**Πολυτεχνείο Κρήτης**
**Τμήμα Μηχανικών Παραγωγής και Διοίκησης**

**<u>Διπλωματική εργασία με θέμα</u>:**

# Πρόβλεψη τιμής πετρελαίου χρησιμοποιώντας το νεύρο-ασαφές σύστημα (ANFIS).

**Σπυριδάκος Γρηγόρης**

**Επιβλέπων καθηγητής: Ατσαλάκης Γιώργος**

Τριμελής Επιτροπή: Διδάσκων του Πολυτεχνείου Ατσαλάκης Γεώργιος
Αναπληρωτής Καθηγητής Σκιαδάς Χρήστος
Αναπληρωτής Καθηγητής Μουστάκης Βασίλειος

ΧΑΝΙΑ, ΙΟΥΝΙΟΣ 2010

***<u>Ευχαριστίες</u>***

Θα ήθελα να ευχαριστήσω τον καθηγητή κ. Γεώργιο Ατσαλάκη για την πολύτιμη καθοδήγηση και βοήθεια που μου προσέφερε για την ολοκλήρωση αυτής της διπλωματικής εργασίας.

# ΠΕΡΙΕΧΟΜΕΝΑ

# ΚΕΦΑΛΑΙΟ 1. ΕΙΣΑΓΩΓΗ

Το πετρέλαιο αποτελεί μια από τις σημαντικότερες μορφές ενέργειας στον πλανήτη. Είναι τόσο σημαντικό αν σκεφτεί κανείς ότι εξαιτίας του έχουν γίνει πόλεμοι μεταξύ κρατών. Αποτελεί βασικό μοχλό της παγκόσμιας οικονομίας και η τιμή του επηρεάζει τις παγκόσμιες αγορές. Παράδειγμα αποτελεί η μεγάλη αύξηση της τιμής του 2-3 μήνες πριν και τα επακόλουθα που αυτή έφερε στην παγκόσμια οικονομία. Μέχρι σήμερα έχουν εφαρμοστεί αρκετές μέθοδοι (Κεφ.2) για την πρόβλεψη της τιμής του πετρελαίου. Σκοπός αυτής της διπλωματικής εργασίας είναι η πρόβλεψη της τιμής του πετρελαίου με τη χρήση ενός προσαρμοστικού δικτύου (ANFIS) με δεδομένα παλαιότερες ετήσιες τιμές του από το 1861 μέχρι σήμερα. Ποιο αναλυτικά η δομή της εργασίας αυτής είναι η εξής: στο 2$^{ο}$ κεφάλαιο παρουσιάζονται κάποιες μέθοδοι που έχουν εφαρμοστεί μέχρι σήμερα για την πρόβλεψη της τιμής του πετρελαίου και γίνεται σύγκριση μεταξύ τους. Στο κεφάλαιο 3 περιγράφεται η θεωρία της ασαφούς λογικής αφού πάνω σ’ αυτή βασίζεται το ANFIS,καθώς είναι ένα νεύρο-ασαφές μοντέλο. Στο κεφάλαιο 4 περιγράφονται συστήματα ασαφούς λογικής και νεύρο-ασαφή συστήματα και στο κεφάλαιο 5 γίνεται περιγραφή του νεύρο-ασαφούς συστήματος που εφαρμόσαμε στην εργασία (ANFIS). Τέλος στο κεφάλαιο 6 παρουσιάζονται και αναλύονται όλα τα αποτελέσματα που μας έδωσε το ANFIS.

# ΚΕΦΑΛΑΙΟ 2. ΜΕΛΕΤΕΣ ΠΡΟΒΛΕΨΗΣ ΤΙΜΩΝ ΠΕΤΡΕΛΑΙΟΥ ΠΟΥ ΕΧΟΥΝ ΔΗΜΟΣΙΕΥΤΕΙ ΧΡΗΣΙΜΟΠΟΙΩΝΤΑΣ ΜΟΝΤΕΛΑ ΠΡΟΒΛΕΨΗΣ

Πληθώρα από μοντέλα πρόβλεψης προτείνονται και παρουσιάζονται σε πολλές μελέτες και έρευνες. Όλα τα μοντέλα που παρουσιάζονται επόμενα χρησιμοποιούν σαν δεδομένα εισόδου ετήσιες, τρίμηνες και μηνιαίες τιμές πετρελαίου από περιόδους προηγούμενες από την περίοδο που έγινε η κάθε μελέτη. Σε πολλά από αυτά τα δεδομένα προ-επεξεργάστηκαν πριν χρησιμοποιηθούν. Τα μοντέλα που έχουν χρησιμοποιηθεί είναι τα εξής:

- Support Vector Machines model (SVM)
- Οικονομετρικά μοντέλα
- Γενετικοί Αλγόριθμοι (Genetic Programming)
- Τεχνητά Νευρωνικά Δίκτυα (Artificial Neural Networks, ANN)
- TEI@I Methodology
- Method of Exponential Smoothing
- MLP Based Data Visualization model
- Wavelet-Based Forecasting Procedure
- Relative Inventory (RTSK) Model
- Generalized Pattern Matching Based on Genetic Algorithm (GPMGA)
- Hybrid Artificial Intelligence Model
- Empirical Mode Decomposition Model
- Fuzzy Neural Networks
- Artificial Intelligence Systems
- Semi-parametric Forecasting Methodology Based on the Bootstrap Approach
- Adaptive Neural Fuzzy Inference Systems (ANFIS)
- Financial and Time Series model
- Structural and Time Series Model
- Principal Components Analysis (PCA)
- ARCO1, ARCO2 Models
- Synergetic Models (DWT-LSSVM)

Παρακάτω παρουσιάζεται ο πίνακας που περιέχει όλα τα μοντέλα που αναφέρθηκαν προηγούμενα καθώς και τα αποτελέσματα σύγκρισης μεταξύ των διαφόρων μοντέλων

Πίνακας 1: Μεταβλητές που χρησιμοποιούνται κατά άρθρο

| **Άρθρα (συγγραφέας)** | **Μεταβλητές εισόδου** |
|---|---|
| A New Method for Crude Oil Price Forecasting Based on Support Vector Machines. **Wen Xie, Lean Yu, Shanying Xu, and Shouyang Wang**. (2006) | The crude oil price data used in this study are of West Texas Intermediate (WTI) |
| An Econometrics Forecasting Model of Short Term Oil Spot Price. **Mehrzad Zamani** (2004) | Data has been selected from IEA (International Energy Agency). **WR**: real price of WTI (West Texas Intermediate) **QU**: is OPEC quota **OV**: is OPEC overproduction **SIS**: is deviation from normal level in OECD (Organization for Economic Co-operation and \|Development) industrial stock level **SGS**: is deviation from normal level in SPR and D90 is a dummy variable for Iraq war in third and fourth quarter of 1990. |
| Compumetric Forecasting of Crude oil Prices. **M.A.Kaboudan** (2001) | All data were obtained from the US Department of Energy and Eromagic.com: Economic Time Series. |
| Crude Oil Price forecasting With TEI@I Methodology. **Wang Shouyang,Yu Lean, K.K.Lai** (2005) | Spot prices of WTI crude oil |
| Forecasting Crude Oil spot Price Using OECD Petroleum Inventory Levels. **Michael Ye, Joanne Shore, John Zyren.** (2001) | Spot prices of WTI crude oil |
| Oil Price forecasting with an EMD-Based Multiscale Neural Network Learning Paradigm. Lean Yu, Kin Keung Lai, Shouyang Wang, Kaijian He. (2007) | Spot prices of WTI crude oil |

Πίνακας 1: Μεταβλητές που χρησιμοποιούνται κατά άρθρο

| **Άρθρα (συγγραφέας)** | **Μεταβλητές εισόδου** |
|---|---|
| The forecasting of the world oil price by summing up linear trend and periodic functions.<br>**E.S.Dokuchaev, A.M.Rogacheva, E.V.Evtushenko.**<br>(2006) | Data taken from Energy Information Administration Monthly Review |
| Forecasting short-run crude oil price using high- and low-inventory variables.<br>**Michael Ye, Joanne Shore, John Zyren.**<br>(2005) | Data taken from International Energy Agency (IEA) |
| Accuracy of MLP based data visualization used in oil prices forecasting task.<br>**Aistis Raudys**<br>(2005) | Data taken from:<br>1. WTI<br>2. Natural Gas-Henry hub.<br>3. Fuel Oil<br>4. Gasoline, Unld. Reg. Non-Oxy(NY)<br>5. American Stock Exchange (AMEX) |
| A comparison of var and neural networks with genetic algorithm in forecasting price of oil.<br>**Sam Mirmirani, Hsi Cheng Li**<br>(2004) | 1. Past oil prices<br>2. Petroleum consumption<br>3. Oil supply<br>4. Money supply |
| Wavelet-based prediction of oil prices.<br>**Shahriar Yousefi, Ilona Weinreich, Dominic Reinnarz**<br>(2004) | Spot prices of WTI crude oil |
| A monthly crude oil spot price forecasting model using relative inventories.<br>**Michael Ye, John Zyren, Joanne Shore**<br>(2005) | Spot prices of WTI crude oil and OECD petroleum inventories |
| A generalized pattern matching approach for multi-step prediction of crude oil price.<br>Ying Fan, Qiang Liang, Yi-Ming Wei.<br>(2006) | Daily Brent and WTI crude oil prices, received from IEA website |

Πίνακας 1: Μεταβλητές που χρησιμοποιούνται κατά άρθρο

| **Άρθρα**<br>**(συγγραφέας)** | **Μεταβλητές εισόδου** |
|---|---|
| A hybrid artificial intelligence approach to monthly forecasting of crude oil price time series.<br>**Mohammad Reza Amin-Naseri, Ehsan Ahmadi Gharacheh** | Dataset for WTI is extracted from EIA Website |
| A new approach for crude oil price analysis based on empirical mode decomposition.<br>**Xun Zhang, K.K.Lai, Shou-Yang Wang**<br>(2007) | Data of West Texas Intermediate (WTI) crude oil spot price |
| A New Approach to Forecast Crude Oil Price Based on Fuzzy Neural Network.<br>**Jinlan Liu, Yin Bai, Bin Li**<br>(2007) | RBF neural network, Markov chain based semi parametric model and wavelet analysis based forecasting model are defined as the input of fuzzy neural Network. |
| A Novel Hybrid AI System Framework for crude oil price forecasting.<br>**Shouyang Wang, Lean Yu, K.K. Lai**<br>(2004) | The crude oil price data are spot prices of West Texas Intermediate (WTI) crude oil |
| A semiparametric approach to short-term oil price forecasting.<br>**Claudio Morana**<br>(2001) | GARCH properties of oil price changes |
| A Small Forecasting Model of the world oil market.<br>**Akihiro Amano**<br>(1987) | Data obtained from British Petroleum Company |
| Crude oil price forecasting with ANFIS.<br>B. Zimberg | Past data from EIA (Energy Information Administration), inventory level, volatility index. |
| Do oil prices help to predict future oil prices?<br>FRBSF Economic Letter<br>(2005) | Past oil prices |
| Evaluating the empirical performance of alternative econometric model for oil price forecasting.<br>Chiara Longo, Matteo Manera, Anil Markandya<br>and Elisa Scarpa<br>(2007) | Crude oil prices as reported by EIA. |

Πίνακας 1: Μεταβλητές που χρησιμοποιούνται κατά άρθρο

| **Άρθρα (συγγραφέας)** | **Μεταβλητές εισόδου** |
|---|---|
| Forecasting crude oil price using neural networks.<br>**Komsan Suriya**<br>(2006) | The input information was the oil price range |
| Forecasting the price of crude oil via convenience yield predictions.<br>**Thomas A. Knetsch**<br>(2006) | Data on actual transaction prices of Brent oil |
| Modeling and forecasting oil prices: the role of asymmetric cycles.<br>**Jesus Crespo Cuaresma, Adusei Jumah and Sohbet Karbuz**<br>(2007) | Data from Energy Information Administration, U.S. Department of Energy |
| On the predictive accuracy of crude oil futures prices.<br>Salah Abosedra , Hamid Baghestani<br>(2003) | This study utilizes spot and futures prices of West Texas Intermediate (WTI) crude oil |
| Probabilistic forecasts from probabilistic models: a case study<br>in the oil market<br>Bruce Abramson, Anthony Finizza<br>(1995) | Oil prices refer to West Texas Inter-<br>Mediate (WTI), the US benchmark crude. |
| Using belief networks to forecast oil prices.<br>Bruce Abramson<br>(1991) | Historical, annual, tax, demand, supply, politics, and price variables. |
| Can the dynamics of the term structure of petroleum<br>Futures are forecasted? Evidence from major markets.<br>Thalia Chantziara, George Skiadopoulos.<br>(2007) | Prices on the West Texas Intermediate (WTI) crude<br>oil, heating oil ,and gasoline |
| Crude Oil Price Prediction Based On Multi-scale Decomposition.<br>Yejing Bao, Xun Zhang, Lean Yu, and Shouyang Wang.<br>(2007) | WTI spot price ($/bbl) |

Πίνακας 2: Τεχνικές μοντελοποίησης, επεξεργασίας και συχνότητας δεδομένων

| **Άρθρα** | **Επεξεργασία δεδομένων** | **Συχνότητα δεδομένων** | **Μοντέλα** |
|---|---|---|---|
| A New Method for Crude Oil Price Forecasting Based on Support Vector Machines. **Wen Xie, Lean Yu, Shanying Xu, and Shouyang Wang**. (2006) | — | **Monthly** From January 1970 to December 2003 | Support Vector Machines (SVM) |
| An Econometrics Forecasting Model of Short Term Oil Spot Price. **Mehrzad Zamani** (2004) | — | **Quarterly** For period 1988-2004 | Econometric Model |
| Compumetric Forecasting of Crude oil Prices. **M.A.Kaboudan** (2001) | **Yes** | **Monthly** From January 1993 to December 1998 | Genetic programming, Artificial Neural Networks (ANN) |
| Crude Oil Price forecasting With TEI@I Methodology. **Wang Shouyang,Yu Lean, K.K.Lai** (2005) | — | **Monthly** From January 1970 to December 1999 | TEI@I methodology |
| Forecasting Crude Oil spot Price Using OECD Petroleum Inventory Levels. **Michael Ye, Joanne Shore, John Zyren.** (2001) | — | **Monthly** From January 1992 to February 2001 | A method uses OECD Petroleum Inventory Levels |

Πίνακας 2: Τεχνικές μοντελοποίησης, επεξεργασίας και συχνότητας δεδομένων

| **Άρθρα (Συγγραφέας)** | **Επεξεργασία δεδομένων** | **Συχνότητα δεδομένων** | **Μοντέλα** |
|---|---|---|---|
| Oil Price forecasting with an EMD-Based Multiscale Neural Network Learning Paradigm. **Lean Yu, Kin Keung Lai, Shouyang Wang, Kaijian He.** (2007) | — | **Daily** From January 1, 1998 to October 30, 2006 | EMD-Based Multiscale Neural Network |
| The forecasting of the world oil price by summing up linear trend and periodic functions. **E.S.Dokuchaev, A.M.Rogacheva, E.V.Evtushenko.** (2007) | — | **Monthly** From 1974 to 2005 | Method of exponential smoothing |
| Forecasting short-run crude oil price using high- and low-inventory variables. **Michael Ye, Joanne Shore, John Zyren.** (2005) | **Yes** | **Monthly and Daily** January 1992 to October 2003 | Forecast model for WTI crude oil spot price using total OECD1* industrial inventory as the explanatory variable |
| Accuracy of MLP based data visualization used in oil prices forecasting task. **Aistis Raudys** (2005) | — | **Daily** November 1,1993 to January 2005 | MLP based data visualization |
| A comparison of var and neural networks with genetic algorithm in forecasting price of oil. **Sam Mirmirani, Hsi Cheng Li** (2004) | — | **Monthly** 1980:01 to 2002:11 | Var and neural networks |

1*: Organization for Economic Cooperation and Development

Πίνακας 2: Τεχνικές μοντελοποίησης, επεξεργασίας και συχνότητας δεδομένων

| **Άρθρα (Συγγραφέας)** | **Επεξεργασία δεδομένων** | **Συχνότητα δεδομένων** | **Μοντέλα** |
|---|---|---|---|
| Wavelet-based prediction of oil prices. **Shahriar Yousefi, Ilona Weinreich, Dominic Reinnarz** (2004) | **Yes** | **Monthly** 2 January 1986 to 31 January 2003 | Wavelet-based forecasting procedure |
| A monthly crude oil spot price forecasting model using relative inventories. **Michael Ye, John Zyren, Joanne Shore** (2005) | — | **Monthly** January 1992 to April 2003 | Relative inventory (RSTK) model |
| A generalized pattern matching approach for multi-step prediction of crude oil price. **Ying Fan, Qiang Liang, Yi-Ming Wei.** (2006) | **Yes** | **Daily** From 5/20/1987 to 7/26/2005 | Generalized pattern matching based on genetic algorithm (GPMGA) |
| A hybrid artificial intelligence approach to monthly forecasting of crude oil price time series. **Mohammad Reza Amin-Naseri, Ehsan Ahmadi Gharacheh** | **Yes** | **Monthly** From January 1983 to December 2006 | A hybrid artificial intelligence model. |

