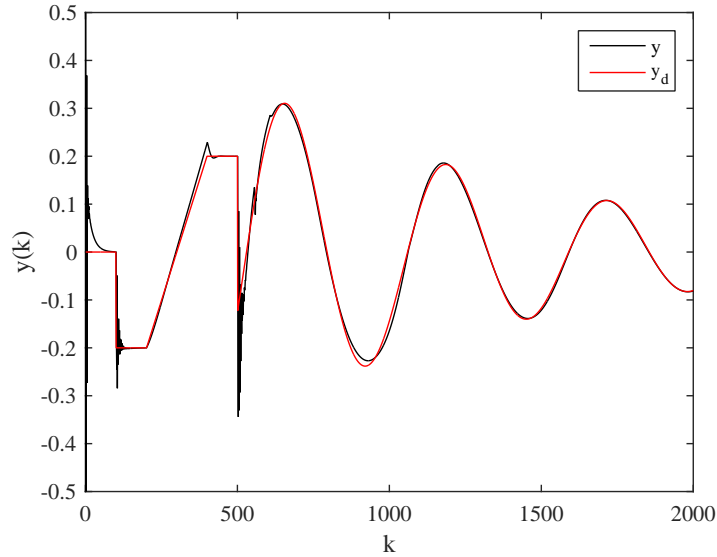
Έχει ιδιαίτερο ενδιαφέρον να φανεί η απόδοση της παρούσας μεθοδολογίας όταν υπάρχουν διαφορετικά είδη επιθυμητών τιμών που πρέπει να ακολουθήσει το σύστημα, αντί της ημοτονοειδούς συνάρτησης που εφαρμόστηκε ως τώρα. Για αυτό το λόγο χρησιμοποιείται η μεταβλητή slope για να αναγνωρίζει τον τρόπο που αλλάζουν οι επιθυμητές τιμές.

Σε αυτή τη περίπτωση οι μεταβλητές σχεδιασμού αλλάζουν μερικώς όπως φαίνονται στον πίνακα 7.3

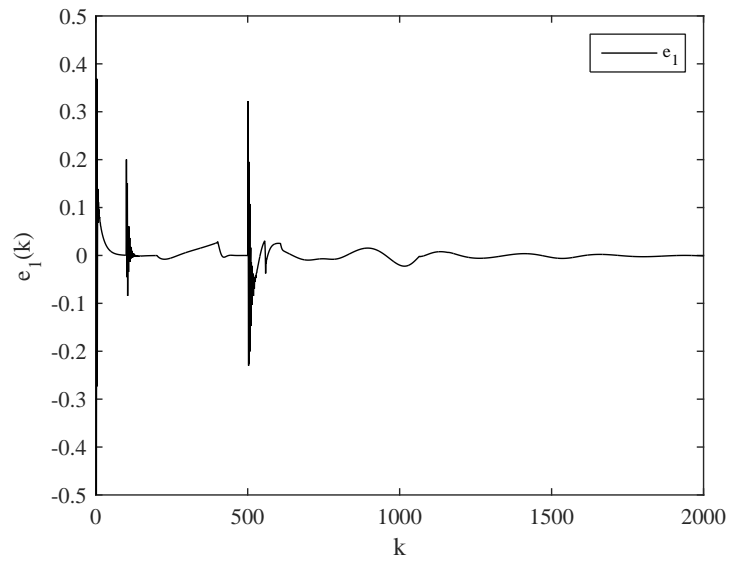
Στο σχήμα 7.8 φαίνεται το επιθυμητό προφίλ της μεταβλητής εξόδου καθώς και η απόκριση που προκύπτει από το προτεινόμενο σύστημα ρύθμισης. Όπως φαίνεται στο σχήμα 7.8, κατά την αλλαγή του τρόπου που μεταβάλλονται οι επιθυμητές τιμές το σφάλμα είναι μεγάλο γεγονός λογικό καθώς ο ρυθμιστής πρέπει να προσαρμοστεί στις νέες αλλαγές. Ενώ στο σχήμα 7.9 παρουσιάζεται το σφάλμα της μεταβλητής εξόδου.

Πίνακας 7.3: Τιμές Σχεδιασμού Περίπτωση 1.3

Τιμές Σχεδιασμού Περίπτωση 1.3	
λ_1	1
λ_2	0.25
N_d	550
N_s	350
N_{max}	2000
s_α	[11 10]
s_c	[7 12 8]
a	1
l_c	0.1
l_α	0.1

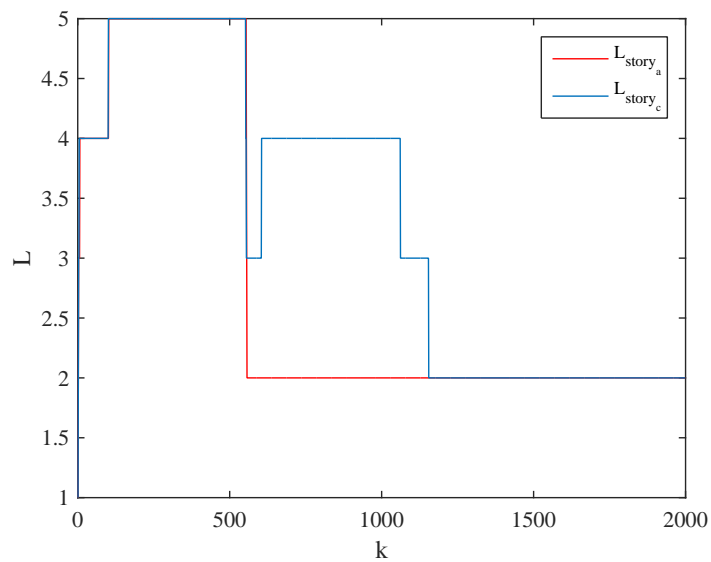


Σχήμα 7.8: Η πορεία της μεταβλητής εξόδου και της επιθυμητής τιμής σε συνάρτηση με το χρόνο, περίπτωση 1.3



Σχήμα 7.9: Η πορεία του σφάλματος σε συνάρτηση με το χρόνο, περίπτωση 1.3

Είναι φανερό από το σχήμα 7.10 ότι δεν χρειάζονται πάνω από 6 κέντρα, γεγονός που σημαίνει ότι μειώνεται αρκετά το υπολογιστικό κόστος.



Σχήμα 7.10: Η εξέλιξη του αριθμού των κέντρων των δύο δικτύων, περίπτωση 1.3

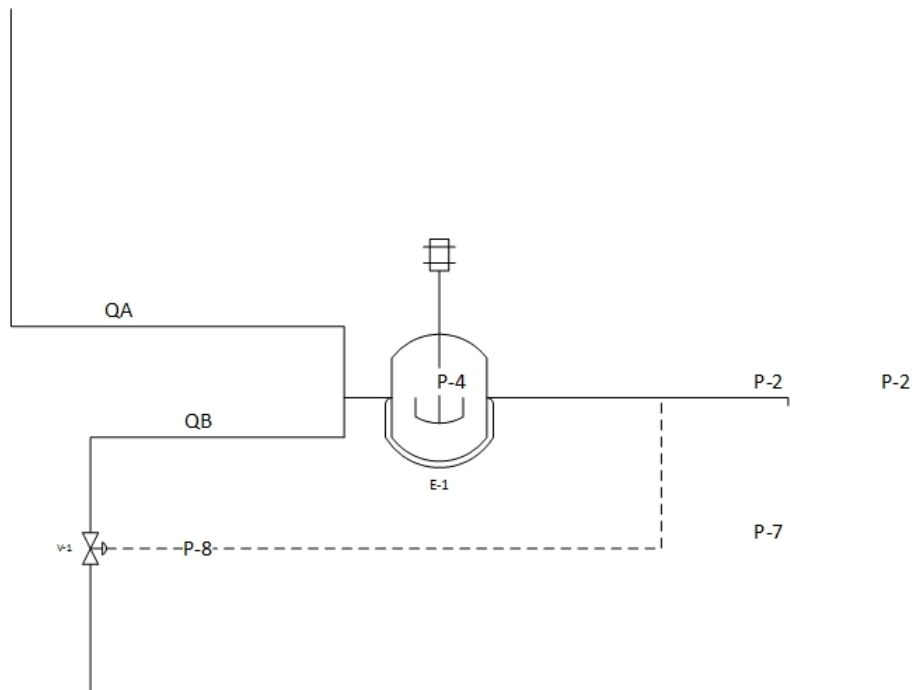
Σημειώνεται ότι το σφάλμα SSE είναι ίσο με 1.19.

7.5.2 Εφαρμογή 2-Αντιδραστήρας με παράλληλες Αντιδράσεις

Είναι συχνό φαινόμενο να πραγματοποιούνται περισσότερες από μια αντιδράσεις στο εσωτερικό ενός αντιδραστήρα. Το φαινόμενο αυτό είναι αρκετά ενδιαφέρον καθώς πολύ συχνά είναι επιθυμητό να ληφθεί το ενδιάμεσο προϊόν. Έτσι, θεωρώντας ότι το ενδιάμεσο προϊόν είναι παρατηρήσιμο, δηλαδή μπορεί να μετρηθεί με κάποιο χρόνο δειγματοληψίας, το σύστημα ρύθμισης καλείται να μεταβάλλει δυναμικά την ογκομετρική παροχή του ρεύματος εισόδου του συστατικού B ώστε να πετύχει την επιθυμητή τιμή της συγκέντρωσης του ενδιάμεσου συστατικού στο ρεύμα εξόδου. Αυτό φαίνεται και στο σχήμα 7.11, με τον αντιδραστήρα να είναι CSTR. Οι αντιδράσεις φαίνονται στην εξίσωση 7.25



με το ενδιάμεσο προϊόν να εμφανίζεται και στην αντίδραση παραγωγής του.



