

ΠΟΛΥΤΕΧΝΕΙΟ ΚΡΗΤΗΣ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΑΡΑΓΩΓΗΣ & ΔΙΟΙΚΗΣΗΣ
ΤΟΜΕΑΣ ΕΠΙΧΕΙΡΗΣΙΑΚΗΣ ΕΡΕΥΝΑΣ

**Πρόβλεψη επιτοκίων κεντρικής τράπεζας με
χρήση Νευρωνικών δικτύων και γενετικών
αλγορίθμων**

Διατριβή που υπεβλήθη για την μερική ικανοποίηση των
απαιτήσεων για την απόκτηση Μεταπτυχιακού Διπλώματος
Ειδίκευσης υπό,

Παπιλοπούλου Ιωάννα Α.Μ. 2007019044

Χανιά, Νοέμβριος 2009



Η Μεταπτυχιακή διατριβή της Παπιλοπούλου Ιωάννας εγκρίνεται από τους :

κ. Ζοπουνίδη Κωνσταντίνο - Καθηγητής

κ. Δούμπο Μιχάλη - Επίκουρος Καθηγητής

κ. Γρηγορούδη Ευάγγελο - Επίκουρος Καθηγητής

Επιβλέπων καθηγητής:

Κ. Ατσαλάκης Γεώργιος

Πρόλογος

Η παρούσα διπλωματική εργασία εκπονήθηκε κατά το ακαδημαϊκό έτος 2008-2009 στο Τμήμα Μηχανικών Παραγωγής και Διοίκησης του Πολυτεχνείου Κρήτης, υπό την επίβλεψη του Καθηγητή κ. Ατσαλάκη Γεώργιου. Το αντικείμενο που πραγματεύεται είναι η πρόβλεψη επιτοκίων κεντρικής τράπεζας με χρήση Νευρωνικών δικτύων και γενετικών αλγορίθμων.

Στο σημείο αυτό, εκτός από τη γενική παρουσίαση της μελέτης που ακολουθεί, θα ήθελα να ευχαριστήσω όλους όσους με στήριξαν κατά τη διάρκεια εκπόνησης της εργασίας για τη γόνιμη συνεργασία και τις παρατηρήσεις τους. Θα ήθελα επίσης να ευχαριστήσω ιδιαίτερα τον επιβλέποντα Καθηγητή μου κ. Ατσαλάκη Γεώργιο για την εμπιστοσύνη που μου έδειξε, την ευγενική του καθοδήγηση που συνέβαλε σημαντικά στο ξεπέρασμα των συναντώμενων δυσκολιών και τη δυνατότητα που μου έδωσε να ασχοληθώ με το συγκεκριμένο, εξαιρετικά ενδιαφέρον θέμα.

ΠΕΡΙΕΧΟΜΕΝΑ

ΠΡΟΛΟΓΟΣ.....	3
1^ο ΚΕΦΑΛΑΙΟ-ΕΙΣΑΓΩΓΗ.....	6
1.1 ΑΝΑΓΚΕΣ ΠΟΥ ΟΔΗΓΗΣΑΝ ΣΤΗΝ ΠΡΑΓΜΑΤΟΠΟΙΗΣΗ ΤΗΣ ΕΡΕΥΝΑΣ	6
1.2 ΔΟΜΗ ΤΗΣ ΜΕΛΕΤΗΣ.....	7
2^ο ΚΕΦΑΛΑΙΟ-ΑΝΑΣΚΟΠΗΣΗ ΒΙΒΛΙΟΓΡΑΦΙΑΣ	8
3^ο ΚΕΦΑΛΑΙΟ-ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ.....	30
3.1 ΕΥΦΥΗ ΣΥΣΤΗΜΑΤΑ	30
3.2 ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ.....	33
3.2.1. Εισαγωγή	33
3.2.2. Τεχνητά Νευρωνικά Δίκτυα (ΤΝΔ).....	34
3.2.3. Τύποι Τεχνητών Νευρωνικών Δικτύων.....	37
3.2.4. Στάδια Διαδικασίας Ανάπτυξης ενός ΤΝΔ.....	43
3.2.5. Μάθηση και Ανάκληση	46
3.3 ΘΕΩΡΙΑ ΑΣΑΦΟΥΣ ΛΟΓΙΚΗΣ (FUZZY LOGIC THEORY)	49
3.3.1 Θεμέλια ασαφούς λογικής.....	50
3.3.2. Συναρτήσεις Συμμετοχής	52
3.3.3 Ιδιότητες των ασαφών συνόλων.....	57
3.3.4. Λογικές πράξεις στα ασαφή σύνολα	60
3.3.5. Ασαφείς σχέσεις	66
3.3.6. Ασαφής συλλογιστική	68

3.4 ΝΕΥΡΟΑΣΑΦΗ ΣΥΣΤΗΜΑΤΑ	71
3.4.1 Εισαγωγή	71
3.4.2 Αρχιτεκτονικές νευρο-ασαφών ελεγκτών	72
 4^ο ΚΕΦΑΛΑΙΟ-ΑΝΑΠΤΥΞΗ ΛΟΓΙΣΜΙΚΟΥ ΓΙΑ ΤΗΝ ΠΡΟΒΛΕΨΗ ΤΗΣ ΤΙΜΗΣ ΤΩΝ ΕΠΙΤΟΚΙΩΝ ΜΕ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ.....	78
4.1 ΠΡΟΔΙΑΓΡΑΦΕΣ ΤΟΥ ΠΡΟΣ ΑΝΑΠΤΥΞΗ ΛΟΓΙΣΜΙΚΟΥ	78
4.2 MATLAB	78
4.2.1 Εργαλείο εκπαίδευσης νευρωνικών δικτύων του MATLAB	78
4.3 ΠΕΡΙΓΡΑΦΗ ΤΟΥ ΝΕΥΡΩΝΙΚΟΥ ΜΟΝΤΕΛΟΥ	79
4.4 ΑΠΟΤΕΛΕΣΜΑΤΑ ΔΟΚΙΜΩΝ	83
 5^ο ΚΕΦΑΛΑΙΟ-ΣΥΜΠΕΡΑΣΜΑΤΑ ΕΡΓΑΣΙΑΣ	90
 ΒΙΒΛΙΟΓΡΑΦΙΑ	91

ΕΙΣΑΓΩΓΗ

Η παρούσα μεταπτυχιακή διατριβή παρουσιάζει την ανάπτυξη μεθοδολογίας για τη πρόβλεψη της τιμής των επιτοκίων με χρήση νευρωνικών δικτύων. Στόχος της εργασίας είναι η εφαρμογή ενός μοντέλου που θα παρέχει αξιόπιστες προβλέψεις αλλά και η εξερεύνηση των δυνατοτήτων που μας δίνονται μέσω νευρωνικών δικτύων και γενετικών αλγορίθμων στον τομέα της πρόβλεψης των επιτοκίων.

Στα πλαίσια της εργασίας αναπτύχθηκε κώδικας σε γλώσσα προγραμματισμού MATLAB όπου το νευρωνικό δίκτυο εκπαιδεύεται και παράγει προβλεπόμενες τιμές επιτοκίου. Παράλληλα η προτεινόμενη μεθοδολογία νευρωνικού δικτύου συγκρίθηκε με άλλες μεθοδολογίες και τα αποτελέσματα που προέκυψαν αποδεικνύουν την αξιοπιστία της προτεινόμενης τεχνικής πρόβλεψης.

1.1 ΑΝΑΓΚΕΣ ΠΟΥ ΟΔΗΓΗΣΑΝ ΣΤΗΝ ΠΡΑΓΜΑΤΟΠΟΙΗΣΗ ΤΗΣ ΕΡΕΥΝΑΣ

Το επιτόκιο είναι μια βασική οικονομική μεταβλητή που επηρεάζει τις αποφάσεις των καταναλωτών, των επιχειρήσεων, των χρηματοδοτικών οργανισμών, των επαγγελματιών επενδυτών και των σχεδιαστών πολιτικής. Οι μεταβολές στα επιτόκια έχουν σημαντικές επιπτώσεις στον επιχειρηματικό κύκλο της οικονομίας και είναι κρίσιμα για την κατανόηση των οικονομικών εξελίξεων και αλλαγών στην οικονομική πολιτική. Οι έγκαιρες προβλέψεις των επιτοκίων μπορούν επομένως να παρέχουν πολύτιμες πληροφορίες στους συμμετέχοντες στην χρηματοοικονομική αγορά και τους σχεδιαστές πολιτικής.

Οι προβλέψεις των επιτοκίων μπορούν επίσης να βοηθήσουν να μειωθεί ο κίνδυνος επιτοκίου που αντιμετωπίζουν τα άτομα και οι εταιρίες. Η πρόβλεψη των επιτοκίων είναι επίσης πολύ χρήσιμη για τις κεντρικές τράπεζες στην αξιολόγηση του γενικού αντίκτυπου (συμπεριλαμβανομένων των αποτελεσμάτων ανατροφοδότησης και προσδοκίας) των πολιτικών της αλλαγών και στην λήψη κατάλληλων διορθωτικών μέτρων, εάν είναι απαραίτητο. Στην πραγματικότητα, η χρησιμότητα των πληροφοριών που περιλαμβάνονται στα επιτόκια αυξάνεται πολύ μετά από τη φιλελευθεροποίηση του οικονομικού τομέα.

Βλέπουμε, επομένως, πόσο σημαντική είναι η πρόβλεψη των επιτοκίων σε διάφορους χρονικούς ορίζοντες. Υπάρχει, επομένως, η ανάγκη για μία κατά το δυνατόν

αξιόπιστη και - με μικρές αποκλίσεις από την πραγματικότητα - πρόβλεψη του επιτοκίου, με στόχο τα αποτελέσματα αυτής να είναι τα καλύτερα δυνατά ούτως ώστε να μπορούν να χρησιμοποιηθούν για το σχεδιασμό των απαραίτητων ενεργειών.

1.2 ΔΟΜΗ ΤΗΣ ΜΕΛΕΤΗΣ

Τα κεφάλαια της εργασίας έχουν την ακόλουθη δομή:

Στο **Κεφάλαιο 2** γίνεται μια συνοπτική παρουσίαση των διαφόρων υπάρχουσών προσεγγίσεων στην πρόβλεψη των επιτοκίων με την έκθεση των κύριων χαρακτηριστικών τους.

Το **Κεφάλαιο 3** ξεκινάει με μια εισαγωγή στο πεδίο που αφορά τα ευφυή συστήματα και ακολουθεί μια παρουσίαση τριών κατηγοριών τους που είναι τα ασαφή συστήματα, τα τεχνητά νευρωνικά δίκτυα και τα νευρο-ασαφή συστήματα. Περιγράφονται βασικές έννοιες, ορισμοί, χαρακτηριστικά, δομή, στάδια ανάπτυξης κλπ.

Στο **Κεφάλαιο 4** γίνεται μια αναλυτική περιγραφή του λογισμικού του νευρωνικού δικτύου που αναπτύχθηκε για την πρόβλεψη της τιμής των επιτοκίων. Παρουσιάζονται επίσης τα αποτελέσματα των δοκιμών δηλαδή τα σφάλματα και οι γραφικές παραστάσεις που αφορούν τα μοντέλα νευρωνικό δίκτυο, ANFIS και γραμμικό μοντέλο, με τελικό στόχο τη σύγκριση των σφαλμάτων από το σύνολο των μοντέλων.

Τέλος, στο **Κεφάλαιο 5** παρατίθενται τα γενικά συμπεράσματα και οι παρατηρήσεις που προέκυψαν από τις αναλύσεις και εφαρμογές των προηγούμενων κεφαλαίων.

ΑΝΑΣΚΟΠΗΣΗ ΒΙΒΛΙΟΓΡΑΦΙΑΣ

Στο σημείο αυτό κρίνουμε χρήσιμο να παρέχουμε μια ανασκόπηση των διάφορων υπαρχουσών προσεγγίσεων στην πρόβλεψη των επιτοκίων. Δεν σκοπεύουμε εδώ να παρέχουμε μια λεπτομερή επεξεργασία όλων αυτών των προσεγγίσεων, αλλά να μεταβιβάσουμε μια ευρετική και διαισθητική έκθεση των κύριων χαρακτηριστικών τους.

Accuracy of International Interest Rate Forecasts

Οι Thomas F. Gosnell" and Robert W. Kolb" σ' αυτή τη μελέτη αξιολογούν τις ομόφωνες προβλέψεις ενάντια σε δύο αφελή πρότυπα: ένα μοντέλο μη-αλλαγής (no-change model) και μια πρόβλεψη μελλοντικής τιμής, χρησιμοποιώντας ως μέτρα της ακρίβειας της πρόβλεψης το Μέσο Τετραγωνικό Σφάλμα (MSE) της πρόβλεψης και το Μέσο Απόλυτο ποσοστιαίο σφάλμα (MAPE). Επιπλέον, εξέτασαν το μέσο όρο των προβλέψεων και τη διάμεση πρόβλεψη δέκα τραπεζών ως δύο διαφορετικές προβλέψεις ομοφωνίας.

Αυτή η μελέτη εξέτασε την ακρίβεια της πρόβλεψης επιτοκίου δέκα τραπεζών που συμμετέχουν ενεργά στις ευρωπαϊκές χρηματαγορές από τον Μάρτιο του 1990 μέχρι τον Αύγουστο του 1992 δίνοντας 25 διαφορετικές ημέρες πρόβλεψης και παράγοντας συνολικά 95 προβλέψεις για κάθε επιτόκιο. Οι τράπεζες προβλέπουν τα τριμηνιαία επιτόκια για πέντε διαφορετικές χώρες, τη Γερμανία, την Ιαπωνία, την Ελβετία, το Ηνωμένο Βασίλειο, και τις Ηνωμένες Πολιτείες, κάνοντας τις προβλέψεις την ίδια ημέρα κάθε μήνα και πρόσφεραν προβλέψεις για τους ορίζοντες 3, 6, 9, και 12 μηνών.

Τα αποτελέσματα έδειξαν ότι οι συνολικές προβλέψεις των τραπεζών και οι προβλέψεις ομοφωνίας αποδείχθηκαν ανώτερες από την πρόβλεψη της μη-αλλαγής. Ενώ η διαφορά δεν ήταν δραματική, ήταν σαφής. Έναντι της πρόβλεψης μελλοντικού επιτοκίου, οι ατομικές τράπεζες και τα μέτρα ομοφωνίας ήταν κατώτερα. Αυτό ισχύει ακόμη και εξετάζοντας τη δραματική επιτυχία των τραπεζών στην πρόβλεψη των αμερικανικών επιτοκίων σε αυτήν την περίοδο. Εντούτοις, εάν τα θεαματικά, και ίσως μεμονωμένα επιτυχή αμερικανικά αποτελέσματα αποκλείονται από την εκτίμηση, οι τράπεζες αποδεικνύονται εντυπωσιακά κατώτερες από το μελλοντικό επιτόκιο για την

πρόβλεψη των επιτοκίων. Κατά συνέπεια, η δυνατότητα αυτών των τραπεζών να προβλέπουν τα τριμηνιαία επιτόκια για αυτές τις πέντε χώρες υπερβαίνει εκείνη της πρόβλεψης της μη-αλλαγής αλλά μειονεκτεί της δυνατότητα πρόβλεψης του μελλοντικού επιτοκίου.

How Well Do Analysts Forecast Interest Rates?

Σ' αυτό το έγγραφο οι R. A. Kolb και H. O. Stekler εξετάζουν τις προβλέψεις επιτοκίου ενός ουσιαστικού αριθμού οικονομικών αναλυτών που γίνονται για την περίοδο 1982-90. και εξετάζει τρία ζητήματα: α) εάν υπάρχει μια γενική συμφωνία μεταξύ των αναλυτών για το επίπεδο των επιτοκίων έξι μήνες στο μέλλον, β) εάν είναι όλοι οι ειδικοί στις προβλέψεις εξίσου καλοί και γ) εάν είναι οι προβλέψεις πολύτιμες σε μελλοντικούς χρήστες. Χρησιμοποιούν τις κατανομές των διατομών των προβλέψεων, την στατιστική του Friedman για την ανάλυση της διακύμανσης κατά τάξη, και δοκιμές ανεξαρτησίας μεταξύ των προβλέψεων και των εκβάσεων για να εξεταστούν αυτές τις ερωτήσεις.

Τα δεδομένα που χρησιμοποιήθηκαν σ' αυτή τη μελέτη προήλθαν από έρευνες που δημοσιεύθηκαν από το περιοδικό Γουόλ Στρητ. Αυτές οι έρευνες εμφανίζουν εξαμηνιαίες (νωρίς τον Ιανουάριο και τον Ιούλιο) και παρούσες εκτιμήσεις των επιπέδων δύο επιτοκίων που αναμένεται να επικρατήσουν 6 μήνες στο μέλλον. Αυτά τα επιτόκια είναι το 90 ημερών επιτόκιο κρατικού έντοκου γραμματίου (Treasury Bill rate) και η απόδοση σε κυβερνητικά ομόλογα 30 ετών. Εξετάζονται λοιπόν προβλέψεις και για βραχυπρόθεσμα και για μακροπρόθεσμα επιτόκια.

Τα αποτελέσματα έδειξαν ότι στις περισσότερες έρευνες, υπήρξε μια ομοφωνία μεταξύ αυτών των αναλυτών ως προς την κατεύθυνση της μετακίνηση και των βραχυπρόθεσμων και των μακροπρόθεσμων επιτοκίων. Επίσης δείξαμε ότι δεν υπήρξε σημαντική διαφορά στην ικανότητα αυτών των οικονομικών αναλυτών να προβλέπουν τα βραχυπρόθεσμα επιτόκια αλλά ότι υπήρχε μια διαφορά όσον αφορά τις μακροπρόθεσμες προβλέψεις, και ότι αυτές οι προβλέψεις δεν ήταν σημαντικά καλύτερες από τις τυχαίες προβλέψεις.

Forecasting Interest Rates: an application for Brazil

Οι Eduardo J. A. Lima, Felipe Luduvise και Benjamin M. Tabak εξετάζουν τη σχέση μεταξύ των βραχυπρόθεσμων και μακροπρόθεσμων επιτοκίων για την οικονομία της Βραζιλίας. Για τον σκοπό αυτό υπολόγισαν το διάνυσμα αυτοπαλινδρόμησης (VAR) και μια διανυσματική διόρθωση σφάλματος (VEC). Εστιάζουν επίσης σε δύο παραλλαγές του μοντέλου: μια στην οποία οι ειδικοί στις προβλέψεις (forecasters) δεν έχουν τις πληροφορίες σχετικά με τα βραχυπρόθεσμα επιτόκια, και άλλη, στην οποία υποθέτουν ότι οι ειδικοί στις προβλέψεις ξέρουν τη μελλοντική πορεία των βραχυπρόθεσμων επιτοκίων.

Τα δεδομένα για τα μοντέλα πρόβλεψης αφορούν την περίοδο από τον Ιανουάριο του 1995 μέχρι τον Νοέμβριο του 2005, λήφθηκαν από το Bloomberg και αποτελούνταν από καθημερινές παρατηρήσεις. Επειδή σε αυτήν την μελέτη οι συγγραφείς υιοθετούν μηνιαίες παρατηρήσεις, έλαβαν τη μέση μηνιαία αξία της χρονικής σειράς, η οποία μείωσε το δείγμα εισροών σε 131 παρατηρήσεις. Υπολόγισαν τα διανύσματα VAR και VEC μεταξύ του βραχυπρόθεσμου επιτοκίου SELIC, το οποίο καθορίζεται από την Επιτροπή νομισματικής πολιτικής (COPOM), και το μεσοπρόθεσμο επιτόκιο, εκτελώντας out-of-sample προβλέψεις αυτών των μεταβλητών. Έπειτα σύγκριναν τις προβλέψεις από αυτά τα μοντέλα με τα αποτελέσματα που παρουσιάστηκαν από έναν τυχαίο περίπατο για να αξιολογήσουν την ακρίβειά τους.

Τα εμπειρικά αποτελέσματα που επιτεύχθηκαν προτείνουν ότι τα πρότυπα VAR/VEC έχουν κακή απόδοση στην πρόβλεψη των μετατοπίσεων στα επιτόκια από την άποψη της πρόβλεψης της ακρίβειας. Ωστόσο, αυτά τα πρότυπα αποδεικνύεται ότι είναι χρήσιμα δεδομένου ότι βοηθούν να προβλέψουν την κατεύθυνση των μετατοπίσεων στα μεσοπρόθεσμα επιτόκια. Επίσης η χρήση των προτύπων VAR/VEC στη διαμόρφωση της νομισματικής πολιτικής φαίνεται να είναι χρήσιμη δεδομένου ότι μπορεί να βοηθήσει τους σχεδιαστές πολιτικής να δημιουργήσουν σενάρια για τη δομή όρου των επιτοκίων στο μέλλον, και να αξιολογήσουν ποιος θα ήταν ο αντίκτυπος στις μακροοικονομικές μεταβλητές σε αυτά τα σενάρια.

Interest Rate Modeling and Forecasting in India

Ο στόχος αυτής της μελέτης των Pami Dua, Nishita Raje, και Satyananda Sahoo είναι η ανάπτυξη μοντέλων που να μπορούν να προβλέπουν βραχυπρόθεσμα και μακροπρόθεσμα επιτόκια. Χρησιμοποιούν μονομεταβλητά μοντέλα που περιλαμβάνουν τα αυτοπαλινδρομούμενα ενσωματωμένα μοντέλα κινούμενου μέσου όρου (Autoregressive Integrated Moving Average models-ARIMA), και μοντέλα ARIMA με αποτελέσματα αυτοπαλινδρομούμενης εξαρτώμενης ετεροσκεδαστικότητας (Autoregressive Conditional Heteroscedasticity (ARCH)/Generalised Autoregressive Conditional Heteroscedasticity (GARCH)), αλλά και πολυμεταβλητά μοντέλα που περιλαμβάνουν διανυσματικά αυτοπαλινδρομούμενα πρότυπα (Vector Autoregressive models-VAR) που εξειδικεύονται στα επίπεδα, τα διανυσματικά πρότυπα διορθώσεων σφάλματος (Vector Error Correction Models-VECM), και τα Μπεϋζιανά διανυσματικά αυτοπαλινδρομούμενα πρότυπα (Bayesian Vector Autoregressive Models-BVAR). Στα πολυμεταβλητά πρότυπα, εξετάζονται παράγοντες όπως η ρευστότητα, το ποσοστό τραπεζών, το ποσοστό γερο, η έκταση της απόδοσης, ο πληθωρισμός, η πίστωση, και τα ξένα επιτόκια. Επίσης χρησιμοποιείται το μοντέλο τυχαίου περιπάτου ως μέτρο σύγκρισης για την αξιολόγηση της επίδοσης της πρόβλεψης κάθε μοντέλου.

Οι μεταβλητές που υιοθετούνται στα πολυμεταβλητά πρότυπα είναι τα εξής: call money rate, 15-91 ημερών επιτόκιο κρατικού έντοκου γραμματίου (15-91 days Treasury bill rate) και επιτόκια στο 1-έτους, 5-ετών, και 10-ετών κυβερνητικά αξιόγραφα (rates on 1-year, 5-years and 10-years government securities). Για να επιλέξουν το καλύτερο μοντέλο, τα εναλλακτικά πρότυπα υπολογίζονται αρχικά χρησιμοποιώντας εβδομαδιαία στοιχεία κατά τη διάρκεια της περιόδου τον από Απρίλιο του 1997 μέχρι τον Δεκέμβριο του 2001 και προβλέψεις εκτός-δείγματος μέχρι 36-εβδομάδες-μπροστά γίνονται από τον Ιανουάριο μέχρι τον Σεπτέμβριο του 2002.

Η μελέτη διαπιστώνει ότι τα πολυμεταβλητά μοντέλα ξεπερνούν γενικά τα μονομεταβλητά για μακρύτερους ορίζοντες πρόβλεψης. Συνολικά, η μελέτη καταλήγει στο συμπέρασμα ότι η προβλεπτική επίδοση των μοντέλων VAR είναι ικανοποιητική για τα περισσότερα επιτόκια και η ανωτερότητά τους στην απόδοση σημειώνεται σε μακρύτερους ορίζοντες πρόβλεψης.

Forecasting the past: the case of US interest rate forecasts

Οι Markus Spiwoks, Nils Bedke και Oliver Hein εξετάζουν τα στοιχεία πρόβλεψης όσον αφορά την αμεροληψία, την αποδοτικότητα, και την ακρίβεια. Με τη βοήθεια του τροποποιημένου test των Diebold-Mariano για την πρόβλεψη που καλύπτει, συγκρίνουν την ποιότητα της πρόβλεψης με τέσσερα διαφορετικά μέτρα σύγκρισης: 1) η 'αφελής' πρόβλεψη (naive forecast), 2) ένα απλό μοντέλο ARIMA, 3) οι προσδοκίες επιτοκίων της κεφαλαιαγοράς υπό μορφή υποκρυπτόμενων μελλοντικών επιτοκίων, 4) οι μέσες προσδοκίες επιτοκίων των αναλυτών της κεφαλαιαγοράς (προβλέψεις ομοφωνίας). Τέλος, ο συντελεστής TOTA χρησιμοποιείται για να προσδιορίσει την τοπικά-προσανατολισμένη συμπεριφορά ρύθμισης τάσης.

Σε αυτήν την μελέτη, αναλύσαμε 136 προβλεπόμενες χρονικές σειρές για 10-ετών αποδόσεις αμερικανικών κυβερνητικών ομολόγων (10-year US government bond yields) και 3-μηνών επιτόκια αμερικανικών κρατικών έντοκων γραμμάτων (3-month US Treasury bill rates) από 34 τράπεζες, ασφαλιστικές εταιρείες, εταιρίες οικονομικών υπηρεσιών, ερευνητικά και συμβουλευτικά ιδρύματα, ομίλους, και τις βιομηχανικές επιχειρήσεις. Η εκτίμηση έγινε για την περίοδο μεταξύ του Οκτωβρίου του 1989 και του Δεκεμβρίου του 2004. Η συντομότερη χρονική σειρά που εξετάστηκε είναι 50 στοιχεία πρόβλεψης, ενώ η πιο μακροχρόνια 171 στοιχεία πρόβλεψης. Κατά μέσον όρο, κάθε μια από τις 136 χρονοσειρές παρέχει 101 στοιχεία πρόβλεψης. Με το σύνολο 13.798 στοιχείων των δεδομένων πρόβλεψης, αυτή η μελέτη είναι κατά πολύ η πιο εκτενής που έχει γίνει από προβλέψεις αμερικανικού επιτοκίου.

Καμία από τις 136 χρονοσειρές πρόβλεψης που αναλύονται δεν μπορεί να θεωρηθεί αμερόληπτη. Επιπλέον, το 70,6% χρησιμοποιούν ανεπαρκώς πληροφορίες από το πολύ πρόσφατο παρελθόν, ενώ το 80,2% δεν προβλέπουν την τάση του επιτοκίου (αύξηση ή πτώση) καθόλου καλύτερα από μια πρόβλεψη τυχαίου περίπατου. Όλες οι χρονοσειρές πρόβλεψης αποκαλύπτουν μια σαφή τοπικά-προσανατολισμένη ρύθμιση τάσης. Επομένως τείνουν να συμφωνήσουν περισσότερο με το παρόν (ή με το πολύ πρόσφατο παρελθόν) παρά με το μέλλον. Το τροποποιημένο test Diebold-Mariano αποκαλύπτει ότι μόνο το 6,6% των χρονοσειρών πρόβλεψης εκθέτουν ένα περιεχόμενο πληροφοριών που υπερβαίνει σημαντικά αυτήν και των τεσσάρων μέτρων σύγκρισης.

Οι προβλέψεις για την 10-ετών απόδοση αμερικανικών κυβερνητικών ομολόγων ήταν γενικά λάθος, όπως ήταν εκείνοι όσον αφορά τα 3-μηνών επιτόκια αμερικανικών κρατικών έντοκων γραμμάτων με έναν ορίζοντα πρόβλεψης 13 μηνών. Μόνο η βραχυπρόθεσμη (τεσσάρων μηνών) πρόβλεψη για τα 3-μηνών επιτόκια αμερικανικών κρατικών έντοκων γραμμάτων παρουσίασε ακρίβεια.

Nearest-Neighbor Forecasts of U.S. Interest Rates

Οι John Barkoulas, Christopher F. Baum και Atreya Chakraborty χρησιμοποιούν μια μη γραμμική, μη-παραμετρική μέθοδο για να μοντελοποιήσουν την στοχαστική συμπεριφορά των αλλαγών σε διάφορα βραχυπρόθεσμα και μακροπρόθεσμα αμερικανικά επιτόκια. Εφαρμόζουν μια μη γραμμική αυτοπαλινδρόμηση στις σειρές χρησιμοποιώντας την τοπικά σταθμισμένη μέθοδο εκτίμησης παλινδρόμησης (LWR), μια μέθοδο του πλησιέστερου γείτονα, και αξιολογούν την απόδοση της πρόβλεψης με ένα μέτρο της ρίζας του μέσου τετραγωνικού σφάλματος (RMSE). Επίσης συγκρίνουν την απόδοση της πρόβλεψης της μη-παραμετρικής μεθόδου με την απόδοση δύο γραμμικών μοντέλων: ένα αυτοπαλινδρομούμενο μοντέλο (AR) και ένα random-walk-with- drift μοντέλο (RW).

Το σύνολο στοιχείων μας αποτελείται από μια ποικιλία βραχυπρόθεσμων και μακροπρόθεσμων αμερικανικών κρατικών επιτοκίων: μηνιαίες παρατηρήσεις στο επιτόκιο ομοσπονδιακών κεφαλαίων (Federal Funds rate), τριμηνιαίες, εξαμηνιαίες και δωδεκάμηνες αποδόσεις αμερικανικού κρατικού έντοκου γραμματίου (Treasury bill yields), και αποδόσεις σε 5-ετών και 10-ετών κρατικά ομόλογα (Treasury bonds). Η περίοδος δειγμάτων είναι από Ιανουάριο του 1957 έως Δεκέμβριο του 1993 για όλες τις σειρές εκτός από την απόδοση του εξαμηνιαίου Treasury bill, για το οποίο είναι από 01/1959 έως 12/1993. Οι τριμηνιαίες και εξαμηνιαίες αποδόσεις είναι ποσοστιαία ετήσια επιτόκια που λαμβάνονται από τη δευτεροβάθμια αγορά. Οι αποδόσεις 1, 5 και 10 ετών είναι σταθερά ετήσια ώριμα επιτόκια ποσοστά. Όλες οι σειρές στοιχείων λαμβάνονται από τη βάση δεδομένων Citibank.

Οι συγγραφείς παρέχουν ενδείξεις ότι η μη-παραμετρική εφαρμογή που παράγεται από τη μέθοδο LWR οδηγεί σε σημαντικές βελτιώσεις, έναντι των (γραμμικών) μοντέλων συγκριτικής μέτρησης επιδόσεων, στις out-of-sample προβλέψεις αμερικανικών βραχυπρόθεσμων και μακροπρόθεσμων επιτοκίων. Τα στοιχεία που παρέχονται καθιερώνουν την παρουσία σημαντικής μη γραμμικής μέσης προβλεψιμότητας στα αμερικανικά επιτόκια, όπως προτείνεται από θεωρητικά συμπεράσματα. Αν και η μέθοδος εκτίμησης LWR έχει αποτύχει να προβλέψει επιτυχώς τις αποδόσεις αποθεμάτων (Hsieh (1991), LeBaron (1988)) και τις συναλλαγματικές ισοτιμίες (Diebold και Nason (1990), Meese και Rose (1990, 1991), Mizrach (1992)), εμφανίζεται να είναι πολύ χρήσιμη στη μοντελοποίηση των υποθετικών μέσων αλλαγών στη σειρά επιτοκίου. Εκτός από τις μη γραμμικότητες στις διακυμάνσεις και τις ενδεχομένως υψηλότερες στιγμές, σημαντικές μη γραμμικότητες στο μέσο όρο υπάρχουν σαφώς για τα αμερικανικά επιτόκια, και η LWR εμφανίζεται να συλλαμβάνει επιτυχώς αυτές τις μη γραμμικότητες.

Improving Forecasts of the Federal Funds Rate in a Policy Model

Πρόσφατες κριτικές για τα μοντέλα VAR θέτουν υπό αμφισβήτηση την ακρίβεια των προβλέψεων VAR, και ως εκ τούτου, προκάλεσαν ανησυχία για τη βιωσιμότητα της χρήσης αυτών των μοντέλων στον χώρο της ‘πολιτικής’.