Πίνακας 2: Τεχνικές μοντελοποίησης, επεξεργασίας και συχνότητας δεδομένων

| **Άρθρα (Συγγραφέας)** | **Επεξεργασία δεδομένων** | **Συχνότητα δεδομένων** | **Μοντέλα** |
|---|---|---|---|
| A new approach for crude oil price analysis based on empirical mode decomposition. **Xun Zhang, K.K.Lai, Shou-Yang Wang** (2007) | — | **Monthly** From January 1946 to May 2006 | Empirical Mode Decomposition |
| A New Approach to Forecast Crude Oil Price Based on Fuzzy Neural Network. **Jinlan Liu, Yin Bai, Bin Li** (2007) | — | May 20, 1987- August 30, 2006 | Oil price forecasting model based on fuzzy neural network. |
| A Novel Hybrid AI System Framework for crude oil price forecasting. **Shouyang Wang, Lean Yu, K.K. Lai** (2004) | — | **Monthly** From January 1970 to December 2002 | Hybrid AI System |
| A semiparametric approach to short-term oil price forecasting. **Claudio Morana** (2001) | **Yes** | **Daily** November 21, 1998 January 21, 1999 | A semiparametric forecasting Methodology based on the bootstrap approach. |
| A Small Forecasting Model of the world oil market. **Akihiro Amano** (1987) | — | **Monthly** 1986-1991 | An econometric model |
| Crude oil price forecasting with ANFIS. B. Zimberg | **Yes** | **Daily** 1991-2003 | Anfis |

Πίνακας 2: Τεχνικές μοντελοποίησης, επεξεργασίας και συχνότητας δεδομένων

| **Άρθρα (Συγγραφέας)** | **Επεξεργασία δεδομένων** | **Συχνότητα δεδομένων** | **Μοντέλα** |
|---|---|---|---|
| Do oil prices help to predict future oil prices? **FRBSF Economic Letter** (2005) | — | **Monthly** **2 Months** **Quarterly** **Yearly** 1987-2005 | Random walk model, Hotelling's model, futures model, futures-spot spread model |
| Evaluating the empirical performance of alternative econometric model for oil price forecasting. **Chiara Longo, Matteo Manera, Anil Markandya and Elisa Scarpa** (2007) | **Yes** | **Daily, weekly, monthly, and quarterly** 2 January 1986 - 31December 2005 | Financial and time series models, Structural and time series models, |
| Forecasting crude oil price using neural networks. **Komsan Suriya** (2006) | — | **Daily** December 27,2002 to March 18,2005 | Neural network |
| Forecasting the price of crude oil via convenience yield predictions. Thomas A. Knetsch (2006) | — | **Monthly** | An oil price forecasting technique which is based on the present value Model of rational commodity pricing. |
| Modeling and forecasting oil prices: the role of asymmetric cycles. Jesus Crespo Cuaresma, Adusei Jumah and Sohbet Karbuz (2007) | — | **Monthly** January 1983 to August 2007 | Simple time series model with asymmetric cyclical behavior |

Πίνακας 2: Τεχνικές μοντελοποίησης, επεξεργασίας και συχνότητας δεδομένων

| **Άρθρα (Συγγραφέας)** | **Επεξεργασία δεδομένων** | **Συχνότητα δεδομένων** | Μοντέλα |
|---|---|---|---|
| On the predictive accuracy of crude oil futures prices. **Salah Abosedra , Hamid Baghestani** (2003) | — | **Monthly** 1991.01-2001.12 | An empirical method |
| Probabilistic forecasts from probabilistic models: a case study in the oil market **Bruce Abramson, Anthony Finizza** (1995) | — | ***1992** | ARCO2 |
| Using belief networks to forecast oil prices. Bruce Abramson (1991) | **Yes** | **Quarterly 1989** | ARCO1 |
| Can the dynamics of the term structure of petroleum Futures be forecasted? Evidence from major markets | — | **Daily 1/1/1993 to 23/6/2006** | Principal Components Analysis (PCA) |
| Crude Oil Price Prediction Based On Multi-scale Decomposition. Yejing Bao, Xun Zhang, Lean Yu, and Shouyang Wang. (2007) | — | **Monthly and weekly.** training sets **(**2004.01.02-2005.12.30) and testing sets (2006.01.06--2006.12.15) for the weekly data and (training sets: 1999.01-2004.12, testing sets: 2005.01-2006.11) for the monthly data. | A synergetic model (DWT-LSSVM) |

Πίνακας 3: Σύγκριση με άλλα μοντέλα

| Άρθρα | Μοντέλα με τα οποία συγκρίνονται | Καλύτερο μοντέλο | Μέτρα απόδοσης των μοντέλων |
|---|---|---|---|
| A New Method for Crude Oil Price Forecasting Based on Support Vector Machines.<br>**Wen Xie, Lean Yu, Shanying Xu, and Shouyang Wang**.<br>(2006) | ARIMA model<br>BPNN model | Support Vector Machines (SVM) | RMSE<br>Dstat (%) |
| An Econometrics Forecasting Model of Short Term Oil Spot Price.<br>**Mehrzad Zamani**<br>(2004) | - | - | Uses a coefficient of error correction |
| Compumetric Forecasting of Crude oil Prices.<br>**M.A.Kaboudan**<br>(2001) | Random Walk Forecast | Genetic Programming (GP) | Mean Square Errors (MSE) |
| Crude Oil Price forecasting With TEI@I Methodology.<br>**Wang Shouyang,Yu Lean, K.K.Lai**<br>(2005) | ARIMA model<br>ANN model<br>Simple integrated forecasting method | Proposed model | RMSE<br>Dstat (%) |
| Forecasting Crude Oil spot Price Using OECD Petroleum Inventory Levels.<br>**Michael Ye, Joanne Shore, John Zyren.**<br>(2001) | — | — | Adjusted $R^2$<br>Standard error of regration<br>Akaike information criterion<br>RMSE |
| Oil Price forecasting with an EMD-Based Multiscale Neural Network Learning Paradigm.<br>**Lean Yu, Kin Keung Lai, Shouyang Wang, Kaijian He.**<br>(2007) | — | — | Normalized mean squared error (NMSE)<br>$R^2$ |
| The forecasting of the world oil price by summing up linear trend and periodic functions.<br>**E.S.Dokuchaev, A.M.Rogacheva, E.V.Evtushenko.**<br>(2006) | — | — | Accidental impulses $a_t$<br>(the difference between the actual price and the forecasted price) |

Πίνακας 3: Σύγκριση με άλλα μοντέλα

| **Άρθρα** | **Μοντέλα με τα οποία συγκρίνονται** | **Καλύτερο μοντέλο** | **Μέτρα απόδοσης των μοντέλων** |
|---|---|---|---|
| Forecasting short-run crude oil price using high- and low-inventory variables.<br>**Michael Ye, Joanne Shore, John Zyren.**<br>(2005) | — | — | Standard Deviation<br>RMSE<br>MAE<br>Range |
| Accuracy of MLP based data visualization used in oil prices forecasting task.<br>**Aistis Raudys**<br>(2005) | — | — | |
| A comparison of var and neural networks with genetic algorithm in forecasting price of oil.<br>**Sam Mirmirani, Hsi Cheng Li**<br>(2004) | Genetic Algorithm (BPN-GA) | Proposed model (BPN-GA) | RMSE<br>MAE |
| Wavelet-based prediction of oil prices.<br>**Shahriar Yousefi, Ilona Weinreich, Dominic Reinnarz**<br>(2004) | — | — | The correlation coefficient between the predicted and the actual prices |
| A monthly crude oil spot price forecasting model using relative inventories.<br>**Michael Ye, John Zyren, Joanne Shore**<br>(2005) | Naïve autoregressive forecast (NAIV) model and Modified Alternative (MALT) model. | Proposed model | RMSE<br>MAE<br>MAPE<br>The Theil inequality coefficient<br>The bias, variance and covariance |
| A generalized pattern matching approach for multi-step prediction of crude oil price.<br>**Ying Fan, Qiang Liang, Yi-Ming Wei.**<br>(2006) | Elman Network and PMRS method | Proposed model | RMSE<br>MAPE |

Πίνακας 3: Σύγκριση με άλλα μοντέλα

| **Άρθρα** | **Μοντέλα με τα οποία συγκρίνονται** | **Καλύτερο μοντέλο** | **Μέτρα απόδοσης των μοντέλων** |
|---|---|---|---|
| A hybrid artificial intelligence approach to monthly forecasting of crude oil price time series. **Mohammad Reza Amin-Naseri, Ehsan Ahmadi Gharacheh** | Short-Term energy outlook (STEO) model of EIA, genetic programming model of Kaboudan (KAB)AI framework of Wang et al. (WANG) and ANN model of Amin-Naseri and Esfahanian (AMIN) | Proposed model | MAE<br>MAPE<br>Max AE<br>SSE<br>MSE<br>RMSE<br>DA<br>U2-Theil<br>$R^2$ |
| A new approach for crude oil price analysis based on empirical mode decomposition. **Xun Zhang, K.K.Lai, Shou-Yang Wang** (2007) | – | – | Variance |
| A New Approach to Forecast Crude Oil Price Based on Fuzzy Neural Network. **Jinlan Liu, Yin Bai, Bin Li** (2007) | RBF Neural network, Markov chain based Semiparametric model and Wavelet analysis based forecasting model. | Proposed model | MSE |
| A Novel Hybrid AI System Framework for crude oil price forecasting. **Shouyang Wang, Lean Yu, K.K. Lai** (2004) | ANN method | Proposed model | RMSE<br>Dstat (%) |
| A semiparametric approach to short-term oil price forecasting. **Claudio Morana** (2001) | – | – | MSFE, Theil, RMSFE, test of West-Who, *p*-value |
| A Small Forecasting Model of the world oil market. **Akihiro Amano** (1987) | – | – | MAPE |

Πίνακας 3: Σύγκριση με άλλα μοντέλα

| **Άρθρα** | **Μοντέλα με τα οποία συγκρίνονται** | **Καλύτερο μοντέλο** | **Μέτρα απόδοσης των μοντέλων** |
|---|---|---|---|
| Crude oil price forecasting with ANFIS.<br>**B. Zimberg** | Econometric model | Proposed model | RMSE |
| Do oil prices help to predict future oil prices?<br>**FRBSF Economic Letter**<br>(2005) | — | — | Futures-spot spread model |
| Evaluating the empirical performance of alternative econometric model for oil price forecasting.<br>**Chiara Longo, Matteo Manera, Anil Markandya and Elisa Scarpa**<br>(2007) | Mixed models (proposed) | Proposed model | MAE<br>MAPE<br>Theil Inequality Coefficient (Theil)<br>RMSE |
| Forecasting crude oil price using neural networks.<br>**Komsan Suriya**<br>(2006) | — | — | MAPE |
| Forecasting the price of crude oil via convenience yield predictions.<br>**Thomas A. Knetsch**<br>(2006) | — | — | RMSE<br>MEAN ERROR<br>Direction of Change (DOC) |
| Modeling and forecasting oil prices: the role of asymmetric cycles.<br>**Jesus Crespo Cuaresma, Adusei Jumah and Sohbet Karbuz**<br>(2007) | Models using the Diebold-Mariamo test | Proposed model | RMSFE<br>Direction of Change (DOC) |
| On the predictive accuracy of crude oil futures prices.<br>Salah Abosedra , Hamid Baghestani<br>(2003) | — | — | |
| Probabilistic forecasts from probabilistic models: a case study in the oil market<br>Bruce Abramson, Anthony Finizza<br>(1995) | — | — | Standard Deviation |

Πίνακας 3: Σύγκριση με άλλα μοντέλα

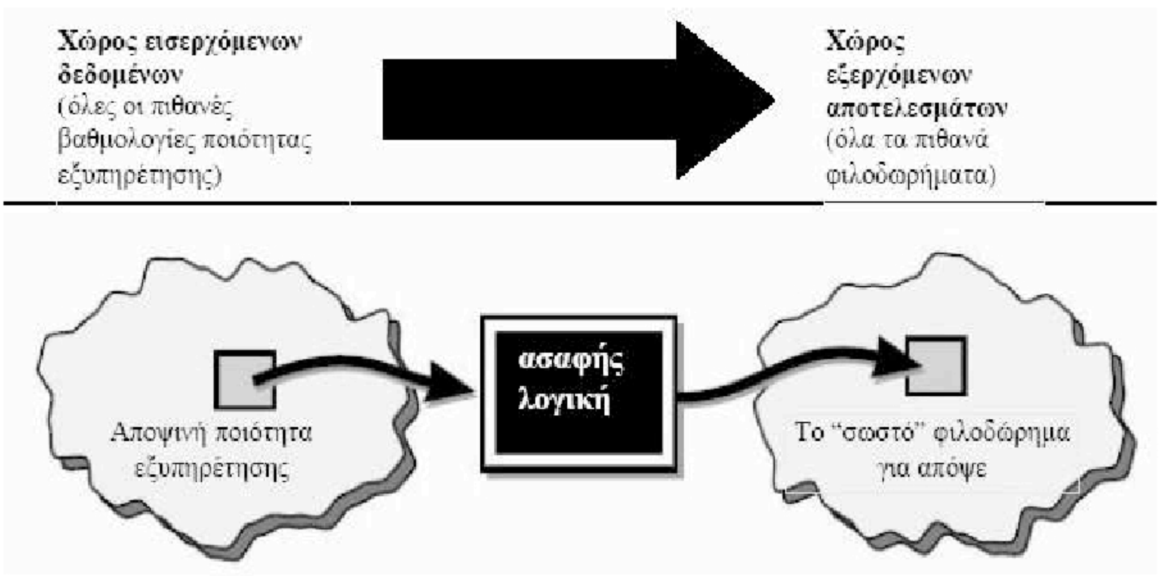
| **Άρθρα** | **Μοντέλα με τα οποία συγκρίνονται** | **Καλύτερο μοντέλο** | **Μέτρα απόδοσης των μοντέλων** |
|---|---|---|---|
| Using belief networks to forecast oil prices.<br>**Bruce Abramson**<br>(1991) | — | — | Means and standard deviations with Monte Carlo analysis |
| Can the dynamics of the term structure of petroleum Futures be forecasted? Evidence from major markets | AR(1) Model, VAR (1) model and ARMA(1,1) | Proposed model is equal to the Var model | RMSE<br>MAE<br>U:Theil Inequality Coefficient (Theil)<br>MCP: Mean Correct Prediction |
| Crude Oil Price Prediction Based On Multi-scale Decomposition.<br>Yejing Bao, Xun Zhang, Lean Yu, and Shouyang Wang.<br>(2007) | — | — | MARE<br>NMSE<br>Dstat |

# ΚΕΦΑΛΑΙΟ 3: ΑΣΑΦΗΣ ΛΟΓΙΚΉ

## *3.1. ΘΕΩΡΊΑ ΤΗΣ ΑΣΑΦΟΥΣ ΛΟΓΙΚΗΣ*

Η ασαφής λογική είναι μια μεθοδολογία που ασχολείται με την έννοια της ακρίβειας.


Σχήμα 3.1:Ακρίβεια και σημαντικότητα




Σχήμα 3.2. Χάρτης Εισόδου-Εξόδου

Η έννοια της ακρίβειας απεικονίζεται στην παραπάνω εικόνα 3.1. 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.

Όπως στο παράδειγμα, πόσο φιλοδώρημα να αφήσουμε σε ένα μαγαζί, ένας χάρτης εισόδου-εξόδου φαίνεται στο σχήμα 3.2.

## *3.2 ΘΕΜΕΛΙΑ ΑΣΑΦΟΥΣ ΛΟΓΙΚΗΣ*

### 3.2.1 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 γίνεται υπόθεση ενός βαθμού.*

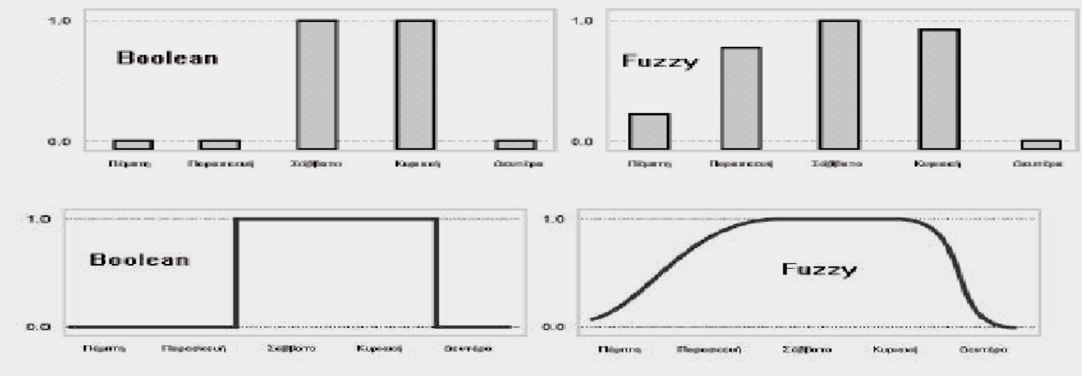
### 3.2.2 ΑΣΑΦΕΙΑ

Το 1963 ο Zadeh θεμελίωσε τη θεωρία των ασαφών συνόλων σα 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-ΟΧΙ 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.




Σχήμα 3.3: Ημέρες που θεωρούμε ότι ανήκουν στο σαββατοκύριακο.