Σχήμα 7.11: Αντιδραστήρας παραγωγής ενδιάμεσου προϊόντος B

Στον αντιδραστήρα εισέρχονται με διαφορετικές ροές το ενδιάμεσο προϊόν (B) και το αντιδρόν (A), ενώ η μεταβλητή εκ χειρισμού είναι η ογκομετρική παροχή του B, με τις τιμές των παραμέτρων των ισοζυγίων να φαίνονται στη συνέχεια.

$$\begin{aligned} \frac{dc_A(t)}{dt} &= -k_1 \frac{c_A(t)c_B(t)}{1+c_B(t)} - \frac{(Q_B(t)+Q_A)c_A(t)}{V} + \frac{Q_A c_{A0}}{V} \\ \frac{dc_B(t)}{dt} &= k_1 \frac{c_A(t)c_B(t)}{1+c_B(t)} - k_2 c_B(t) - \frac{(Q_B(t)+Q_A)c_B(t)}{V} + \frac{Q_B(t)c_{B0}}{V} \end{aligned} \quad (7.26)$$

Τα παραπάνω ισοζύγια καταστρώθηκαν σε συνεχή χρόνο, και για αυτό πραγματοποιείται διακριτοποίηση με τη μέθοδο zero order hold με χρόνο δειγματοληψίας 0.1 sec. Οι σταθερές τιμές στις παραπάνω εξισώσεις είναι $k_1=0.5 \text{ sec}^{-1}$, $k_2=1 \text{ sec}^{-1}$, $Q_A=10 \text{ L/sec}$, $C_{A0}=5 \text{ mol/L}$, $C_{B0}=10 \text{ mol/L}$, $V = 1 \text{ L}$ ενώ αρχικά οι συγκεντρώσεις στον αντιδραστήρα είναι $C_A(0)=0 \text{ mol/L}$, $C_B(0)=5 \text{ mol/L}$. Σε αυτή την περίπτωση γίνονται συνεχόμενες βηματικές μεταβολές της επιθυμητής τιμής που συμβολίζεται με C_{Bd} , ενώ οι τιμές σχεδιασμού παρουσιάζονται στον πίνακα 7.4.

Σε αυτό το πρόβλημα επειδή πραγματοποιήθηκαν συνεχόμενες βηματικές επιβολές στις επιθυμητές τιμές, όταν αφαιρείται κάποιο κέντρο, γίνεται χρήση της μεθόδου gradient decent, ενώ όταν προστίθεται νευρώνας, αυτός αποκτά αρχικά βάρος σύνδεσης με τη στοιβάδα εξόδου, που αντιστοιχεί στο αμέσως προηγούμενο κέντρο που προστέθηκε στο δίκτυο.

Πίνακας 7.4: Τιμές Σχεδιασμού Περίπτωση 2.1

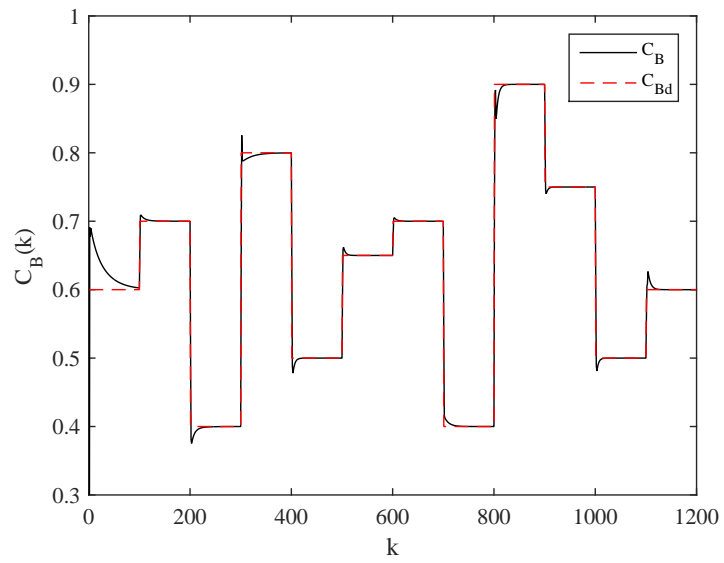
Τιμές Σχεδιασμού Περίπτωση 2.1	
λ_1	0.7
N_d	300
N_{max}	1200
s_α	[8 9]
s_c	[10]
a	1
l_c	0.1
l_α	0.1

Περίπτωση 2.1

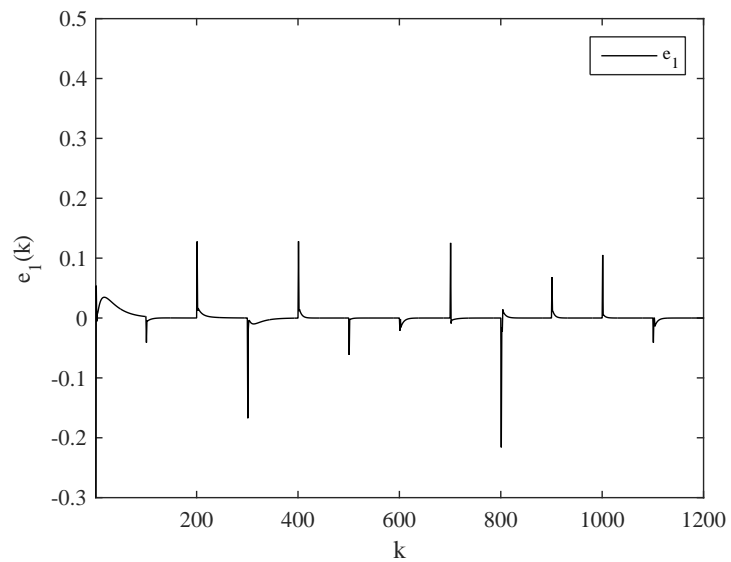
Επίσης μελετήθηκαν 2 περιπτώσεις σχετικά με κάποια πιθανή διαταραχή. Αρχικά θεωρήθηκε ότι υπάρχει άγνωστη διαταραχή που συμπεριφέρεται σύμφωνα με την εξίσωση 7.27. Με αυτό το τρόπο εκφράζεται η διαταραχή στη παροχή του αντιδρώντος A.

$$d(k) = \cos(0.001k) \quad (7.27)$$

Τα αποτελέσματα είναι πολύ ικανοποιητικά καθώς καταφέρνει να ακολουθεί την πορεία των επιθυμητών τιμών με πολύ μικρή υπέρβαση, όπως φαίνεται στο σχήμα 7.12, και αντίστοιχα το σφάλμα είναι πολύ μικρό και προσεγγίζει το μηδέν σε πολύ μικρό χρονικό διάστημα, σύμφωνα με το σχήμα 7.13, με το σφάλμα SSE να είναι 0.786.

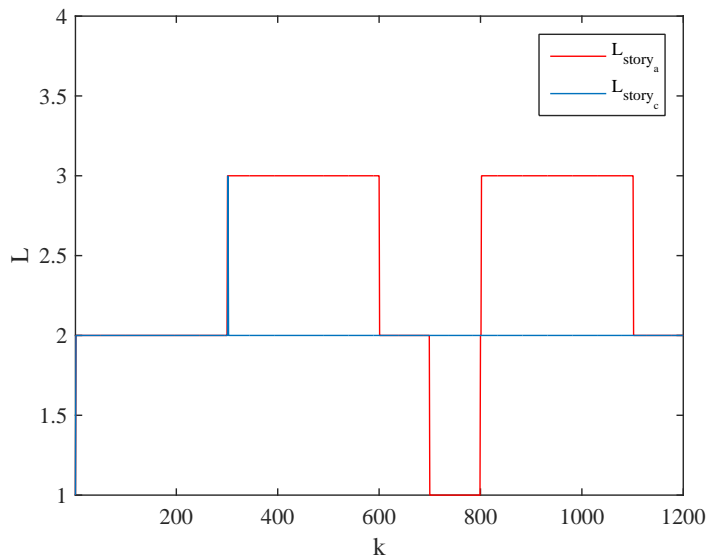


Σχήμα 7.12: Η πορεία της μεταβλητής εξόδου και της επιθυμητής τιμής σε συνάρτηση με το χρόνο, περίπτωση 2.1



Σχήμα 7.13: Η πορεία του σφάλματος σε συνάρτηση με το χρόνο, περίπτωση 2.1

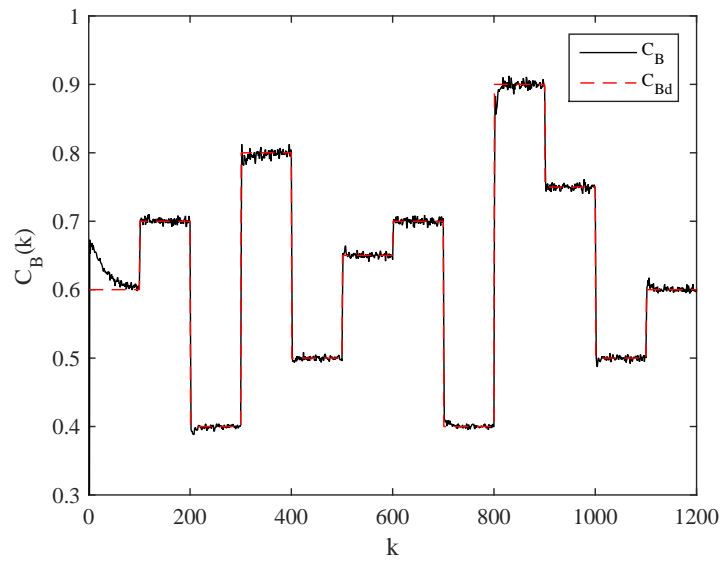
Είναι φανερό από το σχήμα 7.14 ότι δεν χρειάζονται περισσότερα από 3 κέντρα, και άρα το υπολογιστικό κόστος είναι πολύ χαμηλό.