Οι John C. Robertson και Ellis W. Tallman σε αυτό το έγγραφο, εξετάζουν την επίδοση της πρόβλεψης του ενός και δύο μηνών-μπροστά επιτοκίου χρεογράφων ενός διανύσματος Αυτόπαλινδρόμησης (VAR) χρησιμοποιώντας δύο σύνολα δεδομένων που λαμβάνονται από τη βιβλιογραφία. Για κάθε σύνολο δεδομένων υπολογίζουν τους συντελεστές ενός VAR με Μέσα Ελάχιστα Τετράγωνα (OLS) και επιβεβαιώνουν την κακή απόδοση πρόβλεψης που έχει αναφερθεί προηγουμένως. Πειραματίζονται έπειτα με διάφορους περιορισμούς αποκλεισμού στους συντελεστές VAR πριν από την εκτίμηση από OLS και διαπιστώνουν ότι η επιβολή ακριβών περιορισμών βελτιώνει κάπως την απόδοση της πρόβλεψης ενός απεριόριστου VAR, αν και τα συνοπτικά μέτρα πρόβλεψης σφάλματος είναι σημαντικά μεγαλύτερα από τη μελλοντική αγορά σε όλες τις περιπτώσεις. Τέλος, υπολογίζουν τους συντελεστές VAR χρησιμοποιώντας την Μπεϋζιανή μέθοδο που παρουσιάζεται στους Kadiyala και Karlsson (1997), και Sims και Zha (1998). Χρησιμοποιούν αυτήν την μέθοδο επειδή είναι ρητά πολυμεταβλητή, και ως εκ τούτου μπορεί να εφαρμοστεί σε προσδιορισμένα μοντέλα VAR.

Για να εξετάσουν τον αντίκτυπο της τεχνικής εκτίμησης στις προβλέψεις VAR, χρησιμοποιούν δύο σύνολα μηνιαίων δεδομένων, τα οποία αποτελούνται από μεταβλητές που θα ήταν ρητής ενδιαφέροντος για τους νομισματικούς φορείς χάραξης πολιτικής. Το πρώτο είναι ένα σύνολο έξι μεταβλητών που χρησιμοποιούνται από τους Waggoner και Zha (1999) και Zha (1998). Οι σειρές δεδομένων είναι τα επίπεδα του αποτελεσματικού επιτοκίου ομοσπονδιακών χρεογράφων (federal funds rate) και του ποσοστού ανεργίας, μαζί με τους φυσικούς λογάριθμους του δείκτη τιμών προϊόντων, του δείκτη τιμών του καταναλωτή, του πραγματικού ΑΕΠ (μηνιαίας διανομής), και του νομισματικού συνόλου M2. Το δεύτερο σύνολο δεδομένων είναι συγκρίσιμο με αυτό που χρησιμοποιείται από τους Christiano, Eichenbaum και τον Evans (1997), και αποτελείται από το επιτόκιο χρεογράφων, και τους φυσικούς λογαρίθμους των συνολικών αποθεμάτων, M1, της απασχόλησης μισθοδοτικών καταστάσεων, τον προσωπικό αποπληθωριστή δαπανών κατανάλωσης, και των τιμών των προϊόντων. Οι μόνες σειρές κοινές και για τα δύο σύνολα δεδομένων είναι το επιτόκιο ομοσπονδιακών χρεογράφων και ο δείκτης τιμών προϊόντων. Εντούτοις, πολλές από τις σειρές στα δύο σύνολα δεδομένων παρουσιάζουν παρόμοιες τάσεις, και έτσι κάθε σύνολο δεδομένων δεν αντιπροσωπεύει απολύτως ανεξάρτητες πληροφορίες.

Τα αποτελέσματά τους εμφανίζονται να αντικρούουν το επιχείρημα ότι τα μοντέλα VAR κάνουν κακή πρόβλεψη, και ότι δεν είναι καλά εργαλεία για την πολιτική

ανάλυση. Η βελτιωμένη ακρίβεια πρόβλεψης για τα BVAR είναι γερή για τα VAR που υπολογίζονται χρησιμοποιώντας δύο διαφορετικά σύνολα δεδομένων. Για ένα σύνολο δεδομένων, το BVAR παράγει προβλέψεις των οποίων η ακρίβεια είναι κατά ένα μεγάλο μέρος όμοια με την ακρίβεια των προβλέψεων που είναι υποκρυπτόμενες σε δεδομένα ομοσπονδιακών κεφαλαίων μελλοντικών αγορών. Για ένα άλλο σύνολο δεδομένων οι προβλέψεις BVAR είναι πιο ακριβείς από αυτές των εναλλακτικών προδιαγραφών, αλλά ακόμα είναι στατιστικά κατώτερες από τις προβλέψεις που υπονοούνται από τα στοιχεία μελλοντικής αγοράς. Δηλαδή η καλή επίδοση στην πρόβλεψη της BVAR δεν είναι απολύτως ανεξάρτητη από το σύνολο δεδομένων που χρησιμοποιείται στο μοντέλο. Τα εμπειρικά συμπεράσματα σε αυτό το έγγραφο προτείνουν ότι η χρήση των πιθανολογικών προγενέστερων πληροφοριών μπορεί να παρέχει έναν αποτελεσματικό τρόπο παρουσίασης ενός βαθμού φειδωλότητας σε μια υπερδιάστατη προδιαγραφή VAR χωρίς συγχρόνως να επιβάλει πάρα πολλούς ισχυρούς, και ενδεχομένως ψεύτικους περιορισμούς στο μοντέλο.

Forecasting Changes in UK Interest Rates

Οι Tae-Hwan Kim, Paul Mizen and Thanaset Chevapatrakul εκτιμούν την προβλεπτική δύναμη διαφορετικών συνόλων πληροφοριών για τη βρετανική διαδικασία καθορισμού επιτοκίου χρησιμοποιώντας ένα δυναμικό διαταγμένο probit μοντέλο για να συλλάβουν την ιδιαίτερη φύση των αποφάσεων της επιτροπής. Καθορίζουν εάν οι πληροφορίες προβλέπουν ακριβώς την επόμενη πλέον πιθανή αλλαγή χρησιμοποιώντας μηνιαία στοιχεία. Τα αποτελέσματά τους βασίζονται στις προβλέψεις από ορισμένα διαφορετικά σύνολα πληροφοριών συμπεριλαμβάνοντας ένα μοντέλο τυχαίου περιπάτου, ένα τυχαίο περίπατο σε πρώτες διαφορές, το σύνολο πληροφοριών του κανόνα Taylor και ένα ευρύτερο σύνολο πληροφοριών που αυξάνει τις πληροφορίες του κανόνα Taylor με ποσοστά αύξησης στη νομισματική, συναλλαγματικές ισοτιμίες, αποδοχές και μεταβλητές δαπανών. Κάνουν προβλέψεις της κατεύθυνσης της αλλαγής στο επιτόκιο in-sample και out-of-sample χρησιμοποιώντας μια σειρά συνοπτικών στατιστικών για να αξιολογήσουν το ποσοστό των σωστών προβλέψεων από κάθε σύνολο πληροφοριών.

Τα δεδομένα τους είναι μηνιαία στοιχεία των επιτοκίων από την τράπεζα της Αγγλικής Επιτροπής νομισματικής πολιτικής (Monetary Policy Committee-MPC) για τη δειγματική περίοδο 2/1993 έως 7/2003.

Τα αποτελέσματά τους έδειξαν τα εξής. Κατ' αρχάς, η χρήση ευρύτερων συνόλων πληροφοριών βελτιώνει τις in-sample και out-of-sample προβλέψεις σχετικές με το σύνολο πληροφοριών του κανόνα Taylor όταν βασίζονται στο δυναμικό διαταγμένο

probit μοντέλο, το οποίο απεικονίζει την ιδιαίτερα επίμονη και σταδιακή φύση της σειράς επιτοκίου. Δεύτερον, τα περισσότερα αφελή πρότυπα προβλέπουν άσχημα σε σύγκριση με τα δυναμικά διαταγμένα probit μοντέλα χρησιμοποιώντας παρόμοιες τιμές κατωφλίων, και αυτά περιλαμβάνουν έναν τυχαίο περίπατο, ένα μοντέλο AR (1), και δύο γραμμικές προβλέψεις βασισμένες στον ίδιο κανόνα του Taylor και σε ευρύτερα σύνολα πληροφοριών που χρησιμοποιούνται στα δυναμικά διαταγμένα probit μοντέλα. Τρίτον, ένα απλό μοντέλο τυχαίου περιπάτου πρώτων-διαφορών προβλέπει καλά out-of-sample, προσφέροντας ανώτερη απόδοση σε μερικές περιπτώσεις όταν καθορίζονται λεπτομερώς τα κατώφλια έναντι του δυναμικού διαταγμένου probit μοντέλου. Η δυνατότητα αυτού του μοντέλου να αξιοποιήσει την αυτοσυσχέτιση στην αλλαγή στα επιτόκια όταν ρυθμίζονται τα ποσοστά, ενώ συγχρόνως συλλαμβάνοντας την τάση για τα επιτόκια να παραμείνουν αμετάβλητα για μεγάλες περιόδους, δίνει μια απλή αλλά αποτελεσματική πρόβλεψη για την αλλαγή στα βρετανικά επιτόκια. Τα γραμμικά μοντέλα επίσης μερικές φορές προσφέρουν καλές προβλέψεις, αλλά δεν μπορούμε να κάνουμε άμεσες συγκρίσεις με τις δυναμικές διαταγμένες probit προβλέψεις επειδή οι τιμές κατωφλίων διαφέρουν.

Non-linear interest rate dynamics and forecasting: evidence for US and Australian interest rates

Ο στόχος του David G. McMillan είναι να συγκρίνει την επεξηγηματική και προφητική δύναμη διάφορων εμπειρικών μοντέλων με σκοπό να συλλάβουν τη δυναμική μη γραμμικής συνολοκλήρωσης (cointegration) και διόρθωσης-σφάλματος και στην αμερικανική και την αυστραλιανή σειρά επιτοκίου ενάντια στην εναλλακτική τυπικής γραμμικής διόρθωσης-σφάλματος. Πιο συγκεκριμένα, χρησιμοποιεί μια ποικίλα μοντέλων τύπου-κατώφλια που εξετάζονται αλλού στη βιβλιογραφία και τα οποία είναι σε θέση να συλλάβουν συμπεριφορά που είναι σύμφωνη με την παρουσία τριβής στην αγορά, εμπορών θορύβου και ασυμμετρικής πολιτικής επέμβασης. Κατά συνέπεια, εκτός από το απλό μοντέλο των Escibano και Granger (1998) οι οποίοι αυξάνουν τον όρο διόρθωσης-σφάλματος με τετραγωνικούς και κυβικούς όρους, εξετάζει μια ποικίλα μοντέλων κατωφλιού, περιλαμβάνοντας προσαρμογές του αυτοπαλινδρομούμενου μοντέλου κατωφλιού (TAR, Tong, 1983), του momentum-TAR (MTAR, Enders και Granger, 1998; Enders και Siklos, 2001), και των μοντέλων ομαλής-μετάβασης (STR, βλέπε Terasvirta, 1998).

Για τα δεδομένα, λαμβάνει δύο μηνιαίες αμερικανικές και αυστραλιανές σειρές επιτοκίου, συγκεκριμένα, ένα βραχυχρόνιο επιτόκιο (το Federal Funds rate για τις ΗΠΑ

και το overnight interbank rate για την Αυστραλία) και ένα μακροχρόνιο επιτόκιο (η απόδοση στο 10-year T-bond για τις ΗΠΑ και την Αυστραλία) κατά τη διάρκεια των περιόδων 07/1954 έως 10/2002 για τις ΗΠΑ και 07/1969 έως 10/2002 για την Αυστραλία. Τα στοιχεία λαμβάνονται από τις Ομοσπονδιακές αποθεματικές τράπεζες των ΗΠΑ και της Αυστραλίας, αντίστοιχα.

Τα αποτελέσματα συντελεστή προτείνουν βραχυπρόθεσμη αιτιότητα Granger που τρέχει από το μακροχρόνιο στο βραχυχρόνιο ποσοστό και για τα αμερικανικά και τα αυστραλιανά ποσοστά, με μακροπρόθεσμες εξαρτήσεις και στις δύο σειρές για τις ΗΠΑ, ενώ βρίσκονται στοιχεία αδύναμης εξωγένειας για το αυστραλιανό μακροχρόνιο ποσοστό. Περαιτέρω, για τις αμερικανικές σειρές, μετά από την εκτίμηση του μοντέλου LSTR επίσης αποδεικνύεται ότι η βραχυπρόθεσμη αιτιότητα Granger τρέχει από το βραχυχρόνιο στο μακροχρόνιο ποσοστό όταν ο όρος διόρθωσης-σφάλματος είναι πάνω από την ισορροπία, δηλ. το βραχυχρόνιο ποσοστό είναι 'πολύ υψηλότερο' σχετικά με το μακροχρόνιο ποσοστό. Και τα test προδιαγραφών και οι μετρικές πρόβλεψης παρέχουν υποστήριξη για τα μη γραμμικά μοντέλα πέρα από το γραμμικό μοντέλο, με, ειδικότερα, το μοντέλο LSTR να προτιμάται κατά ένα μεγάλο μέρος από όλες τις εναλλακτικές λύσεις, αν και για το μακροχρόνιο ποσοστό οποιοδήποτε όφελος πρόβλεψης είναι οριακό. Τα αποτελέσματα συντελεστών από τα μοντέλα LSTR υποστηρίζουν την ακόλουθη δυναμική ρύθμισης για τις ΗΠΑ, πρώτον, η επαναφορά στην ισορροπία είναι γρηγορότερη όταν το βραχυχρόνιο ποσοστό υπερβαίνει το μακροχρόνιο, που σημαίνει θετικές αποκλίσεις από την ισορροπία, δεύτερον, ότι η επαναφορά μετά από έναν θετικό κλονισμό εμφανίζεται πιο γρήγορα στο βραχυχρόνιο ποσοστό απ' ότι στο μακροχρόνιο, αλλά ότι και οι δύο σειρές παρουσιάζουν επαναφορά. Εντούτοις, για τις αρνητικές αποκλίσεις η επαναφορά εμφανίζεται μόνο μέσω του μακροχρόνιου ποσοστού, και τελικά, η μετάβαση μεταξύ των καθεστώτων συμπεριφοράς είναι γρηγορότερη στο βραχυχρόνιο απ' ότι στο μακροχρόνιο ποσοστό, αν και οι δύο σειρές παρουσιάζουν σχετικά γρήγορη δυναμική ρύθμισης. Για τα αυστραλιανά επιτόκια όλη η μακροπρόθεσμη δυναμική ρύθμισης εμφανίζεται μέσω του βραχυχρόνιου ποσοστού, με το μακροχρόνιο ποσοστό να παρουσιάζει αδύναμης εξωγένειας. Περαιτέρω, εκείνη η γρηγορότερη επαναφορά εμφανίζεται όταν το μακροχρόνιο ποσοστό υπερβαίνει το βραχυχρόνιο.

Why Central Bankers Should Disclose Interest Rate Forecast

Οι Jan Filáček, Luboš Komárek και Petr Král εξετάζουν τη θεωρητική και εμπειρική διάσταση της δημοσίευσης των προβλέψεων των επιτοκίων από τις κεντρικές

τράπεζες. Ο πρώτος στόχος τους είναι να επανεξετάσουν τα επιχειρήματα υπέρ και ενάντια αυτής της απόφασης και να επεξηγήσουν τη συζήτηση χρησιμοποιώντας την εμπειρία τεσσάρων κεντρικών τραπεζών που δημοσιεύουν ή συνήθιζαν να δημοσιεύουν την πρόβλεψη επιτοκίων. Ο δεύτερος στόχος είναι να αξιολογηθεί η ικανότητα της Τσεχικής Εθνικής Τράπεζας να χρησιμοποιεί αποτελεσματικά την δημοσίευση των προβλέψεων των επιτοκίων της για να αυξήσει περαιτέρω τη διαφάνεια, την υπευθυνότητα και την αξιοπιστία της.

Υποστηρίζουν ότι:

- i. υπάρχουν διάφορα σοβαρά θεωρητικά και πρακτικά επιχειρήματα για τη δημοσιοποίηση ή όχι των προβλέψεων των επιτοκίων, χωρίς όμως η ζυγαριά να γέρνει σε μια πλευρά μόνο. Παρ' όλα αυτά, τα πλεονεκτήματα της αποκάλυψης της πορείας των επιτοκίων είναι μιας κάπως μεγαλύτερης σχετικότητας σε σχέση με τα μειονεκτήματα, αν και πρακτικές ρυθμίσεις σε μια συγκεκριμένη κεντρική τράπεζα διαδραματίζουν έναν σημαντικό ρόλο. Το αρχείο διαδρομής των κεντρικών τραπεζών που δημοσιεύουν ή συνήθιζαν να δημοσιεύουν τις ενδογενείς τροχιές επιτοκίου προτείνει ότι αυτό είναι πραγματικά ένα υψηλού επιπέδου μοντέλο (ή ακόμα και κατάσταση προόδου) για τη διαφάνεια νομισματικής πολιτικής. Ένα σύστημα με καλή λειτουργία πρόβλεψης και πολιτικής ανάλυσης, η αποδοτική επικοινωνία νομισματικής πολιτικής και ένα γενικό υψηλό επίπεδο αξιοπιστίας των κεντρικών τραπεζών είναι μεταξύ των απαραίτητων προϋποθέσεων που πρέπει να υπάρξουν για να αρχίσουν τη δημοσίευση μιας πρόβλεψης επιτοκίου.
- ii. η Τσεχική Εθνική Τράπεζα καλύπτει όλες τις προϋποθέσεις που απαιτούνται για την δημοσιοποίηση των προβλέψεων της για τα επιτόκια
- iii. πιο συγκεκριμένα, δεν βρίσκουν κανένα στοιχείο που να αποδεικνύει σημαντικά χειρότερη απόδοση της CNB στην πρόβλεψη επιτοκίων σε σύγκριση με άλλα ιδρύματα
- iv. ο τρέχων τρόπος έκφρασης της πιθανής μελλοντικής ανάπτυξης των επιτοκίων έχει αποδειχθεί αποδοτικός να επηρεάσει, μέχρι ένα σημείο, τις προσδοκίες αγοράς της μελλοντικής νομισματικής πολιτικής. Όμως, όπως δείχνουν τα εμπειρικά στοιχεία, δεν είναι σαφές εάν η αξιολόγηση του κινδύνου σε συνδυασμό με τη λεκτική περιγραφή της πρόβλεψης επιτοκίου που

χρησιμοποιείται αυτήν την περίοδο είναι πάντα αρκετά ακριβής καθοδήγηση για τους συμμετέχοντες της αγοράς

- v. αφού αρχίζει η CNB να δημοσιοποιεί τις προβλέψεις επιτοκίων το 2008, οι προσδοκίες της αγοράς θα κινηθούν πιθανόν πιο κοντά στην άποψη της CNB το οποίο πρέπει στη συνέχεια να ενισχύσει την υπευθυνότητα και την αξιοπιστία των προβλέψεων της CNB και να βελτιώσει την αποδοτικότητα της συμπεριφοράς νομισματικής πολιτικής της.

Further Evidence on the Forecasting Performance of Two Factor Continuous Time Interest Rate Models in International and Asia-Pacific Financial Markets

ΟΙ S. L. Byers και K. Ben Nowman υπολογίζουν εκδοχές δύο παραγόντων των πολύ γνωστών μοντέλων επιτοκίου CKLS, Vasicek και CIR. Κάθε μοντέλο διαφέρει στο βαθμό σχέσης μεταξύ της αστάθειας των επιτοκίων και του επιπέδου επιτοκίων. Υπολογίζουν τα μοντέλα επιτοκίου συνεχούς χρόνου χρησιμοποιώντας διακριτά στοιχεία και τις γκαουσιανές μεθόδους εκτίμησης του Bergstrom (1983, 1985, 1986) που χρησιμοποίησε ο Nowman (2000) πρόσφατα.

Τα βραχυπρόθεσμα επιτόκια που χρησιμοποιούνται σε αυτήν την μελέτη είναι μηνιαία Interbank και Euro-currency επιτόκια για την Αυστραλία, το Βέλγιο, τον Καναδά, τη Δανία, τη Γαλλία, τη Γερμανία, το Χονγκ Κονγκ, την Ιταλία, τη Μαλαισία, την Ολλανδία, τη Σιγκαπούρη, την Ελβετία, την Ταϊλάνδη, τις Ηνωμένες Πολιτείες και τη Νέα Ζηλανδία που λαμβάνονται από την *Datastream*.

Οι διαπιστώσεις που κάνουν οι συγγραφείς στο τέλος είναι ότι ο βαθμός της επίδρασης του επιπέδου-αστάθειας ποικίλλει στις χώρες και στους δύο παράγοντες, καθώς επίσης και ότι η απόδοση της πρόβλεψης ποικίλλει στα μοντέλα και αυτές τις αγορές.

Can Forecasting Performance be improved by considering the Steady State? An Application to Swedish Inflation and Interest Rate

Ο Pär Österholm ερευνά εάν η απόδοση της πρόβλεψης των Bayesian AR και VAR μοντέλων μπορεί να βελτιωθεί με την ενσωμάτωση προηγούμενων πεποιθήσεων

στη σταθερή κατάσταση των χρονοσειρών στο σύστημα. Αυτό το ζήτημα ερευνήθηκε διεξάγοντας μια άσκηση πρόβλεψης out-of-sample χρησιμοποιώντας σουηδικά στοιχεία. Οι πληροφορίες που ενσωματώθηκαν στα μοντέλα σε αυτήν την εφαρμογή προήλθαν από το συνδυασμό (i) της γνώσης ότι η Riksbank είχε εγκαταλείψει τη σταθερή συναλλαγματική ισοτιμία και είχε υιοθετήσει αντ' αυτού έναν στόχο πληθωρισμού με (ii) την υπόθεση Fisher η οποία συνδέει το ονομαστικό επιτόκιο με τον πληθωρισμό.

Τα στοιχεία που χρησιμοποιήσαν ήταν τριμηνιαία σουηδικά δεδομένα από το 1980 Q2 έως το 2004 Q4 στο year-ended CPI inflation και το τριμηνιαίο treasury bill επιτόκιο και προέρχονται από την Sveriges Riksbank.

Τα αποτελέσματα δείχνουν ότι η ικανότητα πρόβλεψης out-of-sample των μοντέλων είναι σχεδόν αμετάβλητη για τον πληθωρισμό, αλλά το κέρδος στην απόδοση των προβλέψεων των επιτοκίων είναι μεγάλο και στατιστικά σημαντικό όταν παρέχονται πληροφοριακές προγενέστερες διανομές στη σταθερή κατάσταση. Τα συμπεράσματα σε αυτό το έγγραφο υπονοούν ότι αυτή η νέα μεθοδολογία θα μπορούσε να είναι χρήσιμη δεδομένου ότι μας επιτρέπει να 'ακονίσουμε' τις προβλέψεις μας στην παρουσία πιθανών παγίδων όπως οι διαδικασίες κοντινής μοναδιαίας ρίζας (near unit root processes) και τα δομικά 'σπασίματα' (structural breaks), ειδικότερα όταν βασιζόμαστε σε μικρά δείγματα.

Forecasting the South African economy with GIBBS sampled BVECMs

Ο Rangan Gupta χρησιμοποιεί την τεχνική δειγματοληψίας Gibbs για να υπολογίσει ένα Μπεϋζιανό διανυσματικό μοντέλο διόρθωσης σφάλματος (BVECM) της νοτιοαφρικανικής οικονομίας για να προβλέψει το ΑΕΠ, την κατανάλωση, την επένδυση, βραχυπρόθεσμα και μακροπρόθεσμα επιτόκια, και το ΔΤΚ. Η ακρίβεια της out-of-sample πρόβλεψης ως αποτέλεσμα του BVECM συγκρίνεται με εκείνα που παράγονται από ένα BVECM που υπολογίζεται χρησιμοποιώντας τη μικτή μέθοδο εκτίμησης του Theil (1971) και ένα μοντέλο κλασσικό VECM. Σημειώνουμε ότι τα πρότυπα BVAR και BVEC έχουν υπολογιστεί χρησιμοποιώντας μικτή τεχνική εκτίμησης του Theil (1971), όπου η δομή σφάλματος υποτίθεται ότι ακολουθεί τις Gauss-Markov υποθέσεις. Εντούτοις, σε αυτό το έγγραφο, η εκτίμηση του BVECM που βασίζεται στη δειγματοληψία Gibbs, μας επιτρέπει να ενσωματώσουμε outliers¹ ή μη-σταθερές διακυμάνσεις.

¹Στην στατιστική outlier είναι μια παρατήρηση που φαίνεται να παρεκκλίνει έντονα από άλλα στοιχεία του δείγματος στο οποίο συντελείται.

Υπολογίζεται λοιπόν ένα μοντέλο BVECM χρησιμοποιώντας τη δειγματοληψία Gibbs για τη νοτιοαφρικανική οικονομία για την περίοδο στις 1/1970 ως 4/2000, βασισμένη σε τριμηνιαία στοιχεία. Έπειτα υπολογίζει out-of-sample ένα— έως οκτώ- τρίμηνα-μπροστά προβλέψεις για την περίοδο 1/2001 ως 4/2005, και συγκρίνει μετά την ακρίβεια της πρόβλεψης σχετικά με τις προβλέψεις που παράγονται από ένα κλασσικό VECM, και ένα BVECM που υπολογίζεται χρησιμοποιώντας τη μικτή τεχνική εκτίμησης του Theil (1971). Οι μεταβλητές που συμπεριλαμβάνονται είναι το πραγματικό ΑΕΠ, η κατανάλωση, η επένδυση, 91 ημερών επιτόκιο κρατικού έντοκου γραμμάτιου (Treasury Bill) , 10 χρόνων και περισσότερο επιτόκιο κυβερνητικών ομολόγων, και ο Δείκτης Τιμών Καταναλωτή (CPI). Όλα τα δεδομένα λαμβάνονται από το Τριμηνιαίο Δελτίο της Κεντρικής Τράπεζας της Νότιας Αφρικής.

Συνοψίζοντας, διαπιστώνουμε ότι τα BVECM_G με πιο ‘σφιχτά’ priors² ταιριάζουν καλύτερα για την πρόβλεψη του μοντέλου μας των έξι-μεταβλητών της νοτιοαφρικανικής οικονομίας κατά τη διάρκεια της περιόδου 1/2001 έως 4/2005. Εντούτοις, είναι δύσκολο να βρεθεί ένα μοναδικό BVECM_G που να ξεπερνά με συνέπεια το VECM και το βέλτιστο BVECM που λαμβάνονται στο Gupta (2006). Αντίθετα από το βέλτιστο BVECM που παρήγαγε χαμηλότερα σφάλματα πρόβλεψης κατά μέσον όρο για 5 μεταβλητές (δαπάνες κατανάλωσης, ΑΕΠ, επιτόκιο Treasury Bill , 10 χρόνων και περισσότερο επιτόκιο κυβερνητικών ομολόγων) όταν συγκρίνεται με το VECM, το BVECM_G με τα πιο ‘σφιχτά’ priors μπορεί μόνο να ξεπεράσει το τελευταίο στην περίπτωση τεσσάρων μεταβλητών (δαπάνες κατανάλωσης, ΔΤΚ, ΑΕΠ, δαπάνες επένδυσης και το Treasury bill rate). Αλλά πρέπει να υπογραμμιστεί ότι είναι πάντα δυνατό να βρούμε ένα BVECM_G, βασισμένο σε εναλλακτικές προγενέστερες προδιαγραφές, το οποίο παράγει χαμηλότερα σφάλματα πρόβλεψης out-of-sample σε σύγκριση με το κλασσικό VECM.

Εντούτοις, όταν συγκρίνεται με το βέλτιστο BVECM, το BVECM_G, με τα πιο ‘σφιχτά’ priors, ξεπερνά το πρώτο για μόνο τρεις μεταβλητές (δαπάνες οικιακής κατανάλωσης, ΑΕΠ και 10 ετών και περισσότερο επιτόκιο κυβερνητικών ομολόγων). Έτσι εάν μεταχειριζόμαστε το VECM ως μέτρο σύγκρισης, τότε σαφώς το ‘βέλτιστο’ BVECM του Gupta (2006) ταιριάζει καλύτερα για την πρόβλεψη της νοτιοαφρικανικής οικονομίας των έξι μεταβλητών κατά τη διάρκεια της περιόδου 1/2001 ως 4/2005, καθώς ξεπερνά το κλασσικό VECM για όλες τις μεταβλητές εκτός από το μέτρο του μακροπρόθεσμου επιτοκίου.

² Στο Μπεϋζιανό συμπέρασμα, το ‘prior’ είναι μια κατανομή πιθανότητας που αντιπροσωπεύει γνώση ή πεποίθηση για μια άγνωστη ποσότητα πριν παρατηρηθούν οποιαδήποτε δεδομένα. Η άγνωστη ποσότητα μπορεί να είναι μια παράμετρος, υπόθεση ή λανθάνουσα μεταβλητή.

Optimizing time-series forecasts for inflation and interest rates using simulation and model averaging

Παρακινημένοι από έννοιες της οικονομικής θεωρίας- η όπως υπόθεση Fisher και η θεωρία της δομής του όρου - οι Adusei Jumah και Robert M. Kunst εξετάζουν ένα μικρό σύνολο απλών διμεταβλητών μοντέλων closed-loop χρονοσειρών για την πρόβλεψη του πληθωρισμού των τιμών και μακροπρόθεσμων και βραχυπρόθεσμων επιτοκίων. Το σύνολο περιλαμβάνει διανυσματικές αυτοπαλινδρομήσεις (VAR) σε επίπεδα και σε διαφορές, ένα VAR συνολοκλήρωσης, και ένα μη γραμμικό VAR με κατώφλι συνολοκλήρωσης.

Τα στοιχεία που χρησιμοποιούν είναι δείγματα κατά τη διάρκεια αρκετών δεκαετιών στη Γερμανία, την Ιαπωνία, τη Βρετανία, και τις ΗΠΑ. Αυτές είναι τέσσερις μεγάλες οικονομίες διαδεδομένης σχετικότητας. Τα δείγματα καλύπτουν το επεισόδιο του υψηλού πληθωρισμού και επιτοκίου που ακολούθησε μετά από τους κλονισμούς του ΟΠΕΚ της δεκαετίας του '70 και είναι επομένως αντιπροσωπευτικά της χαρακτηριστικής συμπεριφοράς στις πιο μακροπρόθεσμες χρονικές εκτάσεις που θα παρουσιάζουν πάντα ήρεμες 'φάσεις' και 'άγρια' επεισόδια που δεν αντιστοιχούν στη φρόνηση εγχειριδίων.

Καταλήγουν ότι το απλό μοντέλο dVAR χωρίς περιορισμούς εξουσιάζει σε πολύ μικρότερους ορίζοντες πρόβλεψης. Στους μακρύτερους ορίζοντες, η μοντελοποίηση της διόρθωσης-σφάλματος στη βάση της επαναφοράς σε ένα φυσικό πραγματικό ποσοστό σύμφωνα με την επίδραση του Fisher, και σε μια ισορροπία διασποράς απόδοσης αξίζει να εκτιμηθεί, ενώ το απλό dVAR παράγει ακόμα την πιο γερή επίδοση. Σε αντίθεση, το οπτικά εντυπωσιακό χαρακτηριστικό γνώρισμα του περιορισμού όλων των υπό εξέταση μεταβλητών δεν βοηθά στη βελτίωση της πρόβλεψης. Ένα απλό VAR που τείνει να θεωρεί όλες τις μεταβλητές στάσιμες αποτυγχάνει, και ένα περίπλοκο VAR κατωφλιού είναι ανίκανο να πείσει, παρουσιάζοντας μερικά πλεονεκτήματα στους μικρούς ορίζοντες για μερικές περιπτώσεις αλλά αποτυγχάνει για άλλα παραδείγματα και στους μακρύτερους ορίζοντες, όταν η μη γραμμικότητα του μοντέλου απαιτεί υπολογισμό του μέσου όρου πέρα από πιθανολογικές προβλέψεις.

Μια τροποποίηση 'κατωφλιού' του γραμμικού μοντέλου που επιτρέπει τις μεταβλητές να συμπεριφέρονται διαφορετικά στις κατανεμητές ουρές παράγει προφανώς εύλογες πιο μακροπρόθεσμες τροχιές, δεδομένου ότι είναι συνολικά σταθερό και αποφεύγει το πασίγνωστο χαρακτηριστικό γνώρισμα των ενσωματωμένων μεταβλητών ότι η υποστήριξή τους είναι απεριόριστη.

Shrunken interest rate forecasts are better forecasts

Στο άρθρο των Reid Dorsey-Palmateer και Gary Smith, η στατιστική αρχή της παλινδρόμησης στο μέσο προτείνει ότι η ακρίβεια των προβλεπόμενων αλλαγών στα επιτόκια να μπορούσε να βελτιωθεί χρησιμοποιώντας την ιστορική συσχέτιση μεταξύ των προβλεπόμενων και πραγματικών αλλαγών για να συρρικνώσουν τις προβλεπόμενες αλλαγές προς έναν προγενέστερο μέσο του μηδενός. Η εφαρμογή αυτής της ιδέας στις προβλέψεις επιτοκίου από την Έρευνα των επαγγελματιών Forecasters (SPF-Survey of Professional Forecasters) διαπίστωσε μια συνεπή βελτίωση στην ακρίβεια των προβλέψεών τους.