Όπως για παράδειγμα στο σχήμα 3.3 φαίνονται οι αληθείς τιμές των ημερών που θεωρούμε ότι ανήκουν στο σαββατοκύριακο, θεωρώντας ότι ο χρόνος παίρνει διακεκριμένες τιμές, πρώτα με ΝΑΙ-ΟΧΙ και δίπλα απαντήσεις Ασαφούς /Λογικής με ενδιάμεσες τιμές.

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

## 3.2.3 ΣΥΝΑΡΤΗΣΕΙΣ ΣΥΜΜΕΤΟΧΗΣ.

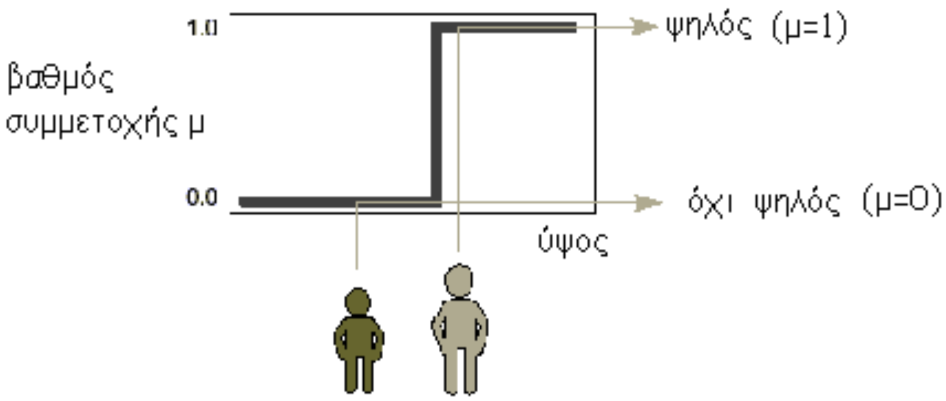
Ένα ασαφές σύνολο Α ορίζεται ως ένα σύνολο διατεταγμένων ζευγών (x, $\mu_A(x)$ όπου $x \in X$ και $\mu_A(x) \in [0,1]$ ). Το σύνολο X αποτελεί ένα ευρύτερο σύνολο αναφοράς (universe of discourse) που περιλαμβάνει όλα τα αντικείμενα στα οποία μπορεί να γίνει αναφορά. Η τιμή $\mu_A(x)$ λέγεται βαθμός αλήθειας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

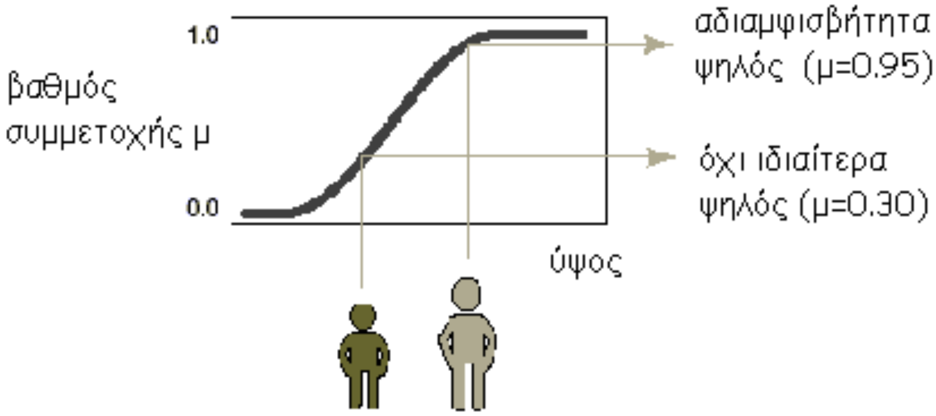
Παρακάτω αναπτύσσεται ένα παράδειγμα με το οποίο μπορεί να γίνει περισσότερο κατανοητή η έννοια της συνάρτησης συμμετοχής.

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 2.30 . 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 1.73m. τότε ένας άνθρωπος με ύψος 1.74 θα χαρακτηρίζεται κοντός ενώ ένας άνθρωπος με ύψος 1.76 θα χαρακτηρίζεται ψηλός. Ο παραπάνω διαχωρισμός φαίνεται παράλογος αφού έχουμε αντιστοιχήσει σε δύο ανθρώπους με αμελητέα διαφορά ύψους δύο αντίθετες μεταξύ τους έννοιες.




Σχήμα 3.4: Διακριτές συναρτήσεις συμμετοχής, για ψηλούς.

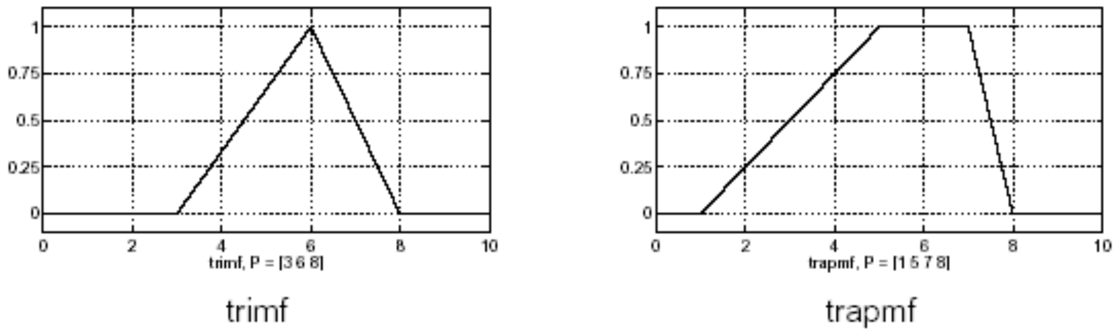
Ένας άλλος τρόπος να ορίσουμε την έννοια ψηλός είναι μέσω μιας καμπύλης που έχει ομαλή διακύμανση και μεταβαίνει από την έννοια ψηλός στην έννοια κοντός. Αυτή η καμπύλη είναι η συνάρτηση συμμετοχής του ασαφούς συνόλου των ψηλών ανθρώπων. Με άλλα λόγια δεχόμαστε ότι όλοι οι άνθρωποι είναι σε κάποιο βαθμό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




Σχήμα 3.5: 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. Αυτές οι δύο συναρτήσεις εξασφαλίζουν την απαίτηση για απλότητα.




Σχήμα 3.6: Τριγωνική και τραπεζοειδής συνάρτηση συμμετοχής.

Η μαθηματική έκφραση της τριγωνικής συνάρτησης συμμετοχής είναι η εξής:

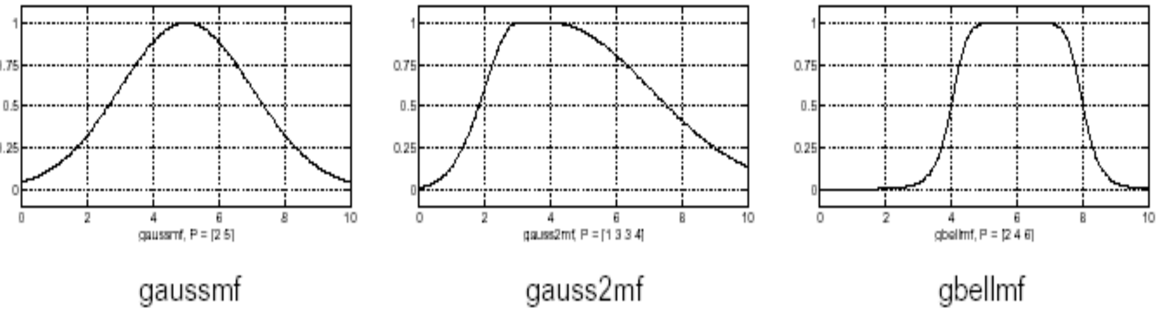
$$A = \begin{cases} 0, x \leq a \\ (x-a)/(b-a), x \in (a,b) \\ (c-x)/c-b), x \in (b,c) \\ 0, x \geq c \end{cases}$$

Παρακάτω δίδεται η μαθηματική έκφραση της τραπεζοειδούς συνάρτησης συμμετοχής:

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

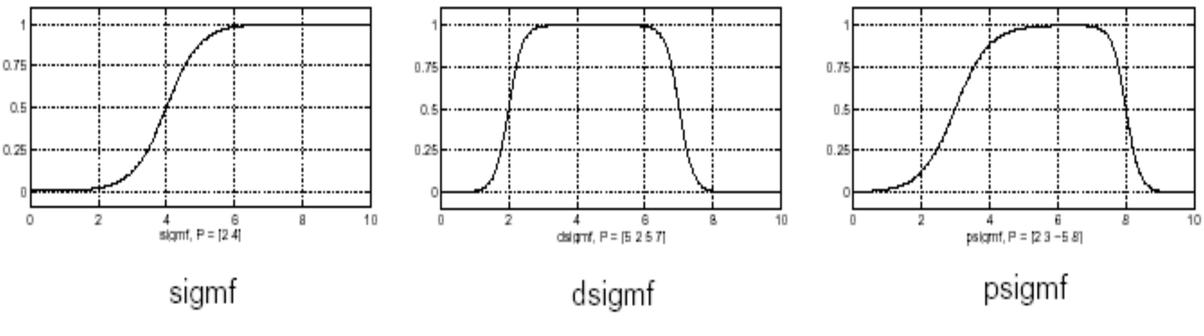
Δύο συναρτήσεις συμμετοχής που είναι δομημένες πάνω στη μορφή της κατανομής Gauss είναι μια απλή γκαουσιανή και μια σύνθεση δύο διαφορετικών γκαουσιανών.

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



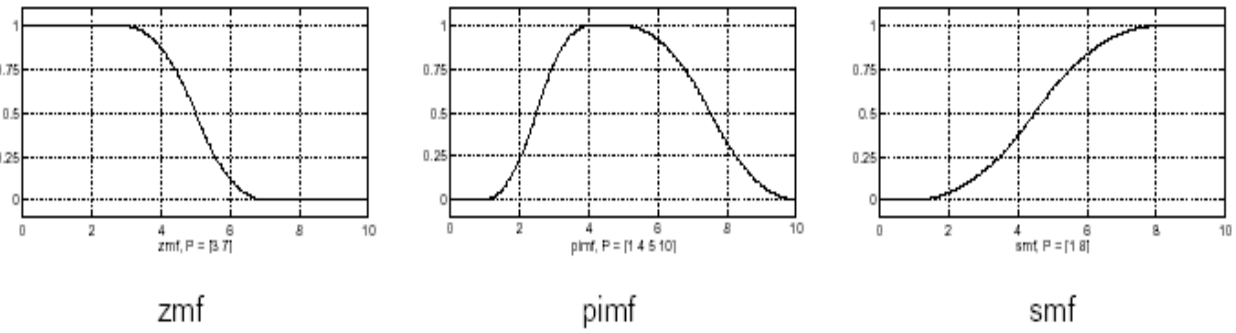

Σχήμα 3.7: Γκαουσιανές και καμπανοειδείς συναρτήσεις συμμετοχής.

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 3.8: Σιγμοειδείς συναρτήσεις συμμετοχής.

Επίσης υπάρχουν πολλές πολυωνυμικές καμπύλες που τις χρησιμοποιούμε σα 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 πολυωνυμική καμπύλη που είναι ανοικτή στα αριστερά, η S είναι η κατοπτρική της Z και η Π είναι μια ασύμμετρη κλειστή καμπύλη σχήματος Π.




Σχήμα 3.9: Πολυωνυμικές καμπύλες.

## *3.3 ΙΔΙΟΤΗΤΕΣ ΤΩΝ ΑΣΑΦΩΝ ΣΥΝΟΛΩΝ*

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 $\mu_A$ από το X στο διάστημα [0 1]. Το A είναι ένα ασαφές υποσύνολο του X , το οποίο όμως δεν έχει αυστηρά καθορισμένα σύνορα. $\mu_A$ είναι ο βαθμός συμμετοχής του στοιχείου x στο A. Όσο πιο κοντά στο 1 είναι το $\mu_A$ τόσο πιο πολύ ανήκει το x στο A.

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

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

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.

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

Το σύνολο A μπορεί να γραφεί και ως:

$\mathbf{A=\mu_1/x_1+\mu_2/x_2+..}$

**$A=\Sigma\mu_i/x_i$**

Στο συνεχή χώρο αναφοράς :

$$A = \int_x \mu_{\mathrm{A}}(x) / x$$

***Ασαφές δυναμοσύνολο***

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

***Υποσύνολο***

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

$$\mu_{\mathrm{A}}(x) \leq \mu_B(x), \forall x \in X$$

Αν ταυτόχρονα τα Α και Β δεν είναι ίσα , τότε το Α θα ονομάζεται γνήσιο υποσύνολο του Β

***Ασαφής διαμέριση***

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

$$A_j \neq A_i, \forall i, j \in N_n (i \neq j)$$

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

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 Α είναι μηδέν.

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

***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= κανονικό αν $\exists x_1 : \mu_{\mathrm{A}}(x_1) = 1$

***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 αν $\mu_{\mathrm{A}}(x) = \mu_{\mathrm{B}}(x) \quad \forall x \in X$

***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

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

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:

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

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

## *3.4 ΛΟΓΙΚΕΣ ΠΡΑΞΕΙΣ ΣΤΑ ΑΣΑΦΗ ΣΥΝΟΛΑ*

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 3.10: Λογικοί κανόνες.

Μεταβαίνοντας στον χώρο της ασαφούς λογικής πρέπει να έχομε 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 |

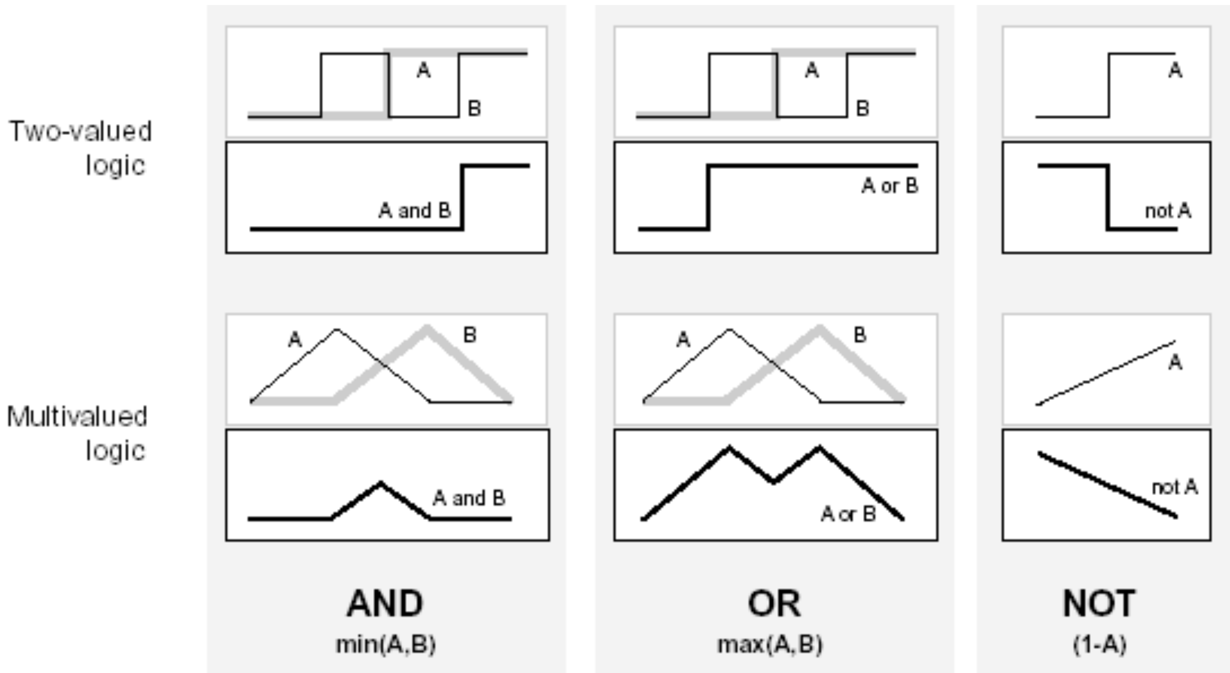
**OR**

| A | 1 - A |
|---|---|
| 0 | 1 |
| 1 | 0 |

**NOT**

Σχήμα 3.11: 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 3.12: 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.