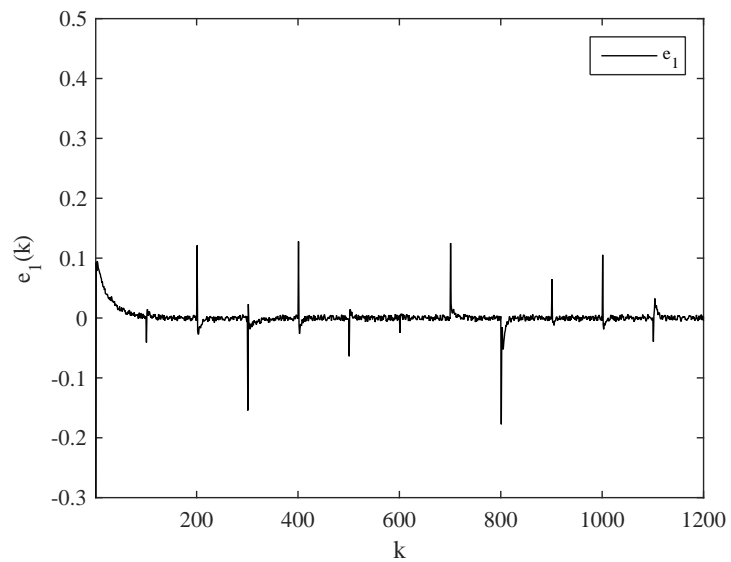
Σχήμα 7.14: Η εξέλιξη του αριθμού των κέντρων των δύο δικτύων, περίπτωση 2.1

Περίπτωση 2.2

Στην περίπτωση αυτή αντικαταστάθηκε η διαταραχή της προηγούμενης περίπτωσης από άγνωστες τυχαίες διαταραχές, τόσο στην παροχή του A αλλά και στη μέτρηση των συγκεντρώσεων. Οι διαταραχές επιλέγονται τυχαία από ομοιόμορφες κατανομές με όρια $[0, 1]$ στην περίπτωση της παροχής και $[0, 0.01]$ στην περίπτωση των μετρήσεων των συγκεντρώσεων. Όπως φαίνεται από το σχήμα 7.15 και πάλι η συγκέντρωση σταθεροποιείται αρκετά ικανοποιητικά δεδομένου των εντελώς τυχαίων διαταραχών.

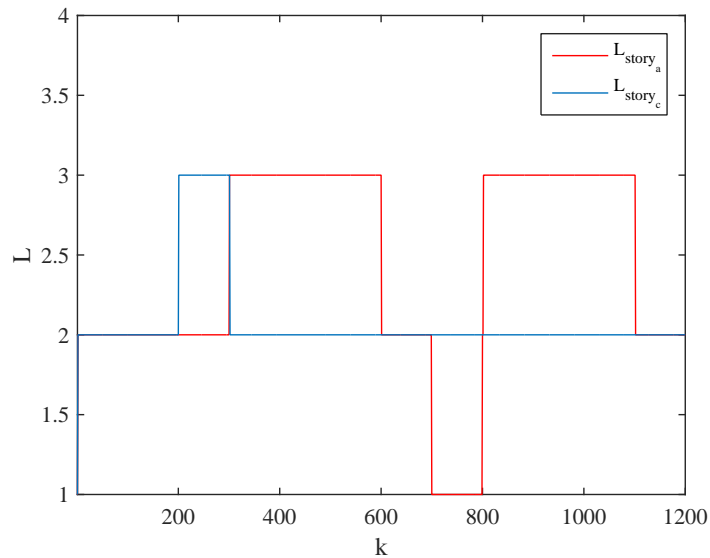


Σχήμα 7.15: Η πορεία της μεταβλητής εξόδου και της επιθυμητής τιμής σε συνάρτηση με το χρόνο, περίπτωση 2.2



Σχήμα 7.16: Η πορεία του σφάλματος σε συνάρτηση με το χρόνο, περίπτωση 2.2

Είναι φανερό από το σχήμα 7.14 ότι δεν χρειάζονται πάνω από 3 κέντρα, και άρα το υπολογιστικό κόστος είναι πολύ χαμηλό.



Σχήμα 7.17: Η εξέλιξη του αριθμού των κέντρων των δύο δικτύων, περίπτωση 2.2

Το σφάλμα σε αυτή τη περίπτωση είναι ελάχιστα πιο υψηλό (0.799) κάτι που είναι αναμενόμενο, καθώς πλέον υπάρχουν 2 εντελώς τυχαίες διαταραχές τόσο στο σύστημα όσο και στις μετρήσεις.

Κεφάλαιο 8

Τροποποίηση της συνάρτησης κόστους στην μεθοδολογία ενισχυτικής μάθησης

Στο προηγούμενο κεφάλαιο στη μεθοδολογία ρύθμισης, χρησιμοποιήθηκε σαν συνάρτηση κόστους η εξίσωση 7.10 και μάλιστα η $r(k)$ ορίσθηκε από την εξίσωση 7.11. Σε αυτή την ενότητα θα χρησιμοποιηθεί διαφορετική συνάρτηση κόστους. Επίσης, είναι φανερό ότι με τον πρώτο ορισμό 7.11 της, ότι οι αποκρίσεις χωρίζονται απλά σε ικανοποιητικές και μη ικανοποιητικές. Για αυτό το λόγο όπως θα φανεί στην επόμενη ενότητα πραγματοποιήθηκε αλλαγή στη συνάρτηση $r(k)$ με σκοπό να είναι συνεχείς, αναδεικνύοντας με καλύτερο τρόπο το πόσο καλή είναι η απόκριση του συστήματος. Ακόμα, στο προηγούμενο κεφάλαιο χρειάστηκε ο αλγόριθμος να αντιλαμβάνεται το είδος του των επιθυμητών τιμών και να αλλάζει έτσι τον τρόπο εκπαίδευσης, πλέον δεν χρειάζεται να συμβαίνει αυτό, με τον αλγόριθμο να αποδίδει με πολύ καλό τρόπο.

8.1 Τροποποίηση του δικτύου Κριτής

Η νέα συνάρτηση κόστους που εφαρμόζεται σε αυτό το κεφάλαιο παρουσιάζεται στην εξίσωση 8.1 [22],[2]

$$J(\mathbf{x}(k), u(k)) = \sum_{i=t_0}^{\infty} \gamma^i r(\mathbf{x}(k+i), u(k+i)) \quad (8.1)$$

Όπου $0 < \gamma \leq 1$ αποτελεί τον εκπτώτικο παράγοντα για προβλήματα απείρου ορίζοντα, ενώ για τη συνάρτηση r ισχύει η εξίσωση 8.2

$$r(\mathbf{x}(k), u(k)) = \mathbf{e}(k)^T Q \mathbf{e}(k) + (u(k) - u(k-1))^T R (u(k) - u(k-1)) \quad (8.2)$$

Στην εξίσωση 8.2, Q είναι ένας θετικά ημιορισμένος πίνακας και R ένας θετικά ορισμένος πίνακας. Από την εξίσωση 8.1 προκύπτει η εξίσωση *Bellman*: [22],[2]

$$J(k) = \gamma J(k-1) - r(\mathbf{x}(k), u(k)) \quad (8.3)$$

Έτσι, πλέον το σφάλμα μπορεί να ορισθεί ως εξής:

$$e_c(k) = \hat{J}(k) - (\gamma \hat{J}(k-1) + r(\mathbf{x}(k), u(k))) \quad (8.4)$$

Όπου $\hat{J}(k)$ η έξοδος του δικτύου κριτή. Σε αυτή την περίπτωση όπως και στο προηγούμενο κεφάλαιο ως συνάρτηση ακτινικής βάσης χρησιμοποιήθηκε η συνάρτηση *Gauss*, άρα το δίκτυο κριτής περιγράφεται από την εξίσωση 7.14. Στόχος της εκπαίδευσης του δικτύου είναι η ελαχιστοποίηση του σφάλματος:

$$E_c(k) = \frac{1}{2} e_c^T(k) e_c(k) \quad (8.5)$$

Η ελαχιστοποίηση γίνεται όπως πριν μέσω της μεθόδου *gradient decent* (7.16). Εξαιτίας όμως της αλλαγής που πραγματοποιήθηκε στη συνάρτηση κόστους προκύπτει το εξής αποτέλεσμα.

$$\hat{w}_c(k+1) = \hat{w}_c(k) - l_c \gamma g_c(\mathbf{x}(k)) [\gamma \hat{J}(k) + r(\mathbf{x}(k), u(k)) - \hat{J}(k-1)] \quad (8.6)$$

Επίσης όπως είναι αναμενόμενο πλέον προστέθηκε και η μεταβλητή u ως είσοδος στο νευρωνικό δίκτυο κριτής.