Οι συγγραφείς χρησιμοποιούν τις προβλέψεις επιτοκίου της SPF. Περίπου 35 επαγγελματίες ειδικοί στις προβλέψεις ερευνούνται κάθε τρίμηνο. Οι τριμηνιαίες προβλέψεις είναι διαθέσιμες για αυτά τα τρία επιτόκια:

1. Επιτόκιο κρατικού έντοκου γραμματίου (Treasury Bill Rate), τριών-μηνών, βάση προεξόφλησης (μη προσαρμοσμένα δεδομένα, εκατοστιαία αναλογία, ημερησίως). Αυτές οι προβλέψεις είναι διαθέσιμες από το τρίτο τρίμηνο του 1981.

2. Επιτόκιο κρατικού έντοκου ομολόγου (Treasury Bond Rate), 10-ετών, σταθερή προθεσμία, (μη προσαρμοσμένα δεδομένα, εκατοστιαία αναλογία, ημερησίως). Αυτές οι προβλέψεις είναι διαθέσιμες από το πρώτο τρίμηνο του 1992.

3. Απόδοση εταιρικού ομολόγου (Moody's AAA Corporate Bond Yield), (μη προσαρμοσμένα δεδομένα, εκατοστιαία αναλογία, ημερησίως). Αυτές οι προβλέψεις είναι διαθέσιμες από το τρίτο τρίμηνο του 1981.

Σε κάθε περίπτωση, οι τριμηνιαίες τιμές είναι μέσοι όροι των μηνιαίων στοιχείων και χρησιμοποιούν το μέσο των επαγγελματικών προβλέψεων. Για κάθε επιτόκιο, εξετάζουν την προβλεπόμενη αλλαγή: τη διαφορά μεταξύ της προβλεπόμενης μελλοντικής αξίας και της πραγματικής αξίας κατά τη διάρκεια του τριμήνου όταν γίνεται η πρόβλεψη. Εργάζονται με την προβλεπόμενη αλλαγή στα επιτόκια επειδή το μηδέν είναι ένας φυσικός προγενέστερος μέσος για μια ομάδα προβλεπόμενων αλλαγών που γίνονται κατά τη διάρκεια του χρόνου.

Η ακρίβεια της πρόβλεψης μετριέται με τρεις τρόπους: ο αριθμός των φορών στους οποίους οι προσαρμοσμένες ή όχι προβλέψεις είναι πιο κοντά στις πραγματικές τιμές, το μέσο απόλυτο σφάλμα (MAE) και η ρίζα του μέσου τετραγωνικού σφάλματος (RMSE).

Comparing forecast performance of exchange rate models

Η μετακίνηση της τιμής συναλλάγματος ελέγχεται τακτικά από τις κεντρικές τράπεζες για λόγους μακροοικονομικής ανάλυσης και επιτήρησης της αγοράς. Παρά την πρωτοποριακή μελέτη των Meese και Rogoff (1983), που παρουσιάζει την ανωτερότητα του μοντέλου τυχαίου-περίπατου στην out-of-sample πρόβλεψη της τιμής συναλλάγματος, υπάρχουν κάποια στοιχεία ότι η μετακίνηση τιμής συναλλάγματος μπορεί να είναι προβλέψιμη στους μακρύτερους χρονικούς ορίζοντες. Αυτή η μελέτη των Lillie Lam, Laurence Fung and Ip-wing Yu, συγκρίνει την απόδοση της πρόβλεψης των μοντέλων i)PPP (Purchasing Power Parity model), ii)UIP (Uncovered Interest Rate Parity model), iii)SP (Sticky Price Monetary model), iv)του μοντέλου που βασίζεται στην τεχνική BMA (Bayesian Model Averaging technique) και v)μια σύνθετη προδιαγραφή που ενσωματώνει τα παραπάνω τέσσερα μοντέλα, με κάποια μέτρα σύγκρισης που δίνονται από το μοντέλο του τυχαίου περίπατου και της ιστορικής μέσης απόδοσης.

Τρία νομίσματα εξετάζονται σε αυτό το τμήμα και είναι το ευρώ (EUR), η βρετανική λίβρα (GBP) και το ιαπωνικό γεν (GEN) ενάντια στο αμερικανικό δολάριο (USD). Για τις εκτιμήσεις υιοθετούνται τριμηνιαία δεδομένα του μέσου όρου της περιόδου των ΗΠΑ, της Βρετανίας, της Γερμανίας και της Ιαπωνίας από το Q1 του 1973 έως το Q4 του 2007. Τα στοιχεία των τεσσάρων χωρών που χρησιμοποιούνται για τις εμπειρικές εκτιμήσεις είναι οι εξής δέκα έξι μεταβλητές:

1. τιμή αποθεμάτων
2. αλλαγή στην τιμή αποθεμάτων
3. μακροπρόθεσμο επιτόκιο
4. βραχυπρόθεσμο επιτόκιο
5. χρονική διασπορά
6. τιμή του πετρελαίου
7. αλλαγή στην τιμή του πετρελαίου
8. απόδοση συναλλαγματικής ισοτιμίας της προηγούμενης περιόδου
9. σημάδι της απόδοσης συναλλαγματικής ισοτιμίας της προηγούμενης περιόδου
10. εποχιακά ρυθμισμένο πραγματικό ΑΕΠ
11. αλλαγή στο εποχιακά ρυθμισμένο πραγματικό ΑΕΠ
12. εποχιακά ρυθμισμένη προσφορά χρήματος
13. αλλαγή στην εποχιακά ρυθμισμένη προσφορά χρήματος
14. επίπεδο τιμών καταναλωτή
15. ποσοστό πληθωρισμού
16. αναλογία του τρέχοντος λογαριασμού στο ΑΕΠ

Από την άποψη της προβλεψιμότητας της τιμής συναλλάγματος, τα εμπειρικά αποτελέσματα προτείνουν ότι τα μοντέλα PPP, UIP και SP είναι γενικά ικανά να ξεπεράσουν το μοντέλο του τυχαίου-περίπατου καθώς επίσης και την ιστορική μέση απόδοση για την πρόβλεψη των συναλλαγματικών ισοτιμιών των EUR/USD και YEN/USD, αλλά όχι αυτή του GBP/ USD. Τα αποτελέσματα είναι σύμφωνα με τον Cheung et al (2004) ο οποίος καταλήγει στο συμπέρασμα ότι ένα συγκεκριμένο μοντέλο μπορεί να αποδίδει καλά για μια συναλλαγματική ισοτιμία αλλά όχι για άλλες. Για τη σύγκριση της απόδοσης της πρόβλεψης των πέντε μοντέλων, τα αποτελέσματα προτείνουν ότι κανένα μοντέλο μόνο του σε αυτήν την μελέτη δεν ξεπερνά με συνέπεια τα άλλα. Η ικανότητα πρόβλεψης ενός συγκεκριμένου μοντέλου εξαρτάται από τη συναλλαγματική ισοτιμία του επιτοκίου και τον ορίζοντα πρόβλεψης. Η συνδυασμένη πρόβλεψη είναι γενικά καλύτερη από την πρόβλεψη που βασίζεται σε ένα μόνο μοντέλο όταν χρησιμοποιούνται ως κριτήρια η ρίζα του μέσου τετραγωνικού σφάλματος πρόβλεψης και η κατεύθυνση των στατιστικών αλλαγής. Λαμβάνοντας υπόψη τον περιορισμό των μεμονωμένων μοντέλων, οι προβλέψεις που βασίζονται σε αυτά πρέπει να χρησιμοποιούνται με την προσοχή.

Have increases in Federal Reserve transparency improved private sector interest rate forecasts?

Ο Eric T. Swanson αποδεικνύει ότι οι προβλέψεις ιδιωτικού τομέα των βραχυπρόθεσμων επιτοκίων στις ΗΠΑ έχουν παρουσιάσει θεαματικές βελτιώσεις κατά τη διάρκεια των προηγούμενων 15-20 ετών, όπως αποδεικνύεται από (1) μια μείωση των σφαλμάτων πρόβλεψης του επιτοκίου ομοσπονδιακών κεφαλαίων σε ορίζοντες αρκετών μηνών, (2) μια πτώση στις εκπλήξεις χρηματιστηριακής αγοράς σε απάντηση στις ανακοινώσεις της FOMC (Federal Open Market Committee), (3) μια μείωση της *ex ante* αβεβαιότητας της χρηματιστηριακής αγοράς για τη μελλοντική πορεία των επιτοκίων, που προέρχεται από τις επιλογές επιτοκίου, και (4) μια πτώση στη διασπορά των προβλέψεων ιδιωτικού τομέα των βραχυπρόθεσμων επιτοκίων. Παρά την πρόσφατη βελτίωση στα σφάλματα πρόβλεψης ιδιωτικού τομέα και στην αβεβαιότητα από τον Ιανουάριο του 2001, μια γενική βελτίωση στις προβλέψεις επιτοκίου ιδιωτικού τομέα εμφανίζεται να είναι ένα γερό χαρακτηριστικό γνώρισμα του στοιχείου που παραμένει ακόμα και μετά από τον έλεγχο για τις αλλαγές στην ορμή του επιτοκίου ομοσπονδιακών κεφαλαίων και την αβεβαιότητα για την κατάσταση της αμερικανικής οικονομίας που πραγματοποιήθηκε κατά τη διάρκεια της περιόδου.

Οι ελλοχεύουσες αιτίες της βελτιωμένης επίδοσης πρόβλεψης δεν μπορούν να καθοριστούν με βεβαιότητα, αλλά δύο στοιχεία προτείνουν έντονα ότι αυξήσεις στη διαφάνεια της Federal Reserve έχουν διαδραματίσει έναν ρόλο. Κατ' αρχάς, τα σφάλματα πρόβλεψης της αγοράς και η διατομική διασπορά πρόβλεψης για τα επιτόκια έχουν μειωθεί ουσιαστικά, ενώ εκείνα για το ΑΕΠ και τον πληθωρισμό όχι, δείχνοντας μια βελτίωση στην ικανότητα του ιδιωτικού τομέα να προβλέπει τα επιτόκια πάνω και πέρα από οποιεσδήποτε βελτιώσεις στην πρόβλεψη άλλων μακροοικονομικών μεταβλητών. Δεύτερον, η αβεβαιότητα της αγοράς για τη μελλοντική πορεία των επιτοκίων ακόμη και 6-9 μήνες μπροστά υποχωρεί ουσιαστικά αφότου έχουν γίνει οι ρητές πολιτικές ανακοινώσεις και οι συνοδευτικές αιτιολογικές εκθέσεις από την Fed – που έχουν γίνει από το Φεβρουάριο του 1994– αλλά δεν παρουσιάζει καμία σημαντική απάντηση στις υπονοούμενες, ανεξήγητες πολιτικές ανακοινώσεις που έγιναν από την Fed πριν από εκείνη την ημερομηνία.

Forecasting with latent structure time series models: an application to nominal interest rates

Οι Sridhar Iyer και Rick L. Andrews αναπτύσσουν μια λανθάνουσα επέκταση ενός συχνά χρησιμοποιημένου δομικού μοντέλου χρονοσειρών και χρησιμοποιούν το μοντέλο ως βάση για προβλέψεις. Κάθε απαρατήρητο καθεστώς έχει τη δική του μοναδική κλίση και διακυμάνσεις για να περιγράψει τη διαδικασία που παράγει τα δεδομένα, και κάθε δεδομένη χρονική περίοδο το μοντέλο προβλέπει a priori ποιο καθεστώς χαρακτηρίζει καλύτερα τα δεδομένα. Αυτό επιτυγχάνεται με τη χρησιμοποίηση ενός πολυωνυμικού logit μοντέλου στο οποίο η αρχική επεξηγηματική μεταβλητή είναι ένα μέτρο για το πόσο συνεπές είναι κάθε καθεστώς τις πρόσφατες παρατηρήσεις. Το μοντέλο εξυπηρετεί ιδιαίτερα καλά την πρόβλεψη των σειρών που υπόκεινται σε συχνούς ή/και σοβαρούς κλονισμούς.

Τα δεδομένα που χρησιμοποιούνται στην ανάλυσή μας είναι το ονομαστικό επιτόκιο σε ένα τριμηνιαίο αμερικανικό κρατικό έντοκο γραμμάτιο (Treasury bill) από τον Απρίλιο του 1951 μέχρι τον Φεβρουάριο του 1991, συνολικά 479 παρατηρήσεις

Μια εφαρμογή στα ονομαστικά επιτόκια δείχνει ότι η συμπεριφορά του επιτοκίου του τριμηνιαίου αμερικανικού κρατικού έντοκου γραμμάτιου εξηγείται επαρκώς από το μοντέλο των τριών καθεστώτων. Η ακρίβεια πρόβλεψης είναι ανώτερη από αυτήν που παράγεται από ένα παραδοσιακό single-regime μοντέλο και ένα standard ARIMA μοντέλο με ένα υπό όρους ετεροσκεδαστικό σφάλμα.

Forecasting the FOMC'S interest rate setting behavior: a further analysis

Σκοπός των Hyeongwoo Kim, John Jackson και Richard Saba είναι να εξετάσουν τα δύο πιθανά μειονεκτήματα της μελέτης για την Ομοσπονδιακή Επιτροπή Ελεύθερης Αγοράς (Federal Open Market Committee) από τους Hu και Phillips (2004a) και με αυτό τον τρόπο να παρουσιάσουν μια εκσυγχρονισμένη αξιολόγηση της συμπεριφοράς της FOMC για τον καθορισμό του επιτοκίου: το αντικείμενό τους είναι να αναπτύξουν ένα μοντέλο που μπορεί να χρησιμεύσει ως ένας καλύτερος ειδικός στις προβλέψεις των μελλοντικών αλλαγών επιτοκίου.

Διατυπώνουν έπειτα πέντε εναλλακτικά μοντέλα της συμπεριφοράς της FOMC για τον καθορισμό του επιτοκίου: τα δύο καθοδηγούνται από δεδομένα (δηλ., προέρχονται από μια αναζήτηση προδιαγραφών των πολυάριθμων πιθανών μέτρων), δύο βασίζονται στη θεωρία (δηλ., εναλλακτικές εκδόσεις μιας προδιαγραφής κανόνα Taylor), και το πέμπτο προέρχεται από μια αποκαλούμενη πλαστή παλινδρόμηση κάτω από μια συνεχή χρονική ρύθμιση.

Διαπιστώνουν ότι εάν το κριτήριο επιλογής μοντέλου είναι αυστηρά εμπειρικό, η διόρθωση της στατικότητας είναι εξαιρετικά σημαντική. Επίσης διαπιστώνουν ότι τα συνηθισμένα, εμπειρικά βασισμένα, in-sample κριτήρια επιλογής μοντέλου και οι σημαντικές μεταβλητές δεν είναι ιδιαίτερα κατάλληλα για την επιλογή ενός καλύτερου μοντέλου για να προβλέψει τη συμπεριφορά καθορισμού του επιτοκίου της FOMC. Όταν αξιολογούν αυτά τα μοντέλα από την άποψη της extra-sample ικανότητας πρόβλεψης, διαπιστώνουν ότι εκείνα που ευνοούνται από τα in-sample κριτήρια αποδίδουν χειρότερα, και ότι τα μοντέλα που έχουν κάποιο θεωρητικό υπόβαθρο αποδίδουν καλύτερα, πιθανώς επειδή εμφανίζουν κάποια γενικότητα στην εφαρμογή πέρα από το ιδιαίτερο δείγμα εκτίμησης που χρησιμοποιείται. Στο τέλος, βρίσκουν ότι το καλύτερο μοντέλο για την πρόβλεψη της συμπεριφοράς της FOMC για τον καθορισμό του επιτοκίου είναι ένα φειδωλά διευκρινισμένο, προνοητικό μοντέλο του κανόνα Taylor.

Forecasting interest rate swap spreads using domestic and international risk factors: evidence from linear and non-linear models

Οι Ilias Lekkos, Costas Milas και Theodore Panagiotidis ερευνούν την ικανότητα των μοντέλων συντελεστή να προβλέπουν τη δυναμική των αμερικανικών και

βρετανικών εκτάσεων ανταλλαγής επιτοκίων (interest rate swap spreads³) μέσα σε ένα γραμμικό και μη γραμμικό πλαίσιο. Απορρίπτουν τη γραμμικότητα για τα αμερικανικά και βρετανικά swap spreads υπέρ ενός STVAR μοντέλου εναλλαγής καθεστώτων, όπου η εναλλαγή μεταξύ των καθεστώτων ελέγχεται από την κλίση της αμερικανικής δομής όρου των επιτοκίων. Συγκρίνουν τη ικανότητα του μοντέλου STVAR να προβλέπει swap spreads με αυτή ενός μη γραμμικού μοντέλου του πλησιέστερου-γείτονα (Nearest Neighbour) καθώς επίσης και με αυτή των γραμμικών μοντέλων AR και VAR.

Υιοθετούν μια διεθνή ρύθμιση, όπου οι δομές όρου των αμερικανικών και βρετανικών εκτάσεων ανταλλαγής επιτοκίων εξαρτώνται από τις εκτάσεις των εταιρικών ομολόγων των δύο χωρών, τα διαφορικά επιτοκίου μεταξύ των αμερικανικών και βρετανικών κυβερνητικών ομολόγων και των κλίσεων (10-ετών επιτόκιο πλην 3-μηνών T-Bill επιτόκιο) των δομών όρου των αποδόσεων ομολόγων μηδενικού τοκομερίδιου (zero-coupon bond yields) των δύο χωρών. Το σύνολο των δεδομένων αποτελείται από εβδομαδιαίες παρατηρήσεις από τον Ιούνιο του 1991 μέχρι τον Ιούνιο του 2001.

Διαπιστώνουν ότι στους μικρούς ορίζοντες πρόβλεψης η ευελιξία του μοντέλου του πλησιέστερου-γείτονα (NN) παρέχει ένα πλεονέκτημα έναντι όλων των μοντέλων παράγοντα STVAR και VAR και την προδιαγραφή AR. Το πλεονέκτημα πρόβλεψης του NN έναντι του μοντέλου STVAR είναι πιο εμφανές για αμερικανικά swap spreads σε περιόδους κατάστασης αυξανόμενου κινδύνου και για βρετανικά swap spreads σε περιόδους κατάστασης μειωμένου κινδύνου. Αφ' ετέρου, η επίδοση των μοντέλων NN παρακμάζει μάλλον γρήγορα με τον ορίζοντα πρόβλεψης. Για προβλέψεις πιο μακρινού ορίζοντα η προδιαγραφή STVAR αυξάνει τη ικανότητα πρόβλεψής της έναντι των υπόλοιπων γραμμικών και μη γραμμικών μοντέλων, ενώ το μοντέλο NN κατατάσσεται στο τέλος.

Combination forecasting for directional accuracy: an application to survey interest rate forecasts

Χρησιμοποιώντας δημοσιευμένες προβλέψεις επιτοκίων που εκδίδονται από επαγγελματίες οικονομολόγους, ο Mark R. Greer κατασκευάζει δύο προβλέψεις συνδυασμού με σκοπό να βελτιωθεί την κατευθυντήρια ακρίβεια της πρόβλεψης επιτοκίου. Η πρώτη πρόβλεψη συνδυασμού παίρνει έναν σταθμισμένο μέσο όρο των

³ Το Swap spread είναι η διαφορά μεταξύ του διαπραγματευόμενου και σταθερού επιτοκίου μιας ανταλλαγής. Η έκταση καθορίζεται από χαρακτηριστικά της προσφοράς της αγοράς και την φερεγγυότητα των πιστωτών.

ατομικών προβλέψεων των forecasters. Όσο πιο επιτυχημένος ήταν ο ειδικός στις προβλέψεις στις προηγούμενες προβλέψεις για την κατεύθυνση της αλλαγής στα επιτόκια, τόσο μεγαλύτερο είναι το βάρος που δίνεται στην τρέχουσα πρόβλεψή του/της. Η δεύτερη πρόβλεψη συνδυασμού είναι απλά η πρόβλεψη που εκδίδεται από τον ειδικό στις προβλέψεις που είχε το μέγιστο ποσοστό επιτυχίας στην πρόβλεψη της κατεύθυνσης της αλλαγής στα επιτόκια στις προηγούμενες προβλέψεις. Σε περιπτώσεις όπου δύο ή περισσότεροι ειδικοί στις προβλέψεις θέλουν το καλύτερο ιστορικό κατευθυντήριο αρχείο διαδρομής ακρίβειας, χρησιμοποιείται ο αριθμητικός μέσος αυτών των ειδικών στις προβλέψεις.

Στην αρχή του Ιανουαρίου και πάλι στην αρχή του Ιουλίου κάθε έτους μεταξύ 1984 και 1998, η Wall Street Journal δημοσίευε προβλέψεις ενός έτους για τα βραχυπρόθεσμα και μακροπρόθεσμα επιτόκια στις ΗΠΑ. Τα βραχυπρόθεσμα επιτόκια αξιολογήθηκαν από την απόδοση στο τριμηνιαίο αμερικανικό κρατικό έντοκο γραμμάτιο (Treasury bill), ενώ τα μακροχρόνια επιτόκια αξιολογήθηκαν από την απόδοση στο τριάντα-ετών αμερικανικό κρατικό ομόλογο (Treasury bond). Σε κάθε έρευνα, η δημοσίευση ανέφερε την πρόβλεψη κάθε forecaster για κάθε επιτόκιο ένα έτος στο μέλλον. Μια επιτροπή επαγγελματιών οικονομολόγων που ερευνήθηκε από το περιοδικό εξέδωσε αυτές τις προβλέψεις. Η επιτροπή ήταν στην αρχή αρκετά μικρή, με μόνο 24 οικονομολόγους που συμμετέχουν στην έρευνα του Ιανουαρίου 1984. Εντούτοις, ο αριθμός των συμμετεχόντων αυξήθηκε σταθερά κατά τη διάρκεια του χρόνου, και μέχρι τον Ιούλιο του 1998, 55 ειδικοί στις προβλέψεις συμμετείχαν στην έρευνα. Συνολικά 1272 ατομικές προβλέψεις για κάθε ένα από τα δύο επιτόκια συμπεριλαμβάνονται στη βάση δεδομένων που χρησιμοποιείται εδώ.

Η μελέτη αυτή διαπιστώνει ότι καμία τέτοια μέθοδος συνδυασμού δεν εκτελεί καλύτερα από την ρίψη του νομίσματος στην πρόβλεψη της κατεύθυνσης της αλλαγής στα επιτόκια. Επίσης καμία μέθοδος δεν ξεπερνά τον απλό αριθμητικό μέσο των προβλέψεων όλων των ειδικών στις προβλέψεις που ερευνούνται στην πρόβλεψη της κατεύθυνσης της αλλαγής στα επιτόκια.

Can forward rates be used to improve interest rate forecasts?

Οι Emilio Domínguez και Alfonso Novales αξιολογούν το βαθμό στον οποίο η επεξηγηματική δύναμη που ανιχνεύεται στη δομή όρου σε διαφορετικές αγορές και χώρες μπορεί πραγματικά να χρησιμοποιηθεί για να παράγει λογικές προβλέψεις των μελλοντικών βραχυπρόθεσμων επιτοκίων. Συγκεκριμένα, αυτή είναι η πρώτη συστηματική προσπάθεια να μετρηθεί η πραγματική προβλεπτική δύναμη της δομής

όρου κάτω από τους περιορισμούς της *Υπόθεσης Προσδοκιών* στα διεθνή στοιχεία. Η *Υπόθεση Προσδοκιών* στο σχηματισμό της δομής όρου των επιτοκίων οδηγεί σε παλινδρομήσεις που εμφανίζονται να εξηγούν τα μελλοντικά βραχυπρόθεσμα επιτόκια ως συναρτήσεις των τρεχόντων μπροστινών επιτοκίων. Συνήθως, μια καλή τακτοποίηση αυτών των παλινδρομήσεων έχει ερμηνευθεί ως το μπροστινό επιτόκιο που έχει σημαντική προβλεπτική δύναμη για τα μελλοντικά βραχυπρόθεσμα επιτόκια, το οποίο θεωρείται ότι είναι η ουσία της *Υπόθεσης Προσδοκιών*. Στην πραγματικότητα, μια μηδενική τομή μαζί με μια μοναδιαία κλίση σε εκείνη την πρόβλεψη των μελλοντικών επιτοκίων στα τρέχοντα μπροστινά επιτόκια είναι γνωστή ως το μπροστινό επιτόκιο που είναι *αμερόληπτος προγνώστης* των μελλοντικών επιτοκίων.

Χρησιμοποιούμε μηνιαία δεδομένα για το διάστημα 1978-1998 στα επιτόκια σε Eurodeposits στο αμερικανικό δολάριο, το γεν, τα Γερμανικά Μάρκα, τη βρετανική λίβρα, την ισπανική πεσέτα, το γαλλικό φράγκο, την ιταλική λιρέτα και το ελβετικό φράγκο, συγκρίνοντας τις προβλέψεις που λαμβάνονται από τα μπροστινά επιτόκια με εκείνες που λαμβάνονται από τις μονομεταβλητές αυτοπαλινδρομήσεις.

Καταλήγουν στο συμπέρασμα ότι μόνα τους τα μπροστινά επιτόκια, παράγουν καλύτερες one-step ahead προβλέψεις, καθώς επίσης και καλύτερες once-and-for all προβλέψεις των ενός μηνός επιτοκίων σε έναν πλήρη ορίζοντα έτους από εκείνες που λαμβάνονται από το παρελθόν των επιτοκίων. Το κέρδος στην one-step ahead πρόβλεψη εξαφανίζεται για τις πιο μακροχρόνιες ωριμότητες, αν και τα μπροστινά επιτόκια ακόμα παράγουν καλύτερες once-and-for all προβλέψεις σε τρία και έξι μηνών επιτόκια από ότι οι μονομεταβλητές αυτοπαλινδρομήσεις για διάφορα νομίσματα.

A Survey on Interest Rate Forecasting

Αυτή η έρευνα των Yvon Fauvel, Alain Paquet και Christian Zimmermann προσπαθεί να καλύψει τις σημαντικότερες μεθόδους που χρησιμοποιούνται για να προβλέψουμε τα επιτόκια. Τα θεωρητικά θεμέλια παρουσιάζονται και συζητιούνται στην προοπτική της ακρίβειας της πρόβλεψης χρησιμοποιώντας αποτελέσματα που δημοσιεύονται στη λογοτεχνία. Παρέχεται επίσης μια βασική αναθεώρηση της μοντελοποίησης επιτοκίου. Επειδή λόγω του όγκου της βιβλιογραφίας, είναι δύσκολο να καλυφθούν όλα όσα έχουν γραφτεί για ένα θέμα, γι αυτό εδώ εστιάζουν στην πρόβλεψη επιτοκίου, κατευθύνουν την προσοχή τους στη λογοτεχνία που έχει δημοσιευθεί τα τελευταία δέκα χρόνια, και τέλος η έρευνά τους προέρχεται κατά ένα μεγάλο μέρος από την ακαδημαϊκή λογοτεχνία.

Αρχικά παρέχουν μια επισκόπηση των διάφορων υπαρχουσών προσεγγίσεων στην πρόβλεψη, με ιδιαίτερη προσοχή στα επιτόκια. Αυτές οι μέθοδοι είναι οι εξής:

- μονομεταβλητές μέθοδοι όπως τα μοντέλα ARIMA
- προσεγγίσεις μονής εξίσωσης (Single Equation Approach)
- δομικά συστήματα ταυτόχρονων εξισώσεων (STRM)
- πολυμεταβλητά συστήματα δυναμικών εξισώσεων: διανυσματικά αυτοπαλινδρομούμενα μοντέλα (VAR) και διανυσματικά μοντέλα διόρθωσης σφάλματος (VECM)
- μη γραμμικά νευρικά δίκτυα (NNM)
- μοντέλα εναλλαγής καθεστώτων (Regime-switching Models)
- ομόφωνες προβλέψεις και συγκέντρωση

Βάσει της εξέτασής τους αυτής, φαίνεται ότι προτιμητέο να εστιάσουμε στις μονομεταβλητές και πολυμεταβλητές προσεγγίσεις. Παρά την προφανή απλότητά τους, τα μονομεταβλητά μοντέλα τείνουν να αποδίδουν αρκετά καλά στην πράξη σε θέματα πρόβλεψης. Ομοίως, η φυσική επέκτασή τους σε μια πολυμεταβλητή ρύθμιση (δηλ. VAR και VECM) αποτελεί μια ενδιαφέρουσα προσέγγιση για μια ενσωματωμένη επεξεργασία διάφορων επιτοκίων, συμπεριλαμβανομένης της βραχυπρόθεσμης δυναμικής τους και οποιωνδήποτε υπαρχουσών μακροπρόθεσμων σχέσεων. Η προγνωστική τους απόδοση είναι επίσης συνήθως καλή. Όταν οι μακροπρόθεσμες σχέσεις ισορροπίας υπάρχουν μεταξύ των μεταβλητών του συστήματος, το VECM μπορεί να είναι ανώτερο από το VAR στην πρόβλεψη σε μακρύτερους ορίζοντες. Επίσης, τα μη γραμμικά μοντέλα των νευρικών δικτύων ή των προσεγγίσεων εναλλαγής καθεστώτων μπορεί να είναι ελπιδοφόρα για τη μοντελοποίηση των σειρών που εκθέτουν ικανοποιητική μη γραμμικότητα. Όσον αφορά την απόδοσή τους για την πρόβλεψη των επιτοκίων, εντούτοις, η αρχική άποψη από αυτήν την νέα λογοτεχνία δεν φαίνεται να αιτιολογεί τη σχετικά σύνθετη επεξεργασία που πρέπει να υπόκειται στις κατάλληλες εκτιμήσεις και προβλέψεις με αυτές τις μεθόδους. Δεν πιστεύουν ότι αυτές οι προσεγγίσεις θα παρέδιδαν εύκολα χρησιμοποιήσιμα ή ανώτερα αποτελέσματα από αυτά άλλων προσεγγίσεων.

ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ

3.1 ΕΥΦΥΗ ΣΥΣΤΗΜΑΤΑ

Τις τελευταίες δεκαετίες οι έρευνες πολλών επιστημόνων από διάφορους τομείς έχουν εστιαστεί στην κατανόηση και τον σχεδιασμό μίας κατηγορίας συστημάτων που αποκαλούνται ‘ευφυή συστήματα’. Αυτά τα συστήματα λειτουργούν με τρόπο που προϋποθέτει ότι έχουν κάποια μορφή ευφυΐας. Βέβαια επειδή ακριβής ορισμός της ευφυΐας δεν υπάρχει σε αυτήν την κατηγορία ανήκουν πολλά συστήματα πολύ διαφορετικά όσον αφορά την αρχή λειτουργίας τους και με πολύ διαφορετικές δυνατότητες.

Ένας ορισμός που θα μπορούσε να δοθεί στα ευφυή συστήματα είναι ο ακόλουθος: είναι συστήματα που έχουν την ικανότητα να θυμούνται αλλά και να χρησιμοποιούν αυτήν την μνήμη ώστε να παίρνουν αποφάσεις για το μέλλον με την προοπτική να ικανοποιήσουν κάποιους στόχους.

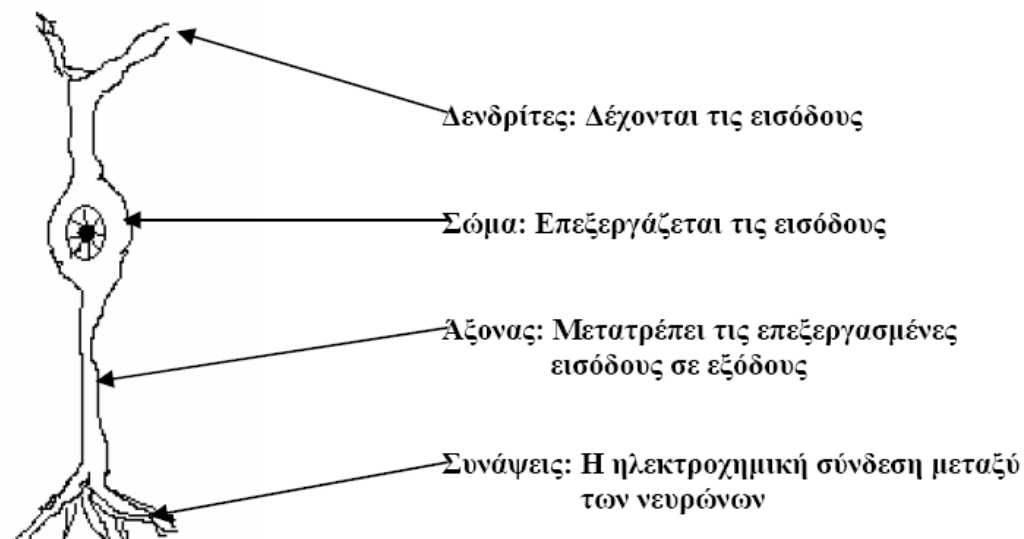
Με τρεις κατηγορίες ευφυών συστημάτων θα ασχοληθούμε βασικά σε αυτήν την εργασία: Τα ασαφή συστήματα, τα τεχνητά νευρωνικά δίκτυα και τα νευρο-ασαφή συστήματα. Και οι τρεις αυτές κατηγορίες συστημάτων εμπίπτουν στον ορισμό που δώσαμε για τα ευφυή συστήματα. Η κάθε μία όμως από αυτές έχει τα δικά της πλεονεκτήματα και μειονεκτήματα. Τα ασαφή συστήματα έχουν μεγάλες ικανότητες εξαγωγής συμπερασμάτων από κανόνες κατανοητούς στους ανθρώπους αλλά έχουν και μεγάλες δυσκολίες όσον αφορά την μάθηση τους. Τα νευρωνικά δίκτυα αντιθέτως έχουν το πλεονέκτημα ότι μαθαίνουν εύκολα και γρήγορα αλλά ο τρόπος που βγάζουν τα συμπεράσματα είναι ακατανόητος. Τα νευροασαφή συστήματα τέλος είναι μία επιτυχημένη προσπάθεια να γεφυρωθούν οι δύο διαφορετικές προσεγγίσεις και να προκύψουν συστήματα που να κρατούν τα πλεονεκτήματα μόνο της κάθε προσέγγισης.