Έχομε ορίσει την τομή ,την ένωση και το συμπλήρωμα ενός ασαφούς συνόλου με τους τελεστές min max 1-A οι όποιοι 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 Τ 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 Τ μπορεί να αντιπροσωπεύει τον πολλαπλασιασμό των $m_A(x)$ και $m_B(x)$. 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: $T(a, b) <= T(c, d)$ αν $a <= c$ και $b <= d$
- Αντιμεταθετικότητα: $T(a, b) = T(b, a)$
- Προσεταιριστικότητα: $T(a, T(b, c)) = T(T(a, b), c)$

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

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

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

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

Παραδείγματα τ-νορμών είναι:

***Συνήθης τομή*** 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 $m_A$(x) και $m_B$ (*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)* <= S(c, d) αν *a <= c καιb <= 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 Α' ενός ασαφούς συνόλου Α δίνεται από τη σχέση:

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

όπου η συνάρτηση c πρέπει να ικανοποιεί τις παρακάτω συνθήκες;

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

Το σύνηθες συμπλήρωμα δίνεται από τη σχέση:

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

## *3.5 ΑΣΑΦΕΙΣ ΣΧΕΣΕΙΣ*

Οι ασαφείς σχέσεις (fuzzy relations) είναι ασαφή σύνολα ορισμένα σε πεδία αναφοράς ανώτερης διάστασης (π.χ. X x X ,X x Y x 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= «x είναι βαρύτερο από y» x∈X, y∈Y και R∈X x 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:

$$R = \begin{bmatrix} \mu_R(x_{1,} y_1) & \mu_R(x_1, y_2) & \cdots & \mu_R(x_1, y_n) \\ \mu_R(x_2, y_1) & \mu_R(x_2, y_2) & \cdots & \mu_R(x_2, y_n) \\ \cdot & & & \\ \cdot & & & \\ \mu_R(x_m, y_1) & \mu_R(x_m, y_2) & \cdots & \mu_R(x_m, y_n) \end{bmatrix}$$

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 x Y με την ασαφή σχέση $R_2(y,z)$ ορισμένη στο Y x Z τότε θα προκύψει μία ασαφής σχέση R(x,z) η οποία θα ορίζεται στο σύνολο X x 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. O 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 x Y και Y x Z αντίστοιχα , τότε η σύνθεση τους δίνει μια νέα σχέση $R_1 \circ R_2$

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

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

Οι περισσότερο 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:

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

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

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

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

## *3.6 ΑΣΑΦΗΣ ΣΥΛΛΟΓΙΣΤΙΚΗ*

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:

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

O **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 Α συνεπάγεται το γεγονός Β και έχομε ως υπόθεση ότι ισχύει το Α σε κάποιο βαθμό , τότε θα ισχύει και το Β σε κάποιο βαθμό.

Πρέπει ,λοιπόν να προσδιοριστεί ο βαθμός στον οποίο πληρείται το γεγονός Β. Αυτός εξαρτάται από το βαθμό στον οποίο πληρείται το γεγονός Α και από το είδος της συνεπαγωγής που εφαρμόζουμε. Η πράξη της ασαφούς συνεπαγωγής υλοποιεί τη μαθηματική σχέση Α ==> Β, όταν τα Α και Β είναι ασαφή γεγονότα..

Το σχήμα που προτείνεται από τη θεωρία των ασαφών συνόλων για την εξαγωγή του γεγονότος Β’ από τα γεγονότα Α ,Α’,Β 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 Α=>Β , δηλαδή το γεγονός Β. η απαίτηση αυτή είναι εύλογη αφού όταν δεν υπάρχει αβεβαιότητα η ασαφής συλλογιστική οφείλει να ταυτίζεται με την κλασσική συλλογιστική.

# ΚΕΦΑΛΑΙΟ 4. ΣΥΣΤΗΜΑΤΑ ΑΣΑΦΟΥΣ ΣΥΛΛΟΓΙΣΤΙΚΗΣ ΚΑΙ ΝΕΥΡΟ-ΑΣΑΦΗ ΣΥΣΤΗΜΑΤΑ

## *4.1 ΕΙΣΑΓΩΓΗ*

Στο προηγούμενο κεφάλαιο μελετήσαμε τα ασαφή σύνολα και τις πράξεις που γίνονται πάνω σ' αυτά. Στο παρόν κεφάλαιο εξετάζουμε τον τρόπο που μπορούν να εφαρμοσθούν τα παραπάνω σε ένα σύστημα λήψης αποφάσεων.

Θα εξετάσουμε ένα απλοϊκό παράδειγμα λήψης αποφάσεων και θα δούμε πως μπορεί να λυθεί με τη βοήθεια της ασαφούς λογικής. Εξετάζουμε το πρόβλημα του φιλοδωρήματος, πόσο δηλαδή είναι το φιλοδώρημα που πρέπει να δώσει ένας πελάτης σε ένα εστιατόριο. Οι παράμετροι του προβλήματος είναι η ποιότητα του φαγητού και η εξυπηρέτηση.

Είναι προφανές για παράδειγμα ότι αν το φαγητό είναι άριστης ποιότητας και η εξυπηρέτηση καλή τότε το φιλοδώρημα θα είναι υψηλό. Αντίστοιχα αν η ποιότητα της εξυπηρέτησης είναι χαμηλή και το φαγητό κακό τότε το φιλοδώρημα θα είναι χαμηλό. Τι γίνεται όμως αν έχουμε μια ενδιάμεση κατάσταση, π.χ. μέτριο φαγητό και καλή εξυπηρέτηση. Επιπλέον παραπάνω έχουμε αναφέρει προσδιορισμούς όπως καλό , καλό , μέτριο που εμπεριέχουν κάποια ασάφεια.

Αν προσπαθήσουμε να λύσουμε το πρόβλημα γραμμικά θα πρέπει να δώσουμε καθορισμένους κανόνες για κάθε μια περίπτωση. Μέσω όμως της ασαφούς λογικής το πρόβλημα αυτό μπορεί να λυθεί ευκολότερα. Πριν προχωρήσουμε στη λύση του προβλήματος θα δούμε πως μπορούμε να υλοποιήσουμε τους κανόνες αν-τότε (if-then rules) που προαναφέραμε (αν το φαγητό είναι κακό κτλ)

## *4.2 Κανόνες αν-τότε (ιf-then rules)*

Τα ασαφή σύνολα και οι ασαφείς τελεστές είναι τα υποκείμενα και τα ρήματα του ασαφούς λογισμού. Οι προτάσεις if-then χρησιμοποιούνται για να σχηματίσουν τις συνθήκες εκείνες που συνιστούν την ασαφή λογική. Ένας απλός κανόνας if-then έχει τη μορφή:

***if x is A then y is B***

όπου A και B είναι οι γλωσσικές μεταβλητές που προσδιορίζονται από ασαφή σύνολα με χώρο αναφοράς X και Y αντίστοιχα. Το πρώτο κομμάτι του κανόνα 'if *x* is *A'* ονομάζεται υπόθεση (*antecedent ή* premise) ενώ το δεύτερο κομμάτι 'then *y* is *B'* ονομάζεται συνέπεια ή συμπέρασμα (*consequent* ή conclusion). Ένα παράδειγμα ενός τέτοιου κανόνα μπορεί να είναι το εξής:

***if service is good then tip is average***

Η λέξη ***good*** αντιπροσωπεύεται από έναν αριθμό που κυμαίνεται ανάμεσα στο 0 και το 1. Έτσι η υπόθεση είναι μια μετάφραση που επιστρέφει μια τιμή από το 0 έως το 1. Από την άλλη μεριά η λέξη ***average*** αποτελεί ένα ασαφές σύνολο και έτσι η συνέπεια του κανόνα είναι μια συσχέτιση της εξόδου y στο ασαφές σύνολο B.

Στους κανόνες if-then η λέξη 'is' έχει διαφορετική έννοια ανάλογα με το αν εμφανίζεται στο πρώτο ή στο δεύτερο μέρος του κανόνα. Έτσι όταν εμφανίζεται στο πρώτο μέρος του κανόνα έχει την έννοια του ελέγχου ισότητας (δηλαδή την ίδια έννοια που δίνουμε στο σύμβολο = =). Αντίθετα όταν εμφανίζεται στο δεύτερο μέρος του κανόνα έχει την έννοια της καταχώρησης (δηλαδή την έννοια που δίνουμε στο σύμβολο =). Ο παραπάνω κανόνας λοιπόν θα μπορούσε να γραφεί συμβολικά και ως εξής:

***if service == good then tip = average***

Γενικά η είσοδος σε έναν κανόνα if-then είναι η τρέχουσα τιμή της μεταβλητής εισόδου ενώ η έξοδος του κανόνα είναι ολόκληρο το ασαφές σύνολο. Από αυτό όμως το ασαφές σύνολο θα πρέπει να κρατήσουμε μια τιμή έτσι ώστε να μπορούμε να βγάλουμε κάποια συμπεράσματα. Γι' αυτό πρέπει να γίνει μια διαδικασία που ονομάζεται αποσαφήνιση (defuzzyfication). Για την αποσαφήνιση αναφερόμαστε εκτενέστερα παρακάτω. Για να διερμηνεύσουμε έναν κανόνα If-then πρέπει να ακολουθήσουμε κάποια στάδια:

1. Πρώτα πρέπει να υπολογίσουμε την είσοδο. Για να γίνει αυτό όμως πρέπει να γίνει ασαφοποίηση της εισόδου και να εφαρμοσθούν οι κατάλληλοι ασαφείς τελεστές

2. Να εφαρμόσουμε το παραπάνω αποτέλεσμα στο δεύτερο μέρος του κανόνα

Αν πάρουμε την απλή περίπτωση των δυαδικών τιμών οι κανόνες if-then είναι αρκετά εύκολο να εφαρμοσθούν. Αν η προϋπόθεση είναι αληθής τότε και η συνέπεια θα είναι αληθής. Πώς όμως επεκτείνεται η παραπάνω συλλογιστική για να συμπεριλάβει και ασαφείς τιμές; Όταν η προϋπόθεση είναι αληθής σε κάποιο βαθμό τότε και η συνέπεια είναι αληθής στον ίδιο βαθμό. Με άλλα λόγια:

***Στη δυαδική λογική: p ® q (Τα p και q είναι και τα δύο είτε πλήρως αληθή ή πλήρως ψευδή)***

***Στην ασαφή λογική: 0.5 p ® 0.5 q (Τα p και q είναι μερικώς αληθή και μερικώς ψευδή)***

Η προϋπόθεση ενός κανόνα μπορεί να αποτελείται από περισσότερα του ενός μέρη. Όπως για παράδειγμα ο παρακάτω κανόνας:

***if sky is gray and wind is strong and barometer is falling, then ..***

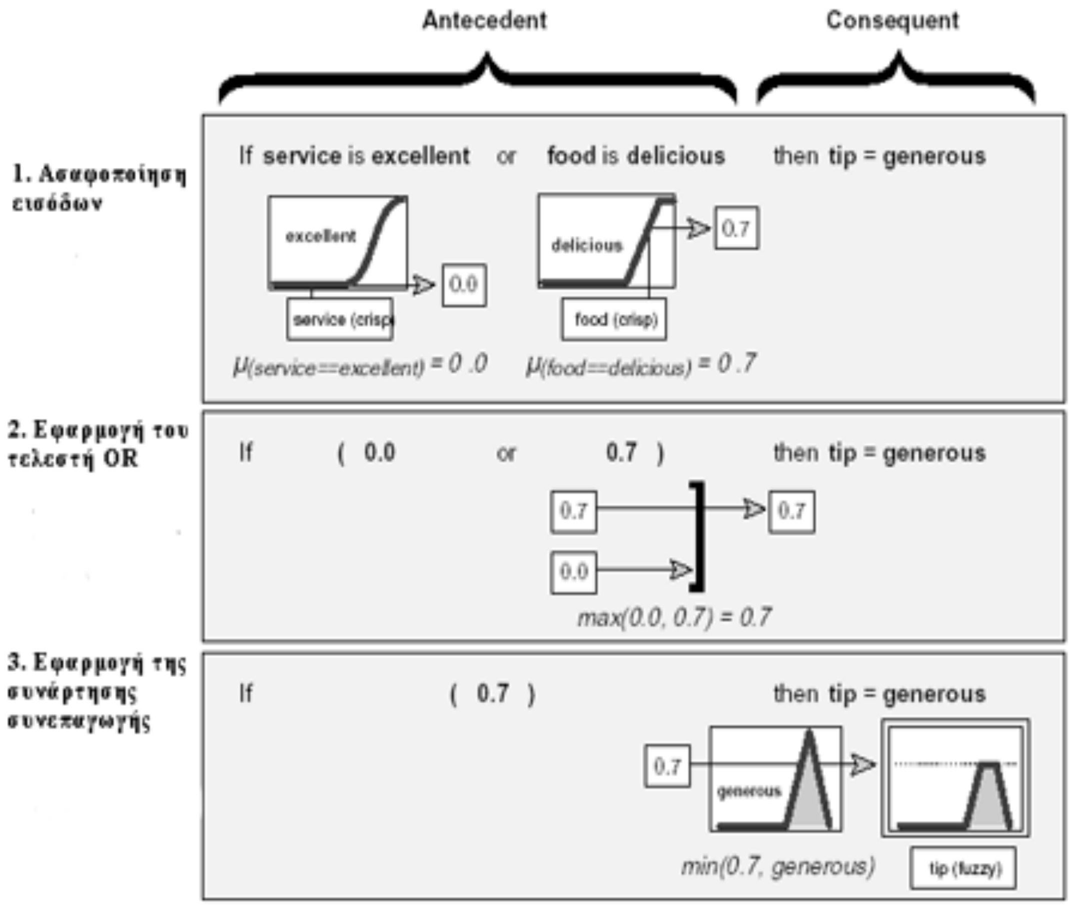
Σ' αυτή την περίπτωση όλα τα μέρη της προϋπόθεσης υπολογίζονται ταυτόχρονα και καταλήγουμε σε ένα μοναδικό νούμερο ανάλογα με τους λογικούς τελεστές που χρησιμοποιούμε.

Αλλά και η συνέπεια ενός κανόνα μπορεί να αποτελείται από περισσότερα του ενός μέρη. Για παράδειγμα ο παρακάτω κανόνας:

***if temperature is cold then hot water valve is open and cold water valve is shut***

Όλα τα μέρη της συνέπειας του κανόνα επηρεάζονται το ίδιο από το αποτέλεσμα της προϋπόθεσης. Η συνέπεια του κανόνα προσδιορίζει ένα ασαφές σύνολο το οποίο συσχετίζεται με την έξοδο. Η συνάρτηση συνεπαγωγής ( *implication function* ) στη συνέχεια τροποποιεί το ασαφές σύνολο στο βαθμό που προσδιορίζεται από την προϋπόθεση του κανόνα.

Μια πολύ συνηθισμένη συνάρτηση για να γίνει αυτό είναι η αποκοπή με τη χρήση του τελεστή min. Στο παρακάτω σχήμα βλέπουμε τον τρόπο με τον οποίο εφαρμόζονται οι κανόνες if-then.




Σχήμα 4.1: Κανόνες if-then στην ασαφή λογική

## *4.3 ΣΥΣΤΗΜΑΤΑ ΕΞΑΓΩΓΗΣ ΣΥΜΠΕΡΑΣΜΑΤΩΝ (fuzzy inference systems)*

Γενικά ένας κανόνας από μόνος του δεν επαρκεί για τα περισσότερα πρακτικά προβλήματα. Αυτό που χρειαζόμαστε είναι περισσότεροι του ενός κανόνες οι οποίοι να αλληλεπιδρούν μεταξύ τους. Η έξοδος κάθε κανόνα θα είναι ένα ασαφές σύνολο. Τα ασαφή σύνολα που προκύπτουν σα συνέπεια των κανόνων συγκεντρώνονται για να σχηματίσουν ένα ασαφές σύνολο εξόδου από το οποίο μέσω της διαδικασίας της αποσαφήνισης θα πάρουμε μια μοναδική τιμή.

Παρακάτω αναπτύσσεται ο τρόπος με τον οποίο εξελίσσεται η παραπάνω διαδικασία μέσα από ένα σύστημα ασαφούς συλλογιστικής (fuzzy inference system ,FIS) που ονομάζεται Mamdani.

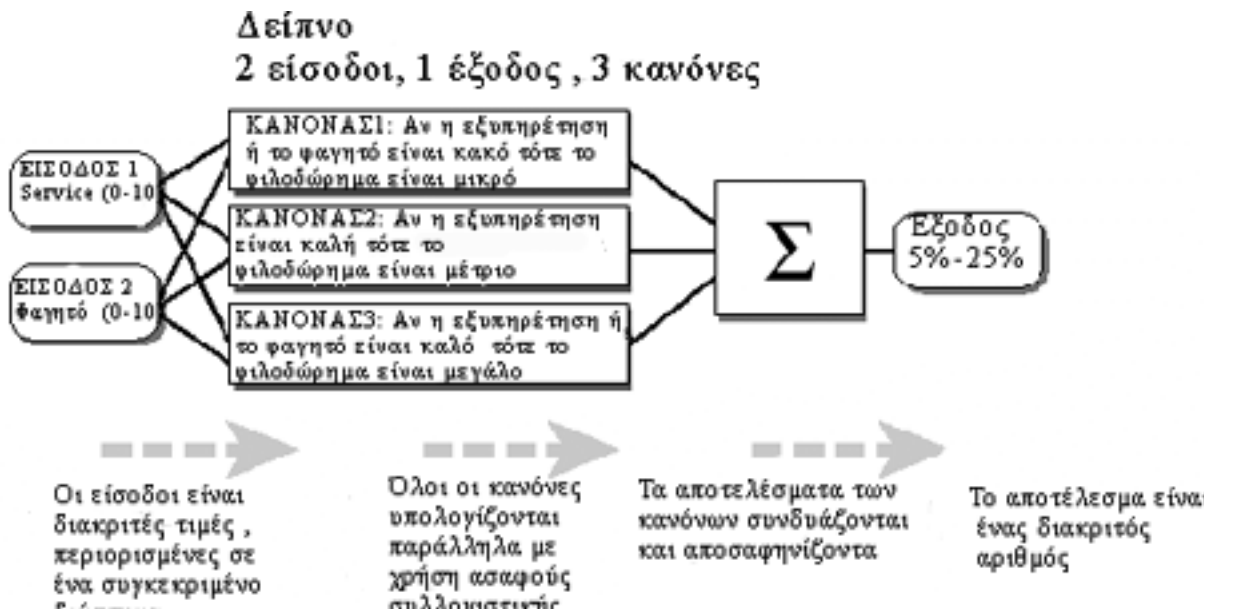
### 4.3.1 ΣΥSΤΗΜΑΤΑ Mamdani

Η ασαφής συλλογιστική είναι η διαδικασία κατά την οποία σχηματίζεται η απεικόνιση από μια δεδομένη είσοδο σε μια έξοδο , χρησιμοποιώντας τις αρχές της ασαφούς λογικής. Η απεικόνιση αυτή θέτει τις βάσεις με τις οποίες μπορεί να ληφθεί μια απόφαση ή να διακριθούν πρότυπα (patterns). Η διαδικασία αυτή της ασαφούς συλλογιστικής περιλαμβάνει όλα εκείνα τα κομμάτια στα οποία αναφερθήκαμε προηγουμένως. Δηλαδή τις συναρτήσεις συμμετοχής, τους ασαφείς λογικούς τελεστές και τους κανόνες if-then. Υπάρχουν δύο είδη ασαφών συστημάτων παρεμβολής τα τύπου-Mamdani και τα τύπου Sugeno. Οι διαφορές ανάμεσα στα δύο αυτά συστήματα έγκεινται κυρίως στον τρόπο με τον οποίο σχηματίζεται η έξοδος του συστήματος. Τα συστήματα Mamdani είναι αυτά που συναντώνται συχνότερα. Η μεθοδολογία των συστημάτων Mamdani εισήχθηκε το 1975 από τον Ebrahim Mamdani

Θα εξετάσουμε τώρα πως λειτουργεί ένα FIS τύπου Mamdani μέσα από το πρόβλημα του φιλοδωρήματος που αναφέραμε παραπάνω .