Τα συναπτικά βάρη, πλέον αναπροσαρμόζονται στα ιστορικά δεδομένα με τη μέθοδο της γραμμικής παλινδρόμησης, και αν πραγματοποιηθεί διαγραφή κάποιου νευρώνα της κρυφής στοιβάδας, εφαρμόζεται επιπλέον η μέθοδος gradient decent, όπως συνέβαινε στη περίπτωση όπου υπήρχε πρόβλημα παρακολούθησης τροχιάς. Ωστόσο πλέον δεν χρειάζεται να γίνεται διαχωρισμός ανάλογα με τη πορεία των επιθυμητών τιμών με τον αλγόριθμο να ανταποκρίνεται πολύ καλά σε κάθε περίπτωση.

8.2 Εφαρμογές

Ο αλγόριθμος αυτός εφαρμόστηκε στα ίδια παραδείγματα που παρουσιάστηκαν στο προηγούμενο κεφάλαιο με σκοπό να συγκριθούν τα αποτελέσματα, αλλά επίσης εφαρμόστηκε σε ένα επιπλέον παράδειγμα χημικού αντιδραστήρα όπου πραγματοποιείται εξώθερμη χημική αντίδραση.

8.2.1 Εφαρμογή 1

Η εφαρμογή του αλγορίθμου θεωρείται άκρως επιτυχής στις περιπτώσεις 1.2 και 1.3, το σφάλμα SSE εμφανίζεται μειωμένο και επίσης δεν χρειάζεται να γίνει καμία αλλαγή ανάλογα με τις επιθυμητές τιμές.

Περίπτωση 1.2

Οι σχεδιαστικές παράμετροι εμφανίζονται στον πίνακα 8.1.

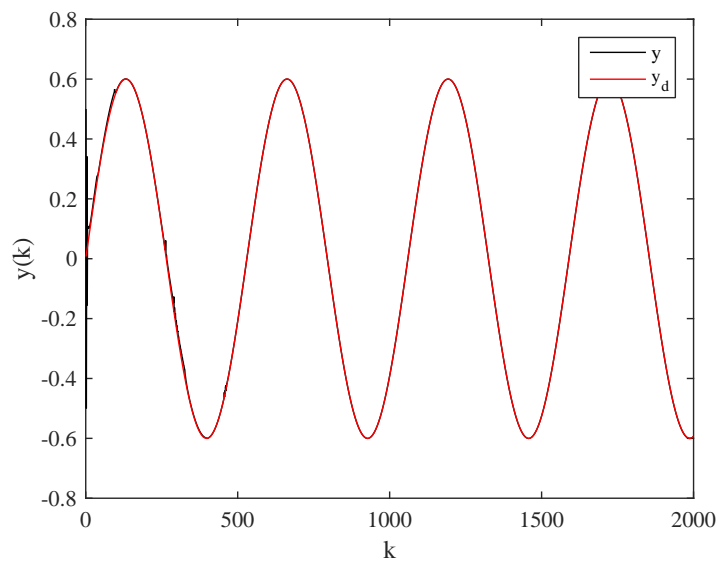
Ενώ για τη συνάρτηση $r(k)$ ισχύει η εξίσωση 8.7

$$r(k) = e_1(k)^2 Q + (u(k) - u(k-1))^2 R \quad (8.7)$$

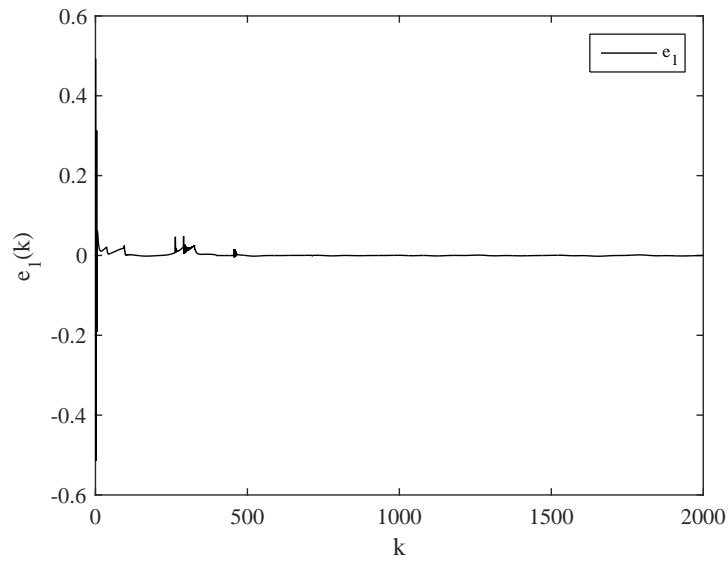
Τα αποτελέσματα της προσομοίωσης παρουσιάζονται στα σχήματα 8.1-8.3

Πίνακας 8.1: Τιμές Σχεδιασμού Περίπτωση 1.2

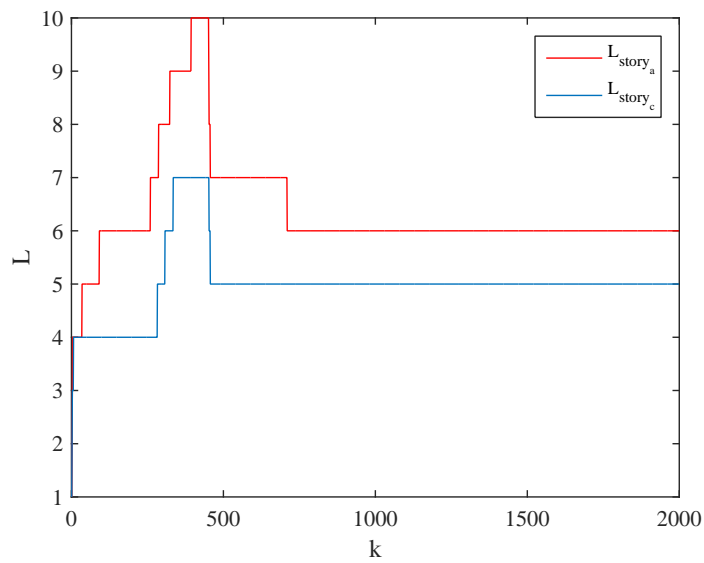
Τιμές Σχεδιασμού Περίπτωση 1.2	
λ_1	0.9
λ_2	0.25
N_d	450
N_d	300
N_{max}	2000
s_α	[10 10 9]
s_c	[12 9]
a	1
l_c	0.1
l_α	0.1
γ	0.01
Q	1
R	0.1



Σχήμα 8.1: Η πορεία της μεταβλητής εξόδου και της επιθυμητής τιμής σε συνάρτηση με το χρόνο, περίπτωση 1.2.



Σχήμα 8.2: Η πορεία του σφάλματος σε συνάρτηση με το χρόνο, περίπτωση 1.2.



Σχήμα 8.3: Η εξέλιξη του αριθμού των κέντρων των δύο δικτύων, περίπτωση 1.2.

Το σφάλμα SSE πλέον υπολογίζεται 0.837 που είναι μικρότερο σε σχέση με αυτό

Κεφάλαιο 8. Τροποποίηση της συνάρτησης κόστους στην μεθοδολογία ενισχυτικής μάθησης

που υπολογίστηκε στο προηγούμενο κεφάλαιο (0.883), ενώ τα κέντρα που χρησιμοποιούνται είναι σχεδόν ίδια, και μάλιστα σε αυτή τη περίπτωση είναι λιγότερα.

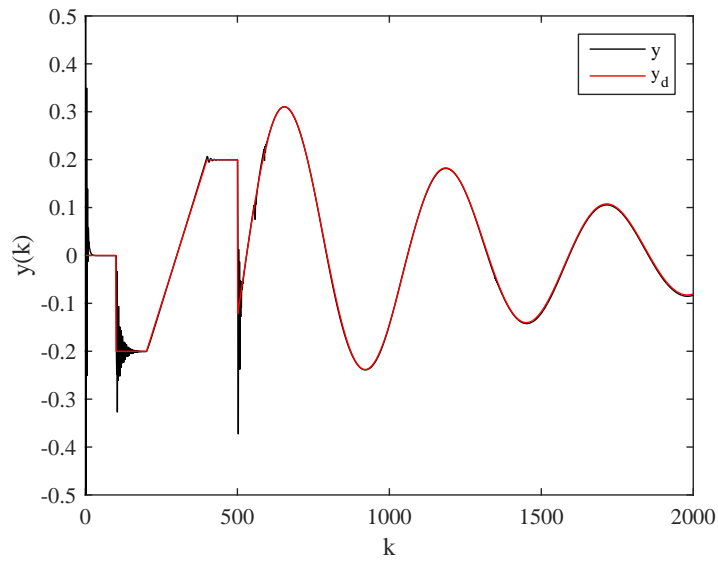
Περίπτωση 1.3

Στο προηγούμενο κεφάλαιο χρειάστηκε ο αλγόριθμος να αντιλαμβάνεται το είδος των επιθυμητών τιμών και να αλλάζει έτσι τον τρόπο εκπαίδευσης, σε αυτή την περίπτωση δεν χρειάζεται να συμβαίνει αυτό, με τον αλγόριθμο να αποδίδει με πολύ καλό τρόπο. Οι παράμετροι φαίνονται στον πίνακα 8.2.