3.2 ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ

3.2.1. Εισαγωγή

Τα τεχνητά νευρωνικά δίκτυα ανήκουν στον τομέα της Μη Συμβολικής Τεχνητής Νοημοσύνης (non symbolic AI), η οποία προσομοιώνει βιολογικές διαδικασίες, όπως τη λειτουργία του εγκεφάλου ή τη διαδικασία εξέλιξης των ειδών, και η οποία διαφέρει από τη Συμβολική Τεχνητή Νοημοσύνη (symbolic AI), που προσομοιώνει τον τρόπο σκέψης του ανθρώπου, χρησιμοποιώντας ως δομικές μονάδες τα σύμβολα (ένα σύμβολο μπορεί να αναπαριστά μια έννοια ή μια σχέση ανάμεσα σε έννοιες).

Η ικανότητα του ανθρώπου να σκέφτεται, να θυμάται και να επιλύει προβλήματα εντοπίζεται στον εγκέφαλό του. Όπως είναι γνωστό από τη Βιολογία, η δομική μονάδα του εγκεφάλου είναι ο νευρώνας. Ένας τυπικός βιολογικός νευρώνας αποτελείται από το σώμα, που συνιστά τον πυρήνα του, τους δενδρίτες, μέσω των οποίων λαμβάνει σήματα από γειτονικούς νευρώνες (σημεία εισόδου), και τον άξονα, που αποτελεί την έξοδο του νευρώνα και το μέσο σύνδεσής του με τους άλλους νευρώνες (βλέπε σχήμα 3.1). Σε κάθε δενδρίτη υπάρχει ένα απειροελάχιστο κενό που ονομάζεται σύναψη. Οι συνάψεις μέσω χημικών διαδικασιών επιταχύνουν ή επιβραδύνουν τη ροή ηλεκτρικών φορτίων προς το σώμα του νευρώνα. Η ικανότητα μάθησης και μνήμης που εμφανίζει ο εγκέφαλος οφείλεται στην ικανότητα των συνάψεων να μεταβάλλουν την αγωγιμότητά τους. Τα ηλεκτρικά σήματα που εισέρχονται στο σώμα μέσω των δενδριτών συνδυάζονται και, εφόσον το αποτέλεσμα ξεπερνά κάποια τιμή κατωφλίου, το σήμα διαδίδεται με τη βοήθεια του άξονα προς άλλους νευρώνες.



Σχήμα 3.1: Απεικόνιση απλού νευρώνα

Ο εγκέφαλος είναι σε θέση να λαμβάνει πολύπλοκες αποφάσεις εκπληκτικά γρήγορα παρά το ότι ο χρόνος απόκρισης των βιολογικών νευρώνων είναι της τάξης των χιλιοστών του δευτερολέπτου. Κατά μια άποψη, αυτό οφείλεται στο ότι η υπολογιστική ικανότητα του εγκεφάλου και η πληροφορία που περιέχει είναι διαμερισμένα σε όλο του τον όγκο. Πρόκειται, δηλαδή, για ένα παράλληλο και κατανεμημένο υπολογιστικό σύστημα. Αυτά τα χαρακτηριστικά διαμορφώνουν το κυριότερο κίνητρο πίσω από την επιθυμία να μοντελοποιηθεί ο ανθρώπινος εγκέφαλος με τα τεχνητά νευρωνικά δίκτυα.

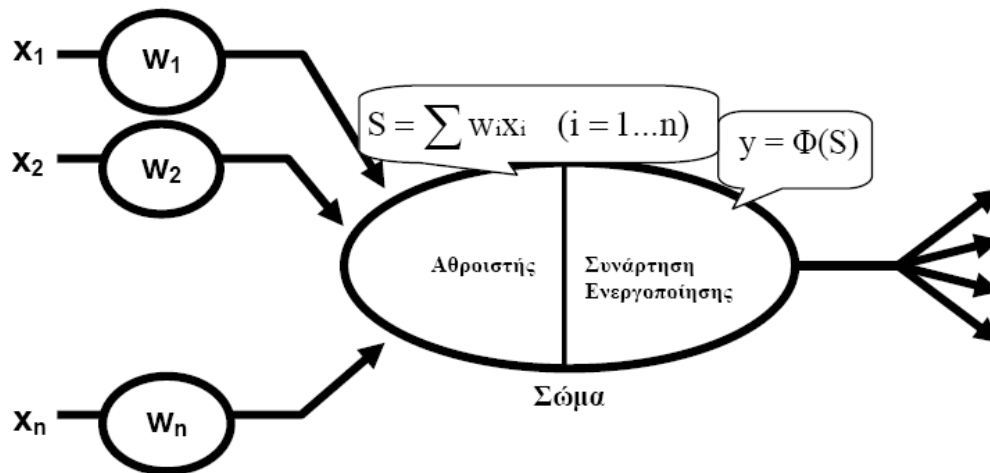
3.2.2. Τεχνητά Νευρωνικά Δίκτυα (ΤΝΔ)

Μοντέλο Τεχνητού Νευρώνα:

Ο τεχνητός νευρώνας (artificial neuron) αποτελεί το πρωταρχικό συστατικό στοιχείο των τεχνητών νευρωνικών δικτύων και είναι ένα υπολογιστικό μοντέλο τα μέρη του οποίου αντιστοιχίζονται άμεσα με αυτά του βιολογικού νευρώνα. Όπως απεικονίζεται στο σχήμα 3.2, ένας τεχνητός νευρώνας δέχεται κάποια σήματα εισόδου $x_0, x_1, \dots, x_n$, τα οποία, σε αντίθεση με τους ηλεκτρικούς παλμούς του εγκεφάλου, αντιστοιχούν σε συνεχείς μεταβλητές. Κάθε τέτοιο σήμα εισόδου σταθμίζεται με κάποιο βάρος w_i (weight), ο ρόλος του οποίου είναι αντίστοιχος με εκείνον της σύναψης του βιολογικού εγκεφάλου.

Η τιμή βάρους μπορεί να είναι θετική ή αρνητική σε αντιστοιχία με την επιταχυντική ή επιβραδυντική λειτουργία της σύναψης. Το σώμα του τεχνητού νευρώνα χωρίζεται σε δύο μέρη:

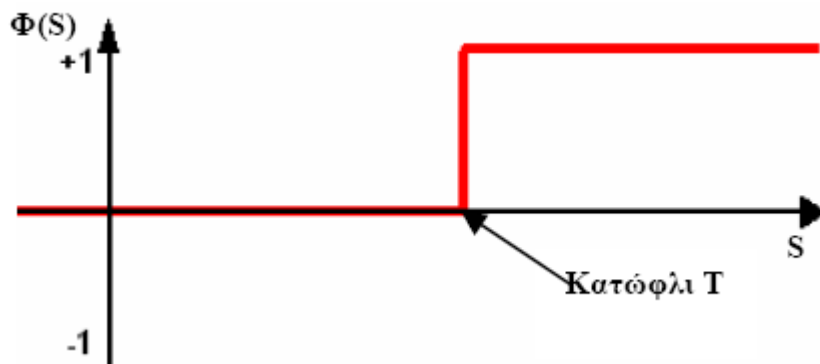
1. τον αθροιστή (sum), ο οποίος προσθέτει τα επηρεασμένα από τα βάρη σήματα εισόδου και παράγει την ποσότητα $S = \sum w_i x_i, i = 1 \dots n$ και
2. τη συνάρτηση ενεργοποίησης ή κατωφλίου (activation ή threshold ή transformation function), που αποτελεί ένα μη γραμμικό φίλτρο το οποίο διαμορφώνει την τελική τιμή του σήματος εξόδου y , σε συνάρτηση με την ποσότητα S . Το παρακάτω σχήμα παρουσιάζει το μοντέλο τεχνητού νευρώνα.



Σχήμα 3.2: Μοντέλο Τεχνητού Νευρώνα

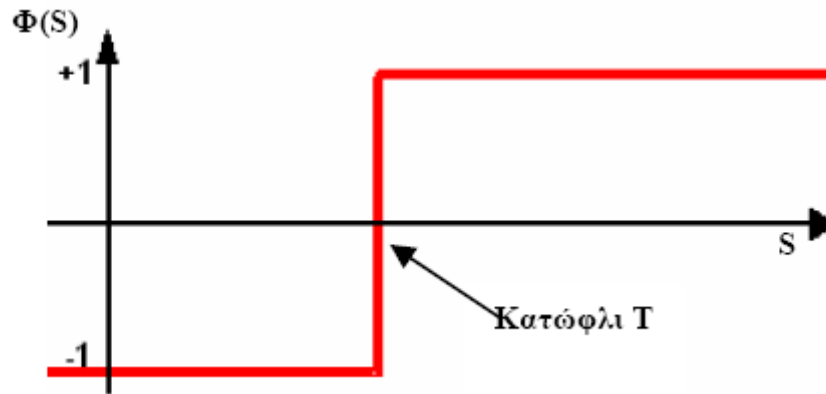
Στα παρακάτω σχήματα παρουσιάζονται τρεις τυπικές μορφές που μπορεί να λάβει η συνάρτηση ενεργοποίησης

- Η βηματική (step) συνάρτηση (βλέπε σχήμα 3.3), η οποία δίνει στην έξοδο αποτέλεσμα (συνήθως 1), εφόσον η τιμή που υπολογίζει ο αθροιστής είναι μικρότερη (ή μεγαλύτερη) από μια τιμή κατώφλιου T .



Σχήμα 3.3: Βηματική Συνάρτηση

- Η συνάρτηση πρόσημου (sign) (βλέπε σχήμα 3.4), η οποία δίνει στην έξοδο αρνητική (ή θετική) πληροφορία, εφόσον η τιμή που υπολογίζει ο αθροιστής είναι μικρότερη (ή μεγαλύτερη) από μια τιμή κατωφλίου T .

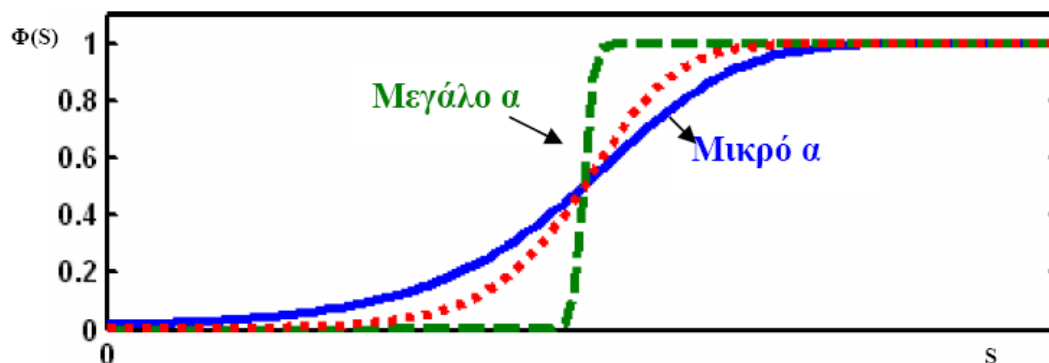


Σχήμα 3.4: Συνάρτηση Προσήμου

- Η σιγμοειδής (sigmoid) συνάρτηση (βλέπε σχήμα 3.5), η οποία εκφράζεται από τη γενική σχέση:

$$\Phi(S) = \frac{1}{1 + e^{-aS}}$$

όπου a είναι ένας συντελεστής ρύθμισης της ταχύτητας μετάβασης μεταξύ των δύο ασύμπτωτων τιμών. Η σιγμοειδής συνάρτηση είναι σημαντική γιατί παρέχει μη γραμμικότητα στο νευρώνα, η οποία είναι απαραίτητη για τη μοντελοποίηση μη γραμμικών φαινομένων.

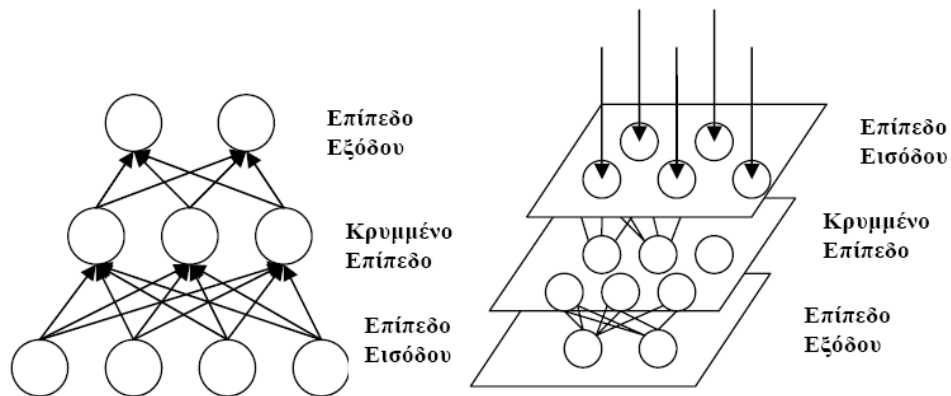


Σχήμα 3.5: Σιγμοειδής Συνάρτηση για διαφορετικές τιμές του συντελεστή a

3.2.3. Τύποι Τεχνητών Νευρωνικών Δικτύων

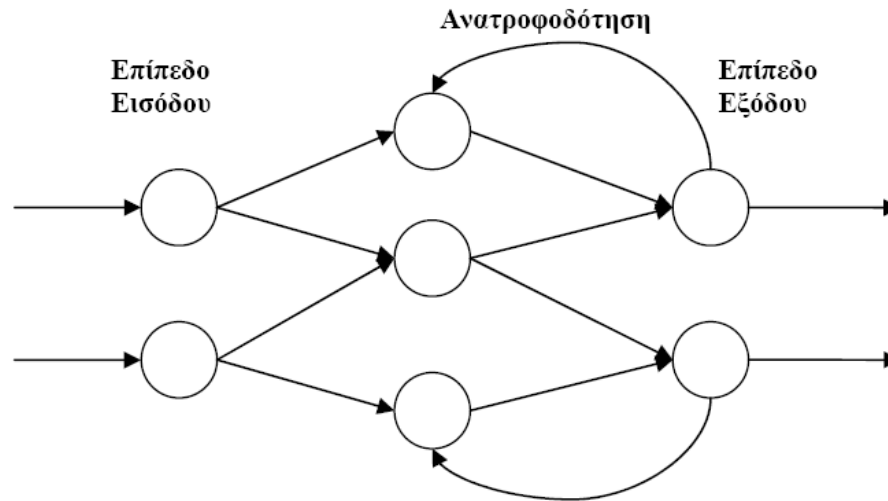
Τα τεχνητά νευρωνικά δίκτυα (artificial neural networks) ή πιο απλά ΤΝΔ, χαρακτηρίζονται ως συστήματα επεξεργασίας δεδομένων που αποτελούνται από ένα πλήθος τεχνητών νευρώνων οργανωμένων σε δομές παρόμοιες με αυτές του ανθρώπινου εγκεφάλου. Συνήθως οι τεχνητοί νευρώνες είναι οργανωμένοι σε μία σειρά από στρώματα ή επίπεδα (layers). Το πρώτο από αυτά τα επίπεδα ονομάζεται επίπεδο εισόδου (input layer) και χρησιμοποιείται για την εισαγωγή των δεδομένων. Τα στοιχεία, που συνιστούν το επίπεδο εισόδου, δεν είναι ουσιαστικά νευρώνες, καθώς δεν εκτελούν κάποιον υπολογισμό (δεν έχουν ούτε βάρη εισόδου, ούτε συναρτήσεις ενεργοποίησης). Στη συνέχεια μπορούν να υπάρχουν, προαιρετικά, ένα ή περισσότερα ενδιάμεσα ή κρυφά επίπεδα (hidden layers). Τέλος ακολουθεί ένα επίπεδο εξόδου (output layer).

Οι νευρώνες των διάφορων στρωμάτων μπορεί να είναι πλήρως ή μερικώς συνδεδεμένοι. Πλήρως συνδεδεμένοι (fully connected) είναι εκείνοι οι οποίοι συνδέονται με όλους τους νευρώνες του επόμενου επιπέδου. Σε κάθε άλλη περίπτωση οι νευρώνες είναι μερικώς συνδεδεμένοι (partially connected). Όταν δεν υπάρχουν συνδέσεις μεταξύ νευρώνων ενός επιπέδου και νευρώνων προηγούμενου επιπέδου (όταν δηλαδή η ροή πληροφορίας είναι μιας κατεύθυνσης) τα ΤΝΔ χαρακτηρίζονται ως δίκτυα με απλή (ή πρόσθια) τροφοδότηση (feedforward). Στην αντίθετη περίπτωση, καθώς και στην περίπτωση συνδέσεων μεταξύ νευρώνων του ίδιου επιπέδου, τα ΤΝΔ χαρακτηρίζονται ως δίκτυα με ανατροφοδότηση (feedback ή recurrent). Ο τύπος του δικτύου με ανατροφοδότηση διαφέρει από τον τύπο της απλής τροφοδότησης στο ότι περιλαμβάνει ένα βρόχο ανάδρασης, όπου κάθε νευρώνας τροφοδοτεί το σήμα της εξόδου του στις εισόδους όλων των άλλων νευρώνων.



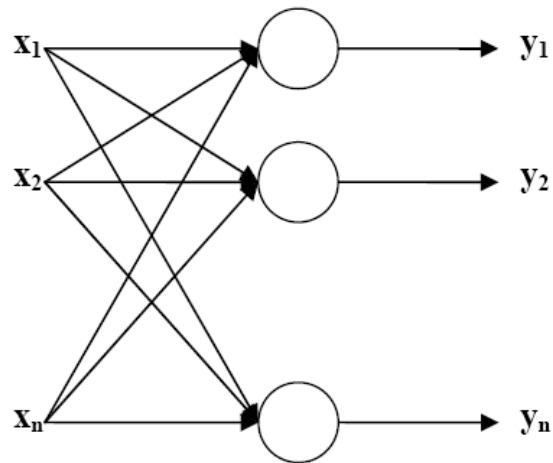
Σχήμα 3.6: Παραδείγματα πλήρως διασυνδεδεμένων ΤΝΔ απλής τροφοδότησης

Τα ΤΝΔ του σχήματος 3.6 αποτελούν παραδείγματα πλήρως διασυνδεδεμένων ΤΝΔ απλής τροφοδότησης, ενώ στο σχήμα 3.7 παραθέτουμε ένα παράδειγμα δικτύου με ανατροφοδότηση:

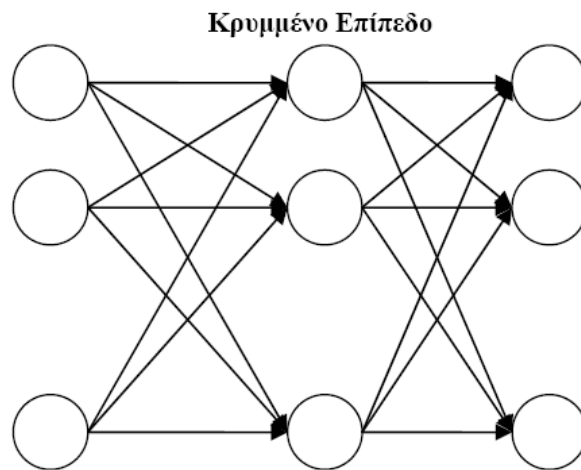


Σχήμα 3.7: Παράδειγμα δικτύου με ανατροφοδότηση

Έχουν αναπτυχθεί συγκεκριμένα ΤΝΔ, πάνω στα πρότυπα ορισμένων βιολογικών λειτουργιών που είναι γνωστά ως paradigms στη διεθνή βιβλιογραφία, και αποτελούν ουσιαστικά τους διαφορετικούς τύπους ΤΝΔ. Τα ΤΝΔ στις περισσότερες περιπτώσεις είναι δομημένα από στρώματα νευρώνων. Κάποια από αυτά έχουν ένα επίπεδο (SLNN=single-layer neural networks) (βλέπε σχήμα 3.8) και κάποια έχουν περισσότερα (MLNN=multi-layer neural networks) (βλέπε σχήμα 3.9).



Σχήμα 3.8: Απεικόνιση Νευρωνικών Δικτύων με ένα επίπεδο



Σχήμα 3.9: Απεικόνιση Νευρωνικών Δικτύων με πολλά επίπεδα

Τα πρώτα χρησιμοποιούν ως κανόνα μάθησης κυρίως τον κανόνα Hebb (Hebb rule) και τον κανόνα Δέλτα (Delta rule). Όσον αφορά τα MLNN οι Rumelhart, Hinton, και Williams (1986) εισήγαγαν τον αλγόριθμο εκμάθησης με ανάστροφη διάδοση του σφάλματος (back-propagation training algorithm). Επιπλέον τα MLNN χρησιμοποιούν σε ορισμένες περιπτώσεις την ανταγωνιστική μάθηση (competitive learning), για την οποία υπάρχουν τέσσερα κυρίως υποδείγματα (Wasserman 1989; Freeman και Skapura 1992; Rumelhart και McClelland, 1986).

Παρακάτω παρουσιάζουμε τους πιο γνωστούς τύπους ΤΝΔ, και κάποιους από τους αλγόριθμους εκπαίδευσής τους.

❖ Δίκτυα με Απλή Τροφοδότηση

Τα ΤΝΔ με απλή τροφοδότηση (feedforward) συνιστούν την πιο απλή μορφή νευρωνικών δικτύων και αποτελούνται από ένα επίπεδο εισόδου, ένα επίπεδο εξόδου και προαιρετικά, ένα ή περισσότερα ενδιάμεσα, κρυφά επίπεδα. Δύο είναι τα θέματα τα οποία ανακύπτουν στην υλοποίηση των νευρωνικών δικτύων του τύπου αυτού.

Το πρώτο αφορά στη μάθηση, δηλαδή στον τρόπο που εκπαιδεύεται το δίκτυο για να έχει την επιθυμητή συμπεριφορά. Στα δίκτυα απλής τροφοδότησης χρησιμοποιούνται μέθοδοι μάθησης με επίβλεψη, οι σημαντικότερες από τις οποίες εξετάζονται στη συνέχεια.

Το δεύτερο θέμα αφορά στην τοπολογία του δικτύου, δηλαδή στο πόσα κρυφά επίπεδα θα έχει το δίκτυο και από πόσους νευρώνες θα αποτελείται το κάθε επίπεδο. Θα πρέπει να αναφερθεί ότι έχουν προταθεί αλγόριθμοι οι οποίοι μπορούν να δημιουργήσουν μια τοπολογία ΤΝΔ βάσει των εισόδων και των επιθυμητών εξόδων (όπως για παράδειγμα ο αλγόριθμος Upstart (Frey, 1990) και ο αλγόριθμος Tiling (Mezard και Nadal, 1989)).

Παρακάτω παρουσιάζεται η πιο απλή μορφή δικτύου απλής τροφοδότησης που ονομάζεται perceptron, καθώς και ο ομώνυμος αλγόριθμος εκπαίδευσης.

Perceptron και Multilayer Perceptron (MLP)

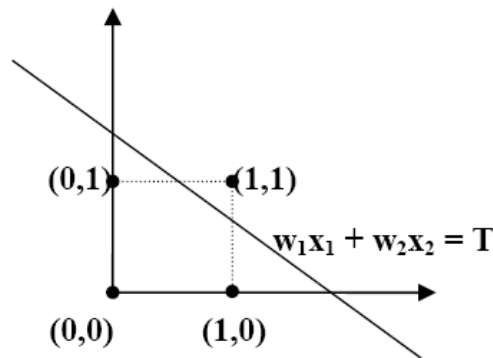
Το perceptron συνιστά την πιο απλουστευμένη τοπολογία δικτύου με απλή τροφοδότηση και αποτελεί ιστορικά μια πρώτη προσέγγιση τεχνητών νευρωνικών δικτύων που την εισήγαγε ο Rosenbalt (1958, 1962). Πρόκειται ουσιαστικά για ένα και μοναδικό τεχνητό νευρώνα, ο οποίος χρησιμοποιεί ως συνάρτηση κατωφλίου τη βηματική συνάρτηση. Η μάθηση στο perceptron συνίσταται στην επιλογή κατάλληλων τιμών βαρών, έτσι ώστε, δεδομένου ενός διανύσματος εισόδου, να παραχθεί η επιθυμητή έξοδος. Πρόκειται δηλαδή για μια απλή μορφή μάθησης υπό επίβλεψη. Ο αλγόριθμος μεταβολής των βαρών έχει ως εξής:

1. Εάν για τη συγκεκριμένη είσοδο παράγεται το επιθυμητό αποτέλεσμα, τότε δεν γίνεται καμία μεταβολή.
2. Εάν το αποτέλεσμα είναι 1, ενώ θα έπρεπε να είναι 0, τότε μειώνονται τα βάρη των ενεργών γραμμών (εκείνων που στο συγκεκριμένο πρότυπο έχουν είσοδο 1) κατά μια τιμή d , η οποία ονομάζεται ρυθμός μάθησης (learning rate).
3. Εάν το αποτέλεσμα είναι 0, ενώ θα έπρεπε να είναι 1, τότε αυξάνονται τα βάρη των ενεργών γραμμών κατά d .

Ένας perceptron με n γραμμές εισόδου μπορεί να θεωρηθεί ότι αναπαριστά ένα υπερεπίπεδο $n-1$ διαστάσεων, που διαχωρίζει τα διανύσματα εισόδου σε δύο ομάδες,

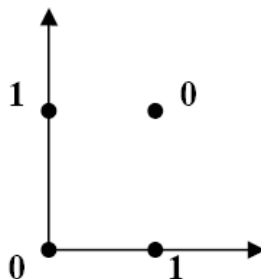
τοποθετώντας από τη μια πλευρά όσα παράγουν έξοδο 1, και από την άλλη όσα παράγουν έξοδο 0. Προβλήματα των οποίων οι τιμές εισόδου-εξόδου υπόκεινται σε αυτόν τον κανόνα ονομάζονται γραμμικώς διαχωρίσιμα (linearly separable) και έχει αποδειχτεί ότι μπορούν να μοντελοποιηθούν με τη χρήση του perceptron. Επομένως μετά από πεπερασμένο χρόνο εκπαίδευσης και εφόσον η συνάρτηση που συνδέει την είσοδο με την έξοδο είναι γραμμική, επέρχεται η σύγκλιση του perceptron.

Στο σχήμα 3.10 απεικονίζεται γραφικά η έννοια του γραμμικώς διαχωρίσιμου για τη συνάρτηση AND η οποία μπορεί να μοντελοποιηθεί με perceptron. Η ευθεία T χωρίζει τα ζευγάρια εισόδου-εξόδου σε δύο περιοχές. Πάνω και δεξιά της T βρίσκονται αυτά που έχουν έξοδο 1 και κάτω αριστερά αυτά που έχουν έξοδο 0. Εκτός από το AND, το perceptron είναι σε θέση να μοντελοποιήσει και τους στοιχειώδεις λογικούς συνδέσμους OR και NOT.



Σχήμα 3.10: Συνάρτηση AND γραμμικώς διαχωρίσιμη

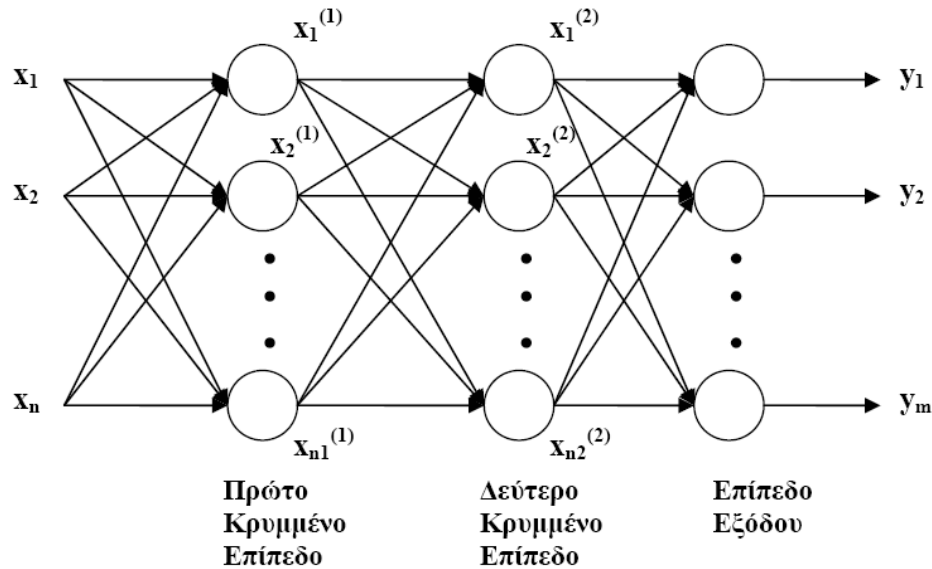
Προφανώς υπάρχουν και προβλήματα στα οποία τα ζευγάρια εισόδου-εξόδου δεν είναι γραμμικώς διαχωρίσιμα. Τέτοια προβλήματα δεν μπορούν να μοντελοποιηθούν με perceptron και απαιτούν τη χρήση ΤΝΔ με ενδιάμεσα κρυφά επίπεδα. Κλασικό παράδειγμα τέτοιας συνάρτησης είναι η XOR, της οποίας ο μη γραμμικώς διαχωρίσιμος χαρακτήρας απεικονίζεται στο σχήμα 3.11:



Σχήμα 3.11: Μη-γραμμικώς διαχωρίσιμη συνάρτηση XOR

Στην κατηγορία των νευρωνικών δικτύων πρόσθιας τροφοδότησης υπάγεται ο Multilayer Perceptron (MLP), ο οποίος δεν χρησιμοποιεί παρελθούσες τιμές των εξόδων του ή άλλου είδους εσωτερικές μεταβλητές για τον υπολογισμό της τρέχουσας εξόδου. Ο MLP αποτελείται από πολλαπλά επίπεδα απλοποιημένων σιγμοειδών κόμβων (επεξεργαστικών στοιχείων) ή νευρώνων, που αλληλεπιδρούν χρησιμοποιώντας σταθμισμένες συνδέσεις.

Η βασική δομή του MLP παρουσιάζεται στο σχήμα 3.12



Σχήμα 3.12: Απεικόνιση ενός πολυεπίπεδου perceptron

Οι κύκλοι αναπαριστούν τους νευρώνες (βάρος, κατώφλι και συνάρτηση ενεργοποίησης) και οι γραμμές αναπαριστούν τις συνδέσεις μεταξύ των εισόδων και των νευρώνων, καθώς και μεταξύ των νευρώνων σε ένα επίπεδο και εκείνων στο επόμενο επίπεδο. Το συγκεκριμένο δίκτυο αποτελεί έναν τριών-επιπέδων perceptron, καθώς αποτελείται από τρία στάδια νευρωνικής επεξεργασίας μεταξύ των εισόδων και των εξόδων. Ο MLP έχει εισόδους τα $x_i, i = 1, 2, \dots, n$ και εξόδους τα $y_j, j = 1, 2, \dots, m$. Ο αριθμός των νευρώνων στο πρώτο κρυμμένο επίπεδο είναι n_1 . Στο δεύτερο κρυμμένο επίπεδο υπάρχουν n_2 νευρώνες, και στο επίπεδο εξόδου υπάρχουν m νευρώνες.