Το σύστημα έχει δύο εισόδους και τρεις κανόνες. Οι είσοδοι του συστήματος είναι η ποιότητα του φαγητού και η εξυπηρέτηση. Θεωρούμε ότι ο πελάτης βαθμολογεί την ποιότητα του φαγητού και την εξυπηρέτηση με μια κλίμακα από το 0 έως το 10. η έξοδος του συστήματος είναι το φιλοδώρημα που θα δώσει ο πελάτης και το οποίο κυμαίνεται από 5% έως 25% της τιμής του λογαριασμού.

Η βασική του δομή του συστήματος φαίνεται στο παρακάτω διάγραμμα:




Σχήμα 4.2: Ο μηχανισμός συμπεράσματος φιλοδωρήματος.

Η πληροφορία οδεύει από τα αριστερά προς τα δεξιά και καταλήγει σε μία μοναδική έξοδο. Ένα σημαντικό χαρακτηριστικό των συστημάτων ασαφούς λογικής είναι η παράλληλη εφαρμογή των κανόνων. Στα κλασσικά συστήματα δυαδικής λογικής έχουμε χρήση διακοπτών που ανοίγουν και κλείνουν ανάλογα με τις τιμές των εισόδων. Στα ασαφή συστήματα μεταβαίνουμε ομαλά από περιοχή σε περιοχή ανάλογα με τις συνέπειες των κανόνων. Η διαδικασία της απεικόνισης των εισόδων στις εξόδου μέσα από ένα σύστημα ασαφούς λογικής ακολουθεί πέντε βήματα:

1. Ασαφοποίηση των εισόδων

2. Εφαρμογή των ασαφών λογικών τελεστών στις προϋπόθεσης των κανόνων

3. Καθορισμός των συνεπαγωγών των κανόνων

4. Συγκέντρωση όλων των συνεπαγωγών όλων των κανόνων

5. Αποσαφήνιση

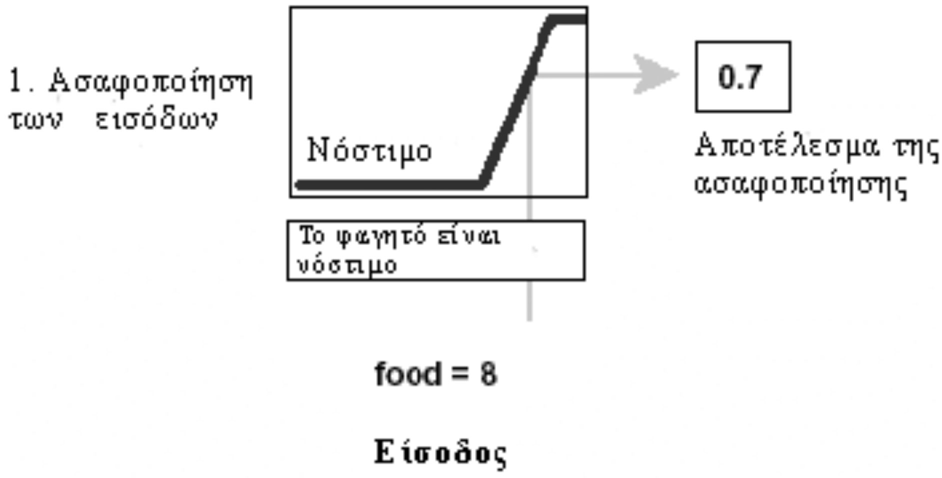
Θα εξετάσουμε κάθε βήμα από τα παραπάνω ξεχωριστά.

## ΒΗΜΑ 1$^{Ο}$ Ασαφοποίηση εισόδων

Το πρώτο βήμα έχει να κάνει με την ασαφοποίηση των εισόδων δηλαδή αυτό που κάνουμε είναι να πάρουμε τις εισόδους και να καθορίσουμε σε ποιο βαθμό κάθε είσοδος ανήκει σε ένα ασαφές σύνολο μέσα από τις συναρτήσεις συμμετοχής. Οι είσοδοι είναι αριθμητικές τιμές μέσα στα όρια του χώρου αναφοράς (στην περίπτωση του προβλήματος που εξετάζουμε από 0 μέχρι 10) και οι έξοδοι είναι βαθμοί συμμετοχής στο προσδιορισμένο ασαφές σύνολο.

Η ασαφοποίηση των εισόδων μπορεί να γίνει είτε με χρήση πίνακα τιμών είτε μέσα από μια συνάρτηση. Το παράδειγμα μας είναι δομημένο σε τρεις κανόνες κάθε ένας από τους οποίους προϋποθέτει την κατάταξη των εισόδων σε διαφορετικές λεκτικές μεταβλητές service is poor, service is good, food is rancid, food is delicious.

Πριν να εφαρμοσθούν οι κανόνες θα πρέπει να ασαφοποιηθούν οι είσοδοι σύμφωνα με τις λεκτικές μεταβλητές. Για παράδειγμα σε ποιο βαθμό το φαγητό είναι νόστιμο. Αν υποθέσουμε ότι το φαγητό έχει αξιολογηθεί με μια συγκεκριμένη βαθμολογία από το 0 έως το 10 (π.χ. 7) τότε πως μπορούμε να συμπεράνουμε σε ποιο βαθμό είναι νόστιμο. Για το λόγο αυτό θα χρησιμοποιήσουμε μια γραφική απεικόνιση που ουσιαστικά αποτελεί τη συνάρτηση συμμετοχής του ασαφούς συνόλου νόστιμο.

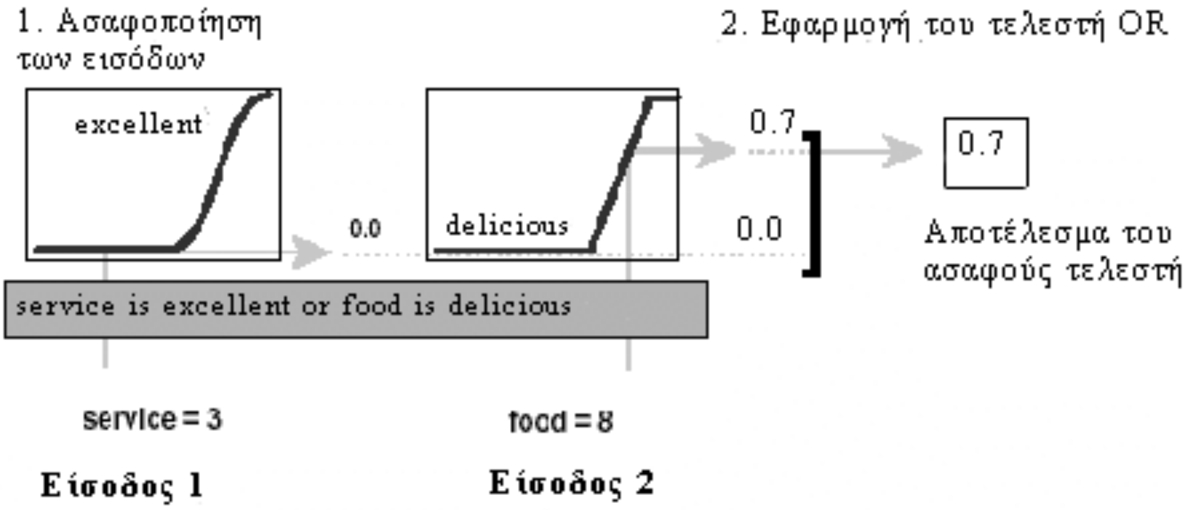



Σχήμα 4.3: Ασαφοποίηση των εισόδων

**ΒΗΜΑ 2ο. Εφαρμογή των ασαφών τελεστών**

Αφού έχουμε ασαφοποιήσει στις εισόδους μπορούμε να γνωρίζουμε σε ποιο βαθμό ικανοποιείται κάθε μέρος των προϋποθέσεων των if-then κανόνων. Αν ένας κανόνας if-then έχει στο σκέλος if περισσότερα του ενός μέρη τότε πρέπει να εφαρμοσθούν οι ασαφείς τελεστές ώστε να καθορισθεί ένα νούμερο που αντιπροσωπεύει το αποτέλεσμα της προϋπόθεσης του κανόνα. Το νούμερο αυτό θα εφαρμοσθεί έπειτα στη συνάρτηση εξόδου. Οι είσοδοι ενός λογικού τελεστή είναι δύο ή περισσότεροι βαθμοί συμμετοχής που έχουν προκύψει από τις ασαφοποιημένες μεταβλητές εισόδου. Η έξοδος του τελεστή είναι απλά ένας βαθμός αλήθειας.

Όπως είδαμε κατά την αναφορά στους ασαφείς τελεστές υπάρχει ένας μεγάλος αριθμός συναρτήσεων που μπορούν να χρησιμοποιηθούν για να ορίσουν τους τελεστές AND και OR. Συνήθως χρησιμοποιούνται ο τελεστής max για το OR και min για το AND. Στο παρακάτω διάγραμμα βλέπουμε πως εφαρμόζονται οι τελεστές.

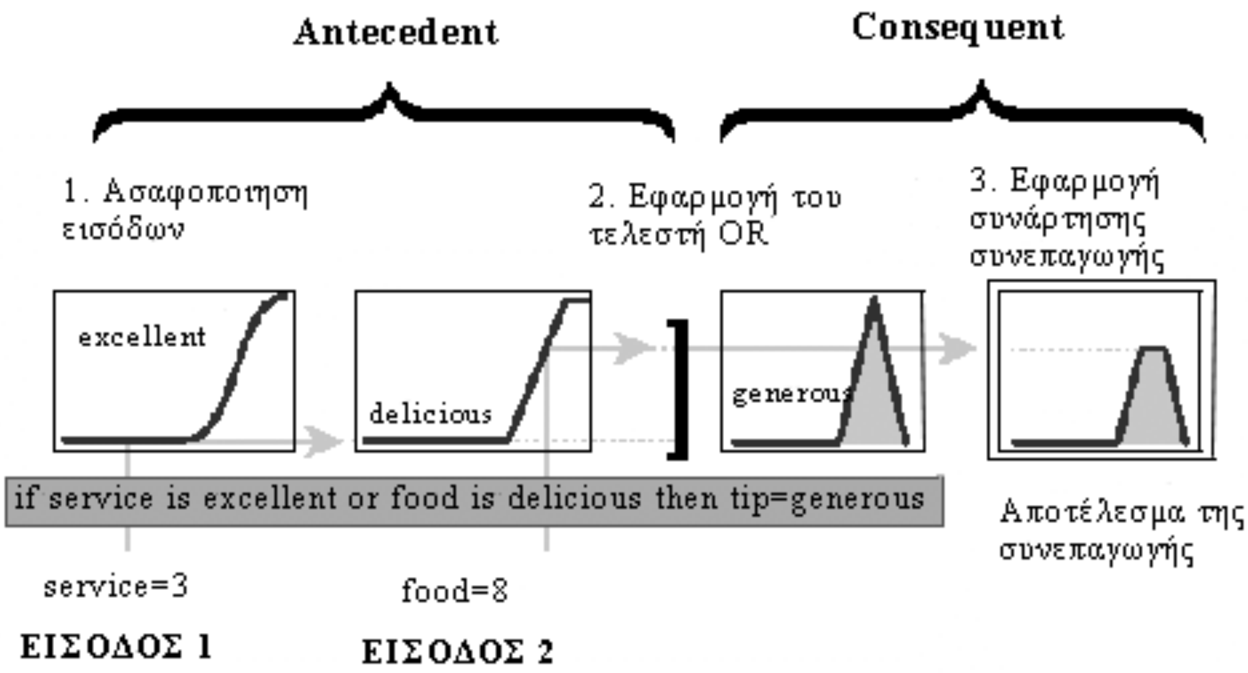



Σχήμα 4.4: Τρεις κανόνες του προβλήματος και ο τρόπος που συνδυάζονται τα αποτελέσματά τους σε ένα ασαφές σύνολο

**ΒΗΜΑ 3$^{o}$ Μέθοδος εφαρμογής της συνεπαγωγής (Apply Implication Method)**

Για να εφαρμόσουμε την συνάρτηση συνεπαγωγής πρέπει να γνωρίζουμε το βάρος του κάθε κανόνα. Οι κανόνες if-then έχουν βάρη τα οποία εφαρμόζονται στο νούμερο που δίνει το κομμάτι της υπόθεσης και που μπορούν να κυμαίνονται από 0 έως 1. Συνήθως τα βάρη αυτά είναι ίσα με 1 , ωστόσο μπορούμε να δώσουμε σ΄ αυτά μια οποιαδήποτε τιμή από 0 έως 1. Για παράδειγμα αν θέλουμε να δώσουμε μικρότερη έμφαση σε έναν κανόνα σε σχέση με κάποιον άλλο , αυτό μπορούμε να το κάνουμε μέσω των τιμών των βαρών τους.

Αφού έχομε αναθέσει στους κανόνες τα βάρη μπορούμε να εφαρμόσουμε την συνάρτηση συνεπαγωγής. Το αποτέλεσμα ενός κανόνα if-then είναι ένα ασαφές σύνολο που αντιπροσωπεύεται από μια συνάρτηση συμμετοχής που σταθμίζει κατάλληλα τα λεκτικά χαρακτηριστικά που έχουν ανατεθεί σ' αυτόν. Το αποτέλεσμα αυτό αναδιαμορφώνεται με τη χρήση μιας συνάρτησης που είναι συνδεδεμένη με το μέρος της υπόθεσης. Η είσοδος για τη διαδικασία συνεπαγωγής είναι ένας αριθμός που δίδεται από την υπόθεση και η έξοδος είναι ένα ασαφές σύνολο. Αυτή η διαδικασία εφαρμόζεται για κάθε κανόνα