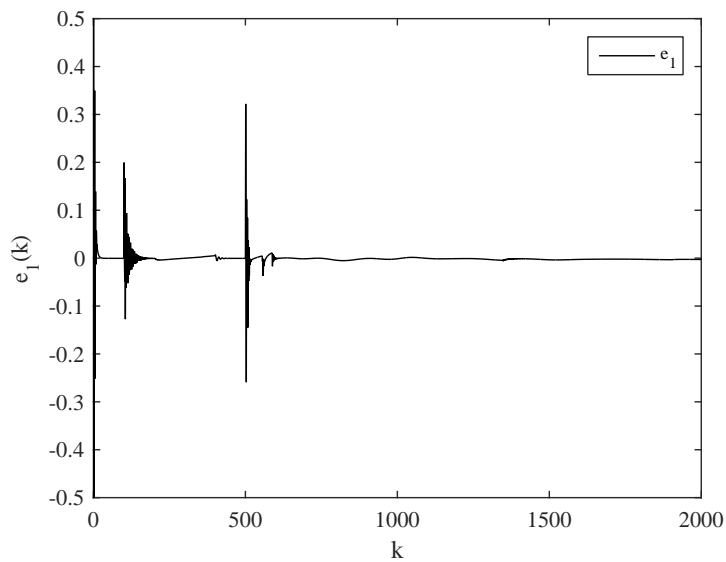
Πίνακας 8.2: Τιμές Σχεδιασμού Περίπτωση 1.3

Τιμές Σχεδιασμού Περίπτωση 1.3	
λ_1	0.9
λ_2	0.25
$N_{d\alpha}$	550
$N_{s\alpha}$	300
N_{dc}	350
N_{sc}	400
N_{max}	2000
s_α	[8 9 8]
s_c	[11 8]
a	1
l_c	0.1
l_α	0.1
γ	0.01
Q	1
R	2

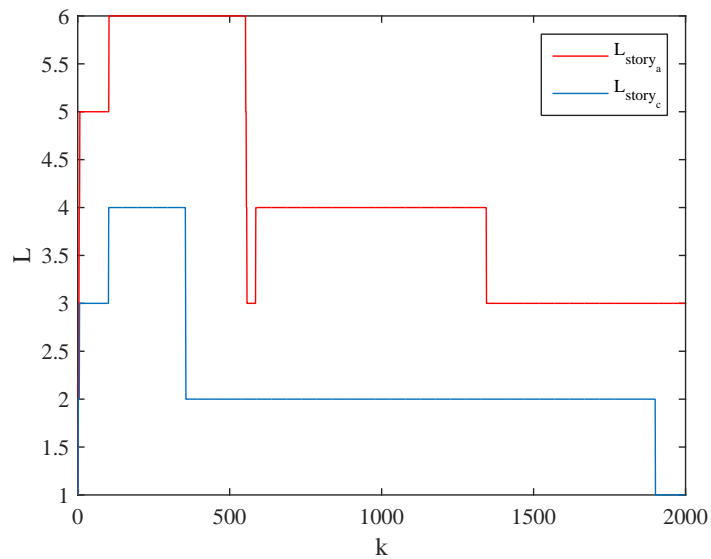
Τα αποτελέσματα της διαδικασίας ρύθμισης εμφανίζονται στα σχήματα 8.4-8.6



Σχήμα 8.4: Η πορεία της μεταβλητής εξόδου και της επιθυμητής τιμής σε συνάρτηση με το χρόνο, περίπτωση 1.3.



Σχήμα 8.5: Η πορεία του σφάλματος σε συνάρτηση με το χρόνο, περίπτωση 1.3.



Σχήμα 8.6: Η εξέλιξη του αριθμού των κέντρων των δύο δικτύων, περίπτωση 1.3.

Τα αποτελέσματα είναι αρκετά ικανοποιητικά έχοντας και πάλι μικρότερο σφάλμα σε σχέση με το προηγούμενο κεφάλαιο καθώς είναι πλέον 1.05 από 1.19. Σημειώνεται ότι σε αυτό το παράδειγμα χρησιμοποιήθηκαν διαφορετικές τιμές για τις παραμέτρους N_s και N_d για τα δύο νευρωνικά δίκτυα.

8.2.2 Εφαρμογή 2-Αντιδραστήρας με παράλληλες Αντιδράσεις

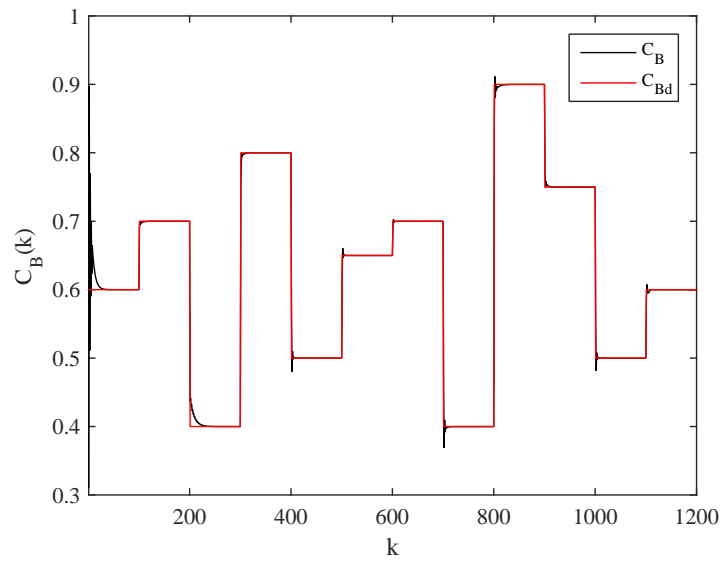
Εφαρμόζοντας την εν λόγω μεθοδολογία στον αντιδραστήρα με τις παράλληλες αντιδράσεις, το αποτέλεσμα είναι και εδώ βελτιωμένο για κάθε μια από τις δύο περιπτώσεις διαταραχών. Στον πίνακα 8.3 παρουσιάζονται οι παράμετροι σχεδιασμού της δεύτερης προτεινόμενης μεθοδολογίας ρύθμισης ενώ τα ισοζύγια παραμένουν ίδια με αυτά του προηγούμενου κεφαλαίου.

Πίνακας 8.3: Τιμές Σχεδιασμού Περίπτωση 2.1,2.2

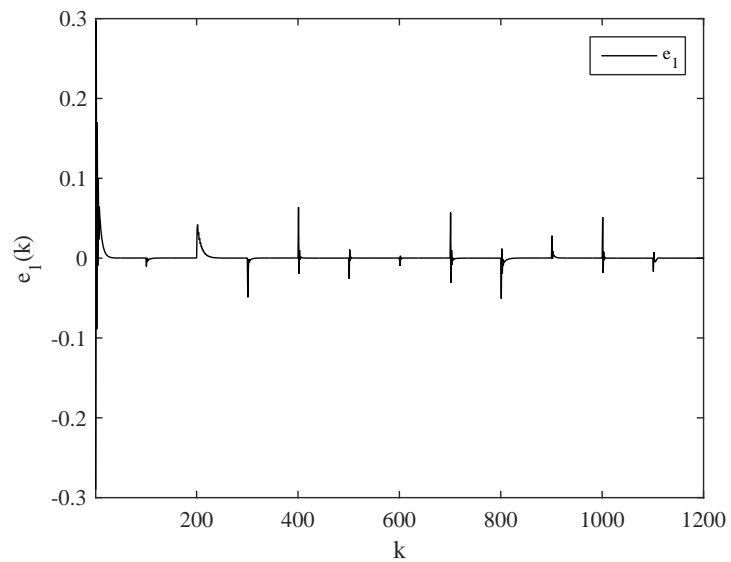
Τιμές Σχεδιασμού Περίπτωση 2.1,2.2	
λ_1	0.9
λ_2	0.25
N_d	700
N_d	300
N_{max}	1200
s_α	[10 9]
s_c	[10 11]
a	1
l_c	0.1
l_α	0.1
γ	0.01
Q	1
R	0.1

Περίπτωση 2.1

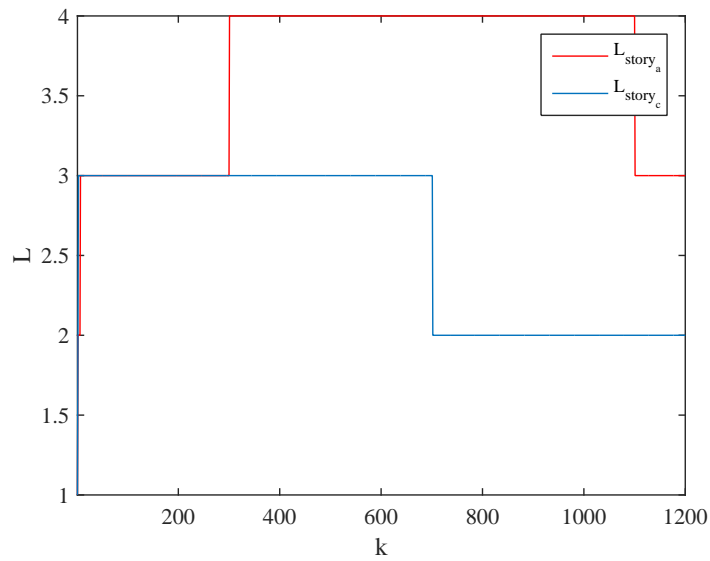
Η εφαρμογή της μεθοδολογίας παράγει τα αποτελέσματα που εμφανίζονται στα σχήματα 8.7-8.9. Το σφάλμα SSE είναι μικρότερο (0.511) σε σχέση με το προηγούμενο κεφάλαιο (0.786) και αναδεικνύεται η υπεροχή της εναλλακτικής μεθοδολογίας.



Σχήμα 8.7: Η πορεία της μεταβλητής εξόδου και της επιθυμητής τιμής σε συνάρτηση με το χρόνο, διαφορετικές επιθυμητές τιμές.