Οι νευρώνες στο πρώτο επίπεδο του MLP εκτελούν υπολογισμούς, των οποίων τα αποτελέσματα δίνονται από τον παρακάτω τύπο.

$$x_j^{(1)} = f_j^{(1)} \left(\left(\sum_{i=1}^n w_{ij}^{(1)} x_i \right) - \theta_j^{(1)} \right)$$

όπου $j = 1, 2, \dots, n_1$. Οι νευρώνες στο δεύτερο επίπεδο του MLP εκτελούν υπολογισμούς, το αποτέλεσμα των οποίων δίνεται από:

$$x_j^{(2)} = f_j^{(2)} \left(\left(\sum_{i=1}^{n_1} w_{ij}^{(2)} x_i^{(1)} \right) - \theta_j^{(2)} \right)$$

όπου $j = 1, 2, \dots, n_2$. Οι νευρώνες στο επίπεδο εξόδου του MLP εκτελούν υπολογισμούς, των οποίων τα αποτελέσματα δίνονται από:

$$y_j = f_j \left(\left(\sum_{i=1}^{n_2} w_{ij} x_i^{(2)} \right) - \theta_j \right)$$

όπου $j = 1, 2, \dots, m$.

Οι παράμετροι $w_{ij}^{(1)}$ ονομάζονται βάρη του πρώτου κρυμμένου επιπέδου. Τα $w_{ij}^{(2)}$ ονομάζονται βάρη του δεύτερου κρυμμένου επιπέδου. Τα w_{ij} ονομάζονται βάρη του επιπέδου εξόδου. Οι παράμετροι $\theta_j^{(1)}$ και $\theta_j^{(2)}$ ονομάζονται κατώφλια του πρώτου και δεύτερου κρυμμένου επιπέδου αντίστοιχα, ενώ τα θ_j ονομάζονται τα κατώφλια του επιπέδου εξόδου. Οι συναρτήσεις f_j (για το επίπεδο εξόδου), $f_j^{(2)}$ (για το δεύτερο κρυμμένο επίπεδο), και $f_j^{(1)}$ (για το πρώτο κρυμμένο επίπεδο) αναπαριστούν τις συναρτήσεις ενεργοποίησης (activation functions). Οι συναρτήσεις ενεργοποίησης μπορούν να διαφοροποιούνται μεταξύ των νευρώνων του MLP.

3.2.4. Στάδια Διαδικασίας Ανάπτυξης ενός ΤΝΔ

Η διαδικασία ανάπτυξης σχηματισμού ενός ΤΝΔ αποτελείται από 9 στάδια (βλέπε σχήμα 3.13). Στο πρώτο στάδιο συλλέγονται τα δεδομένα που θα χρησιμοποιηθούν για την εκπαίδευση και τον έλεγχο του δικτύου. Στο δεύτερο στάδιο το δείγμα εκμάθησης πρέπει να αναγνωριστεί και ένα πλάνο πρέπει να δημιουργηθεί αναφορικά με τον έλεγχο της αποτελεσματικότητας του δικτύου. Επομένως κατά τα δύο πρώτα στάδια τα δεδομένα χωρίζονται σε δείγματα εκμάθησης (training test), που σκοπό έχουν την προσαρμογή των βαρών των συνδέσεων του ΤΝΔ, και σε δείγματα ελέγχου (testing set), που σκοπό έχουν την εξακρίβωση της εγκυρότητας του δικτύου. Γενικά όσο περισσότερα είναι τα δεδομένα, τόσο το καλύτερο, καθώς τα μεγάλα σύνολα δεδομένων, μπορεί από τη μια πλευρά να αυξάνουν τον χρόνο επεξεργασίας του δικτύου,

αλλά από την άλλη βελτιώνουν την ακρίβεια της εκμάθησης και συχνά οδηγούν σε ταχύτερη σύγκλιση σε ένα καλό σύνολο σταθμίσεων των βαρών.

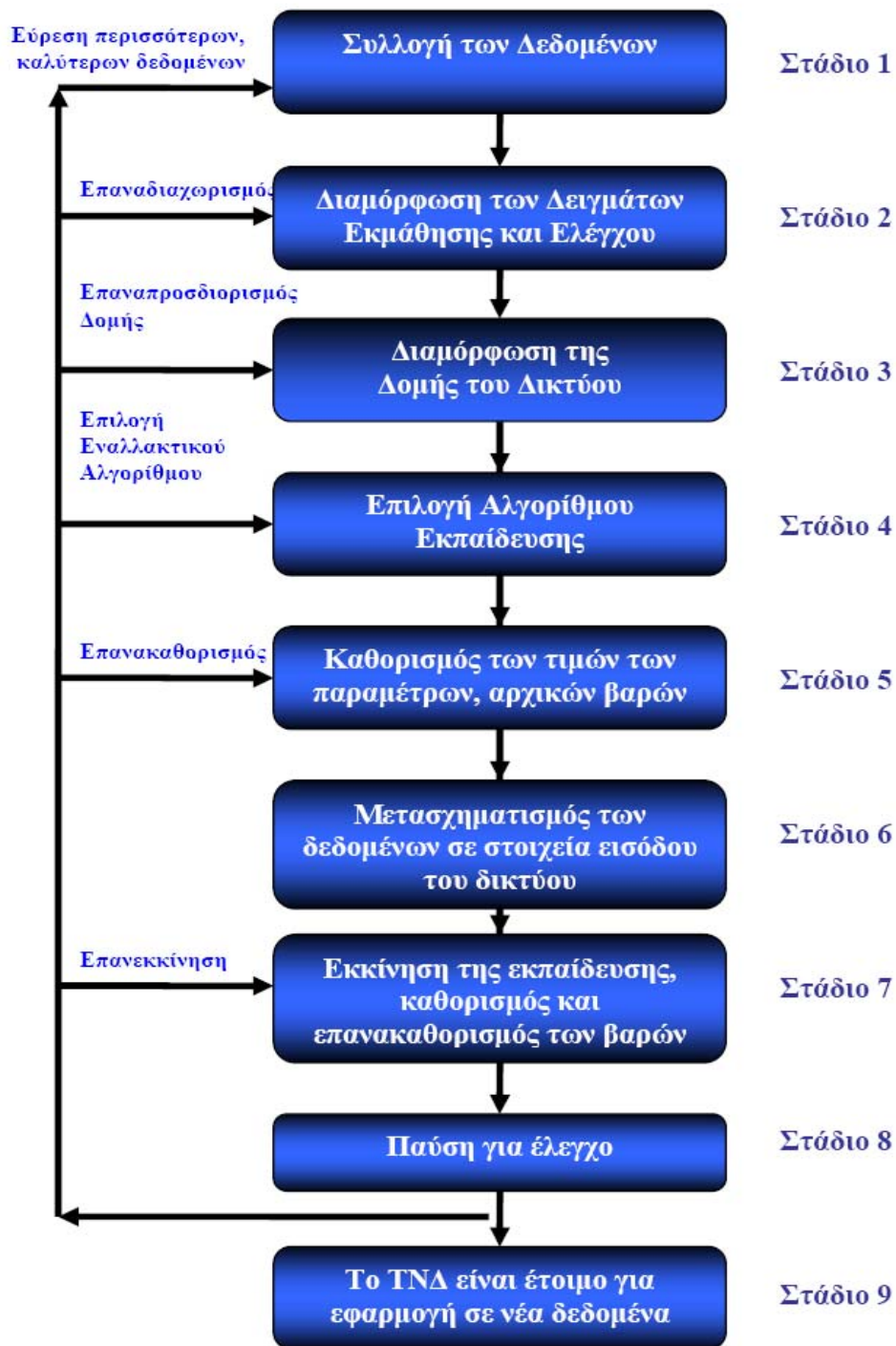
Στα στάδια 3 και 4, επιλέγεται η αρχιτεκτονική του δικτύου και ο αλγόριθμος εκπαίδευσης του δικτύου. Ο ακριβής αριθμός νευρώνων και επιπέδων (layers) θεωρείται σημαντική παράμετρος αυτών των σταδίων.

Μέρος της διαδικασίας του πέμπτου σταδίου αποτελεί να τεθούν τα αρχικά βάρη συνδέσεων των νευρώνων του δικτύου και να καθοριστούν οι παράμετροι για το επιθυμητό επίπεδο αποτελεσματικότητας της διαδικασίας μάθησης.

Το έκτο στάδιο μετατρέπει τα δεδομένα της εφαρμογής στην απαιτούμενη από το νευρωνικό δίκτυο κατάλληλη μορφή. Αυτό μπορεί να σημαίνει ότι απαιτείται λογισμικό για την προεπεξεργασία των δεδομένων. Σε αυτό το στάδιο σχεδιάζονται τεχνικές και διαδικασίες αποθήκευσης και χειρισμού δεδομένων για την αποτελεσματική και ευχερή επανεκπαίδευση του δικτύου, εφόσον βέβαια αυτή απαιτείται.

Στα στάδια 7 και 8 διεξάγονται διαδοχικά η εκπαίδευση και ο έλεγχος του δικτύου με την εισαγωγή των δεδομένων εισόδου και των επιθυμητών ή γνωστών δεδομένων εξόδου σε αυτό. Το δίκτυο υπολογίζει τα αποτελέσματα και προσαρμόζει τα βάρη μέχρι το σφάλμα απόκλισης των αποτελεσμάτων, που προκύπτει ως προς τα δεδομένα εξόδου των δεδομένων περιπτώσεων που χρησιμοποιήθηκαν, να πάρει τιμή στα όρια ανοχής που έχει θέσει ο χρήστης. Οι επιθυμητές εξόδους του ΤΝΔ και οι σχέσεις τους ως προς τις εισόδους του εξάγονται από ιστορικά στοιχεία (που αποτελούν μέρος των δεδομένων που συλλέξαμε στο πρώτο στάδιο).

Το δίκτυο είναι πλέον σε θέση να αναπαράγει τις επιθυμητές εξόδους με βάση εισόδους παρόμοιες με εκείνες του δείγματος εκμάθησης. Το δίκτυο τότε θεωρείται έτοιμο για χρήση.



Σχήμα 3.13: Διάγραμμα ροής της διαδικασίας διαμόρφωσης ενός ΤΝΔ

3.2.5. Μάθηση και Ανάκληση

Τα ΤΝΔ χρησιμοποιούν δύο βασικές λειτουργίες: τη μάθηση και την ανάκληση.

Ως μάθηση (learning) αποκαλείται η διαδικασία της τροποποίησης της τιμής των βαρών του δικτύου, ώστε, δοθέντος ενός συγκεκριμένου διανύσματος εισόδου, να παραχθεί συγκεκριμένο διάνυσμα εξόδου. Η διαδικασία αυτή ονομάζεται επίσης και εκπαίδευση (training) του ΤΝΔ. Ως ανάκληση (recall) αποκαλείται η διαδικασία του υπολογισμού ενός διανύσματος εξόδου για συγκεκριμένο διάνυσμα εισόδου και τιμές βαρών. Ο γενικός τρόπος με τον οποίο γίνεται η τροποποίηση των βαρών ενός ΤΝΔ κατά την εκπαίδευσή του, επιτρέπει τη διάκριση διαφορετικών ειδών μάθησης στα ΤΝΔ, που είναι τα εξής:

Μάθηση Υπό Επίβλεψη (supervised learning)

Στη μάθηση υπό επίβλεψη (supervised learning) το δίκτυο τροφοδοτείται με ζεύγη διανυσμάτων εισόδου – επιθυμητής εξόδου. Το ΤΝΔ, παράγει μία έξοδο με βάση την τρέχουσα κατάσταση βαρών, η οποία αρχικά διαφέρει από την επιθυμητή έξοδο. Αυτή η διαφορά ονομάζεται σφάλμα (error) και βάσει αυτής, καθώς και ενός αλγορίθμου εκπαίδευσης, γίνεται συνήθως η αναπροσαρμογή των βαρών.

Στην πράξη στις περισσότερες εφαρμογές ΤΝΔ χρησιμοποιείται μάθηση υπό επίβλεψη, για την οποία υπάρχουν αρκετοί αλγόριθμοι. Στον αλγόριθμο που βασίζεται στον κανόνα Δέλτα (Delta rule learning), η διαφορά μεταξύ πραγματικής και επιθυμητής εξόδου ελαχιστοποιείται μέσω μιας διαδικασίας ελαχίστων τετραγώνων. Στον αλγόριθμο ανάστροφης διάδοσης του σφάλματος (back propagation) η μεταβολή των βαρών βασίζεται στον υπολογισμό της συνεισφοράς του κάθε βάρους στο συνολικό σφάλμα. Σε αυτήν την περίπτωση το σφάλμα μεταξύ επιθυμητής και πραγματικής εξόδου ανατροφοδοτείται στα βάρη των προηγούμενων επιπέδων του ΤΝΔ, χρησιμοποιώντας τη διαφορική της τιμή ή τις προσεγγίσεις αυτής.

Μάθηση χωρίς επίβλεψη (unsupervised learning)

Στη μάθηση χωρίς επίβλεψη, (unsupervised learning) η απόκριση του δικτύου βασίζεται στην ικανότητά του να αυτό-οργανώνεται με βάση τα διανύσματα εισόδου (π.χ. δίκτυα με βάση τον κανόνα του Kohonen, υπό τον οποίο αναπτύσσονται διανύσματα εξόδου με βάση εκείνα της εισόδου, χωρίς να απαιτείται η ύπαρξη προτύπων εξόδου). Αυτή η εσωτερική οργάνωση γίνεται έτσι ώστε σε συγκεκριμένο σύνολο εισόδων να αντιδρά ισχυρά ένας συγκεκριμένος νευρώνας. Τέτοια σύνολα εισόδων αντιστοιχούν σε έννοιες και χαρακτηριστικά του πραγματικού κόσμου, τα οποία το ΤΝΔ καλείται να μάθει. Καθώς το δίκτυο δεν γνωρίζει τις πραγματικές τιμές εξόδου, εξετάζει τα πρότυπα εισόδου με βάση μετρούμενες ποσότητες ομοιότητας ή βαθμούς ποιότητας των δεδομένων εισόδου. Με αυτόν τον τρόπο διαμερίζεται το σύνολο εισόδων σε έναν

αριθμό από ομάδες οι οποίες αυτό-προσαρμόζονται. Η συνάρτηση σφάλματος των νευρωνικών δικτύων που εκπαιδεύονται με μάθηση χωρίς επίβλεψη βασίζεται στις μετρήσεις ομοιότητας ή ποιότητας (για παράδειγμα αποστάσεις) για όλες τις εισόδους. Το αντικείμενο της εκπαίδευσης είναι η ελαχιστοποίηση της συνάρτησης αυτής, ώστε να δημιουργηθούν ομάδες προτύπων με παρόμοιες ιδιότητες. Οι ομάδες αυτές μπορούν να θεωρηθούν ως κατηγορίες εξόδων προερχόμενες από μια ταξινόμηση προτύπων. Για το σχηματισμό των ομάδων, πάντως, απαιτούνται ορισμένες κατευθυντήριες οδηγίες σχετικά με το ποια χαρακτηριστικά (σχήμα, χρώμα, σύσταση του υλικού ή κάποιο άλλο χαρακτηριστικό του) θα χρησιμοποιηθούν στην ομαδοποίηση αυτή, διαφορετικά η ομαδοποίηση μπορεί να μην είναι επιτυχής.

Ανταγωνιστική μάθηση (competitive learning)

Στην ανταγωνιστική μάθηση (competitive learning), που αποτελεί ιδιαίτερη μορφή μάθησης υπό επίβλεψη, οι τεχνητοί νευρώνες συναγωνίζονται, κατά κάποιο τρόπο, μεταξύ τους και μόνο αυτός με τη μεγαλύτερη απόκριση σε δοθείσα είσοδο τροποποιεί τα βάρη του. Με αυτόν τον τρόπο επιτελείται ένα είδος ανταγωνισμού στους νευρώνες εξόδου ενός δικτύου, που έχει ως αποτέλεσμα σε κάθε χρονική στιγμή να είναι ενεργός μόνο ένας νευρώνας. Ένας τέτοιος νευρώνας εξόδου που κερδίζει τον ανταγωνισμό ονομάζεται «ο νικητής τα παίρνει όλα» (winner-takes-all). Μια ειδικής μορφής κατηγορία νευρωνικών δικτύων προβολής δεδομένων, που βασίζεται στην ανταγωνιστική μάθηση, είναι οι αυτό-οργανωμένοι χάρτες χαρακτηριστικών (self-organizing feature maps). Η κύρια εργασία των δικτύων αυτών είναι η μεταφορά των προτύπων από τη διάσταση που ορίζουν τα χαρακτηριστικά τους σε μια χαμηλότερη διάσταση (συνήθως επίπεδο) μέσω μίας τοπολογικής μορφής απεικόνισης. Για αυτόν το σκοπό χρησιμοποιείται ένας τοπογραφικός χάρτης των προτύπων εισόδου, στον οποίο οι συντεταγμένες των νευρώνων πάνω στο πλέγμα που σχηματίζεται αντιστοιχούν σε χαρακτηριστικά των προτύπων.

Στην γενική του μορφή, ο αλγόριθμος μάθησης που χρησιμοποιούν οι αυτό-οργανωμένοι χάρτες χαρακτηριστικών βασίζεται στους αλγόριθμους εκπαίδευσης του πλησιέστερου γείτονα (k-nearest neighbor) και στη γειτνίαση των κόμβων του πλέγματος που ορίζει ο χάρτης. Έτσι κατά την τροφοδότηση ενός προτύπου x στο δίκτυο, υπολογίζεται ο κοντινότερός του κόμβος j . Θεωρώντας ότι κάθε κόμβος j στο πλέγμα χαρακτηρίζεται από ένα διάνυσμα βάρους w_j , η παρακάτω ενημέρωση του δικτύου λαμβάνει χώρα.

$$w_i \leftarrow w_i + h_{ij}(t)[x - w_i]$$

για όλους τους γείτονες i του κόμβου j . Η συνάρτηση h_{ij} μπορεί είτε να είναι σταθερή (ρυθμός μάθησης) ή να εξαρτάται από τον βαθμό γειτνίασης του κόμβου i ως προς το j .

Ενισχυτική μάθηση (reinforced learning)

Η ενισχυτική μάθηση (reinforced learning) αποτελεί μια διαφορετική προσέγγιση στη διεργασία μάθησης σε σχέση με τη μάθηση υπό επίβλεψη και τη μάθηση χωρίς επίβλεψη. Το κυριότερο χαρακτηριστικό της ενισχυτικής μάθησης βρίσκεται στην ύπαρξη ενός ενισχυτικού σήματος (reinforcement signal) με βάση το οποίο αποτιμάται η αποτελεσματικότητα ενός συστήματος. Το ενισχυτικό σήμα μπορεί να είναι συνεχές ή διακριτό στην περιοχή τιμών $[0,1]$ και απεικονίζει την απόκριση του περιβάλλοντος στην εκτέλεση μιας ενέργειας, με την μορφή ενός βαθμού επιτυχίας (reward) ή αποτυχίας (penalty). Κατά τη διάρκεια της ενισχυτικής μάθησης σε ένα νευρωνικό δίκτυο, οι παράμετροι του δικτύου (βάρη συνάψεων) προσαρμόζονται με τέτοιο τρόπο, ώστε να επιτρέπουν την κατάλληλη επιλογή ενέργειας για κάθε σήμα εισόδου, όπως εκτιμάται από το ενισχυτικό σήμα. Ένα σήμα επιτυχίας έχει με αυτόν τον τρόπο ως αποτέλεσμα την ακόμα μεγαλύτερη ενίσχυση της ενέργειας που επιλέχτηκε, ενώ ένα σήμα αποτυχίας θα αποτρέψει την επανάληψη της ίδιας επιλογής ενέργειας κατά την επόμενη εμφάνιση του αντίστοιχου σήματος εισόδου.

Βαθμολογημένη μάθηση (graded learning)

Στη βαθμολογημένη μάθηση (graded learning) η έξοδος χαρακτηρίζεται ως «καλή» ή «κακή» με βάση μια αριθμητική κλίμακα και τα βάρη αναπροσαρμόζονται με βάση αυτόν τον χαρακτηρισμό.

Τυχαία Μάθηση (random learning)

Στην τυχαία μάθηση (random learning), οι μεταβολές στα βάρη εισάγονται τυχαία και ανάλογα με το αν η έξοδος βελτιώνεται ή όχι με βάση κάποια προκαθορισμένα από το χρήστη κριτήρια, οι μεταβολές αυτές υιοθετούνται ή απορρίπτονται.

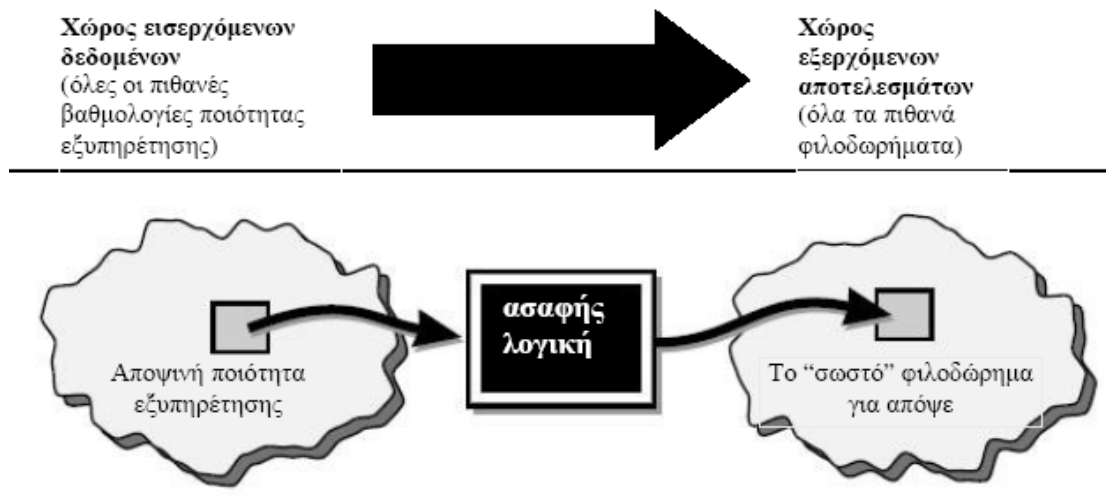
3.3 ΘΕΩΡΙΑ ΑΣΑΦΟΥΣ ΛΟΓΙΚΗΣ (FUZZY LOGIC THEORY)

Η Ασαφής λογική (Α.Λ.) είναι μια μεθοδολογία που ασχολείται με την έννοια της ακρίβειας.



Σχήμα 3.14: Ακρίβεια και σημαντικότητα

Η έννοια της ακρίβειας απεικονίζεται στην παραπάνω εικόνα 3.14. ή διαφορετικά η έννοια της κοινής λογικής. Ένας τρόπος να χαρτογραφήσουμε το χώρο των δεδομένων που εισάγουμε, στον χώρο των αποτελεσμάτων που εξάγουμε. Ο κύριος μηχανισμός για να γίνει κάτι τέτοιο είναι μια λίστα δηλώσεων αν-τότε (if-then) που ονομάζονται κανόνες. Η σειρά δεν έχει καμία σημασία και όλοι οι κανόνες υπολογίζονται παράλληλα.



Σχήμα 3.15. Χάρτης Εισόδου-Εξόδου

Όπως στο παράδειγμα, πόσο φιλοδώρημα να αφήσουμε σε ένα μαγαζί, ένας χάρτης εισόδου-εξόδου φαίνεται στο παραπάνω σχήμα.

3.3.1 Θεμέλια ασαφούς λογικής

ΑΣΑΦΗ ΣΥΝΟΛΑ

Η ασαφής λογική ξεκινάει από τη σύλληψη του ασαφούς συνόλου, το οποίο δεν έχει ξεκάθαρα και σαφώς καθορισμένα όρια και μπορεί να περιέχει στοιχεία με μερικό βαθμό συμμετοχής.

Στην ασαφή λογική, η αλήθεια κάθε υπόθεσης γίνεται υπόθεση ενός βαθμού.

ΑΣΑΦΕΙΑ

Το 1965 ο Zadeh θεμελίωσε τη θεωρία των ασαφών συνόλων σε μια μέθοδο διαχείρισης της ανακρίβειας που αντιμετωπίζεται σε πολλά πρακτικά προβλήματα. Η ανακρίβεια, ή η ασάφεια είναι ο πυρήνας των ασαφών συνόλων και της ασαφούς λογικής. Τα ασαφή σύνολα ουσιαστικά αποτελούν μια γενίκευση των κλασσικών συνόλων. Από τις αρχές της δεκαετίας του 1980 τα ασαφή σύνολα βρήκαν πολλές εφαρμογές, ιδίως σε συστήματα ελέγχου.

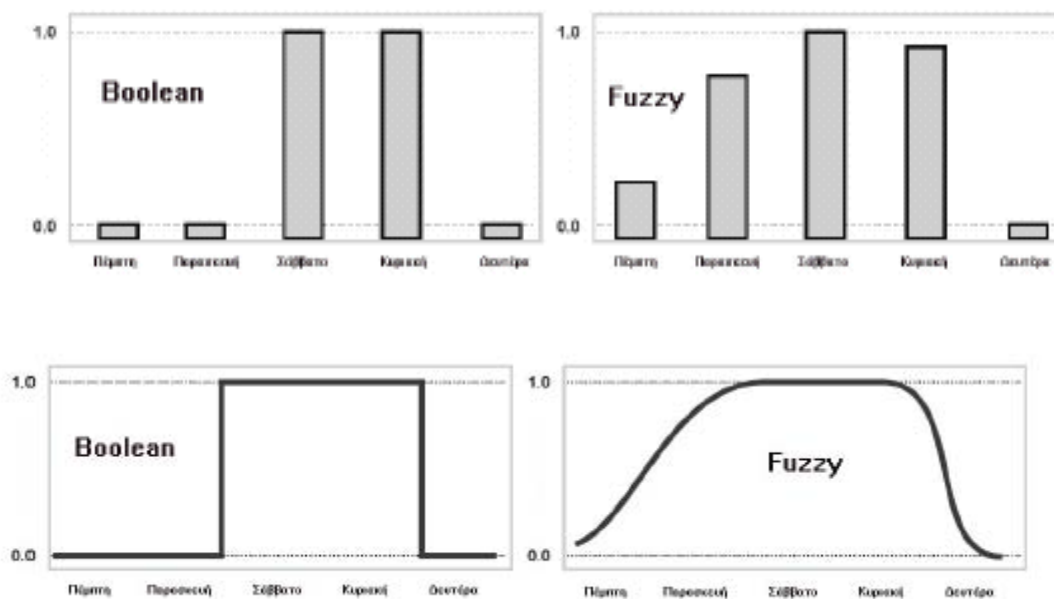
Οι άνθρωποι σκέπτονται συνήθως όχι με όρους ακριβών συμβόλων και αριθμών αλλά με ασαφείς όρους. Αυτοί οι ασαφείς όροι προσδιορίζουν κατηγορίες όχι όμως απόλυτα διαχωρισμένα και σαφώς καθορισμένα σύνολα. Η μετάβαση από τη μια κατηγορία στην άλλη γίνεται σταδιακά, μεταβαίνοντας από καταστάσεις με περισσότερη ή λιγότερη συσχέτιση με την κατηγορία. Με βάση λοιπόν την πρακτική των ανθρώπων να χρησιμοποιούν ελαστικά σύνολα, ο Zadeh πρότεινε την ιδέα των ασαφών συνόλων.

Τα ασαφή σύνολα είναι συναρτήσεις που απεικονίζουν μια τιμή, η οποία μπορεί να είναι μέλος του συνόλου, σε έναν αριθμό από το 0 έως το 1. Ο αριθμός αυτός δηλώνει το βαθμό συσχέτισης της τιμής με το ασαφές σύνολο. Όταν ο βαθμός αυτός είναι 0 τότε αυτό υποδηλώνει ότι η τιμή δεν ανήκει στο σύνολο, ενώ όταν είναι 1 τότε

αυτό σημαίνει ότι η τιμή αυτή αντιπροσωπεύει πλήρως το ασαφές σύνολο. Ο βαθμός αυτός καθορίζεται από τη συνάρτηση συμμετοχής του ασαφούς συνόλου.

Η σκέψη στην ασαφή λογική είναι απλώς ένα πρόβλημα γενίκευσης της γνωστής Ναι-Όχι λογικής Boolean. Αν δώσουμε την τιμή 1 στο «αληθές» και την τιμή 0 στο «ψευδές», τότε λέμε ότι η ασαφής λογική επιτρέπει επίσης ενδιάμεσες τιμές, όπως 0,3 ή 0,452 ή 0,81.

Όπως για παράδειγμα στο σχήμα 3.16 φαίνονται οι αληθείς τιμές των ημερών που θεωρούμε ότι ανήκουν στο σαββατοκύριακο, θεωρώντας ότι ο χρόνος παίρνει διακεκριμένες τιμές, πρώτα με Ναι-Όχι και δίπλα απαντήσεις Ασαφής Λογικής με ενδιάμεσες τιμές.



Σχήμα 3.16: Ημέρες που θεωρούμε ότι ανήκουν στο σαββατοκύριακο.

Στο ίδιο παράδειγμα, θεωρώντας τώρα ότι ο χρόνος παίρνει συνεχείς τιμές, πρώτα με Boolean Ναι-Όχι, και δίπλα απαντήσεις Ασαφούς Λογικής με ενδιάμεσες τιμές. Τώρα καθορίζουμε και το βαθμό στον οποίο κάθε στιγμή ανήκει το σαββατοκύριακο, αντί για ολόκληρη την ημέρα κάθε φορά.

3.3.2. Συναρτήσεις Συμμετοχής

Ένα ασαφές σύνολο A ορίζεται ως ένα σύνολο διατεταγμένων ζευγών $(x, \mu_A(x))$ όπου $x \in X$ και $\mu_A(x) \in [0,1]$. Το σύνολο X αποτελεί ένα ευρύτερο σύνολο αναφοράς (universe of discourse) που περιλαμβάνει όλα τα αντικείμενα στα οποία μπορεί να γίνει αναφορά. Η τιμή $\mu_A(x)$ λέγεται βαθμός αλήθειας, συμβολίζει το βαθμό συμμετοχής του x στο A και παίρνει τιμές στο διάστημα $[0,1]$. Τέλος η συνάρτηση μ_A ονομάζεται συνάρτηση συμμετοχής (membership function). Στην πράξη η συνάρτηση συμμετοχής μπορεί να προέρχεται από:

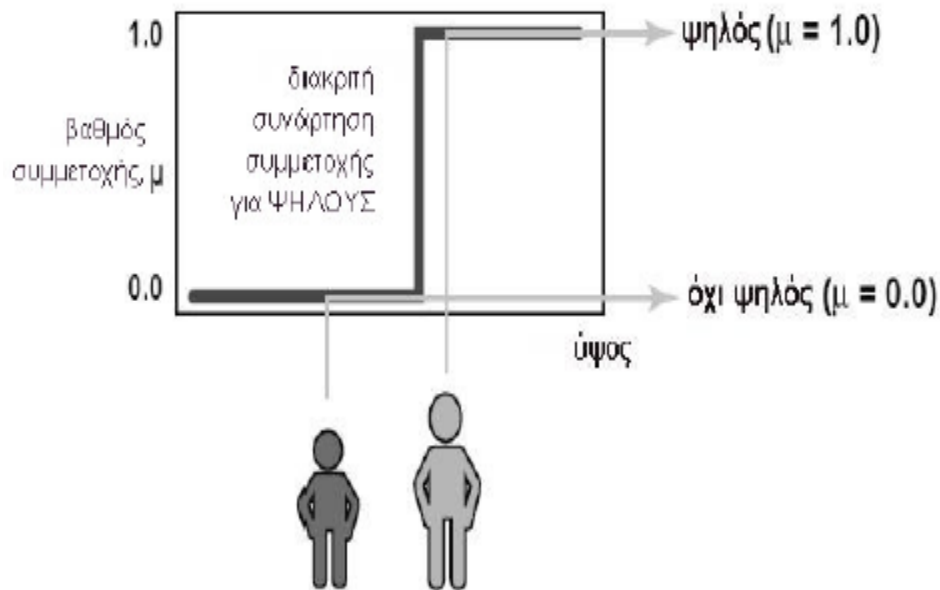
- Υποκειμενικές εκτιμήσεις
- Προκαθορισμένες (ad hoc) και απλοποιημένες μορφές
- Συχνότητες εμφανίσεων και πιθανότητες
- Φυσικές μετρήσεις
- Διαδικασίες μάθησης και προσαρμογής (π.χ. με νευρωνικά δίκτυα)

Η διαφορά των ασαφών συνόλων συγκριτικά με την κλασσική θεωρία συνόλων είναι ότι στην κλασσική θεωρία συνόλων ισχύει $\mu_A(x) \in \{0,1\}$, δηλαδή το x είτε ανήκει στο A ($\mu_A(x)=1$) ή δεν ανήκει ($\mu_A(x)=0$). Άρα η ασαφής θεωρία συνόλων μεταπίπτει στην αντίστοιχη κλασσική, όταν οι δυνατές τιμές της συνάρτησης συμμετοχής είναι μόνο 0 ή 1.

Παρακάτω αναπτύσσεται ένα παράδειγμα με το οποίο μπορεί να γίνει περισσότερο κατανοητή η έννοια της συνάρτησης συμμετοχής.