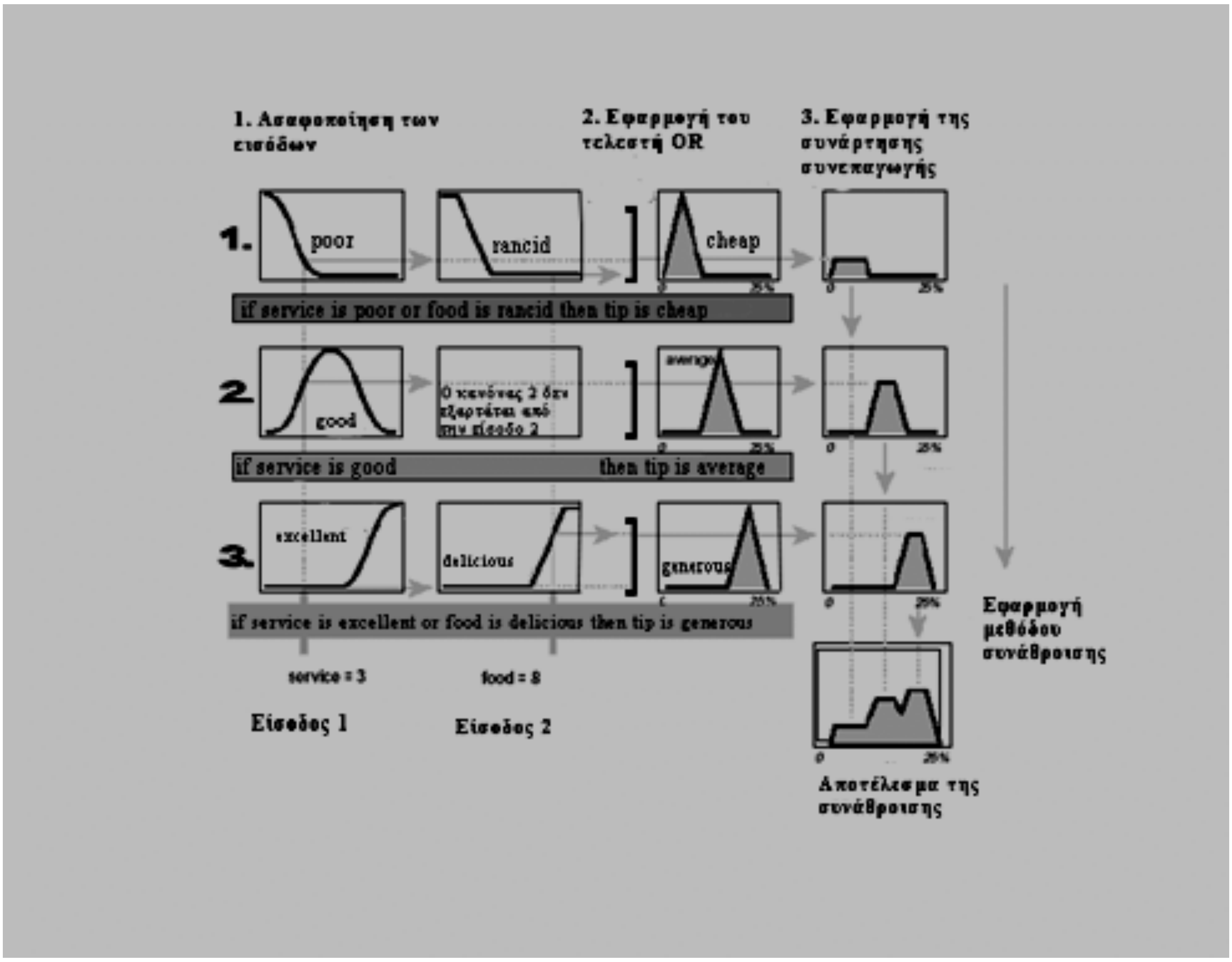

Σχήμα 4.5: Μέθοδος εφαρμογής της συνεπαγωγής

**<u>ΒΗΜΑ 4$^{ο}$ συγκέντρωση όλων των εξόδων</u>**

Σε ένα FIS οι αποφάσεις βασίζονται στον έλεγχο όλων των κανόνων. Επομένως όλοι οι κανόνες πρέπει να συνδυαστούν προκειμένου να καταλήξουμε σε μια απόφαση. Η συγκέντρωση των κανόνων είναι μια διαδικασία κατά την οποία όλα τα ασαφή σύνολα που προκύπτουν ως έξοδοι των κανόνων συνδυάζονται ώστε να σχηματίσουν ένα μοναδικό ασαφές σύνολο. Η συγκέντρωση είναι η διαδικασία που προηγείται της αποσαφήνισης και δέχεται ως είσοδο τις αποκομμένες συναρτήσεις εξόδου του προηγούμενου βήματος. Η διαδικασία της συγκέντρωσης των εξόδων είναι αντιμεταθετική και έτσι η σειρά με την οποία εκτελούνται οι κανόνες δεν έχει σημασία στην έκβαση του αποτελέσματος.

***Στο παρακάτω διάγραμμα βλέπουμε τους τρεις κανόνες του προβλήματος και τον τρόπο που συνδυάζονται τα αποτελέσματά τους σε ένα ασαφές σύνολο.***

.




Σχήμα 4.5: FIS Mandami, για το ποσοστό φιλοδωρήματος

**ΒΗΜΑ 5$^{o}$ Αποσαφήνιση**

Η είσοδος για τη διαδικασία της αποσαφήνισης είναι το ασαφές σύνολο που προέκυψε από το προηγούμενο βήμα. Η έξοδος της αποσαφήνισης είναι ένας αριθμός. Η αποσαφήνιση είναι αναγκαία γιατί επιθυμούμε να έχομε ως αποτέλεσμα έναν αριθμό που θα μας επιτρέψει να πάρουμε συγκεκριμένες αποφάσεις. Υπάρχουν πολλοί τρόποι για να γίνει η αποσαφήνιση. Ένας ιδιαίτερα δημοφιλής είναι ο υπολογισμός του κέντρου βάρους της περιοχής που περικλείεται από την καμπύλη της συνάρτησης συμμετοχής του ασαφούς συνόλου. Άλλες επίσης δημοφιλείς μέθοδοι είναι η μέση τιμή , η μέγιστη η μέγιστη θέση του μεγίστου και η ελάχιστη θέση του μεγίστου.

*Μέθοδος αποσαφήνισης MAXIMUM*

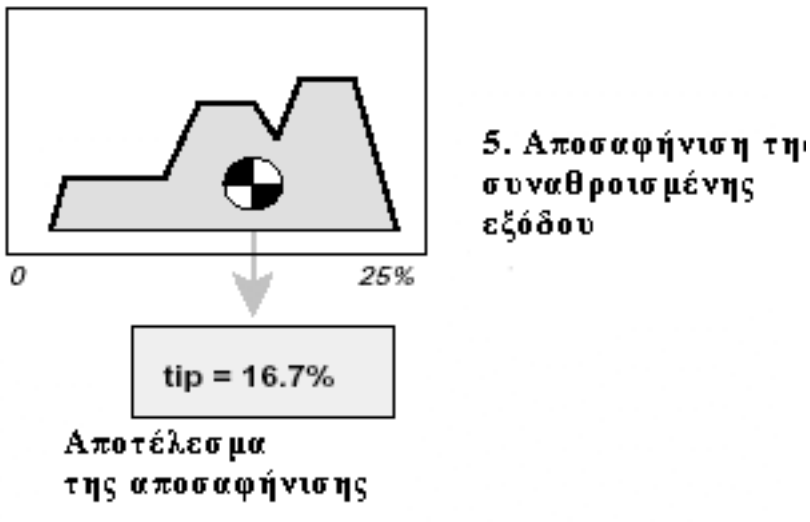
Σύμφωνα με τη μέθοδο αποσαφήνισης maximum , η διακριτή τιμή είναι αυτή που αντιστοιχεί στη μέγιστη τιμή συγγένειας του τελικού αποτελέσματος. Αν υπάρχουν περισσότερες από μία τέτοιες τιμές , τότε λαμβάνεται ανάλογα με την περίπτωση είτε ο μέσος όρος τους (average-of-maximum) ή η μέγιστη τιμή τους (maximum-of-maximum) ή η ελάχιστη τιμή τους (minimum-of-maximum).

***Μέθοδος Cendroid***

Σύμφωνα με τη μέθοδο αποσαφήνισης cendroid , η διακριτή τιμή είναι αυτή που προκύπτει από το κέντρο βάρους της τελικής συνάρτησης συμμετοχής για την ασαφή παράμετρο εξόδου. Το κέντρο βάρους μιας επιφάνειας που ορίζεται από μια συνάρτηση f(t) και τους καρτεσιανούς άξονες , βρίσκεται στη θέση t που ορίζεται από τη σχέση:

$$t_{\kappa\beta} = \frac{\int t \cdot f(t)dt}{\int f(t)dt}$$

***Στο παρακάτω σχήμα βλέπουμε πως εφαρμόζεται η μέθοδος της αποσαφήνισης για το παράδειγμά μας.***




Σχήμα 4.7: Μέθοδος της αποσαφήνισης

***Στην περίπτωση διακριτού συνόλου αναφοράς , τα ολοκληρώματα στην παραπάνω σχέση αντικαθίστανται με διακριτό άθροισμα και γίνεται δειγματοληψία Ν σημείων στο σύνολο αναφοράς.***

Ένα χαρακτηριστικό της μεθόδου αποσαφήνισης cendroid είναι ότι στην περίπτωση που έχει γίνει σύνθεση αποτελεσμάτων από επιμέρους κανόνες και υπάρχουν τυχόν αλληλεπικαλυπτόμενες περιοχές , αυτές λαμβάνονται υπόψη μία μόνο φορά. Επίσης στην περίπτωση που η συνάρτηση συμμετοχής είναι παντού 0 , το αποτέλεσμα της αποσαφήνισης ορίζεται κατά σύμβαση.

*Διαγραμματική μέθοδος επίλυσης*

Μία εναλλακτική προσέγγιση στην επίλυση προβλημάτων ασαφούς συλλογιστικής είναι η διαγραμματική επίλυση. Αυτή επιτρέπει τη γρήγορη , αλλά προσεγγιστική εκτίμηση της τελικής λύσης χωρίς να απαιτεί όλους τους αριθμητικούς υπολογισμούς που παρουσιάστηκαν παραπάνω. Βασική προϋπόθεση για την εφαρμογή της μεθόδου είναι οι συναρτήσεις συμμετοχής των παραμέτρων του προβλήματος να είναι συνεχείς καμπύλες και όχι σύνολο ζευγών.

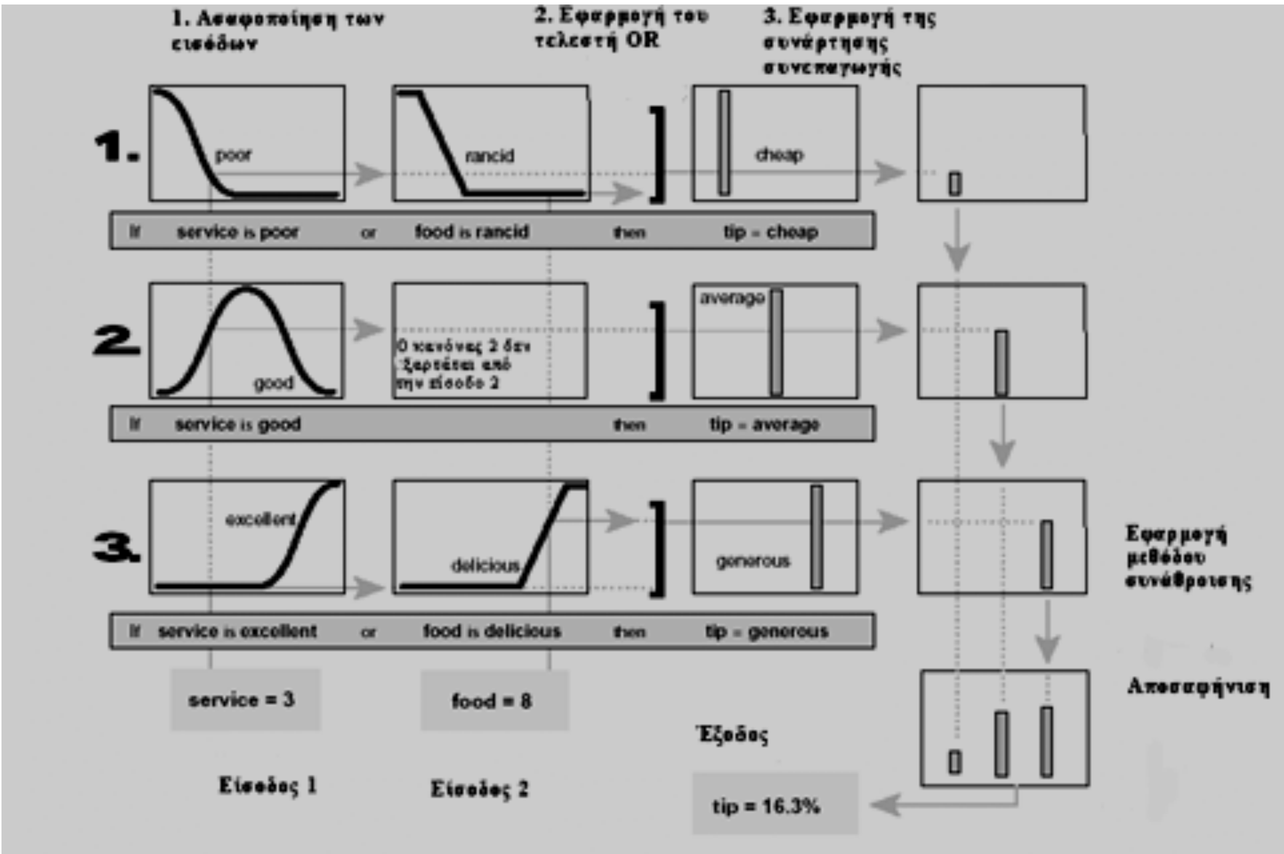
### 4.3.2 Συστήματα τύπου Sugeno

Παραπάνω ασχοληθήκαμε με τα συστήματα Mamdani που είναι και τα πιο ευρέως χρησιμοποιούμενα. Ωστόσο υπάρχει και η μέθοδος Sugeno που εισήχθηκε to 1985 η μέθοδος αυτή έχει ομοιότητες με τη μέθοδο των συστημάτων Mamdani. Για παράδειγμα τα 2 πρώτα βήματά της (ασαφοποίηση των εισόδων και εφαρμογή των τελεστών) είναι ακριβώς τα ίδια. Η κύρια διαφορά ανάμεσα στα δύο συστήματα έγκειται στο ότι οι συναρτήσεις συμμετοχής στην έξοδο των συστημάτων Sugeno είναι μόνο γραμμική ή σταθερά.

Ένας τυπικός ασαφής κανόνας σε συστήματα sugeno μηδενικής τάξης έχει την μορφή:

***if x is A and y is B then z = k***

όπου A και B είναι τα ασαφή σύνολα της προϋπόθεσης ενώ κ είναι μια αριθμητική τιμή. Αφού η συνέπεια του κανόνα είναι μια σταθερά τότε αυτό σημαίνει ότι το βήμα 3 εκφυλίζεται σε ένα απλό πολλαπλασιασμό ενώ το βήμα 4 καταλήγει να είναι η συγκέντρωση όλων των σταθερών.



Σχήμα 4.8: FIS Sugeno

Ένα σύστημα sugeno πρώτης τάξης θα έχει κανόνες με τη γενική μορφή

***if x is A and y is B then z = p*x + q*y + r***

όπου A και B είναι τα ασαφή σύνολα της προϋπόθεσης ενώ τα p,q,r είναι σταθερές.

Ένας τρόπος για να δούμε τα συστήματα πρώτης τάξης είναι να θεωρήσουμε ότι κάθε κανόνας προσδιορίζει τη θέση ενός κινούμενου singleton. Το singleton αυτό μπορεί να κινείται στο χώρο της εξόδου , με γραμμικό τρόπο και η θέση του εξαρτάται από τις τιμές των εισόδων.

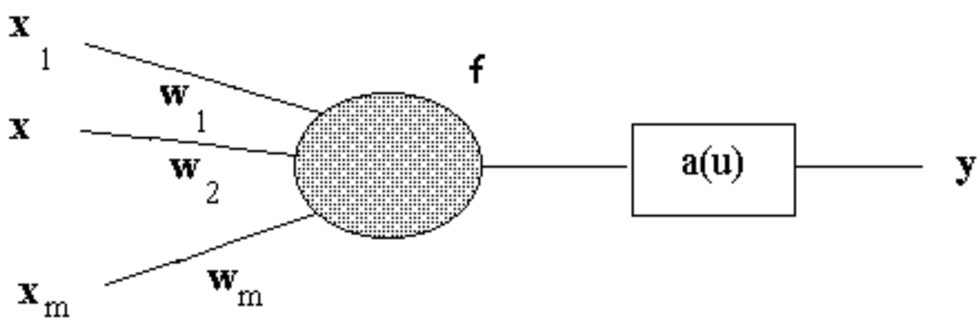
Συστήματα sugeno ανώτερης τάξης είναι εφικτά , όμως δεν προσφέρουν σημαντικές βελτιώσεις και ταυτόχρονα εισαγάγουν σημαντική πολυπλοκότητα.

## *4.4 ΝΕΥΡΟ ΑΣΑΦΗ ΣΥΣΤΗΜΑΤΑ*

Στην ενότητα αυτή εξετάζουμε τα νευροασαφή συστήματα. Η δομική μονάδα των συστημάτων αυτών είναι οι ασαφείς νευρώνες

### 4.4.1 ΑΣΑΦΕΙΣ ΝΕΥΡΩΝΕΣ

Οι ασαφείς νευρώνες έχουν τη μορφή που φαίνεται στο παρακάτω σχήμα:




Σχήμα 4.9: Ασαφείς Νευρώνες

Όπου :

$x_1, x_2, .., x_m$ οι είσοδοι του νευρώνα
$w_1, w_2, .., w_m$ τα βάρη των συνάψεων
f η συνάρτηση συμμετοχής του νευρώνα
α η συνάρτηση ενεργοποίηση του νευρώνα
y η έξοδος του νευρώνα

Οι ασαφείς νευρώνες διακρίνονται σε νευρώνες συμμετοχής και σε λειτουργικούς νευρώνες. Παρακάτω γίνεται αναφορά στα χαρακτηριστικά κάθε κατηγορίας ασαφών νευρώνων.

## 4.4.2 ΝΕΥΡΩΝΕΣ ΣΥΜΜΕΤΟΧΗΣ

Οι συνηθέστερες μορφές συναρτήσεων συμμετοχής είναι η τριγωνική, η γκαουσιανή, η τραπεζοειδής , η κανονική κοκ. Οι παραπάνω συναρτήσεις συμμετοχής μπορούν να υλοποιηθούν με τη χρήση ενός νευρώνα. Η συνάρτηση συμμετοχής μπορεί να υλοποιηθεί θεωρώντας ότι ο νευρώνας δέχεται ως είσοδο το x και δίνει έξοδο A(x) απαιτώντας η συνάρτηση συμμετοχής του να έχει τη μορφή A. αν δηλαδή υποθέσουμε ότι η συνάρτηση συμμετοχής είναι η κανονική , τότε για να υλοποιηθεί πρέπει η συνάρτηση ενεργοποίησης να είναι

$$a(u) = \exp\left( - \frac{(u-m)^2}{\sigma^2} \right)$$

Όπου m κέντρο της συνάρτησης συμμετοχής και σ το εύρος της.