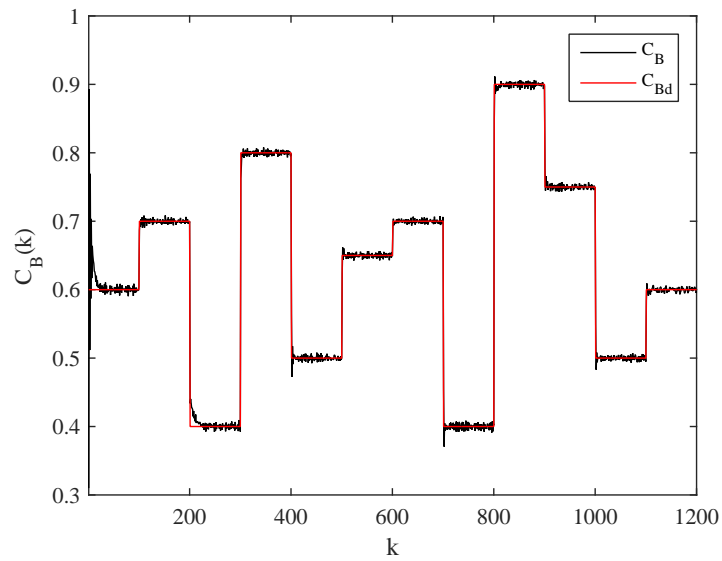
Σχήμα 8.8: Η πορεία του σφάλματος σε συνάρτηση με το χρόνο, περίπτωση 2.1.



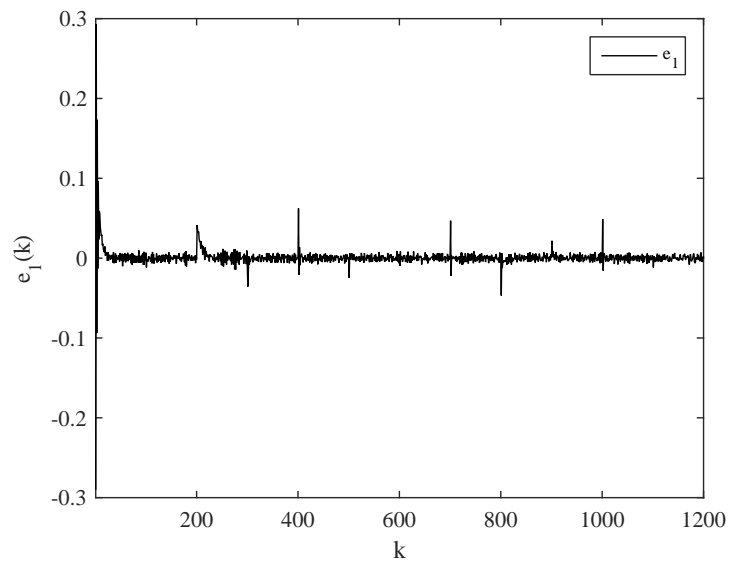
Σχήμα 8.9: Η εξέλιξη του αριθμού των κέντρων των δύο δικτύων, περίπτωση 2.1.

Περίπτωση 2.2

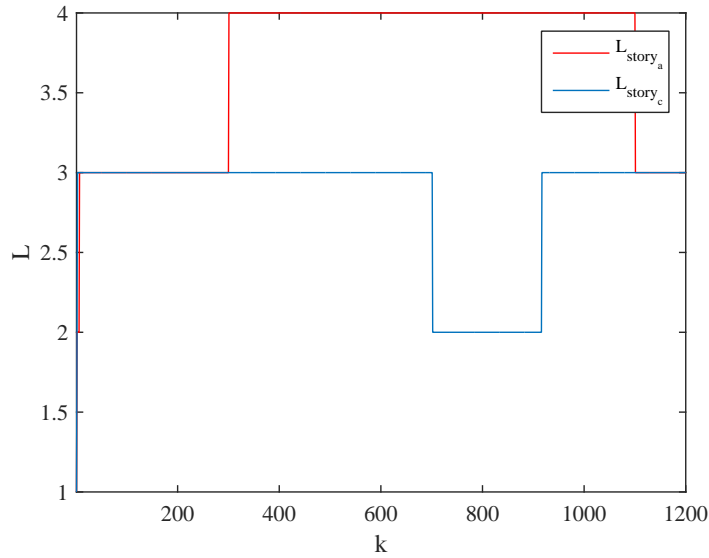
Στην περίπτωση αυτή αντικαταστάθηκε η διαταραχή της προηγούμενης περίπτωσης από άγνωστες τυχαίες διαταραχές με τα χαρακτηριστικά που περιγράφηκαν στην ενότητα 7.5. Και σε αυτήν την περίπτωση υπάρχει σαφής βελτίωση του σφάλματος SSE από 0.799 σε 0.521. Τα αποτελέσματα εμφανίζονται στα σχήματα 8.10-8.12



Σχήμα 8.10: Η πορεία της μεταβλητής εξόδου και της επιθυμητής τιμής σε συνάρτηση με το χρόνο, περίπτωση 2.2.



Σχήμα 8.11: Η πορεία του σφάλματος σε συνάρτηση με το χρόνο, περίπτωση 2.2.



Σχήμα 8.12: Η εξέλιξη του αριθμού των κέντρων των δύο δικτύων, περίπτωση 2.2.

Το σφάλμα είναι και πάλι μικρότερο (0.521) όπως αναμενόταν σε σχέση με το προηγούμενο (0.799) και άρα κρίνεται πιο επιτυχής αυτή η μέθοδος.

8.2.3 Εφαρμογή 3 - Χημικός αντιδραστήρας

Η νέα μεθοδολογία ρύθμισης εφαρμόστηκε και σε ένα τρίτο παράδειγμα, στο οποίο το ζητούμενο είναι η ρύθμιση της συγκέντρωσης του ρεύματος εξόδου ως προς ένα συστατικό χρησιμοποιώντας ως μεταβλητή εκ χειρισμού τη θερμοκρασία της ροής θερμότητας. Η κανονικοποιημένη αδιάστατη μορφή των ισοζυγίων μάζας και ενέργειας παρουσιάζονται στις εξισώσεις 8.8 και 8.9 [27].

$$dx_1/dt = -\alpha x_1 + D_\alpha (1 - x_1)^2 e^{\frac{x_2}{1 + x_2/\gamma}} \quad (8.8)$$

$$dx_2/dt = -\alpha x_2 + BD_\alpha (1 - x_1)^2 e^{\frac{x_2}{1 + x_2/\gamma}} - \beta(x_2 - u) \quad (8.9)$$

Κεφάλαιο 8. Τροποποίηση της συνάρτησης κόστους στην μεθοδολογία ενισχυτικής μάθησης

$$\text{Όπου } x_1 = \frac{C_{Af} - C_A}{C_{Af}}, x_2 = \frac{\gamma(T - T_{f0})}{T_{f0}}, u = \frac{\gamma(T_u - T_{f0})}{T_{f0}}, B = -\frac{\gamma\Delta HC_{Af}}{C_p T_{f0}}, D_\alpha = k_0 e^{-\gamma},$$

$$\alpha = \frac{V}{Q_f}, \beta = \frac{UA}{VC_p}, \gamma = \frac{E_\alpha}{RT_{f0}}.$$

Επίσης η μεταβλητή εξόδου, που επιθυμείται να ρυθμιστεί σε συγκεκριμένη τιμή, μετά την κανονικοποίηση είναι η μετατροπή $y = x_1$.

Στον πίνακα 8.4 εμφανίζονται οι μεταβλητές στις οποίες αντιστοιχούν τα παραπάνω σύμβολα.

Πίνακας 8.4: Παράμετροι αντιδραστήρα

Σύμβολο	Ερμηνεία
A	Εμβαδόν
C_{Af}	Συγκέντρωση εισόδου
C_p	Θερμοχωρητικότητα
E_α	Ενέργεια ενεργοποίησης
ΔH	Θερμότητα αντίδρασης
k_0	Σταθερά αντίδρασης
Q_f	Ογκομετρική παροχή εισόδου
T_{f0}	Θερμοκρασία τροφοδοσίας
U	Ολικός συντελεστής μεταφοράς θερμότητας
V	όγκος αντιδραστήρα
T	θερμοκρασία αντιδραστήρα πλήρους ανάμειξης
T_u	θερμοκρασία utility stream

Οι τιμές για τις παραμέτρους του συστήματος δίνονται στον πίνακα 8.5.

Στόχος της εφαρμογής της μεθοδολογίας είναι χωρίς καμία γνώση για το σύστημα να μπορέσει η συγκέντρωση να ακολουθήσει τις συνεχόμενες βηματικές επιβολές στις επιθυμητές τιμές, θεωρώντας ότι υπάρχουν άγνωστες τυχαίες διαταραχές τόσο στο ίδιο το σύστημα όσο και στη μέτρηση των μεταβλητών κατάστασης. Οι διαταραχές αυτές επιλέγονται τυχαία από ομοιόμορφες κατανομές με όρια [0 0.01] στην περίπτωση του συστήματος και [0 0.001] στην περίπτωση των μετρήσεων των συγκεντρώσεων. Ο χρόνος δειγματοληψίας είναι 0.5 sec. Οι τιμές των παραμέτρων

Πίνακας 8.5: Τιμές παραμέτρων δυναμικής συμπεριφοράς αντιδραστήρα

Παράμετροι	Τιμή
a	1.0
β	0.3
γ	20.0
B	1.0
D_α	0.072

σχεδιασμού για τον ρυθμιστή παρουσιάζονται στον πίνακα 8.6

Πίνακας 8.6: Τιμές Σχεδιασμού Περίπτωση Χημικού αντιδραστήρα με εξώθερμη αντίδραση

Τιμές Σχεδιασμού	
λ_1	0.9
N_d	700
N_d	300
N_{max}	1500
s_α	[9 8]
s_c	[12 11]
a	1
l_c	0.1
l_α	0.1
γ	0.01
Q	1
R	2

Μελετήθηκαν και εδώ 2 περιπτώσεις αυτού του προβλήματος, στην πρώτη όλες οι παράμετροι του συστήματος παρέμειναν σταθερές, ενώ στη δεύτερη θεωρήθηκε ότι η σταθερά β μειώνεται από τη στιγμή $k=500$ μέχρι να φτάσει στο ήμισυ της αρχικής τιμής.