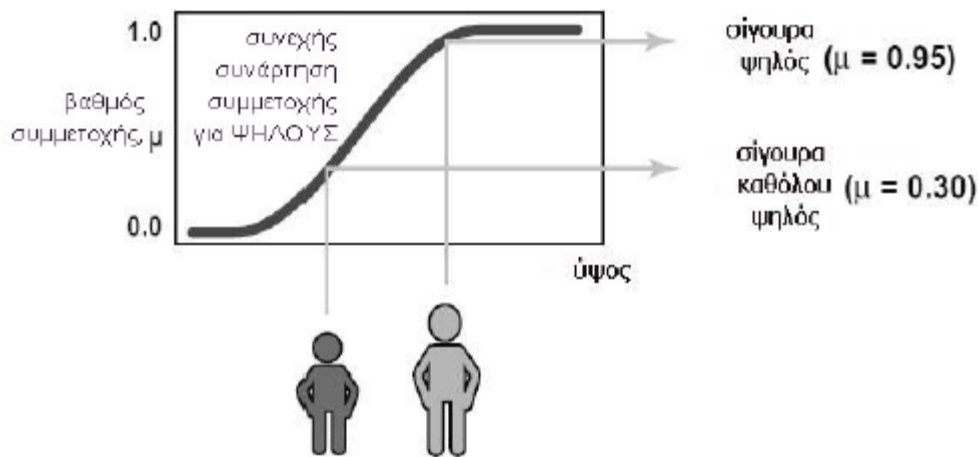
Ας υποθέσουμε ότι ο χώρος αναφοράς X είναι το σύνολο όλων των ανθρώπων. Ένα ασαφές υποσύνολο του χώρου αυτού είναι οι ψηλοί άνθρωποι. Τα πιθανά ύψη έστω

ότι κυμαίνονται από 1.20 μέχρι 2.50. Η λέξη ψηλός μπορεί να συσχετισθεί με μια καμπύλη η οποία δείχνει κατά πόσο ένας άνθρωπος είναι ψηλός ή όχι. Αν χρησιμοποιήσουμε τις αρχές των κλασσικών συνόλων τότε για να ορίσουμε το σύνολο των ψηλών ανθρώπων θα πρέπει να ορίσουμε μια συγκεκριμένη τιμή ύψους, η οποία θα διαχωρίζει τους ανθρώπους σε ψηλούς και κοντούς. Π.χ. ας υποθέσουμε ότι η τιμή αυτού του ύψους είναι 1.75m τότε ένας άνθρωπος με ύψος 1.73 θα χαρακτηρίζεται κοντός ενώ ένας άνθρωπος με ύψος 1.77 θα χαρακτηρίζεται ψηλός. Ο παραπάνω διαχωρισμός φαίνεται παράλογος αφού έχουμε αντιστοιχήσει σε δύο ανθρώπους με αμελητέα διαφορά ύψους δύο αντίθετες μεταξύ τους έννοιες.



Σχήμα 3.17: Διακριτές συναρτήσεις συμμετοχής, για ψηλούς

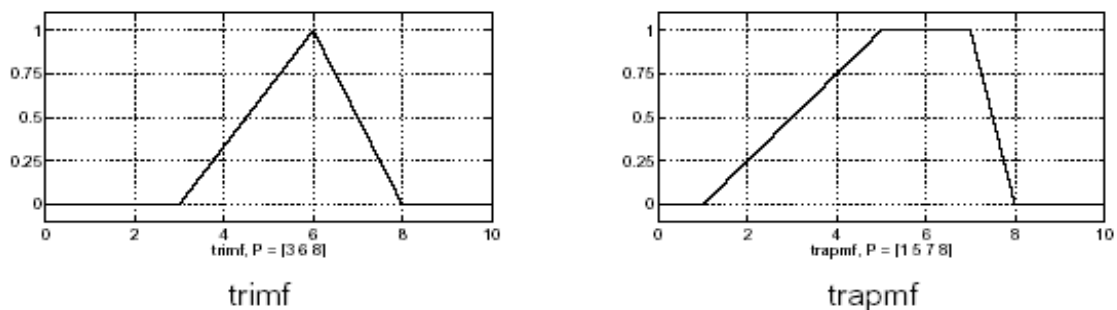
Ένας άλλος τρόπος να ορίσουμε την έννοια ψηλός είναι μέσω μιας καμπύλης που έχει ομαλή διακύμανση και μεταβαίνει από την έννοια ψηλός στην έννοια κοντός. Αυτή η καμπύλη είναι η συνάρτηση συμμετοχής του ασαφούς συνόλου των ψηλών ανθρώπων. Με άλλα λόγια δεχόμαστε ότι όλοι οι άνθρωποι είναι σε κάποιο βαθμό ψηλοί αλλά δεν είναι όλοι στον ίδιο βαθμό.



Σχήμα 3.18 Συνεχείς συναρτήσεις συμμετοχής, για ψηλούς.

Μπορούμε να παρατηρήσουμε ότι υποκειμενικοί παράγοντες ενυπάρχουν στα χαρακτηριστικά της δομής ενός ασαφούς συνόλου. Η μορφή δηλαδή της καμπύλης δεν μπορεί να είναι η ίδια όταν αναφερόμαστε σε ενήλικες και ανήλικες, σε γυναίκες και άντρες κλπ. Η μορφή επίσης της καμπύλης επιλέγεται αυθαίρετα σύμφωνα με την αντίληψη που έχει κάθε άνθρωπος για την έννοια ψηλός. Η μόνη προϋπόθεση που πρέπει να ικανοποιεί μια συνάρτηση συμμετοχής είναι να βρίσκεται στο διάστημα τιμών $[0, 1]$. Το σχήμα της επιλέγεται μεν αυθαίρετα αλλά και με τρόπο που να διασφαλίζει όσο είναι δυνατό την απλότητα.

Οι απλούστερες συναρτήσεις συμμετοχής είναι αυτές που σχηματίζονται από ευθείες γραμμές. Η απλούστερη από αυτές είναι η τριγωνική συνάρτηση συμμετοχής, που δεν είναι τίποτα άλλο από ένα τρίγωνο. Στην ίδια κατηγορία ανήκει και η τραπεζοειδής συνάρτηση συμμετοχής (σχήμα 3.19).



Σχήμα 3.19: Τριγωνική και τραπεζοειδή συνάρτηση συμμετοχής

Η μαθηματική έκφραση της τριγωνικής συνάρτησης συμμετοχής είναι η εξής:

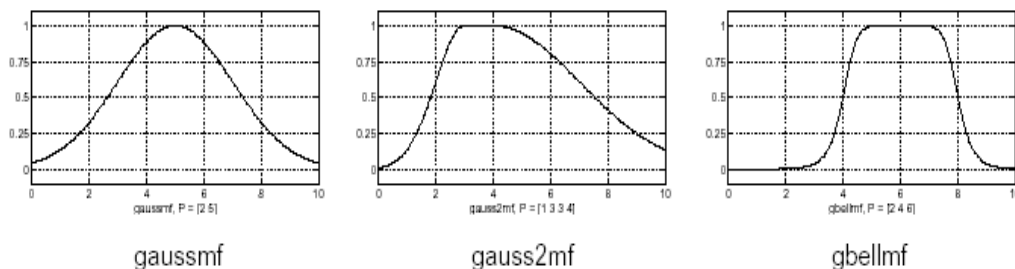
$$A = \begin{cases} 0, x \leq a \\ (x - a) / (b - a), x \in (a, b) \\ (c - x) / (c - b), x \in (b, c) \\ 0, x \geq c \end{cases}$$

Παρακάτω δίδεται η μαθηματική έκφραση της τραπεζοειδούς συνάρτησης συμμετοχής:

$$A = \begin{cases} 0, x \leq a \\ (x - a) / (b - a), x \in (a, b) \\ 1, x \in (b, c) \\ (d - x) / (d - c), x \in (c, d) \end{cases}$$

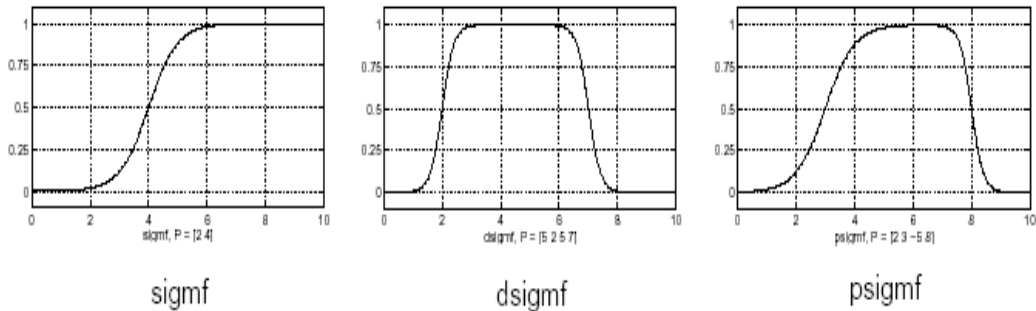
Δύο συναρτήσεις συμμετοχής που είναι δομημένες πάνω στη μορφή της κατανομής Gauss είναι μια απλή γκαουσιανή και μια σύνθεση δύο διαφορετικών γκαουσιανών.

Η γενικευμένη καμπανοειδής συνάρτηση συμμετοχής με μορφή καμπάνας έχει τρεις παραμέτρους, μια παραπάνω από την γκαουσιανή. Η γκαουσιανή και η καμπανοειδής μπορούν να χρησιμοποιούνται συχνά στα ασαφή σύνολα λόγω της ομαλότητάς τους. Έχουν δε το πλεονέκτημα να διατηρούν μη μηδενικές τιμές σε όλα τα σημεία.



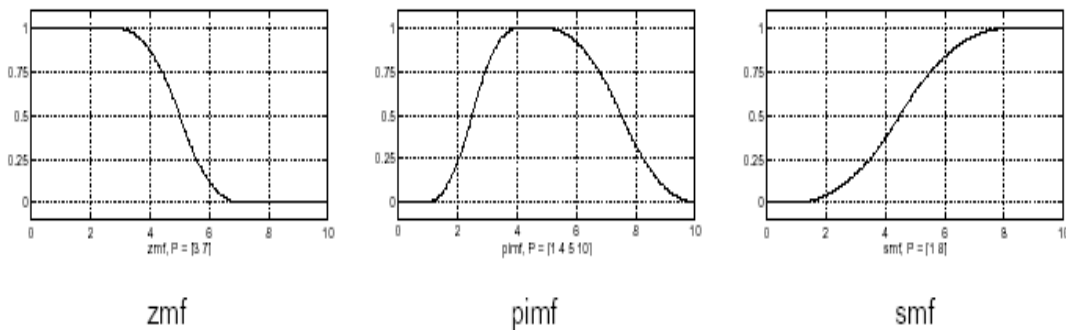
Σχήμα 3.20: Γκαουσιανές και καμπανοειδής συναρτήσεις συμμετοχής

Παρά το γεγονός ότι η γκαουσιανή συνάρτηση συμμετοχής και η συνάρτηση καμπάνας επιτυγχάνουν ομαλή διακύμανση δε μπορούν ωστόσο να ορίσουν ασύμμετρες συναρτήσεις συμμετοχής που είναι χρήσιμες σε πολλά πρακτικά προβλήματα. Για τον λόγο αυτό χρησιμοποιείται η σιγμοειδής συνάρτηση συμμετοχής η οποία είναι ασύμμετρη και ανοικτή είτε προς τα δεξιά είτε προς τα αριστερά. Κλειστές συναρτήσεις συμμετοχής αυτού του τύπου μπορούν να παραχθούν αν συνθέσουμε δύο σιγμοειδείς. Έτσι προκύπτει η διαφορά μεταξύ δύο σιγμοειδών και το άθροισμα τους. (σχήμα 3.21).



Σχήμα 3.21: Σιγμοειδές συναρτήσεις

Επίσης υπάρχουν πολλές πολωνυμικές καμπύλες (σχήμα 3.22) που τις χρησιμοποιούμε σε συναρτήσεις συμμετοχής. Τρεις από αυτές είναι η Z η S και η Π οι οποίες έχουν ονομασθεί έτσι εξαιτίας του σχήματός τους. Η Z είναι μια ασύμμετρη πολωνυμική καμπύλη που είναι ανοικτή στα αριστερά, η S είναι η κατοπτρική της Z και η Π είναι μια ασύμμετρη κλειστή καμπύλη σχήματος Π.



Σχήμα 3.22: Πολυωνυμικές καμπύλες

3.3.3 Ιδιότητες των ασαφών συνόλων

Έστω X ένα σύνολο αντικειμένων, του οποίου τα στοιχεία συμβολίζονται με το γράμμα x . Η συμμετοχή σε ένα υποσύνολο A του συνόλου X είναι μια συνάρτηση συμμετοχής μ_A από το X στο διάστημα $[0, 1]$. Το A είναι ένα ασαφές υποσύνολο του X , το οποίο όμως δεν έχει αυστηρά καθορισμένα σύνορα, μ_A είναι ο βαθμός συμμετοχής του στοιχείου x στο A . Όσο πιο κοντά στο 1 είναι το μ_A τόσο πιο πολύ ανήκει το x στο A .

Το σύνολο A μπορεί να προσδιοριστεί επομένως από το σύνολο των παρακάτω ζευγών:

$$A = \{(x, \mu_A(x)), x \in X\}$$

Κάθε ζευγάρι $(x, \mu_A(x))$ ονομάζεται μονοσύνολο

Χώρος αναφοράς του ασαφούς συνόλου A είναι το σύνολο των στοιχείων του X που έχουν μη μηδενικό βαθμό συμμετοχής στο A .

$$\text{Supp } A = \{x \in X | \mu_A(x) > 0\}.$$

Το σύνολο A μπορεί να γραφεί και ως:

$$A = \mu_1/x_1 + \mu_2/x_2 + \dots$$

$$A = \sum \mu_i/x_i$$

Στο συνεχή χώρο αναφοράς :

$$A = \int \mu_A(x) / x$$

Ασαφές δυναμοσύνολο:

Ασαφές δυναμοσύνολο, $F(x)$, του υπερσυνόλου αναφοράς X , ονομάζεται το σύνολο όλων των ασαφών υποσυνόλων του X .

Υποσύνολο:

Το σύνολο A είναι υποσύνολο του B $A \subseteq B$ αν και μόνο αν

$$\mu_A(x) \leq \mu_B(x), \forall x \in X$$

Αν ταυτόχρονα τα A και B δεν είναι ίσα, τότε το A θα ονομάζεται γνήσιο υποσύνολο του B.

Ασαφής διαμέριση:

Μια οικογένεια ασαφών υποσυνόλων του X, θα λέγεται ασαφής διαμέριση $P^n(X)$ του X τάξης n ($n \in \mathbb{N}$) και θα συμβολίζεται με $A^n = \{A_1, A_2, \dots, A_n\}$, αν και μόνο αν:

$$A_j \neq A_i, \forall i, j \in N_n (i \neq j)$$
$$0 < \sum_{k=1}^m A_i(x_k) < m, \forall i \in N_n$$

Τα στοιχεία A_i $i \in N_n$ της A_n θα λέγονται κλάσεις της ασαφούς διαμέρισης.

Κενό ασαφές σύνολο:

Ένα ασαφές σύνολο με χώρο αναφοράς τον X, λέγεται κενό αν για κάθε στοιχείο x που ανήκει στον X, η συνάρτηση συμμετοχής του A είναι μηδέν.

$$A \equiv 0 \quad \text{αν} \quad \mu_A(x) = 0 \quad \forall x \in X$$

Κανονικό ασαφές σύνολο

Ένα ασαφές σύνολο A που ορίζεται στο χώρο αναφοράς X , λέγεται κανονικό αν υπάρχει τουλάχιστο ένα στοιχείο x του X για το οποίο η συνάρτηση συμμετοχής να παίρνει τιμή ίση με τη μονάδα.

$$A=\text{κανονικό} \quad \text{αν} \quad \exists x_1 : \mu_A(x_1)=1$$

Ισότητα ασαφών συνόλων:

Δύο ασαφή σύνολα A και B που ορίζονται στο χώρο αναφοράς X , λέγονται ίσα αν για κάθε στοιχείο x του X οι συναρτήσεις συμμετοχής των A και B είναι ίσες. Δηλαδή:

$$A=B \quad \text{αν} \quad \mu_A(x) = \mu_B(x) \quad \forall x \in X$$

Συστολή ασαφών συνόλων:

Έστω ασαφές σύνολο A που ορίζεται στο χώρο αναφοράς X . Η συστολή, $CON(A)$ του συνόλου αυτού είναι ένα νέο ασαφές σύνολο με συνάρτηση συμμετοχής που ορίζεται ως εξής

$$\mu_{CON(A)}(x) = (\mu_A(x))^2$$

Η συστολή του ασαφούς συνόλου αντιστοιχεί στην προσθήκη του όρου πολύ, μπροστά από τη λεκτική μεταβλητή που περιγράφει το ασαφές σύνολο. Δηλαδή η συστολή του ασαφούς συνόλου ψηλός είναι ένα νέο ασαφές σύνολο που αντιστοιχεί στην έννοια πολύ ψηλός.

Διαστολή ασαφών συνόλων

Αντιστοίχως η διαστολή ενός ασαφούς συνόλου A , είναι ένα νέο ασαφές σύνολο που συμβολίζεται με $DIL(A)$ και έχει συνάρτηση συμμετοχής που περιγράφεται από την παρακάτω σχέση:

$$\mu_{DIL(A)}(x) = \sqrt{\mu_A(x)}$$

Η διαστολή του ασαφούς συνόλου αντιστοιχεί στην προσθήκη του όρου λίγο, μπροστά από τη λεκτική μεταβλητή που περιγράφει το ασαφές σύνολο. Δηλαδή η διαστολή του ασαφούς συνόλου ψηλός είναι ένα νέο ασαφές σύνολο που αντιστοιχεί στην έννοια λίγο ψηλός.

3.3.4. Λογικές πράξεις στα ασαφή σύνολα

Μέχρι τώρα έχουμε αναφερθεί στα ασαφή σύνολα και καθόλου στην ασαφή λογική. Η ασαφής λογική δεν είναι παρά ένα υπερσύνολο της λογικής Boolean. Αν δηλαδή απομονώσουμε τους ακραίους βαθμούς συμμετοχής 0 (πλήρως ψευδές) και 1 (πλήρως αληθές) τότε οι κλασσικοί λογικοί τελεστές μπορούν να εφαρμοσθούν. Για παράδειγμα ας θεωρήσουμε τον παρακάτω πίνακα αληθείας:

A	B	A and B
0	0	0
0	1	0
1	0	0
1	1	1

AND

A	B	A or B
0	0	0
0	1	1
1	0	1
1	1	1

OR

A	not A
0	1
1	0

NOT

Σχήμα 3.23: Λογικοί κανόνες

Μεταβαίνοντας στον χώρο της ασαφούς λογικής πρέπει να έχουμε κατά νου ότι οι έννοιες αληθές και ψευδές είναι θέμα βαθμού συμμετοχής. Επομένως ο πίνακας αυτός πρέπει να μετατραπεί με τρόπο που να συμπεριλαμβάνει αυτήν την αρχή. Οι τιμές των εισόδων A και B είναι τώρα πραγματικοί αριθμοί από το 0 μέχρι το 1. Πρέπει λοιπόν να ευρεθεί μια συνάρτηση που να διατηρεί τις ιδιότητες της συνάρτησης AND και ταυτόχρονα να μπορεί να επεκτείνεται για πραγματικούς αριθμούς.

Μια πιθανή απάντηση μπορεί να είναι ο τελεστής $\min(A,B)$, η ελάχιστη δηλαδή τιμή των εισόδων A και B. Με βάση το ίδιο σκεπτικό μια συνάρτηση που μπορεί να αντικαταστήσει τον τελεστή OR της Boolean λογικής είναι ο τελεστής $\max(A,B)$. Τέλος ο τελεστής NOT A μπορεί να αντικατασταθεί με την πράξη $1-A$.

Παρακάτω βλέπουμε ότι ο πίνακας αληθείας παραμένει αμετάβλητος αν εφαρμόσουμε τις παραπάνω υποκαταστάσεις.

A	B	$\min(A,B)$
0	0	0
0	1	0
1	0	0
1	1	1

AND

A	B	$\max(A,B)$
0	0	0
0	1	1
1	0	1
1	1	1

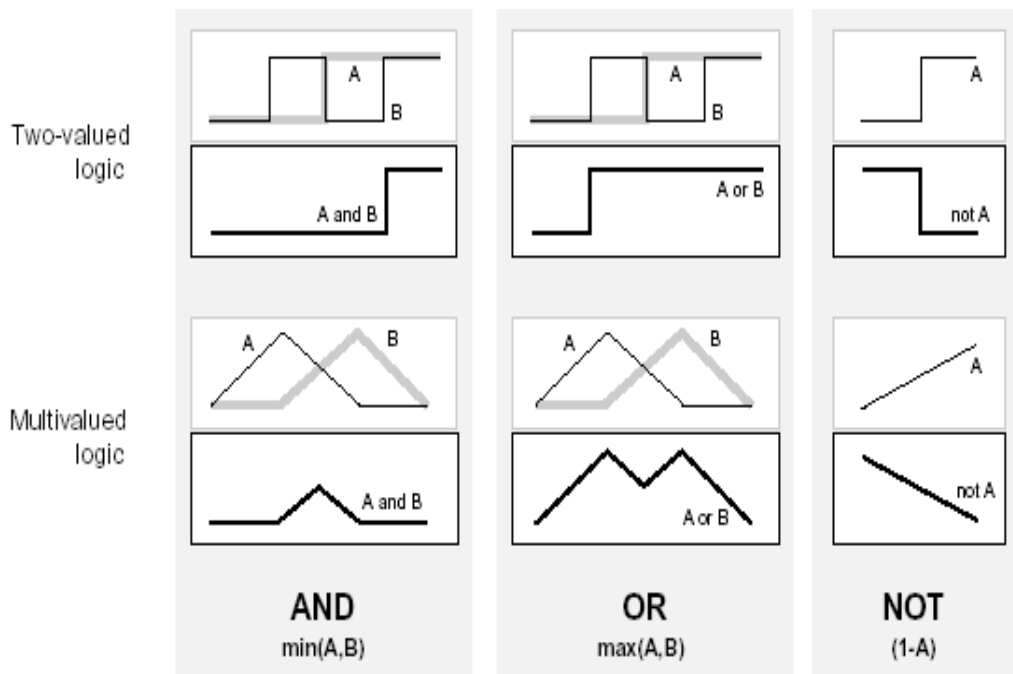
OR

A	$1 - A$
0	1
1	0

NOT

Σχήμα 3.24: Λογικοί κανόνες

Αφού έχουμε ορίσει συναρτήσεις που μπορούν ορίσουν τον πίνακα αληθείας, μπορούμε πλέον να επεκταθούμε και στην περίπτωση πραγματικών αριθμών. Στο παρακάτω σχήμα έχουμε αντικαταστήσει τον πίνακα αληθείας με ένα γράφημα δύο ασαφών συνόλων. Στο πάνω μέρος έχουμε την περίπτωση συνόλων με δύο τιμές ενώ στο κάτω φαίνεται ο τρόπος που φαίνεται πως λειτουργούν οι τελεστές στην περίπτωση που οι τιμές αληθείας A και B μεταβάλλονται συνεχώς από το 0 στο 1.



Σχήμα 3.25: Κατασκευή νέων καμπυλών από ήδη υπάρχουσες, με τη βοήθεια των τελεστών

Δεδομένων των παραπάνω συναρτήσεων μπορούμε να κατασκευάσουμε δομές με βάση ασαφή σύνολα και τους λογικούς κανόνες AND OR και NOT. Βέβαια το γεγονός ότι βρήκαμε συναρτήσεις που να επεκτείνουν τους τελεστές αυτούς από το χώρο της Boolean λογικής στο χώρο της ασαφούς λογικής δε σημαίνει σε καμία περίπτωση ότι οι συναρτήσεις αυτές είναι και οι μοναδικές.

Έχουμε ορίσει την τομή, την ένωση και το συμπλήρωμα ενός ασαφούς συνόλου με τους τελεστές $\min$ $\max$ $1-A$ οι οποίοι είναι οι πιο συχνά χρησιμοποιούμενοι. Ωστόσο μπορούμε να χρησιμοποιήσουμε και άλλες συναρτήσεις για τον ίδιο σκοπό.

Γενικά η τομή δύο ασαφών συνόλων μπορεί να παρασταθεί με μια δυαδική απεικόνιση T που αθροίζει (aggregates) τις δύο συναρτήσεις συμμετοχής ως ακολούθως:

$$\mu_{A \cap B} = T(\mu_A(x), \mu_B(x))$$

Για παράδειγμα ο δυαδικός τελεστής T μπορεί να αντιπροσωπεύει τον πολλαπλασιασμό των $\mu_A(x)$ και $\mu_B(x)$. Αυτού του είδους οι τελεστές τομής αναφέρονται συνήθως ως τ-νόρμες (τριγωνική νόρμα, triangular norm) και πρέπει να ικανοποιούν τις παρακάτω προϋποθέσεις:

- Οριακές συνθήκες: $T(0, 0) = 0$, $T(a, 1) = T(1, a) = a$
- Μονοτονία: $T(a, b) \leq T(c, d)$ αν $a \leq c$ και $b \leq d$
- Αντιμεταθετικότητα: $T(a, b) = T(b, a)$
- Προσεταιριστικότητα: $T(a, T(b, c)) = T(T(a, b), c)$

Η πρώτη απαίτηση επιβάλλει τη γενίκευση σε σύνολα διακριτών τιμών (crisp sets).

Η δεύτερη υπονοεί ότι μια μείωση του βαθμού συμμετοχής στο A ή στο B δεν μπορεί να οδηγήσει σε αύξηση του βαθμού συμμετοχής στην τομή των A και B .

Η τρίτη απαίτηση δείχνει ότι ο τελεστής είναι αδιάφορος στον τρόπο διάταξης με τον οποίο τέμνονται τα δύο σύνολα A και B .

Τέλος η τέταρτη απαίτηση μας επιτρέπει να παίρνουμε την τομή οποιουδήποτε αριθμού συνόλων σε οποιαδήποτε διάταξη ζευγών.

Παραδείγματα τ-νορμών είναι:

$$\text{Συνήθης τομή} \quad T(a, b) = \min(a, b)$$

Αλγεβρικό γινόμενο

$$T(a,b)=ab$$

Φραγμένη διαφορά

$$T(a,b)=\max(0,a+b-1)$$

Συνάρτηση Hamacher

Η συνάρτηση Hamacher δίνεται από τη σχέση:

$$t(a,b) = \frac{ab}{r + (1-r)(a+b-ab)}$$

Όπως η ασαφής τομή έτσι και η ασαφής ένωση προσδιορίζεται με μια δυαδική απεικόνιση S .

$$\mu_{A \cup B}(x) = S(\mu_A(x), \mu_B(x))$$

Για παράδειγμα ο τελεστής S μπορεί να είναι το άθροισμα των $\mu_A(x)$ και $\mu_B(x)$. Οι ασαφείς τελεστές ένωσης αναφέρονται συνήθως ως σ-νόρμες και πρέπει να πληρούν τις παρακάτω προϋποθέσεις:

- Οριακές συνθήκες: $S(1, 1) = 1, S(a, 0) = S(0, a) = a$
- Μονοτονία: $S(a, b) \leq S(c, d)$ αν $a \leq c$ και $b \leq d$
- Αντιμεταθετικότητα: $S(a, b) = S(b, a)$
- Προσεταιριστικότητα: $S(a, S(b, c)) = S(S(a, b), c)$

Παραδείγματα σ-νορμών είναι τα εξής:

Συνήθης ένωση $s(a,b)=\max(a,b)$

Αλγεβρικό άθροισμα $s(a,b)=a+b-ab$

Φραγμένο άθροισμα $s(a,b)=\min(1,a+b)$

Συμπλήρωμα

Το συμπλήρωμα A' ενός ασαφούς συνόλου A δίνεται από τη σχέση:

$$\mu_{A'}(x) = c(\mu_A(x))$$

όπου η συνάρτηση c πρέπει να ικανοποιεί τις παρακάτω συνθήκες;

- Οριακές συνθήκες: $c(0)=1$ και $c(1)=0$
- Μονοτονία: $\forall a, b \in [0,1], \text{ αν } a \leq b \Rightarrow c(a) \geq c(b)$
- Συνέχεια: c συνεχής στο $[0,1]$.
- Εναγωγή $\forall a \in [0,1]$ είναι $c(c(a))=a$

Το σύνηθες συμπλήρωμα δίνεται από τη σχέση:

$$\mu_{A'}(x) = 1 - \mu_A(x)$$

3.3.5. Ασαφείς σχέσεις

Οι ασαφείς σχέσεις (fuzzy relations) είναι ασαφή σύνολα ορισμένα σε πεδία αναφοράς ανώτερης διάστασης (π.χ. $X \times X$, $X \times Y \times Z$ κλπ). Ποιοτικά, μια ασαφής σχέση R θα μπορούσε να είναι μια έκφραση της μορφής «είναι βαρύτερο από» και η οποία θα συνδέει τα στοιχεία δύο άλλων συνόλων:

$$R = \langle x \text{ είναι βαρύτερο από } y \rangle \quad x \in X, y \in Y \text{ και } R \in X \times Y$$

Οι ασαφείς σχέσεις μπορεί να εκφραστούν με αναφορά όλων των ζευγών (τιμή, βαθμός συμμετοχής), δηλαδή ζευγών της μορφής $((x,y), \mu_R(x,y))$. Ένας άλλος τρόπος αναπαράστασης, ιδιαίτερα χρήσιμος σε υπολογισμούς είναι σε μορφή πίνακα:

$$R = \begin{bmatrix} \mu_R(x_1, y_1) & \mu_R(x_1, y_2) & \cdots & \mu_R(x_1, y_n) \\ \mu_R(x_2, y_1) & \mu_R(x_2, y_2) & \cdots & \mu_R(x_2, y_n) \\ \vdots & \vdots & \ddots & \vdots \\ \mu_R(x_m, y_1) & \mu_R(x_m, y_2) & \cdots & \mu_R(x_m, y_n) \end{bmatrix}$$

Οι ασαφείς σχέσεις μπορούν να συνδυαστούν μεταξύ τους μέσω της διαδικασίας της σύνθεσης (composition). Αν για παράδειγμα συνδυαστεί η ασαφής σχέση $R_1(x,y)$ ορισμένη στο $X \times Y$ με την ασαφή σχέση $R_2(y,z)$ ορισμένη στο $Y \times Z$ τότε θα προκύψει μία ασαφής σχέση $R(x,z)$ η οποία θα ορίζεται στο σύνολο $X \times Z$ και θα συσχετίζει άμεσα στοιχεία των συνόλων X και Z . Βέβαια είναι απαραίτητο να προσδιοριστεί επακριβώς η συνάρτηση συμμετοχής $\mu_R(x,z)$ της R με χρήση των συναρτήσεων συμμετοχής των R_1 και R_2 .

Οι βασικές πράξεις που ορίζονται μεταξύ των ασαφών σχέσεων είναι η αντιστροφή και η σύνθεση.

Αντιστροφή

Αντίστροφη σχέση της $R(X,Y)$ είναι η ασαφής σχέση $R^{-1}(Y,X)$ με τύπο: $R^{-1}(y,x)=R(x,y)$ για κάθε x που ανήκει στον X και κάθε y που ανήκει στον Y . Ο πίνακας συμμετοχής που παριστάνει την R^{-1} είναι ο ανάστροφος του R .

Σύνθεση

Η σύνθεση είναι πολύ σημαντική διαδικασία καθώς όπως θα δούμε παρακάτω, οι κανόνες της μορφής if-then αντιστοιχούν σε ασαφείς σχέσεις και το πρόβλημα της ασαφούς συλλογιστικής είναι μαθηματικά ισοδύναμο με τη σύνθεση.

Αν $R_1(x,y)$ και $R_2(y,z)$ είναι δύο ασαφείς σχέσεις ορισμένες στα σύνολα $X \times Y$ και $Y \times Z$ αντίστοιχα, τότε η σύνθεση τους δίνει μια νέα σχέση $R_1 \circ R_2$.

Η sup-t σύνθεση $R: X \times Y \rightarrow [0,1]$ δύο ασαφών σχέσεων $R_1: X \times Y \rightarrow [0,1]$ και $R_2: Y \times Z \rightarrow [0,1]$ ορίζεται από την εξίσωση:

$$R(x, z) = [R_1 \circ^t R_2](x, z) = \sup_{y \in Y} t[R_1(x, y), R_2(y, z)]$$

Οι περισσότεροι γνωστές μέθοδοι σύνθεσης ασαφών συνόλων είναι η σύνθεση max-min και η σύνθεση max-product.

Η συνάρτηση συμμετοχής για την περίπτωση της max-min σύνθεση δίνεται από τη σχέση:

$$\mu_{R_1 \circ R_2}(x, z) = \bigvee_y [\mu_{R_1}(x, y) \wedge \mu_{R_2}(y, z)]$$

Στην περίπτωση της σύνθεσης max-product έχουμε την παρακάτω σχέση:

$$\mu_{R_1 \circ R_2}(x, z) = \bigvee_y [\mu_{R_1}(x, y) \bullet \mu_{R_2}(y, z)]$$

Οι υπολογισμοί στο δεξιό μέρος των παραπάνω σχέσεων είναι παρόμοιοι με του πολλαπλασιασμό των πινάκων.