### 4.4.3 ΛΕΙΤΟΥΡΓΙΚΟΙ ΝΕΥΡΩΝΕΣ

Με τον ίδιο τρόπο μπορούν να υλοποιηθούν οι βασικές πράξεις των ασαφών συνόλων. Οι πράξεις αυτές στηρίζονται σε λειτουργίες όπως η τ- νόρμα η σ-νόρμα , το ασαφές συμπλήρωμα κλπ. Οι παραπάνω λειτουργίες μπορούν να αντικαταστήσουν τη συνάρτηση ενεργοποίησης του νευρώνα αν θεωρήσουμε ότι οι είσοδοι δεν αθροίζονται πολλαπλασιασμένοι με βάρη όπως στους κλασσικούς νευρώνες. Για να μπορούμε όμως να εφαρμόσουμε τους αλγορίθμους μάθησης (πχ τον back propagation αλγόριθμος αντίστροφης διάδοσης) θα πρέπει οι λειτουργίες να είναι παραγωγίσιμες.

Για την υλοποίηση των ασαφών λειτουργιών μπορούμε να επεκτείνουμε την έννοια της συνάρτησης μεταφοράς του νευρώνα. Στην περίπτωση αυτή ο νευρώνας δεν υλοποιεί το αναλυτικό γινόμενο της εισόδου με το διάνυσμα των βαρών αλλά την πράξη της συγκεκριμένης ασαφούς λειτουργίας θεωρώντας ότι όλα τα βάρη είναι 1. ως συνάρτηση συμμετοχής λαμβάνεται η :

$$a(u) = \begin{cases} 1, u > 1 \\ u, 0 \le u \le 1 \\ 0, u < 0 \end{cases}$$

### 4.4.4 ΣΥΝΘΕΤΙΚΟΙ ΝΕΥΡΩΝΕΣ

Παραπάνω αναφερθήκαμε στο στη σύνθεση των ασαφών σχέσεων. Η πράξη της σύνθεσης αποτελεί μια γενίκευση της πράξης του αναλυτικού γινομένου των διανυσμάτων και υλοποιείται από τη συνάρτηση μεταφοράς f των νευρώνων. Η μορφή του νευρώνα είναι η ίδια και ορίζεται από τις εξισώσεις:

$$f = \underset{i \in N_m}{union}\, t(x_t, w_t)$$

$$y = a(u)$$

όπου union είναι μια σ-νόρμα και t είναι μια τ-νόρμα. Ως συνάρτηση ενεργοποίησης θεωρείται η συνάρτηση αναρρίχησης. Παρατηρούμε ότι η ο συνθετικός νευρώνας είναι μια γενίκευση του κλασσικού νευρώνα , αφού η πράξη της φραγμένης άθροισης , η οποία εξασφαλίζεται από τη συνάρτηση ενεργοποίησης είναι μια ειδική περίπτωση της φραγμένης ένωσης, ενώ το γινόμενο είναι μια ειδική περίπτωση της ασαφούς τομής.

## *4.5 ΠΡΟΣΑΡΜΟΣΤΙΚΑ ΔΙΚΤΥΑ (Adaptive Networks)*

Ένα προσαρμοστικό δίκτυο είναι αυτό που στη δομή του περιέχει έναν αριθμό κόμβων συνδεδεμένων μέσω κατευθυντικών συνδέσμων. Κάθε κόμβος αναπαριστά μια μονάδα επεξεργασίας και οι σύνδεσμοι μεταξύ κόμβων προσδιορίζουν την αιτιολογική σχέση (casual relationship) μεταξύ των συνδεδεμένων κόμβων. Προσαρμοστικός είναι ο κόμβος όλος ή μέρος του, που σημαίνει ότι οι έξοδοι αυτών των κόμβων εξαρτώνται από προσαρμοζόμενες – τροποποιήσιμες παραμέτρους που ανήκουν σε αυτούς τους κόμβους. Οι κανόνες εκπαίδευσης προσδιορίζουν πως αυτές οι παράμετροι θα πρέπει να αλλάζουν ώστε να ελαχιστοποιείται το προκαθορισμένο μέτρο σφάλματος (error measure) το οποίο είναι μια μαθηματική έκφραση που μετράει την ασυμφωνία μεταξύ της πραγματικής εξόδου του δικτύου και της επιθυμητής εξόδου.

Τα προσαρμοστικά δίκτυα χρησιμοποιούνται στην ταυτοποίηση συστήματος. Εμείς θα πρέπει να βρούμε μια κατάλληλη αρχιτεκτονική για το δίκτυο και να θέσομε ένα σύνολο παραμέτρων που μπορούν να μοντελοποιήσουν καλύτερα ένα σύστημα-στόχο, το οποίο περιγράφεται από ένα σύνολο ζευγών δεδομένων εισόδου-εξόδου.

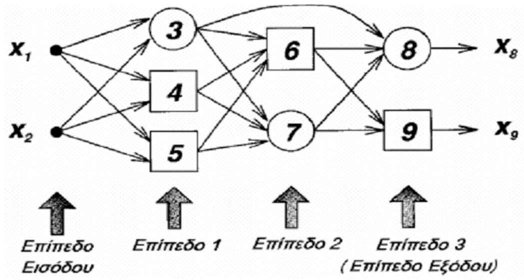
Ο βασικός κανόνας ενός προσαρμοστικού δικτύου είναι η απότομη φθίνουσα μέθοδος, στην οποία το βαθμωτό διάνυσμα προκύπτει από επιτυχείς επικλήσεις του κανόνα της αλυσίδας. Αυτή η μέθοδος χρησιμοποιείται και για την εύρεση του βαθμωτού σ’ ένα νευρωνικό δίκτυο πολλών επιπέδων. Αυτή η μέθοδος ονομάζεται κανόνας οπισθόδρομης διάδοσης (back propagation learning rule).

### 4.5.1 ΠΑΡΑΓΩΓΟΙ ΠΑΡΑΜΕΤΡΟΠΟΙΗΣΙΜΩΝ ΣΥΝΑΡΤΗΣΕΩΝ ΣΥΜΜΕΤΟΧΗΣ

Για να κάνουμε ένα ασαφές σύστημα, προσαρμοστικό, θα πρέπει να ξέρουμε τις παραγώγους μιας συνάρτησης συμμετοχής ως προς το όρισμα εισόδου και τις παραμέτρους. Αυτές οι πληροφορίες από τις παραγώγους παίζουν σημαντικό ρόλο στην μάθηση ή την προσαρμογή ενός ασαφούς συστήματος.

### 4.5.2 ΑΡΧΙΤΕΚΤΟΝΙΚΗ ΠΡΟΣΑΡΜΟΣΤΙΚΩΝ ΔΙΚΤΥΩΝ

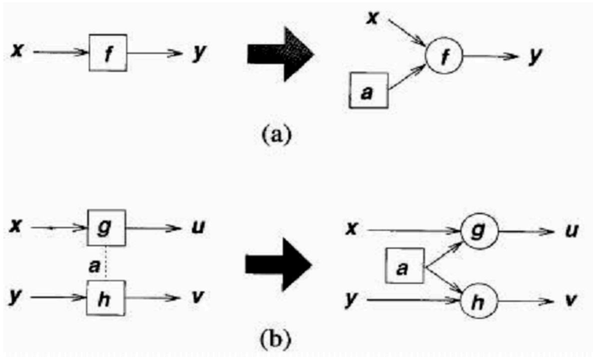
Όπως λέει και η ονομασία, ένα προσαρμοστικό δίκτυο είναι μια δομή δικτύου που η συνολική συμπεριφορά εισόδου-εξόδου διαμορφώνεται από ένα σύνολο τροποποιήσιμων παραμέτρων. Συγκεκριμένα η σύνθεση ενός προσαρμοστικού δικτύου πραγματοποιείται από ένα σύνολο κόμβων συνδεδεμένων με κατευθυνθείς συνδέσμους, όπου κάθε κόμβος εκτελεί μια συγκεκριμένη λειτουργία κόμβου στα εισερχόμενα σήματα, για να παράξει μια μονήρη έξοδο κόμβου και κάθε σύνδεσμος καθορίζει τη κατεύθυνση ροής σήματος από τον ένα κόμβο στον άλλον. Συνήθως μια συνάρτηση κόμβου είναι μια παραμετρική συνάρτηση με τροποποιήσιμες παραμέτρους. Αλλάζοντας τις τελευταίες, μπορούμε να αλλάξουμε τη λειτουργία κόμβου όπως στην συνολική συμπεριφορά του προσαρμοστικού δικτύου.




Σχήμα 4.10: Ένα εμπροσθόδρομο προσαρμοστικό δίκτυο σε αναπαράσταση με επίπεδα.

Οι σύνδεσμοι σε ένα προσαρμοστικό δίκτυο χρησιμοποιούνται απλώς για να προσδιορίσουν την κατεύθυνση διάδοσης των εξόδων κόμβου. Γενικά υπάρχουν βάρη ή παράμετροι σχετιζόμενοι με συνδέσμους. Το σχήμα 4.10 είναι ένα αντιπροσωπευτικό προσαρμοστικό δίκτυο με 2 εισόδους και με 2 εξόδους.

Οι παράμετροι ενός προσαρμοστικού δικτύου είναι κατανεμημένοι μέσα στους κόμβους, ώστε κάθε κόμβος να έχει ένα τοπικό σύνολο παραμέτρων. Η ένωση αυτών των τοπικών παραμέτρων είναι το ολικό σύνολο παραμέτρων του δικτύου. Αν το σύνολο παραμέτρων ενός κόμβου δεν είναι κενό, τότε η λειτουργία του κόμβου εξαρτάται από τις τιμές των παραμέτρων. Στα σχήματα, αναπαριστούμε τον **προσαρμοστικό κόμβο** αυτού του είδους με ένα τετράγωνο. Από την άλλη αν ο κόμβος έχει κενό σύνολο παραμέτρων, τότε η λειτουργία του είναι καθορισμένη. Ακόμα αναπαριστούμε τον τύπο του **προσαρμοστικού κόμβου** με ένα κύκλο. Κάθε προσαρμοστικό κόμβος μπορεί να διασπαστεί σε ένα προκαθορισμένο κόμβο συν ένα ή περισσότερους κόμβους παραμέτρων.

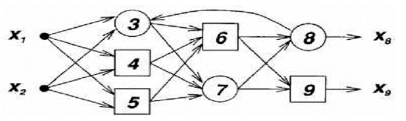



Σχήμα 4.11: Ανάλυση προσαρμοστικών κόμβων: (α) ένας μονός κόμβος (β) ένα πρόβλημα έχει κοινές παραμέτρους.

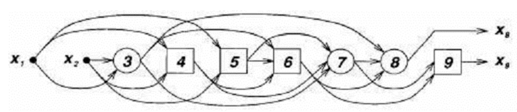
## *4.6 ΚΑΤΑΜΕΡΙΣΜΟΣ ΠΑΡΑΜΕΤΡΩΝ ΣΕ ΠΡΟΣΑΡΜΟΣΤΙΚΑ ΔΙΚΤΥΑ*

Το σχήμα 4.11 δείχνει ένα προσαρμοστικό δίκτυο με έναν μόνο κόμβο, το οποίο μπορεί να αναπαρασταθεί σαν $y = f(x, a)$ όπου x,y είναι η είσοδος και η έξοδος αντιστοίχως και *α* είναι η παράμετρος κόμβου. Μια ισοδύναμη απεικόνιση είναι να μετακινήσουμε μια παράμετρο έξω από τον κόμβο και να την τοποθετήσουμε σε ένα **κόμβο παράμετρο** όπως φαίνεται στο σχήμα 4.11 (α). Αυτός ο κόμβος παράμετρος είναι μια ειδική περίπτωση προσαρμοστικού κόμβου στον οποίο δεν υπάρχουν είσοδοι και η έξοδος είναι η ίδια παράμετρος. Ο κόμβος παράμετρος είναι χρήσιμος στην επίλυση συγκεκριμένων αντιπροσωπευτικών προβλημάτων, όπως στο παράδειγμα καταμερισμού παραμέτρων στο σχήμα 4.11 (α) όπου δυο προσαρμοστικοί κόμβοι $u = g(x, a)$και *u*=h(y, a) μοιράζονται την ίδια παράμετρο α, όπως δηλώνεται και από τη διάστικτη γραμμή που ενώνει τους δυο αυτούς κόμβους. Βγάζοντας έξω την παράμετρο και βάζοντας την μέσα σε ένα κόμβο παράμετρο, μπορούμε να ενσωματώσουμε τις απαιτήσεις του καταμερισμού παραμέτρων μέσα στην αρχιτεκτονική σχεδίαση του δικτύου. Αυτό απλοποιεί την αναπαράσταση του δικτύου όπως και την εφαρμογή του σε λογισμικό.

Τα προσαρμοστικά δίκτυα είναι γενικώς κατηγοριοποιημένα σε δυο κατηγορίες με βάση τον τύπο των διασυνδέσεων που έχουν: εμπροσθόδρομα δίκτυα (feedward) και επαναληπτικά. Το προσαρμοστικό δίκτυο που φαίνεται στο σχήμα 4.13 είναι εμπροσθόδρομο, μιας και η έξοδος κάθε κόμβου διαδίδεται από την πλευρά της εισόδου (αριστερά) προς τη πλευρά της εξόδου (δεξιά) πάντα. Αν υπάρχει σύνδεσμος ανάδρασης που σχηματίζει ένα κυκλικό μονοπάτι στο δίκτυο, τότε το δίκτυο είναι επαναληπτικό. Το σχήμα 4.12 είναι ένα παράδειγμα. Στα γραφήματα, ένα εμπροσθόδρομο δίκτυο αναπαριστάται από ένα κυκλικό κατευθυνόμενο γράφημα που δεν περιέχει κατευθυνόμενους κύκλους, ενώ ένα επαναληπτικό δίκτυο περιέχει πάντα τουλάχιστον ένα κατευθυνόμενο κύκλο.




Σχήμα 4.12: Ένα επαναλαμβανόμενο προσαρμοστικό δίκτυο




Σχήμα 4.13: Ένα εμπροσθόδρομο προσαρμοστικό δίκτυο σε αναπαράσταση τοπολογικής διάταξης.

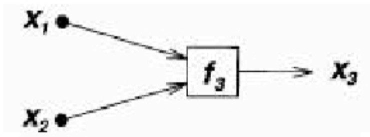
Στην **αναπαράσταση με επίπεδα** του εμπροσθόδρομου προσαρμοστικού δικτύου στο σχήμα 4.11, δεν υπάρχουν σύνδεσμοι μεταξύ κόμβων στο ίδιο επίπεδο και οι έξοδοι των κόμβων σε ένα συγκεκριμένο επίπεδο είναι πάντα συνδεδεμένοι με κόμβους σε διαδοχικά επίπεδα. Αυτή η αναπαράσταση είναι συνήθως προτιμητέα εξαιτίας της εύκολης διαμόρφωσης της επειδή οι κόμβοι στο ίδιο επίπεδο έχουν την ίδια λειτουργία ή γεννούν το ίδιο επίπεδο αφαίρεσης όσων αφορά τα διανύσματα εισόδου.

Μια άλλη αναπαράσταση εμποσθόδρομου δικτύου είναι η αναπαράσταση τοπολογικής ταξινόμησης η οποία ετικετοποιεί τους κόμβους σε μια διατεταγμένη ακολουθία 1,2,3, …, τέτοια που να μην υπάρχουν σύνδεσμοι από τον κόμβο i στον κόμβο j, οποτεδήποτε $i \geq j$. Το σχήμα 4.13 είναι η αναπαράσταση τοπολογικής ταξινόμησης (topological ordering representation) του δικτύου στο σχήμα 4.11. Αυτή η αναπαράσταση είναι λιγότερο διαμορφώσιμη από ότι η αναπαράσταση με επίπεδα ,όμως διευκολύνει τον σχηματισμό κανόνων εκπαίδευσης. Η αναπαράσταση τοπολογικής ταξινόμησης είναι στην πραγματικότητα μια ειδική περίπτωση της αναπαράστασης με επίπεδα, με έναν κόμβο ανά επίπεδο.

Ένα εμπροσθόδρομο προσαρμοστικό δίκτυο είναι στην πραγματικότητα μια στατική αντιστοίχηση μεταξύ των χώρων εισόδου και εξόδου. Αυτή η αντιστοιχία μπορεί να είναι μια απλή γραμμική σχέση ή μη γραμμική, εξαρτώμενη από τη δομή του δικτύου (διάταξη κόμβων και συνδέσεων) και την λειτουργία κάθε κόμβου. Για να κατασκευάσουμε ένα προσαρμοστικό δίκτυο χρησιμοποιούμε ένα σύνολο δεδομένων προς εκπαίδευση και κάποιες διαδικασίες όπως οι κανόνες εκπαίδευσης ή οι προσαρμοστικοί αλγόριθμοι για την τροποποίηση των παραμέτρων για να βελτιώσουμε την απόδοση του δικτύου. Συνήθως η απόδοση ενός δικτύου μετριέται ως η διαφορά μεταξύ της επιθυμητής εξόδου και της πραγματικής κάτω από τις ίδιες συνθήκες. Αυτή η διαφορά ονομάζεται μέτρηση σφάλματος. Γενικότερα, ένας κανόνας εκπαίδευσης εξάγεται από την εφαρμογή μιας συγκεκριμένης τεχνικής βελτιστοποίησης για μια δεδομένη μέτρηση σφάλματος.