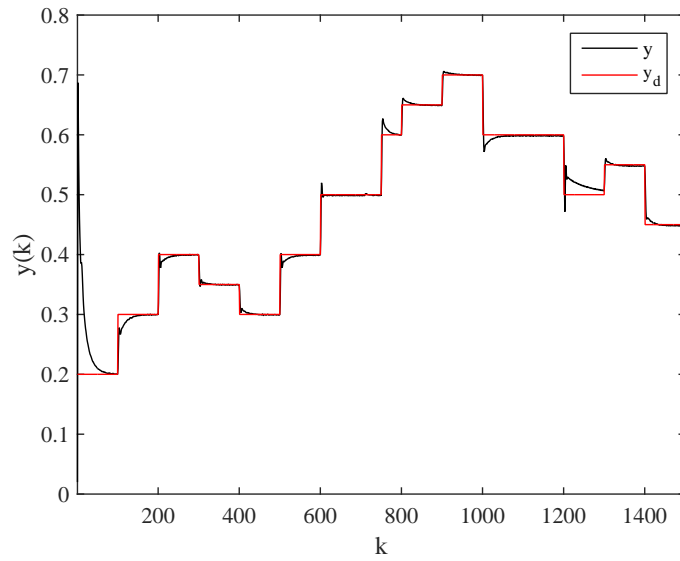
Περίπτωση 3.1

Μετά την εφαρμογή του αλγορίθμου στο σύστημα τα αποτελέσματα φαίνονται στα σχήματα 8.13-8.15 και είναι φανερό ότι ανταποκρίνεται πολύ ικανοποιητικά στις μεταβολές των επιθυμητών τιμών με το συνολικό σφάλμα να προκύπτει 1.12.

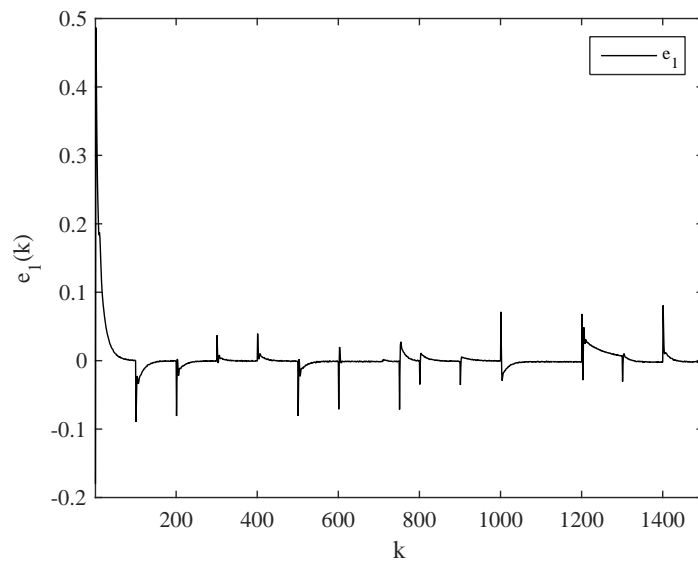
Περίπτωση 3.2

Είναι σύνηθες κατά τη διάρκεια της λειτουργίας τέτοιων αντιδραστήρων να εμφανίζονται σταδιακές μειώσεις στον συντελεστή μεταφοράς θερμότητας, λόγω κυρίων επικαλύψεων στον εναλλάκτη. Ο στόχος είναι το σύστημα να ανταποκρίνεται το ίδιο ικανοποιητικά με πριν.

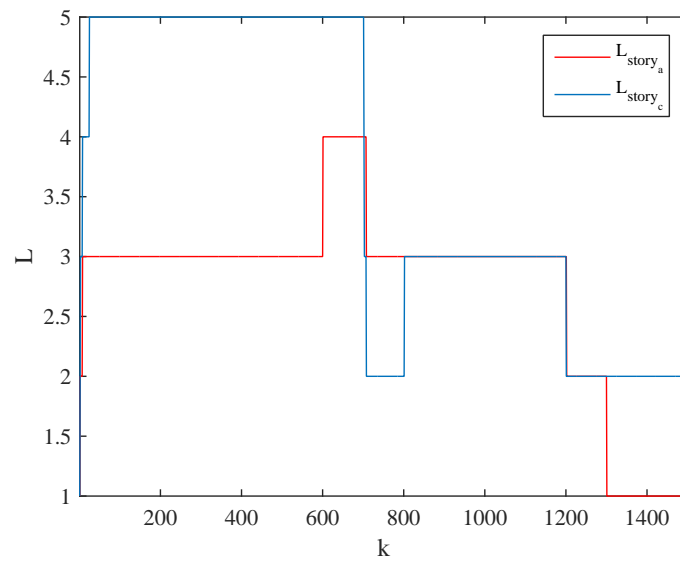
Τα αποτελέσματα εμφανίζονται στα σχήματα 8.13-8.15 ενώ το σφάλμα SSE αυξήθηκε οριακά. Είναι φανερό ότι το προτεινόμενο σύστημα ρύθμισης μπορεί να δράσει ως προσαρμοστικός ρυθμιστής (adaptive controller) αφού ανταποκρίνεται με εξαιρετικό τρόπο και σε αλλαγές της δυναμικής του συστήματος



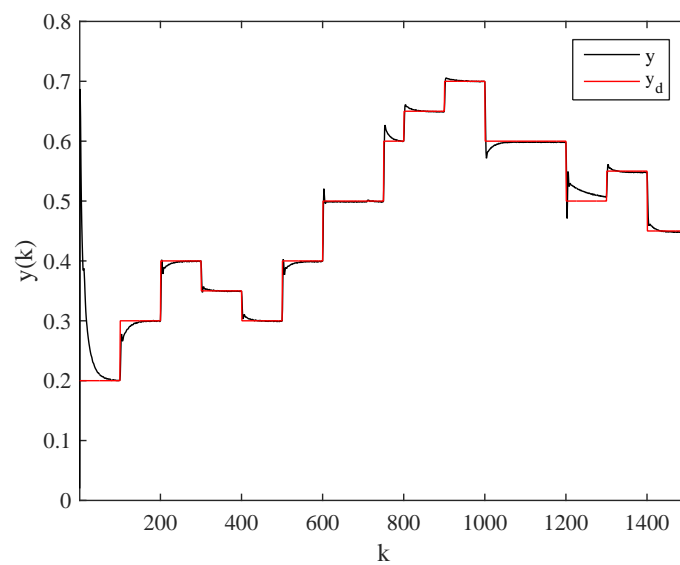
Σχήμα 8.13: Η πορεία της μεταβλητής εξόδου και της επιθυμητής τιμής σε συνάρτηση με το χρόνο, περίπτωση 3.1



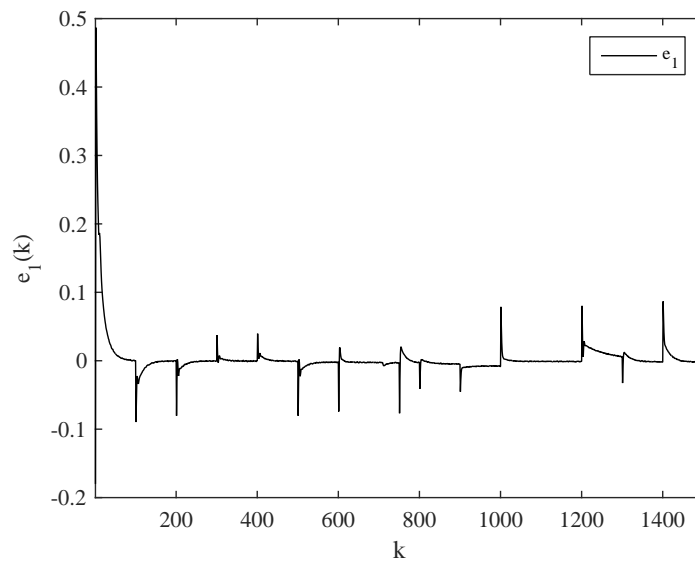
Σχήμα 8.14: Η πορεία του σφάλματος σε συνάρτηση με το χρόνο, περίπτωση 3.1