3.3.6. Ασαφής συλλογιστική

Η βάση στην οποία στηρίζεται η λήψη αποφάσεων είναι η παραγωγή συλλογιστικής. Η ασαφής λογική ασχολείται με την παραγωγή συλλογιστικής σε περιβάλλον αβεβαιότητας. Για το σκοπό αυτό, θεμελιώνεται η δομή και η μαθηματική αναπαράσταση ενός ασαφούς γεγονότος με τον ορισμό των ασαφών συνόλων και καθορίζεται ο τρόπος με τον οποίο συνδυάζουμε τα γεγονότα για να παράγουμε λογικές προτάσεις ή σχέσεις και συνεπώς συμπεράσματα.

Οι συλλογιστικοί τρόποι που κυρίως χρησιμοποιούνται είναι τρεις

- Ο modus ponens (MP)
- Ο modus tolens (MT)
- Ο υποθετικός συλλογισμός (HS)

Ο modus ponens παράγει συμπεράσματα από ένα σύνολο υποθέσεων σύμφωνα με το σχήμα:

$$(A \Rightarrow B) \wedge A \Rightarrow B$$

όπου A και B συγκεκριμένα γεγονότα.

Ο παραπάνω κανόνας ερμηνεύεται ως εξής: Αν το γεγονός A συνεπάγεται το γεγονός B και επιπλέον έχουμε ως υπόθεση ότι ισχύει το A , τότε το συμπέρασμα που

παίρνουμε είναι ότι ισχύει και το B. Όμως σε περιβάλλον ασάφειας τα γεγονότα ισχύουν σε κάποιο βαθμό. Έτσι ο παραπάνω κανόνας πρέπει να τροποποιηθεί για να συμπεριλάβει και την ασάφεια. Καταλήγουμε μ' αυτό τον τρόπο στο γενικευμένο κανόνα modus ponens (generalized modus ponens) ο οποίος έχει την παρακάτω μορφή:

$$(A \Rightarrow B) \wedge A' \Rightarrow B'$$

Αντίστοιχα ο γενικευμένος κανόνας modus tolens (generalized modus tolens) διατυπώνεται ως εξής:

$$(A \Rightarrow B) \wedge B' \Rightarrow A'$$

Ας εξετάσουμε τώρα την ερμηνεία των παραπάνω κανόνων. Για παράδειγμα ο γενικευμένος κανόνας modus ponens ερμηνεύεται ως εξής: Αν το γεγονός A συνεπάγεται το γεγονός B και έχουμε ως υπόθεση ότι ισχύει το A σε κάποιο βαθμό, τότε θα ισχύει και το B σε κάποιο βαθμό.

Πρέπει, λοιπόν να προσδιοριστεί ο βαθμός στον οποίο πληρείται το γεγονός B. Αυτός εξαρτάται από το βαθμό στον οποίο πληρείται το γεγονός A και από το είδος της συνεπαγωγής που εφαρμόζουμε. Η πράξη της ασαφούς συνεπαγωγής υλοποιεί τη μαθηματική σχέση $A \Rightarrow B$, όταν τα A και B είναι ασαφή γεγονότα.

Το σχήμα που προτείνεται από τη θεωρία των ασαφών συνόλων για την εξαγωγή του γεγονότος B' από τα γεγονότα A, A', B περιγράφεται από την παρακάτω σχέση που αποτελεί το συνθετικό κανόνα του Zadeh:

$$B'(y) = \sup_{x \in X} t[A'(x), \sigma(A(x), B(y))]$$

όπου σ είναι μια συνάρτηση που υλοποιεί την πράξη της ασαφούς συνεπαγωγής.

Επομένως η επιλογή της συνάρτησης που υλοποιεί την ασαφή συνεπαγωγή έχει ουσιαστικό ρόλο για την ασαφή συλλογιστική που παράγεται με βάση το παραπάνω σχήμα. Ένα από τα κριτήρια που χρησιμοποιείται γι' αυτή την επιλογή είναι το κριτήριο της ανάκλησης (recall), το οποίο διατυπώνεται ως εξής:

$$B(y) = \sup_{x \in X} t[A(x), \sigma(A(x), B(y))]$$

Το κριτήριο αυτό έχει την εξής ερμηνεία: η ασαφής συνεπαγωγή πρέπει να είναι τέτοια ώστε όταν η υπόθεση πληρείται ακριβώς, τότε να λαμβάνουμε το συμπέρασμα του κανόνα $A \Rightarrow B$, δηλαδή το γεγονός B . Η απαίτηση αυτή είναι εύλογη αφού όταν δεν υπάρχει αβεβαιότητα η ασαφής συλλογιστική οφείλει να ταυτίζεται με την κλασσική συλλογιστική.

3.4 ΝΕΥΡΟΑΣΑΦΗ ΣΥΣΤΗΜΑΤΑ

3.4.1 Εισαγωγή

Τα νευροασαφή συστήματα αποτελούν μια υβριδική κατηγορία του Ευφυούς Ελέγχου, που έχει αναπτυχθεί σε μια προσπάθεια συνδυασμού των καλύτερων στοιχείων διαφόρων κατηγοριών. Τα ασαφή και τα νευρωνικά συστήματα μοντελοποιούν τη συμπεριφορά ενός ειδικού κατά την αντιμετώπιση μιας πολύπλοκης διεργασίας. Αυτό σημαίνει ότι στόχος κατά το σχεδιασμό τέτοιων συστημάτων δεν είναι η μοντελοποίηση του ίδιου του προβλήματος κατασκευάζοντας ένα αναλυτικό μαθηματικό μοντέλο, αλλά η αξιοποίηση της γνώσης και της εμπειρίας ενός ειδικού, κατασκευάζοντας ένα σύστημα ελέγχου που θα ελέγχει τη διεργασία όπως ο ειδικός.

Οι αρχές της ασαφούς λογικής και των ΤΝΔ είναι όμως εντελώς διαφορετικές: η ασαφής λογική επιδιώκει να αναπαραγάγει τους μηχανισμούς της ανθρώπινης σκέψης και την ικανότητα συλλογισμού, ενώ τα νευρωνικά δίκτυα επιχειρούν να μιμηθούν τους μηχανισμούς του ανθρώπινου νου σε βιολογικό επίπεδο. Έτσι, η δομή και ο τρόπος επεξεργασίας των ελεγκτών των δύο τεχνικών διαφέρουν ριζικά.

Το πιο σημαντικό χαρακτηριστικό των νευρωνικών δικτύων είναι η ικανότητά τους να μαθαίνουν από παραδείγματα. Χρησιμοποιώντας τους ήδη γνωστούς αλγόριθμους μάθησης, εκπαιδεύονται επεξεργαζόμενοι ένα σύνολο δεδομένων. Το κύριο μειονέκτημα που εμφανίζουν τα νευρωνικά δίκτυα είναι ότι δεν μπορεί να αποδειχθεί ότι δουλεύουν όπως ήταν αναμενόμενο. Εξαιτίας της κατανεμημένης του φύσης, δεν είναι σαφές το αν το δίκτυο έχει εκπαιδευτεί σωστά. Ένα νευρωνικό δίκτυο μαθαίνει, αλλά ο χρήστης δε μπορεί να μάθει απ' το δίκτυο. Για το χρήστη λειτουργεί ως ένα «μαύρο κουτί». Έτσι, αν υπάρχει κάποια γνώση για τη συμπεριφορά του συστήματος πριν την εκπαίδευση, αυτή δε μπορεί να αξιοποιηθεί στη διαδικασία εκπαίδευσης του νευρωνικού ελεγκτή.

Αντίθετα, η δράση των ασαφών συστημάτων μπορεί εύκολα να κατανοηθεί, καθώς ακολουθούν διαδοχικά βήματα κατά τη διαδικασία συμπερασμού με πλήρη επεξήγηση των κανόνων που έχουν ενεργοποιηθεί. Το κύριο μειονέκτημά τους είναι ότι δεν έχουν τη δυνατότητα να ρυθμιστούν μέσω αλγορίθμων μάθησης. Δημιουργούνται μέσω σαφούς γνώσης που εκφράζεται με τη μορφή λεκτικών (ασαφών) κανόνων, κάποιες φορές όμως η διαθέσιμη γνώση δεν είναι αρκετή για να καθοριστούν με ακρίβεια όλες οι παράμετροί τους. Εάν η απόδοσή τους κρίνεται μη ικανοποιητική, οι παράμετροί τους πρέπει να ρυθμιστούν χειροκίνητα. Η διαδικασία ρύθμισης είναι αρκετά χρονοβόρα και πολλές φορές οδηγεί σε λανθασμένα αποτελέσματα. Θα ήταν λοιπόν επιθυμητό να μπορούσε να εφαρμοσθεί ένας αλγόριθμος εκμάθησης στα ασαφή συστήματα, παρόμοιος με αυτόν της εκπαίδευσης των νευρωνικών συστημάτων.

Η ικανότητα εκπαίδευσης των νευρωνικών συστημάτων είναι ο κύριος λόγος συνδυασμού τους με τα ασαφή συστήματα. Αυτός ο συνδυασμός μπορεί να δημιουργήσει κανόνες για το ασαφές σύστημα ή να βελτιστοποιήσει τους ήδη υπάρχοντες, αυτοματοποιώντας τη ρύθμιση των παραμέτρων του ασαφούς συστήματος. Όσον αφορά στα νευρωνικά συστήματα, ο συνδυασμός τους με τα ασαφή αναιρεί τη συμπεριφορά τους ως «μαύρα κουτιά» κι έτσι οποιαδήποτε εκ των προτέρων γνώση για το σύστημα μπορεί να χρησιμοποιηθεί κατά τη διαδικασία εκμάθησης περιορίζοντας το χρόνο που αυτή απαιτεί.

Συνεπώς, συνδυάζοντας τα νευρωνικά με τα ασαφή συστήματα διατηρούνται τα πλεονεκτήματα και των δύο μεθόδων, ενώ εξαλείφονται κάποια από τα μειονεκτηματά τους.

3.4.2 Αρχιτεκτονικές νευρο-ασαφών ελεγκτών

Γενικά, υπάρχουν δύο κύριοι συνδυασμοί ανάμεσα στα νευρωνικά και στα ασαφή συστήματα. Στην πρώτη περίπτωση, το νευρωνικό και το ασαφές σύστημα δουλεύουν ανεξάρτητα το ένα από το άλλο και το νευρωνικό σύστημα καθορίζει διάφορες παραμέτρους του ασαφούς, σε πραγματικό ή μη πραγματικό χρόνο. Αυτός ο συνδυασμός ονομάζεται *νευρο-ασαφές σύστημα συνεργασίας (cooperative neuro-fuzzy system)*, διότι το νευρωνικό δίκτυο συνεργάζεται με το ασαφές ώστε να το βοηθήσει να βρει τις κατάλληλες παραμέτρους του.

Η δεύτερη περίπτωση συνδυασμού νευρωνικού και ασαφούς συστήματος είναι το υβριδικό νευρο-ασαφές σύστημα. Ο όρος υβριδικό χρησιμοποιείται διότι το σύστημα που προκύπτει μπορεί να αντιμετωπιστεί είτε ως νευρωνικό είτε ως ασαφές σύστημα, αλλά είναι ένα ομογενές σύστημα που δεν μπορεί να διαιρεθεί σε δύο ξεχωριστά υποσυστήματα. Ένα γνωστό παράδειγμα υβριδικού νευρο-ασαφή ελεγκτή είναι το μοντέλο ANFIS (Adaptive Network Based Fuzzy Inference System).

❖ Νευρο-ασαφή συστήματα συνεργασίας

Το Παρακάτω Σχήμα απεικονίζει τις τέσσερις διαφορετικές κατηγορίες νευρο-ασαφών συστημάτων συνεργασίας.

1η κατηγορία:

Το νευρωνικό δίκτυο ρυθμίζει τις συναρτήσεις συμμετοχής του ασαφούς δικτύου χρησιμοποιώντας τα δεδομένα εκπαίδευσης (training data), Σχ.(α). Τα ασαφή σύνολα που ρυθμίζονται με αυτόν τον τρόπο (σε μη πραγματικό χρόνο) συνδυάζονται με τους ασαφείς κανόνες που δίνονται από τον έμπειρο χειριστή ώστε να σχεδιασθεί το τελικό

ασαφές σύστημα. Σε αυτή την περίπτωση, τα δεδομένα εκπαίδευσης πρέπει να αφορούν βαθμούς συμμετοχής που αντιστοιχούν σε δεδομένες τιμές εισόδων.

2η κατηγορία:

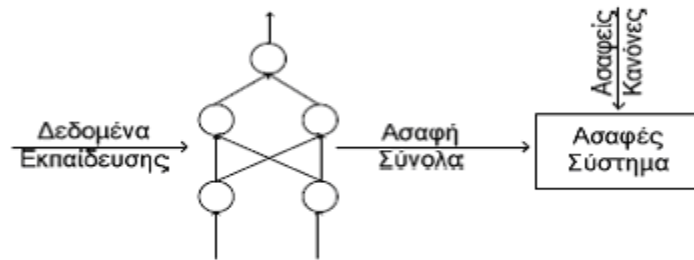
Το νευρωνικό δίκτυο σχηματίζει τους ασαφείς κανόνες από τα δεδομένα εκπαίδευσης, Σχ. (β). Αυτή η διαδικασία γίνεται σε μη πραγματικό χρόνο. Οι ασαφείς κανόνες που προκύπτουν συνδυάζονται με δεδομένες συναρτήσεις συμμετοχής για να προκύψει το τελικό ασαφές σύστημα.

3η κατηγορία:

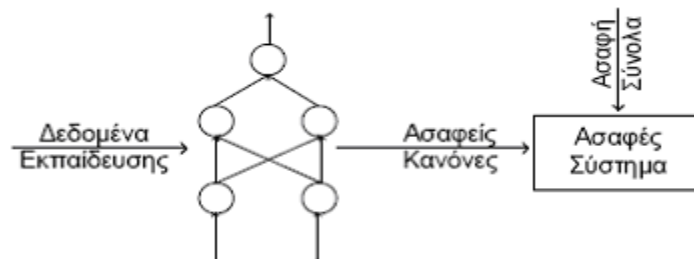
Το σύστημα ρυθμίζεται σε πραγματικό χρόνο, δηλαδή κατά τη διάρκεια λειτουργίας του ασαφούς συστήματος, για να προσαρμοσθούν οι συναρτήσεις συμμετοχής, Σχ. (γ). Σε αυτή την περίπτωση πρέπει να δοθούν εκ των προτέρων οι ασαφείς κανόνες και οι συναρτήσεις συμμετοχής. Επιπρόσθετα, πρέπει να καθορισθεί ένα μέτρο σφάλματος το οποίο οδηγεί τη διαδικασία εκπαίδευσης του νευρωνικού δικτύου. Σε αυτή την περίπτωση συνήθως δεν υπάρχει πραγματικό νευρωνικό δίκτυο, αλλά ένας νευρωνικός αλγόριθμος εκπαίδευσης, όπως αυτός της οπισθόδρομης διάδοσης.

4η κατηγορία:

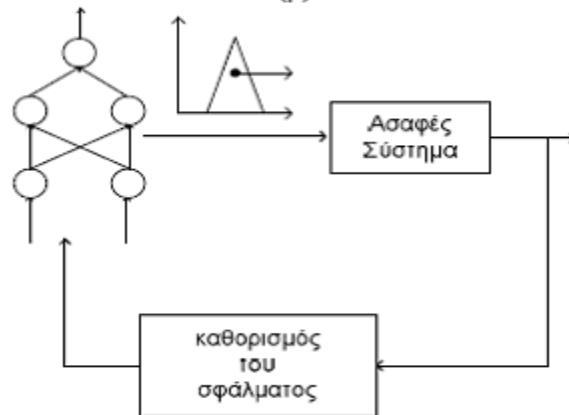
Το νευρωνικό δίκτυο (ή ένας νευρωνικός αλγόριθμος εκπαίδευσης) καθορίζει τα βάρη των κανόνων, είτε σε πραγματικό είτε σε μη πραγματικό χρόνο, Σχ. (δ). Και σε αυτή την περίπτωση, πρέπει να δοθούν εκ των προτέρων οι ασαφείς κανόνες και οι συναρτήσεις συμμετοχής.



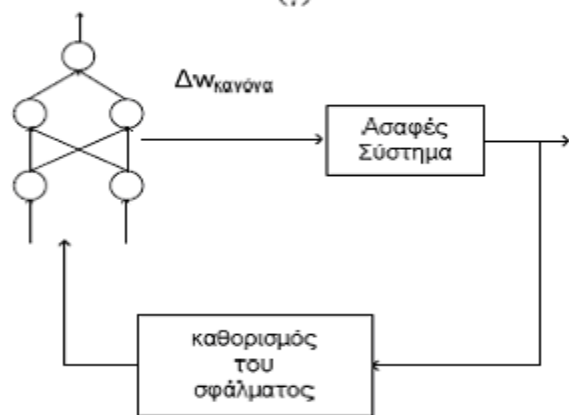
(α)



(β)



(γ)

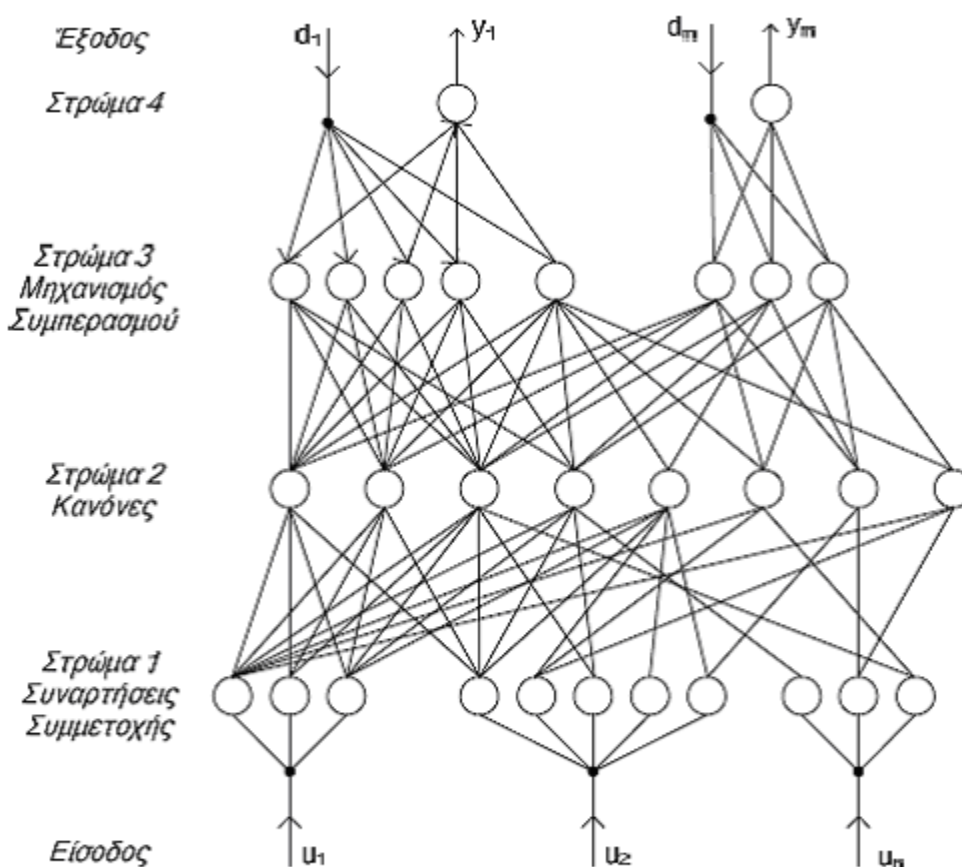


(δ)

Κατηγορίες νευρο-ασαφών συστημάτων.

❖ Υβριδικά νευρο-ασαφή συστήματα

Το παρακάτω σχήμα παρουσιάζει ένα παράδειγμα ενός υβριδικού νευρο-ασαφούς ελεγκτή. Αποτελείται από ένα πολυστρωματικό δίκτυο πρόσθιας τροφοδότησης που υλοποιεί τα βασικά στοιχεία ενός κλασσικού ασαφούς ελεγκτή με αντιστοιχία ένα προς ένα. Τα αίτια και τα συμπεράσματα αντιστοιχούν στους κόμβους εισόδου και εξόδου, ενώ στα κρυφά στρώματα οι νευρώνες εκτελούν τις λειτουργίες υπολογισμού των συναρτήσεων συμμετοχής και βάσης κανόνων. Στην αρχιτεκτονική του ελεγκτή του σχήματος κάθε στρώμα αντιστοιχεί ένα προς ένα στα στάδια υλοποίησης ενός ασαφούς ελεγκτή. Ο ελεγκτής επιτυγχάνει την αντιστοιχία σε τέσσερα στρώματα.



Παράδειγμα υβριδικού νευρο-ασαφής ελεγκτή.

Οι κόμβοι στο πρώτο στρώμα της εισόδου τροφοδοτούν τα επόμενα στρώματα και το τέταρτο στρώμα παράγει την έξοδο. Κάθε νευρώνας στο πρώτο στρώμα αντιστοιχεί στις συναρτήσεις συμμετοχής. Στο δεύτερο στρώμα, κάθε νευρώνας αντιπροσωπεύει έναν κανόνα. Οι διασυνδέσεις μεταξύ του δεύτερου και του τρίτου στρώματος είναι το συνδεδεσμένο ισοδύναμο του μηχανισμού συμπερασμού του ελεγκτή. Οι νευρώνες στο τρίτο στρώμα υπολογίζουν τις συναρτήσεις συμμετοχής των αντίστοιχων γλωσσικών μεταβλητών. Στο τέταρτο και τελευταίο στρώμα υπάρχει ένας κόμβος όπου εμφανίζεται η έξοδος του ελεγκτή και ένας νευρώνας από όπου εισάγονται τα υποδείγματα με τα οποία εκπαιδεύεται το δίκτυο.

Οι νευρώνες του δικτύου έχουν χαρακτηριστικά ανάλογα με το στρώμα στο οποίο ανήκουν. Η γενικευμένη σχέση εισόδου-εξόδου κάθε νευρώνα εκφράζεται ως:

$$\omega = f(\sigma), \quad \text{με} \quad \sigma = \sum_{i=1}^p w_i u_i$$

όπου ω είναι η έξοδος και u οι είσοδοι του νευρώνα αντίστοιχα, σ το σταθμισμένο άθροισμα, w τα βάρη και f η συνάρτηση μεταφοράς ή παραμόρφωσης του νευρώνα. Υπάρχουν p διασυνδέσεις και ισάριθμα βάρη σε κάθε νευρώνα.

Στο πρώτο στρώμα, κάθε νευρώνας παραμορφώνει το σταθμισμένο άθροισμα των εισόδων του μη-γραμμικού στοιχείου f ώστε να αντιστοιχεί στην αντίστοιχη συνάρτηση συμμετοχής. Για παράδειγμα, στην περίπτωση συνάρτησης συμμετοχής της μορφής Gauss:

$$\sigma = -\frac{(x_i - \xi_{ij})^2}{s_{ij}^2} \quad \text{και} \quad f(\sigma) = e^{\sigma}$$

όπου ξ_{ij} είναι το κέντρο και s_{ij} η διασπορά της συνάρτησης συμμετοχής. Στην περίπτωση αυτή, τα βάρη των συνδέσεων του πρώτου στρώματος είναι ίσα με τα κέντρα των συναρτήσεων συμμετοχής, δηλαδή $w_{ij} = \xi_{ij}$.

Στο δεύτερο στρώμα οι νευρώνες εκτελούν τον τελεστή ΚΑΙ της συνεπαγωγής min του Mamdani:

$$\sigma = \min(u_1, u_2, \dots, u_p) \quad \text{και} \quad f(\sigma) = \sigma$$

για να συνδυαστούν οι κανόνες που έχουν ενεργοποιηθεί. Όλα τα βάρη στο στρώμα αυτό έχουν τιμή 1.

Στο τρίτο στρώμα οι νευρώνες εκτελούν τον τελεστή H οπότε:

$$\sigma = \sum_{i=l}^p u_i \quad f(\sigma) = \min(l, \sigma)$$

Στο τέταρτο και τελευταίο στρώμα υπάρχουν δύο ειδών κόμβοι. Το πρώτο είδος κόμβων αφορά στην αποασαφοποίηση με την οποία υπολογίζεται η τελική δράση ελέγχου. Το δεύτερο είδος κόμβων σε αυτό το στρώμα μεταφέρει τα υποδείγματα εκπαίδευσης από το τέλος προς τα εμπρός, οπότε $s_i = d_i$.

ΑΝΑΠΤΥΞΗ ΛΟΓΙΣΜΙΚΟΥ ΓΙΑ ΤΗΝ ΠΡΟΒΛΕΨΗ ΤΗΣ ΤΙΜΗΣ ΤΩΝ ΕΠΙΤΟΚΙΩΝ ΜΕ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ

Στο συγκεκριμένο κεφάλαιο παρουσιάζεται το λογισμικό που αναπτύχθηκε για την πρόβλεψη της τιμής των επιτοκίων με νευρωνικά δίκτυα, καθώς επίσης και τα αποτελέσματα που προκύπτουν από τη χρήση του λογισμικού.

4.1 ΠΡΟΔΙΑΓΡΑΦΕΣ ΤΟΥ ΠΡΟΣ ΑΝΑΠΤΥΞΗ ΛΟΓΙΣΜΙΚΟΥ

Πρωταρχικός σκοπός του λογισμικού είναι η πρόβλεψη του επιτοκίου. Ο χρήστης έχει τη δυνατότητα να επέμβει στα χαρακτηριστικά του νευρωνικού δικτύου ώστε να παράγει καλύτερα αποτελέσματα. Πιο ειδικά, το πρόγραμμα δημιουργήθηκε με τη βοήθεια του λογισμικού πακέτου Matlab.

4.2 MATLAB

Η MATLAB είναι μια υψηλού επιπέδου γλώσσα προγραμματισμού και παράλληλα ένα περιβάλλον εργασίας στο οποίο μπορεί να πραγματοποιηθεί ανάπτυξη αλγορίθμου, απεικόνιση στοιχείων – αποτελεσμάτων, ανάλυση δεδομένων, αριθμητικοί υπολογισμοί. Το λογισμικό πακέτο MATLAB καλύπτει ένα μεγάλο φάσμα εφαρμογών, όπως επεξεργασία εικόνων και σημάτων, επικοινωνίες, σχεδιασμός ελέγχου, μετρήσεις, προσομοιώσεις, οικονομικές αναλύσεις κ.ά.. Οι εφαρμογές του MATLAB επεκτείνονται με τη χρήση των πρόσθετων εργαλειοθηκών (MATLAB toolbox).

4.2.1 Εργαλείο εκπαίδευσης νευρωνικών δικτύων του MATLAB

Το εργαλείο εκπαίδευσης νευρωνικών δικτύων του MATLAB (Neural Network Toolbox, NNET) χρησιμοποιεί τη λογική δημιουργίας των νευρωνικών δικτύων και αυτοματοποιεί τις διαδικασίες. Με αυτό τον τρόπο ο χρήστης μπορεί να χρησιμοποιήσει νευρωνικά δίκτυα σε διάφορες εφαρμογές. Το NNET παρέχει τη δυνατότητα στο χρήστη να

δημιουργήσει το νευρωνικό δίκτυο της επιλογής του είτε με χρήση γραφικού περιβάλλοντος (GUI, Graphical User Interface) είτε με το να ορίζει τα χαρακτηριστικά του (μέθοδος εκπαίδευσης, αριθμός νευρώνων, αριθμός κρυφών στρωμάτων, κ.ά.) μέσα από τη γραμμή εντολών του MATLAB.

4.3 ΠΕΡΙΓΡΑΦΗ ΤΟΥ ΝΕΥΡΩΝΙΚΟΥ ΜΟΝΤΕΛΟΥ

Ο αλγόριθμος όπως ήδη αναφέραμε είναι γραμμένος σε γλώσσα προγραμματισμού MATLAB. Στη συγκεκριμένη παράγραφο θα γίνει ανάλυση των σημαντικότερων στοιχείων του κώδικα με στόχο την επίδειξη του τρόπου λειτουργίας και των δυνατοτήτων του συστήματος.

Το μοντέλο μας είναι ένα νευρωνικό δίκτυο πρόσθιας τροφοδότησης (feedforward) το οποίο εκπαιδεύεται από ένα γενετικό αλγόριθμο. Το μοντέλο αποτελείται από τρία στρώματά:

- 1) ένα στρώμα εισόδου το οποίο έχει τέσσερις εισόδους
- 2) δύο κρυμμένα στρώματα που αποτελούνται από 3 νευρώνες το πρώτο και 3 νευρώνες το δεύτερο
- 3) ένα στρώμα εξόδου το οποίο έχει μια έξοδο

Οι 4 δείκτες που επελέγησαν για να αποτελέσουν τις εισόδους του νευρωνικού συστήματος είναι τα εξής οικονομικά μεγέθη:

- πληθωρισμός (inflation rate)
- απασχόληση (unemployment rate),
- μέση τιμή πώλησης οικιών (median price of houses)
- οι τιμές του χρηματιστηριακού δείκτη Dow Jones

Μεταβλητή εξόδου είναι η πρόβλεψη του επιτοκίου της FED. Τα δεδομένα αφορούν το χρονικό διάστημα από τον Ιανουάριο του 1963 μέχρι τον Ιούλιο του 2008 σε μηνιαία βάση για τις Ηνωμένες Πολιτείες Αμερικής, δηλαδή έχουμε 547 παρατηρήσεις.

Στην Εικόνα 4.1 παρουσιάζεται το φύλλο εργασίας στο οποίο είναι αποθηκευμένα τα δεδομένα. Στις 2 πρώτες στήλες αναγράφεται η ακριβής ημερομηνία των μετρήσεων, και στις άλλες 5 στήλες τα δεδομένα εισόδου και εξόδου που προαναφέραμε.

	A	B	C	D	E	F	G
	YEAR	MONTH	BANK PRIME LOAN	MEDIAN PRICE OF HOUSES	INFLATION RATE	UNEMPLOYMENT RATE	DOW JONES
1							
2	1963	1	2,92	17200	1,33	5,70	682,85
3	1963	2	3,00	17700	1,00	5,90	662,94
4	1963	3	2,98	18200	1,33	5,70	682,52
5	1963	4	2,90	18200	0,99	5,70	717,70
6	1963	5	3,00	17500	0,99	5,90	726,96
7	1963	6	2,99	18000	1,32	5,60	706,88
8	1963	7	3,02	18400	1,32	5,60	695,43
9	1963	8	3,49	17800	1,32	5,40	729,32
10	1963	9	3,48	17900	0,99	5,50	732,79
11	1963	10	3,50	17600	1,32	5,50	755,23
12	1963	11	3,48	18400	1,32	5,70	750,52
13	1963	12	3,38	18700	1,64	5,50	762,95
14	1964	1	3,48	17800	1,64	5,60	785,34
15	1964	2	3,48	18000	1,64	5,40	800,14

Εικόνα 4.1

Στο μοντέλο χρησιμοποιείται ένας γενετικός αλγόριθμος για τη εκπαίδευση του δηλαδή την εξεύρεση των καταλλήλων βαρών του μοντέλου έτσι ώστε να ελαχιστοποιηθεί το σφάλμα πρόβλεψης. Ο αλγόριθμος έχει μέγιστο αριθμό γενεών 100 που σημαίνει ότι θα αλλάξει 100 φορές τα βάρη που συνδέονται οι νευρώνες για να μειώσει το σφάλμα πρόβλεψης. Έτσι κάθε φορά που εκπαιδεύει το μοντέλο , κάνει πρόβλεψη, υπολογίζει τα σφάλματα εκπαίδευσης και πρόβλεψης και επιστρέφει πάλι πίσω δίνοντας νέες τιμές στα βάρη (η επόμενη γενιά) από τα οποία θα πάρει νέα πρόβλεψη και νέα σφάλματα (εκπαίδευσης και πρόβλεψης) τα οποία πρέπει να είναι μικρότερα από τα προηγούμενα, άρα θα έχουμε και καλύτερη πρόβλεψη λόγω καλύτερης εκπαίδευσης.

Οφείλεται να αναφερθεί ότι το δείγμα σε κάθε περίπτωση χωρίστηκε σε δύο τμήματα: δεδομένα εκπαίδευσης (in sample) και δεδομένα επαλήθευσης (out of sample), με αναλογία 50%-50%. Δηλαδή χρησιμοποιήσαμε τα πρώτα 273 δεδομένα για την εκπαίδευση του συστήματος, ενώ τα υπόλοιπα δεδομένα τα χρησιμοποιήσαμε για την αξιολόγηση της προβλεπτικής ικανότητας του μοντέλου.

Τα δεδομένα εισόδου και εξόδου περιγράφονται ως εξής:

```

y=xlsread('FedRate_data2','c3:c549'); %Fed Interest Rate
x1=xlsread('FedRate_data2','d3:d549');%House prices
x2=xlsread('FedRate_data2','e3:e549');%Inflation rate
x3=xlsread('FedRate_data2','f3:f549'); % Unemployment rate
x4=xlsread('FedRate_data2','g3:g549'); % Dow Jones Index

```

Στη συνέχεια ένας βρόχος τρέχει όλο το δείγμα των στοιχείων και διαχωρίζει τις μεταβλητές x και y .

Γίνεται βελτιστοποίηση των βαρών από το γενετικό αλγόριθμο.

Το μοντέλο ρυθμίζει τα στοιχεία συμφωνά με τον αριθμό των καθυστερήσεων των μεταβλητών και τα φέρνει σε μορφή πίνακα.

Έπειτα, όπως αναφέραμε, διαχωρίζει τα στοιχεία σε δυο δείγματα: α) το δείγμα εκπαίδευσης δηλαδή τα στοιχεία που χρησιμοποιούνται για την εκπαίδευση του μοντέλου και β) το δείγμα πρόβλεψης δηλαδή τα στοιχεία που χρησιμοποιούνται για να κάνει πρόβλεψη το μοντέλο συμφωνά πάντα με το ποσοστό που έχει οριστεί και είναι 50% για την εκπαίδευση και 50% για τη πρόβλεψη:

Ακολουθώς αφαιρεί την τάση όλων των στοιχείων x και y και των δυο δειγμάτων και μετά συμπιέζει τις τιμές των στοιχείων στο διάστημα 0 έως 1 με τη βοήθεια της μεθόδου Helga Peterson.

Κατόπιν, καθορίζει τον αριθμό των παραμέτρων σε κάθε στρώμα του νευρωνικού .

Στη συνέχεια εκπαιδεύεται το μοντέλο.

Τέλος θα αντιστρέψει την μετατρεπόμενη έξοδο και τα στοιχεία πάλι στις κανονικές τιμές (από 0 έως 1 που είχαν μετατραπεί προηγουμένως), και θα υπολογίσει το σφάλμα με τις κανονικές τιμές αυτή τη φορά.

Η διαδικασία αυτή τέλος επαναλαμβάνεται και για το δείγμα πρόβλεψης.

Ολοκληρώνοντας τα στάδια εκπαίδευσης και πρόβλεψης θα προχωρήσει στη βελτιστοποίηση των βαρών και τη μείωση των σφαλμάτων. Όταν το μοντέλο φτάσει σε αυτό το σημείο της βέλτιστης λύσης υπολογίζει τα σφάλματα RMSE (root mean square error), MAE (Mean Absolute Error), MAPE (Mean Absolute percentage Error), MSE (mean square error), για την εκπαίδευση και τη πρόβλεψη. Επίσης υπολογίζει και το χρόνο που χρειάζεται το μοντέλο για να φτάσει στη λύση. Τα σφάλματα υπολογίζονται ως εξής:

➤ Μέσο τετραγωνικό σφάλμα (MSE)

$$MSE = \frac{1}{N} \cdot \sum_{t=1}^N e_t^2$$

➤ Ρίζα του μέσου τετραγωνικού σφάλματος (RMSE)

$$RMSE = \sqrt{\frac{\sum_{t=1}^N e_t^2}{N}}$$

➤ Μέσο απόλυτο σφάλμα (MAE)

$$MAE = \frac{1}{N} \cdot \sum_{t=1}^N |e_t|$$

➤ Μέσο απόλυτο ποσοστιαίο σφάλμα (MAPE)

$$MAPE = \frac{100}{N} \sum_{i=1}^s \frac{F_t - A_t}{F_t}$$

όπου

A_t -πραγματική τιμή

F_t

$e_t = F_t - A_t$

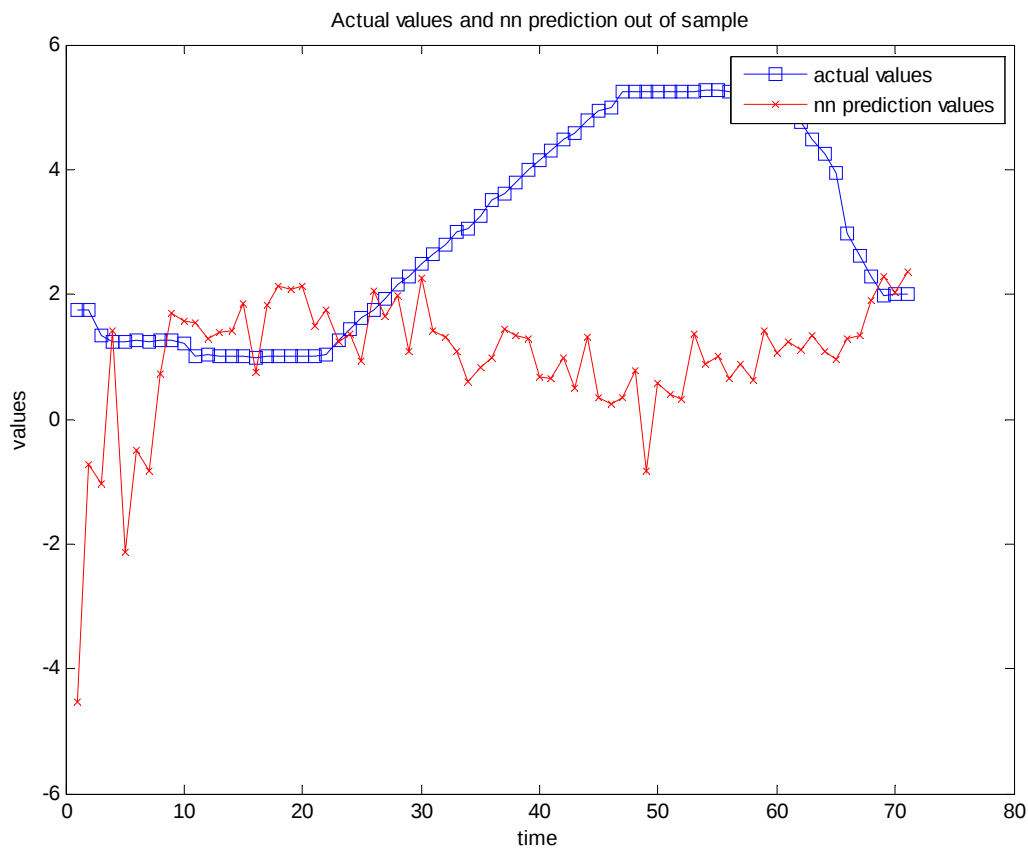
-προβλεπόμενη τιμή

e_t - προβλεπόμενο σφάλμα

4.4 ΑΠΟΤΕΛΕΣΜΑΤΑ ΔΟΚΙΜΩΝ

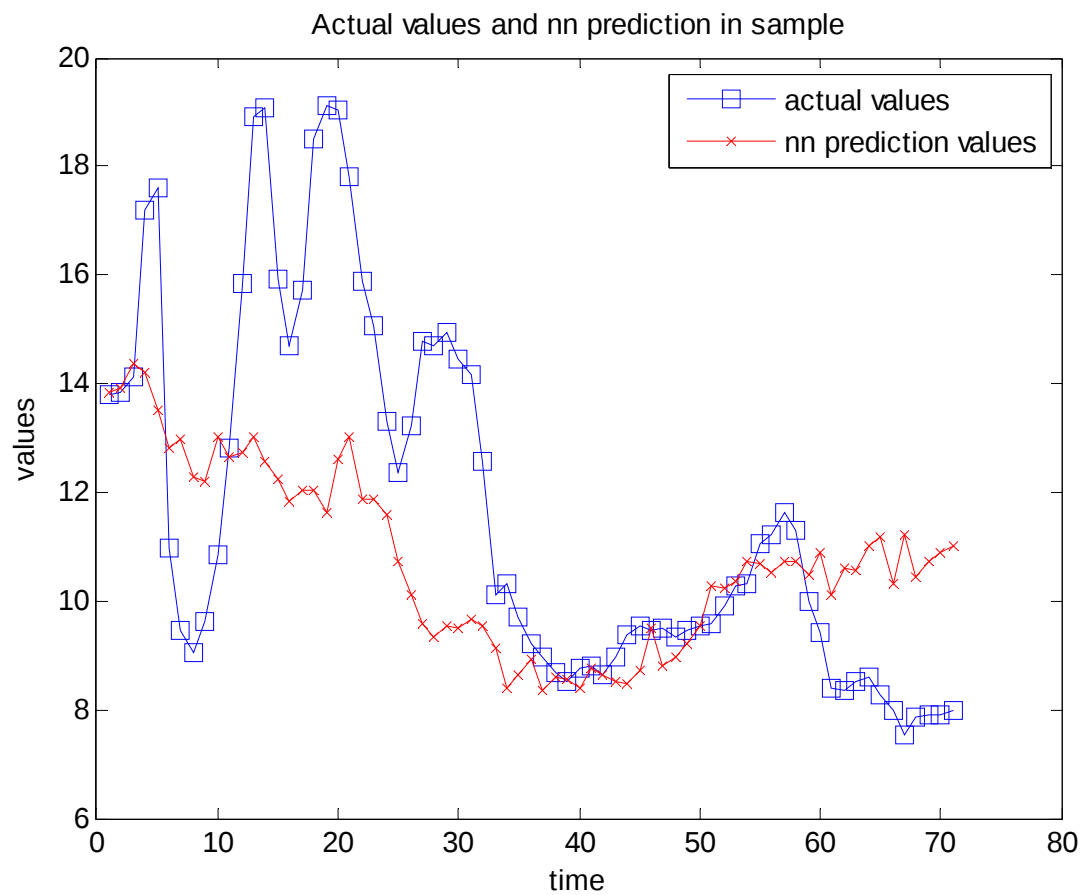
Στην ενότητα αυτή γίνεται παρουσίαση των γραφικών παραστάσεων και αποτελεσμάτων (σφάλματα κ.τ.λ.) του νευρωνικού μας δικτύου. Στη συνέχεια θα γίνει μια σύγκριση αυτών με τα αποτελέσματα δύο άλλων μεθόδων πρόβλεψης, ενός γραμμικού μοντέλου και ενός προσαρμοστικού νεύρο-ασαφές δικτύου (ANFIS) για να διαπιστώσουμε ποια από όλες τις μεθόδους αποδίδει καλύτερα στην πρόβλεψη επιτοκίων.

Στο Σχήμα 4.1 παρουσιάζονται γραφικά οι προβλεπόμενες τιμές των επιτοκίων δηλαδή οι τιμές που υπολογίζει το νευρωνικό δίκτυο (κόκκινη γραμμή) σε σχέση με τις πραγματικές τιμές τους (μπλε γραμμή) από τη χρήση δεδομένων εκτός δείγματος εκπαίδευσης (out of sample).

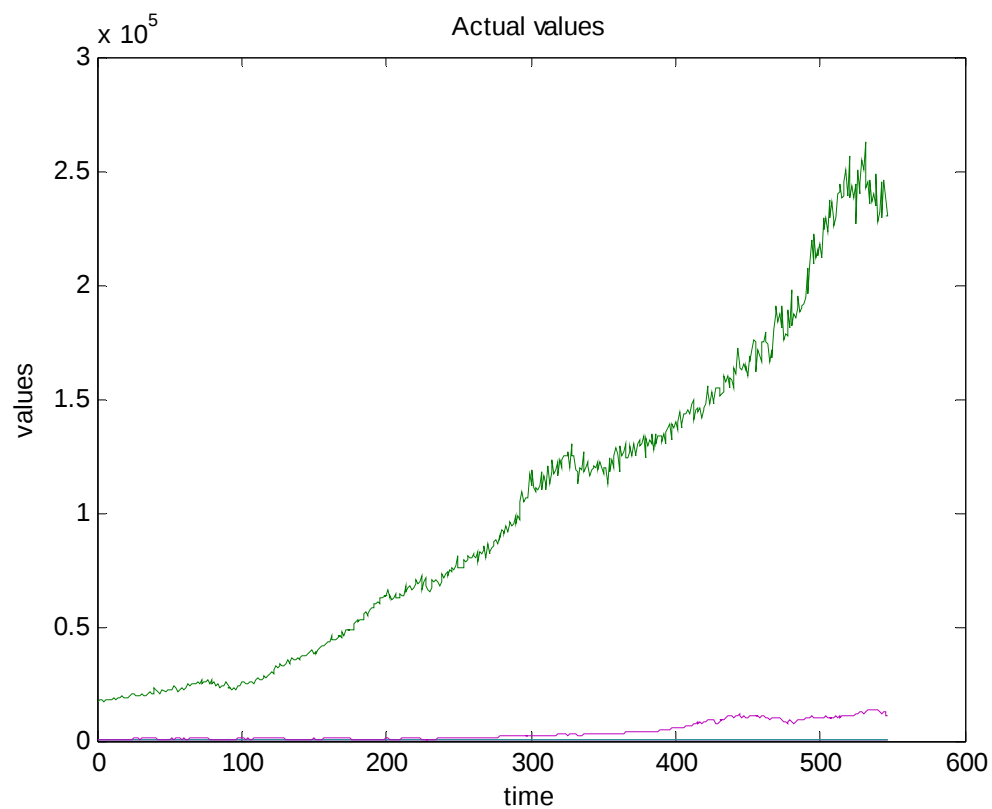


Σχήμα 4.1: Πραγματικές τιμές και πρόβλεψη νευρωνικού δικτύου out of sample

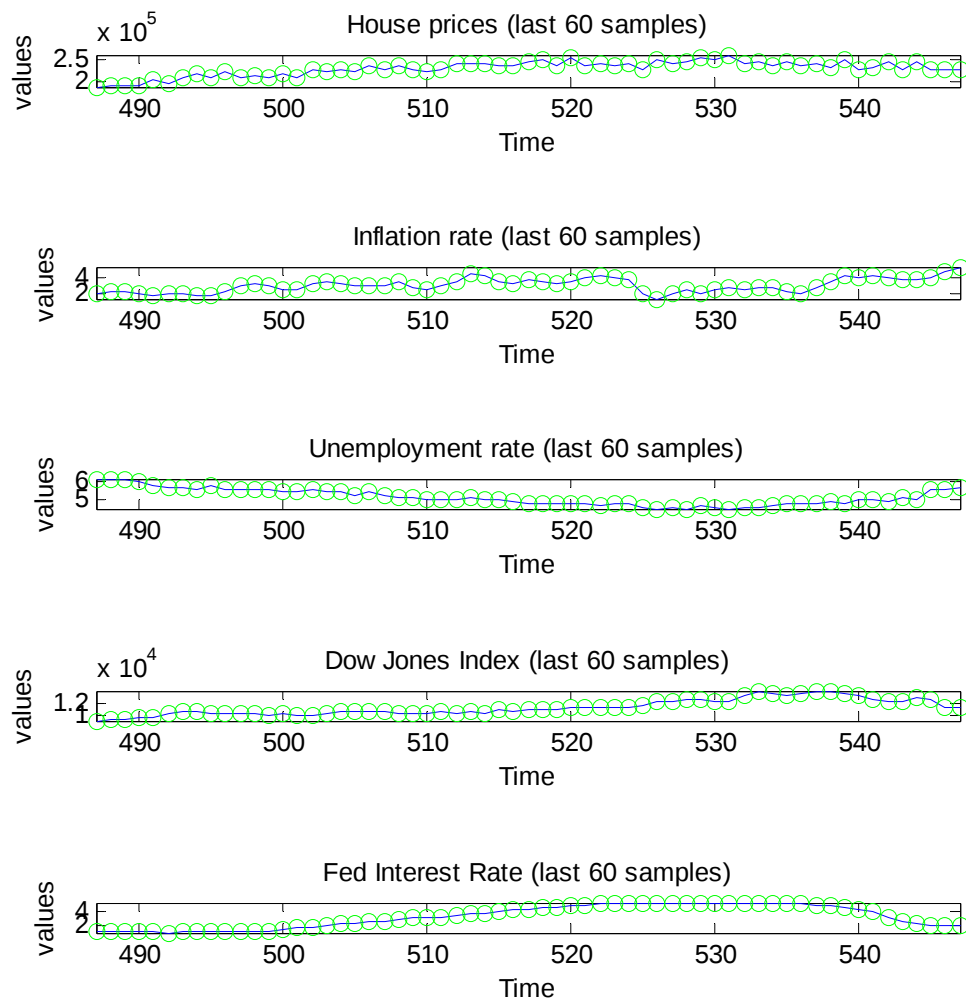
Στο Σχήμα 4.2 βλέπουμε επίσης με κόκκινη γραμμή τις προβλέψεις του νευρωνικού και με μπλε τις πραγματικές τιμές των επιτοκίων, που αφορούν όμως το δείγμα εκπαίδευσης (in sample). Μπορούμε έτσι να δούμε κατά πόσο συγκλίνουν αυτές οι δύο τιμές και συνεπώς πόσο μεγάλο ή μικρό είναι το σφάλμα της εκπαίδευσης του μοντέλου.



Σχήμα 4.2: Πραγματικές τιμές και πρόβλεψη νευρωνικού δικτύου in sample



Σχήμα 4.3: Πραγματικές τιμές συναρτήσει του χρόνου



Σχήμα 4.4: Γραφική παράσταση των μεταβλητών εισόδου

Προκειμένου να υπάρχει ένα κοινό μέτρο σύγκρισης επιλέχθηκε για την περίπτωση των μοντέλων που πραγματεύεται η μελέτη αυτή να υπολογιστούν τα σφάλματα εκπαίδευσης και πρόβλεψης. Χρησιμοποιώντας τα ίδια δεδομένα εισόδου και για τις 3 μεθόδους πρόβλεψης, παραθέτουμε παρακάτω σε πίνακες τα σφάλματα της κάθε μιας. Κρίνουμε σκόπιμο να αναφέρουμε εν συντομία εδώ τη σημασία κάθε σφάλματος έτσι ώστε να

κατανοήσουμε καλύτερα τους δείκτες με τους οποίους θα γίνει σύγκριση της ακρίβειας των μεθόδων.

MSE (Mean Square Error): το μέσο τετραγωνικό σφάλμα δηλαδή παίρνουμε μια μέση τιμή του τετραγώνου του σφάλματος δηλαδή της διαφοράς της πραγματικής από τη προβλεπόμενη τιμή κατά μήκος των δυο δειγμάτων είτε είναι το δείγμα εκπαίδευσης είτε το δείγμα πρόβλεψης οπότε και παίρνουμε και τα αντίστοιχα μέσα τετραγωνικά σφάλματα.

RMSE (Root Mean Square Error): η ρίζα του μέσου τετραγωνικού σφάλματος το οποίο αναλύσαμε προηγουμένως.

MAE (Mean Absolute Error): το μέσο απόλυτο σφάλμα δηλαδή παίρνουμε τη μέση τιμή του απόλυτου σφάλματος δηλαδή της διαφοράς της πραγματικής από τη προβλεπόμενη τιμή κατά μήκος των δυο δειγμάτων εκπαίδευσης και πρόβλεψης και παίρνουμε τα αντίστοιχα μέσα απόλυτα σφάλματα εκπαίδευσης και πρόβλεψης.

MAPE (Mean Absolute percentage Error): μέσο ποσοστιαίο απόλυτο σφάλμα δηλαδή το μέσο απόλυτο σφάλμα εκφρασμένο σε ποσοστό επί τις εκατό.

Σφάλματα νευρωνικού δικτύου (με 5.000 επαναλήψεις) και γραμμικού μοντέλου εντός δείγματος (in sample):

In sample	Γραμμικό	NN
ERRORS		
Root Mean Square Error	2.091380746737908	1.488240314397722
Mean Absolute Error	1.551580843709003	1.078571762181256
Mean Absolute Percentage Error	24.792787754575652	16.209245950602185
Mean Square Error	4.373873427826007	2.214859233398629

Σφάλματα νευρωνικού δικτύου (με 5.000 επαναλήψεις), γραμμικού μοντέλου και ANFIS εκτός δείγματος (out of sample):

Out of sample

	<i>Γραμμικό</i>	<i>NN</i>	<i>ANFIS</i>
ERRORS			
Root Mean Square Error	2.8416192231597	2.7346015177843	6,465025127492
Mean Absolute Error	2.4565584571565	2.1692277625349	6,087217186848
Mean Absolute Percentage Error	133.2854212022169	72.9655005169621	208,732904260571
Mean Square Error	8.0747998094305	7.4780454610683	41,796549899101

Διαπιστώνεται εύκολα από τους παραπάνω πίνακες ότι τα νευρωνικά δίκτυα έδωσαν τα καλύτερα αποτελέσματα, όπως αναμενόταν άλλωστε, τόσο για τα δεδομένα εκπαίδευσης (in sample) όσο και για τα δεδομένα επαλήθευσης (out of sample).

Στην περίπτωση των σφαλμάτων εντός δείγματος το Νευρωνικό δίκτυο δίνει Μέσο Τετραγωνικό Σφάλμα της τάξεως του 2.214, σε σχέση με το γραμμικό μοντέλο που ακολουθεί δεύτερο με εξίσου ικανοποιητικά αποτελέσματα δίνοντας σφάλμα 4.373, αποτελώντας έτσι ένα εναλλακτικό μοντέλο πρόβλεψης σε σχέση με τα νευρωνικά δίκτυα που εφαρμόζονται κυρίως, στην πράξη. Βέβαια το Νευρωνικό Δίκτυο υπερτερεί και στα άλλα 3 σφάλματα με Μέσο απόλυτο σφάλμα 1.078, Ρίζα του μέσου τετραγωνικού σφάλματος 1.488 και Μέσο απόλυτο ποσοστιαίο σφάλμα 16.209, ελαφρώς πιο ικανοποιητικά από τα αντίστοιχα του γραμμικού.

Στην περίπτωση των σφαλμάτων εκτός δείγματος έχουμε και πάλι σταθερή υπεροχή του Νευρωνικού Δικτύου έναντι του γραμμικού, με τις τιμές τους ωστόσο να παρουσιάζουν μικρότερη διαφορά από ότι πριν. Το Μέσο Τετραγωνικό Σφάλμα του πρώτου είναι 7.478 ενώ του δεύτερου 8.074. Στο Μέσο απόλυτο σφάλμα και στη Ρίζα του μέσου τετραγωνικού σφάλματος παρατηρούμε ότι τα αποτελέσματα είναι σχεδόν εφάμιλλα μεταξύ τους, με το νευρωνικό να δίνει 2.169 για το πρώτο και 2.734 για το δεύτερο και το γραμμικό 2.456 και 2.841 αντίστοιχα.

Όσον αφορά το ANFIS που δοκιμάζεται επίσης σε αυτήν την μελέτη, διαπιστώνεται ότι τα σφάλματα των προβλέψεων προκύπτουν αρκετά μεγαλύτερα και άρα τα αποτελέσματα που δίνει δεν είναι ικανοποιητικά. Στην περίπτωση των σφαλμάτων εκτός δείγματος το ANFIS παρουσίασε τη χειρότερη τιμή για το Μέσο Τετραγωνικό Σφάλμα της τάξης του 41,796 σε σχέση με τα άλλα δυο μοντέλα των οποίων τα αποτελέσματα κινήθηκαν σε λογικότερα πλαίσια της τάξης του 7-8. Το ίδιο ισχύει και για το Μέσο απόλυτο ποσοστιαίο σφάλμα (208,732) του οποίου η τιμή είναι

κατά πολύ μεγαλύτερη. Οι τιμές που δίνει το ANFIS στο Μέσο απόλυτο σφάλμα και στη Ρίζα του μέσου τετραγωνικού σφάλματος είναι σαφώς καλύτερες καθώς προσεγγίζουν περισσότερο τις αντίστοιχες του νευρωνικού και του γραμμικού.

Μπορούμε λοιπόν να πούμε με βάση τα αποτελέσματα που μόλις αναλύσαμε ότι τα νευρωνικά δίκτυα αποδείχθηκαν ένα απόλυτα ικανοποιητικό εργαλείο πρόβλεψης της τιμής των επιτοκίων.

5ο ΚΕΦΑΛΑΙΟ

ΣΥΜΠΕΡΑΣΜΑΤΑ ΕΡΓΑΣΙΑΣ

Μέσα στα πλαίσια της παρούσας εργασίας παρουσιάστηκε το θεωρητικό υπόβαθρο των νευρωνικών δικτύων και η εφαρμογή τους σε ένα πρόβλημα πρόβλεψης επιτοκίων.

Σύμφωνα με τα αποτελέσματα της εργασίας, τα νευρωνικά δίκτυα παρουσιάζουν την πιο αξιόπιστη συμπεριφορά συγκριτικά με το γραμμικό μοντέλο και το προσαρμοστικό νεύρο-ασαφές δίκτυο (ANFIS). Σύμφωνα με τα αποτελέσματα που παρουσιάστηκαν στο 3ο Κεφάλαιο, το νευρωνικό δίκτυο έχει:

- Το χαμηλότερο μέσο σφάλμα στο σύνολο ελέγχου, γεγονός που το καθιστά ιδιαίτερα αξιόπιστο μέσο πρόβλεψης και συγκεκριμένα σε ένα δύσκολο πρόβλημα πρόβλεψης όπως αυτή της τιμής των επιτοκίων.
- Μικρός υπολογιστικός χρόνος. Παραγωγή αποτελεσμάτων σε λίγο χρόνο κάτι που καθιστά ιδιαίτερα ευέλικτο το συγκεκριμένο μοντέλο πρόβλεψης.

Αναφερόμενοι πιο συγκεκριμένα στα χαρακτηριστικά του νευρωνικού, το γραφικό περιβάλλον του μοντέλου δίνει τη δυνατότητα στο χρήστη να υπολογίζει τα βέλτιστα χαρακτηριστικά του νευρωνικού δικτύου της επιλογής του. Ακόμη μπορεί να υπολογίζει την επιμέρους βέλτιστη τιμή για το χαρακτηριστικό της επιλογής του χρήστη. Το αρχείο εξόδου του λογισμικού και η παρουσίαση των αποτελεσμάτων του γίνονται σε γραφικό περιβάλλον με το οποίο εντοπίζεται πιο εύκολα τη βέλτιστη συμπεριφορά. Γενικά προγράμματα τα οποία έχουν γραφικό περιβάλλον χρήζουν συνεχούς βελτίωσης. Επίσης το γραφικό περιβάλλον εκτός από το πρακτικό κομμάτι της λειτουργίας του, την εισαγωγή δεδομένων, απεικονίζει τα δεδομένα με λειτουργικό τρόπο και εύχρηστο.

Όσον αφορά το λογισμικό, υπολογίζονται αυτόματα οι τιμές και τα σφάλματα πρόβλεψης και από άλλα μοντέλα πρόβλεψης ώστε τα αποτελέσματα να συγκρίνονται αυτόματα. Παράλληλα μπορεί να γίνει περαιτέρω σύγκριση με άλλες μεθόδους πρόβλεψης, όπου η μέθοδος πρόβλεψης με νευρωνικά δίκτυα θα αξιολογηθεί σε άλλα χαρακτηριστικά πέραν του σφάλματος πρόβλεψης, όπως ο υπολογιστικός χρόνος και η γενική συμπεριφορά του μοντέλου.

ΒΙΒΛΙΟΓΡΑΦΙΑ

Adusei Jumah and Robert M. Kunst (2008), *Optimizing time-series forecasts for inflation and interest rates using simulation and model averaging*, Institute for Advanced Studies, Vienna, Economics Series 231

Atsalakis G., Valavanis K. (2009), *Forecasting Stock Market Short-Term Trends Using a Neuro-Fuzzy Based Methodology*, Journal of Expert Systems With Applications, doi: 10.1016/j.eswa.2009.02.043, Vol **36**, 10696-10707

Atsalakis G., Belonias L. and Zopounidis C. (2008), *Federal funds effective rate forecasting by a neuro-fuzzy model*, submitted in the 7th Hellenic Finance and Accounting Association, December 12-13, 2008, Technical University of Crete, Chania.

Atsalakis, G., K. Valavanis and C. Zopounidis (2008), *Stock Market Prediction Techniques*, (Kleidarithmos Editions, Athens), Greece (in Greek).

David G. McMillan (2007), *Non-linear interest rate dynamics and forecasting: Evidence for US and Australian interest rates*, International Journal of Finance and Economics, Vol **14**, 139-155

Eduardo J. A. Lima, Felipe Luduvise and Benjamin M. Tabak (2006), *Forecasting Interest Rates: an application for Brazil*, Central Bank of Brazil, Research Department, working paper No.120, 1-24

Emilio Domínguez and Alfonso Novales (2002), *Can forward rates be used to improve interest rate forecasts*, Spain, Applied financial economics, vol **12**, 493-504

Eric T. Swanson (2004), *Federal Reserve Transparency and Financial Market Forecasts of Short-Term Interest Rates*, Board of Governors of the Federal Reserve System

Eric T. Swanson (2006), *Have Increases in Federal Reserve Transparency Improved Private Sector Interest Rate Forecasts*, Journal of Money, Credit, and Banking, Vol. **38**, No. 3, 791-819

Hyeonwoo Kim, John Jackson and Richard Saba (2008), *Forecasting the FOMC's Interest Rate Setting Behavior: A Further Analysis*, Journal of Forecasting, 145-165

Ilias Lekkos, Costas Milas and Theodore Panagiotidis (2007), *Forecasting Interest Rate Swap Spreads Using Domestic and International Risk Factors: Evidence from Linear and Non-linear Models*, Journal of Forecasting **26**, 601–619

Jan Filáček, Luboš Komárek and Petr Král (2007), *Why Central Bankers Should Disclose Interest Rate Forecast*, Czech Journal of Economics and Finance, vol **57**, no. 11-12, 558-576

John Barkoulas, Christopher F. Baum and Atreya Chakraborty, *Nearest-Neighbor Forecasts Of U.S. Interest Rates*, Boston College working papers in economics, Vol 313

John C. Robertson and Ellis W. Tallman (1999), *Improving Forecasts of the Federal Funds Rate in a Policy Model*, Federal Reserve Bank of Atlanta, Working Paper 99-3

Lillie Lam, Laurence Fung and Ip-wing Yu (2008), *Comparing forecast performance of exchange rate models*, Hong Kong Monetary Authority, Research Department, Working Paper 08/2008

Mark R. Greer (2005), *Combination Forecasting for Directional Accuracy: an Application to Survey Interest Rate Forecasts*, Journal of Applied Statistics Vol. **32**, No. 6, 607–615

Markus Spiwoks, Nils Bedke and Oliver Hein (2008), *Forecasting the past: the case of US interest rate forecasts*, Financ Mark Portf Manag **22**, 357–379

Pami Dua, Nishita Raje, and Satyananda Sahoo (2003), *Interest Rate Modeling and Forecasting in India*, Reserve Bank of India, Development Research Group Study No. 24

Pär Österholm (2008), *Can Forecasting Performance be Improved by Considering the Steady State? An Application to Swedish Inflation and Interest Rate*, Journal of Forecasting **27**, 41–51

Paul D. McNelis, 2005. “Neural Networks in Finance: Gaining Predictive Edge in the Market”, Elsevier

R. A. Kolb and H. O. Stekler (1996), *How Well Do Analysts Forecast Interest Rates*, Journal of Forecasting, Vol. **15**, 385-394

Rangan Gupta (2007), *Forecasting the South African economy with GIBBS sampled BVECMs*, South African Journal of Economics Vol. **75**

Reid Dorsey-Palmateer and Gary Smith (2007), *Shrunk interest rate forecasts are better forecasts*, Applied Financial Economics, **17**, 425–430\

S. L. Byers and K. Ben Nowman (2001), *Further Evidence on the Forecasting Performance of Two Factor Continuous Time Interest Rate Models in International and Asia-Pacific Financial Markets*, Managerial Finance Vol. **27** No 1/2

Sridhar Iyer and Rick L. Andrews (1999), *Forecasting with Latent Structure Time Series Models: An Application to Nominal Interest Rates*, Journal of Forecasting **18**, 395-409

Tae-Hwan Kim, Paul Mizen and Thanaset Chevapatrakul (2008), *Forecasting Changes in UK Interest Rates*, Journal of Forecasting **27**, 53–74

Thomas F. Gosnell and Robert W. Kolb (1997), *Accuracy of International Interest Rate Forecasts*, The financial review, 431-448

Yvon Fauvel, Alain Paquet and Christian Zimmermann (1999), *A Survey on Interest Rate Forecasting*, Cahier de recherche/Working Paper No. **87**, Université du Québec à Montréal

ΙΣΤΟΣΕΛΙΔΕΣ

<http://prlab.ceid.upatras.gr/index.php?module=courses&coursepg=coursessimeiwseis&courseid=6>

www.economagic.com

<http://www.eia.doe.gov/>

<http://www.inflationdata.com/inflation/>

<http://research.stlouisfed.org/fred2/series/PRIME>

<http://www.wtrg.com/prices.htm>

<http://fraser.stlouisfed.org/>