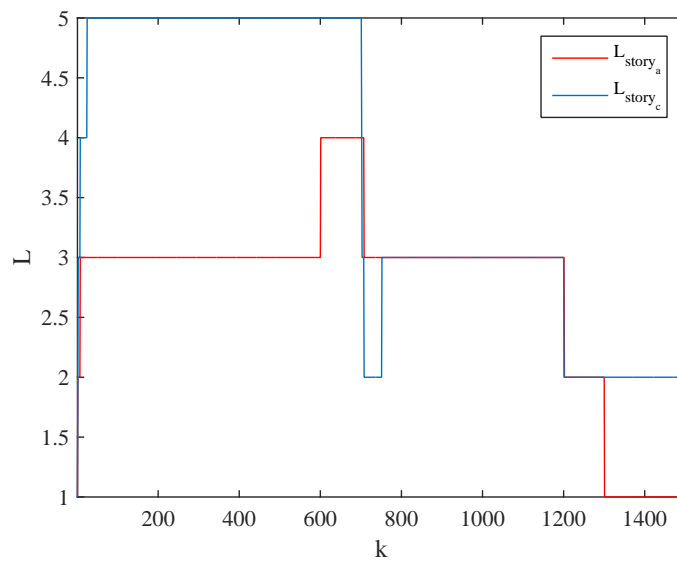
Σχήμα 8.15: Η εξέλιξη του αριθμού των κέντρων των δύο δικτύων, περίπτωση 3.1



Σχήμα 8.16: Η πορεία της μεταβλητής εξόδου και της επιθυμητής τιμής σε συνάρτηση με το χρόνο, περίπτωση 3.2



Σχήμα 8.17: Η πορεία του σφάλματος σε συνάρτηση με το χρόνο, περίπτωση 3.2



Σχήμα 8.18: Η εξέλιξη του αριθμού των κέντρων των δύο δικτύων, περίπτωση 3.2

Κεφάλαιο 9

Συμπεράσματα

Στην παρούσα εργασία παρουσιάστηκαν 2 καινοτόμες μεθοδολογίες για τη ρύθμιση συστημάτων που συνδυάζουν τα τεχνητά νευρωνικά δίκτυα με την αρχή της ενισχυτικής μάθησης. Το σημαντικό πλεονέκτημα των μεθοδολογιών είναι ότι επιτρέπουν την κατασκευή των δικτύων από μηδενική βάση και την πλήρη αναπροσαρμογή τους με βάση τις πληροφορίες που συλλέγονται κατά τη διάρκεια λειτουργίας του συστήματος σε αντίθεση με άλλες μεθοδολογίες που έχουν προταθεί στη βιβλιογραφία, οι οποίες αναπροσαρμόζουν μόνο τα βάρη που συνδέουν την κρυφή στοιβάδα με τη στοιβάδα εξόδου. Η πρώτη μέθοδος παρουσίασε ένα μειονέκτημα, καθώς έπρεπε σε κάθε περίπτωση να αναγνωρίζει ποιός είναι ο τρόπος που αλλάζουν οι επιθυμητές τιμές και αντίστοιχα να εφαρμόζει διαφορετική τεχνική εκπαίδευσης των συναπτικών βαρών. Αυτό το πρόβλημα λύθηκε με την ανάπτυξη της δεύτερης μεθοδολογίας όπου το πεδίο τιμών της συνάρτησης χρησιμότητας r είναι συνεχές, σε αντίθεση με τη πρώτη μέθοδο στην οποία το πεδίο τιμών αποτελείται μόνο από τις τιμές 0 και 1. Ο δεύτερος αλγόριθμος σύμφωνα και με το άθροισμα των τετραγώνων των σφαλμάτων βρέθηκε να έχει καλύτερη συμπεριφορά σε σχέση με τη πρώτη μεθοδολογία σε μια σειρά από εφαρμογές στις οποίες δοκιμάστηκαν και οι δύο αλγόριθμοι.

Στον πίνακα 9.1 παρουσιάζονται συναπτικά τα αποτελέσματα κάθε περίπτωσης που μελετήθηκε και είναι εμφανές ότι στη τροποποιημένη μεθοδολογία τα σφάλματα SSE έχουν μειωθεί.

Πίνακας 9.1: Συνοπτικός πίνακας αποτελεσμάτων (SSE)

	Πρώτη Μεθοδολογία Ρύθμισης	Τροποποιημένη Μεθοδολογία
Εφαρμογή 1, Περίπτωση 1.1	0.897	-
Εφαρμογή 1, Περίπτωση 1.2	0.883	0.837
Εφαρμογή 1, Περίπτωση 1.3	1.19	1.05
Εφαρμογή 2, Περίπτωση 2.1	0.786	0.511
Εφαρμογή 2, Περίπτωση 2.2	0.799	0.521
Εφαρμογή 3, Περίπτωση 3.1	-	1.12
Εφαρμογή 3, Περίπτωση 3.2	-	1.13

Βιβλιογραφία

- [1] S. Haykin. *Neural Networks and Machine Learning*, chapter 1,2,5,7. Prince Hall, 1 edition, 2009.
- [2] Q. Yang and S. Jagannathan. Reinforcement learning controller design for affine nonlinear discrete-time systems using online approximators. *IEEE Trans. Syst. Man Cybern. Part B Cybern.*, 42:377–390, 2012.
- [3] H. Sarimveis, A. Alexandridis, G. Tsekouras, and G. Bafas. 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*, 41:751–759, 2002.
- [4] A. Alexandridis, H. Sarimveis, and G. Bafas. new algorithm for online structure and parameter adaptation of rbf networks. *Neural Networks*, 16:1003–1017, 2003.
- [5] A. Alexandridis, H. Sarimveis, and K. Ninos. A radial basis function network training algorithm using a non-symmetric partition of the input space – application to a model predictive control configuration. *Advances in Engineering Software*, pages 830–837, 2011.
- [6] F. L. Lewis and S. Jagannathan. *Yesildirek, Neural Network Control of Robot Manipulators and Nonlinear Systems*, chapter 1,3. Εκδόσεις Taylor and Francis, London, 3 edition, 2003.
- [7] neural networks history-2. <http://cs.stanford.edu/people/eroberts/>

- courses/soco/projects/neural-networks/History/history2.html.
Accessed: 29-04-2015.
- [8] neural networks history-1. <http://cs.stanford.edu/people/eroberts/courses/soco/projects/neuralnetworks/History/history1.html>.
Accessed: 29-04-2015.
- [9] Διαμαντάρας. *Τεχνικά Νευρωνικά Δίκτυα*, chapter 1,3. Κλειδάριθμος, 3 edition, 2007.
- [10] T. J. Anastasio. *Tutorial on neural systems modeling*, chapter 1. University of Illinois, 2003.
- [11] I. Βλαχαβάς and Π. Κεφάλας. *Τεχνητή Νοημοσύνη*, chapter 1,2. Γκιούρδας, 3 edition, 2000.
- [12] Π. Αργυράκης. *Νευρωνικά Δίκτυα και εφαρμογές*, chapter 1. Ελληνικό Ανοικτό Πανεπιστήμιο, 1 edition, 2001.
- [13] Γ. Ρίζος. *Τεχνικά Νευρωνικά Δίκτυα*, chapter 1,4. Νέων Τεχνολογιών, 1 edition, 1996.
- [14] Α. Κυρίου. Ανάπτυξη νευρωνικού ρυθμιστή για τον αυτόματο έλεγχο ασταθών μη γραμμικών δυναμικών συστημάτων. diploma thesis, Σχολή Χημικών Μηχανικών Ε.Μ.Π., Αθήνα, 2012.
- [15] C. E. Brodely and M. A. Friedl. Identifying and eliminating mislabeled training instances. *Artificial Intelligence Research*, 11:131–167, 1999.
- [16] M. R. Smith and T. Martinez. Improving classification accuracy by identifying and removing instances that could be misclassified. In *Proceedings of the International Joint Conference on Neural Networks*, pages 2690–2697, Iizuka, Japan, August 1-7 2011.
- [17] R. Sutton and A. Barto. *Reinforcement Learning: An Introduction*, chapter 1,2. MIT Press, 2 edition, 2012.

-
- [18] D. Bertsekas. *Dynamic Programming and Optimal Control, vol:I*, chapter 1,2. Athena Scientific, 3 edition, 2005.
- [19] T. Hastie, R. Tibshirani, and J. Friedman. *The elements of Statistical Learning Data Mining, Inference and Prediction*, chapter 1,3. Springer, 2 edition, 2008.
- [20] <http://www.alyuda.com/products/forecaster/neural-network-applications.htm>. Accessed: 29-04-2015.
- [21] X. Cao. *Stochastic Learning and Optimization*, chapter 1. Springer-Verlag, Berlin, 2009.
- [22] F. L. Lewis and D. Vrabie. Vrabiereinforcement learning and adaptive dynamic programming for feedback control. *IEEE circuits and systems magazine*, 2:1531–1536, 2004.
- [23] D. Bertsekas and J. Tsitsiklis. *Neuro-Dynamic Programming*, chapter 1,2. Athena Scientific, 1996.
- [24] Π. Γκίλιας. Ανάπτυξη Αλγορίθμων Δομικής Μάθησης και Εκτίμηση Εμπιστοσύνης Εξόδου σε Νευρωνικά Δίκτυα Ακτινικής Βάσης. diploma thesis, Ε.Μ.Π., Αθήνα, 2003.
- [25] C. Darken and J. Moody. Fast adaptive k-means clustering: some empirical results. In *IEEE INNS international joint conference on neural networks; proceedings*, page 233, Iizuka, Japan, August 1-7 1990.
- [26] X. Yang, D. Liu, D. Wang, and Q. Wei. Discrete-time online learning control for a class of unknown nonaffine nonlinear systems using reinforcement learning. *Neural Networks*, 55:30–41, 2014.
- [27] A. Uppal, W. H. Ray, and A. B. Poore. On the dynamic behavior of continuous stirred tank reactors. *Chemical Engineering Science*, 29:967–985, 1974.