## *4.7 ΠΑΡΑΔΕΙΓΜΑΤΑ ΠΡΟΣΑΡΜΟΣΤΙΚΩΝ ΔΙΚΤΥΩΝ*

### 4.7.1 ΠΡΟΣΑΡΜΟΣΤΙΚΟ ΔΙΚΤΥΟ ΜΕ ΕΝΑΝ ΓΡΑΜΜΙΚΟ ΚΟΜΒΟ (LINEAR NODE)




Σχήμα 4.14: Ένα γραμμικό προσαρμοστικό δίκτυο

Στο σχήμα 4.14 απεικονίζεται ένα προσαρμοστικό δίκτυο με ένα κόμβο. Ορίζεται : $x_3 = f_3\left(x_1, x_2, ; a_1, a_2, a_3\right) = a_1 x_1 + a_2 x_2 + a_3$

Όπου $x_1, x_2$ είναι είσοδοι και $a_1, a_2, a_3$ είναι οι τροποποιημένες παράμετροι.

Η συνάρτηση προσδιορίζει ένα επίπεδο σε ένα χώρο $x_1 - x_2$ και βάζοντας διάφορες τιμές για τις παραμέτρους, μπορούμε να τοποθετήσουμε αυτό το επίπεδο αυθαίρετα όπως εμείς θέλουμε. Χρησιμοποιώντας το τετραγωνικό σφάλμα ως σφάλμα μέτρησης για το δίκτυο, μπορούμε να αναγνωρίσουμε τις βέλτιστες παραμέτρους μέσω της μεθόδου εκτίμησης των ελαχίστων τετραγώνων.

### 4.7.2 ΔΙΚΤΥΟ PERCEPTON

Εάν προσθέσουμε και έναν άλλο κόμβο ώστε να επιτρέψουμε στην έξοδο του δικτύου στο σχήμα 4.14 να παίρνει μόνο δυο τιμές 0 και 1 τότε παίρνουμε το μη γραμμικό δίκτυο στο σχήμα 4.15.

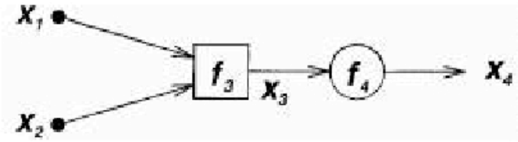
Οι κόμβοι εκφράζονται:

$$x_3 = f_3\left(x_1, x_2, ; a_1, a_2, a_3\right) = a_1 x_1 + a_2 x_2 + a_3$$

και $x_4 = f_4\left(x_3\right) = \begin{cases} 1 & if \quad x_3 \geq 0 \\ 0 & if \quad x_3 < 0 \end{cases}$

όπου $f_3$ είναι μια γραμμική σχέση και $f_4$ είναι μια βηματική συνάρτηση που αντιστοιχεί το $x_3$ είτε στο 0 είτε στο 1.

Την ολική συνάρτηση του δικτύου μπορούμε να την δούμε σαν ένα γραμμικό ταξινομητή. Ο πρώτος κόμβος σχηματίζει ένα όριο απόφασης σαν μια ευθεία γραμμή στον χώρο των $x_1$ - $x_2$ και ο δεύτερος δείχνει σε ποιο ημιεπίπεδο ανήκει το διάνυσμα εισόδου ($x_1, x_2$). Μπορεί να σχηματιστεί ένα ισοδύναμο ημιεπίπεδο με ένα μόνο κόμβο όπου η συνάρτηση των $f_3$ και $f_4$. Ο κόμβος που προκύπτει είναι ένα δομικό υλικό.



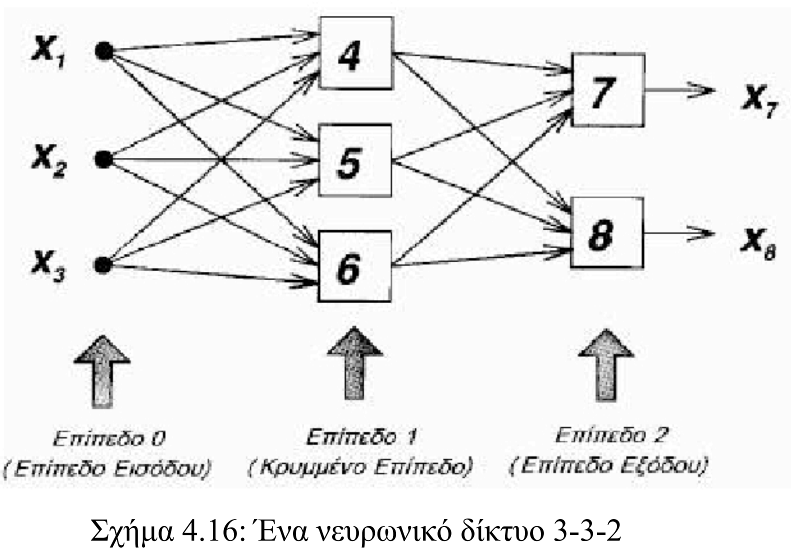

Σχήμα 4.15: Ένα μη γραμμικό προσαρμοστικό δίκτυο με ένα κόμβο.

Μιας και η βηματική συνάρτηση είναι ασυνεχής σε ένα σημείο και επίπεδη σε όλα τα άλλα σημεία. Μια εναλλακτική λύση είναι να χρησιμοποιήσουμε τη σιγμοειδή συνάρτηση σαν μια συθλιπτική συνάρτηση που παίρνει τιμές από 0 έως 1.

$$x_4 = f_4\left(x_3\right) = \frac{1}{1+e^{-x_3}}$$

### 4.7.3 ΕΝΑ ΠΟΛΥΕΠΙΠΕΔΟ PERCEPTON

Στο σχήμα 4.16 βλέπουμε την αρχιτεκτονική για ένα πολυεπίπεδο percepton με τρεις εισόδους, δυο εξόδους και τρεις κρυμμένους κόμβους που δεν συνδέονται απευθείας ούτε σε είσοδο ούτε σε έξοδο.




Σχήμα 4.16: Ένα νευρωνικό δίκτυο 3-3-2

Κάθε κόμβος σε ένα δίκτυο αυτού του τύπου έχει την ίδια συνάρτηση κόμβου, που είναι σύνθεση μιας γραμμικής $f_3$ και μιας σιγμοειδούς $f_4$. Δηλ. η συνάρτηση κόμβου στον κόμβο 7 στο παραπάνω σχήμα είναι:

$$x_7 = \frac{1}{1 + \exp\left[-\left(w_{4,7}x_4 + w_{5,7}x_5 + w_{6,7}x_6 + t_7\right)\right]}$$

όπου $x_4$, $x_5$ και $x_6$ είναι οι έξοδοι από τους κόμβους 4,5 και 6 αντίστοιχα και το σύνολο των παραμέτρων του κόμβου 7 δηλώνεται με το $\{\, w_{4,7}\ w_{5,7}\ w_{6,7}\ t_7 \,\}$. Συνήθως το $w_{i,j}$ είναι το βάρος που σχετίζεται με το σύνδεσμο που ενώνει τον *ι-οστό* κόμβο και τον *j-οστό* κόμβο, και το $t_j$ σαν το κατώφλι (threshold) που συνδέεται με τον κόμβο j.

Γενικά, ένας σύνδεσμος δείχνει μόνο την κατεύθυνση της ροής σήματος μεταξύ των διασυνδεμένων κόμβων.

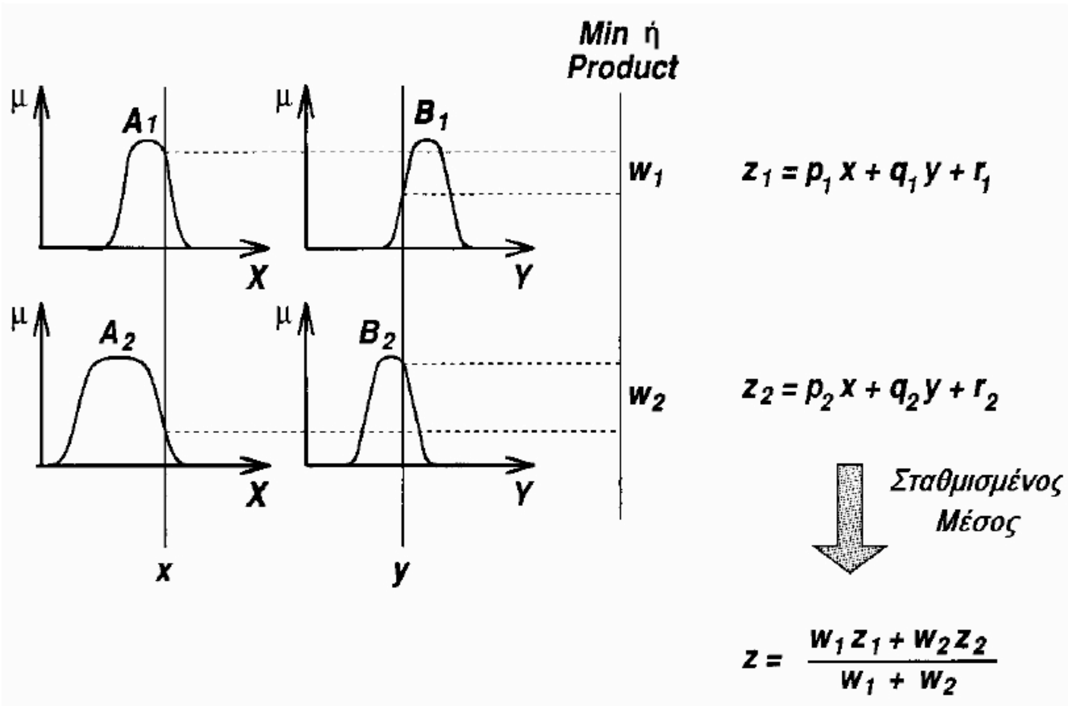
# ΚΕΦΑΛΑΙΟ 5. ΠΡΟΣΑΡΜΟΣΤΙΚΟ ΝΕΥΡΟ-ΑΣΑΦΕΣ ΣΥΣΤΗΜΑ ΕΞΑΓΩΓΗΣ ΣΥΜΠΕΡΑΣΜΑΤΟΣ (ADAPTIVE NEYRO-FUZZY INFERENCES SYSTEM-ANFIS)

## *5.1 ΑΡΧΙΤΕΚΤΟΝΙΚΗ ΤΟΥ ANFIS*

Ο αλγόριθμος ANFIS είναι από τους κυριότερους αλλά και ταυτόχρονα από τους πρώτους που εφαρμόστηκαν στο πεδίο της νεύρο-ασαφούς προσέγγισης προβλημάτων. Παρακάτω περιγράφεται το δίκτυο, υποθέτοντας ότι το πρόβλημα που θα αναλυθεί έχει 2 εισόδους **x** και **y** και μια έξοδο **z**. Υποθέτοντας για ένα 1$^{ης}$ τάξης μοντέλο Sugeno, μια τυπική βάση κανόνων (rule base) θα μπορούσε να είναι η εξής:

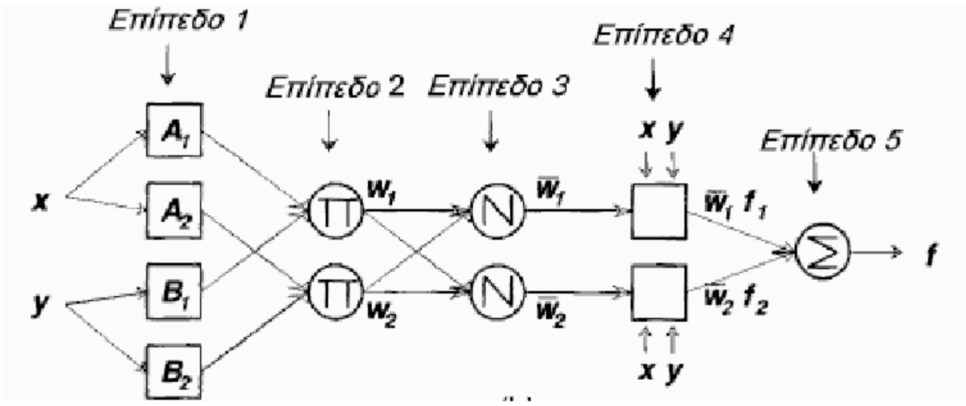
*Κανόνας* 1: Εάν χ είναι $A_1$ και y είναι $B_1$ τότε $f_1 = p_1 \times x + q_1 \times y + r_1$

*Κανόνας* 2: Εάν χ είναι $A_2$ και y είναι $B_2$ τότε $f_2 = p_2 \times x + q_2 \times y + r_2$




Σχήμα 5.1: Ασαφής Συλλογιστική (fuzzy reasoning)

Όπως είναι φανερό, η έξοδος z του πρωτοβάθμιου μοντέλου Sugeno είναι ένας σταθμισμένος μέσος όρος. Η αντίστοιχη αναπαράσταση του δικτύου ANFIS παρουσιάζεται στο παρακάτω σχήμα 5.2.




Σχήμα: 5.2 Η αρχιτεκτονική του ANFIS

Το παραπάνω σχήμα απεικονίζει τον συλλογιστικό μηχανισμό (reasoning) για αυτό το μοντέλο Sugeno και η αντίστοιχη ισοδύναμη αρχιτεκτονική του ANFIS όπου οι κόμβοι του ίδιου επιπέδου έχουν παρόμοιες συναρτήσεις. Παρακάτω παρουσιάζεται πιο αναλυτικά η διεργασία που εκτελείται σε κάθε επίπεδο.

**Επίπεδο 1:** Κάθε κόμβος $i$ σε αυτό το επίπεδο είναι ένας προσαρμόσιμος (adaptive) κόμβος με συνάρτηση κόμβου

$$O_{1,i} = \mu_{A_i}(\chi) \text{ για i=1,2 ή}$$
$$O_{1,i} = \mu_{B_{i-2}}(y) \text{ για i=3,4 ή}$$
$$O_{1,i} = \mu_{C_j}(y) \text{ για i= j+4 και j=1,2,}$$

όπου χ (ή y ή z) – η είσοδος στον κόμβο $i$ , $A_i$ (ή $B_{i-2}$ ή $C_j$ )- η γλωσσική μεταβλητή (small, large, κλπ.) που σχετίζεται με αυτή τη συνάρτηση κόμβου. Με άλλα λόγια, το $O_{1,i}$ είναι ο βαθμός συμμετοχής του $A (= A_1, A_2 \text{ ή } B_1, B_2, C_1 \text{ ή } C_2)$ και καθορίζει το βαθμό στον οποίο η είσοδος $x$ (ή $y$ ή $z$) ικανοποιεί τον ποσοτικοποιητή $A$. Εδώ η συνάρτηση συμμετοχής για το A μπορεί να είναι οποιαδήποτε κατάλληλη παραμετρική συνάρτηση συμμετοχής όπως η καμπανοειδής για παράδειγμα:

$$\mu_{A_i}(x) = e^{\frac{1}{2}\left(\frac{x-c_i}{a_i}\right)^2}$$

όπου $a_i$ , $c_i$ είναι το σύνολο των παραμέτρων.

Καθώς οι τιμές των παραμέτρων αλλάζουν, οι συναρτήσεις ποικίλουν ανάλογα, παρουσιάζοντας έτσι διάφορες μορφές της συνάρτησης συμμετοχής για το ασαφές σύνολο *A*. Οι παράμετροι σε αυτό το επίπεδο αναφέρονται ως ***αρχικοί παράμετροι*** **(premise parameters).**

**Επίπεδο 2:** Κάθε κόμβος σε αυτό το επίπεδο είναι ένας σταθερός (fixed) κόμβος Π, του οπίου η έξοδος είναι το γινόμενο όλων των εισερχόμενων σημάτων:

$$O_{2,i} = w_i = \mu_{A_1}(x) * \mu_{B_i}(y) * \mu_{C_1}(z) \text{ για } i=1,2$$

$$O_{2,i} = w_i = \mu_{A_1}(x) * \mu_{B_{i-2}}(y) * \mu_{C_2}(z) \text{ για } i=3,4$$

$$O_{2,i} = w_i = \mu_{A_2}(x) * \mu_{B_2}(y) * \mu_{C_j}(z) \text{ για } i= j+4,\ j=1,2$$

$$O_{2,i} = w_i = \mu_{A_2}(x) * \mu_{B_2}(y) * \mu_{C_{j-2}}(z) \text{ για } i= j+4,\ j=3,4$$

**Επίπεδο 3:** Κάθε κόμβος σε αυτό το επίπεδο είναι ένας σταθερός κόμβος N. O *i-οστός* κόμβος υπολογίζει το λόγο του βαθμού ενεργοποίησης του *i- οστού* κανόνα στο άθροισμα των βαθμών ενεργοποίησης όλων των κανόνων:

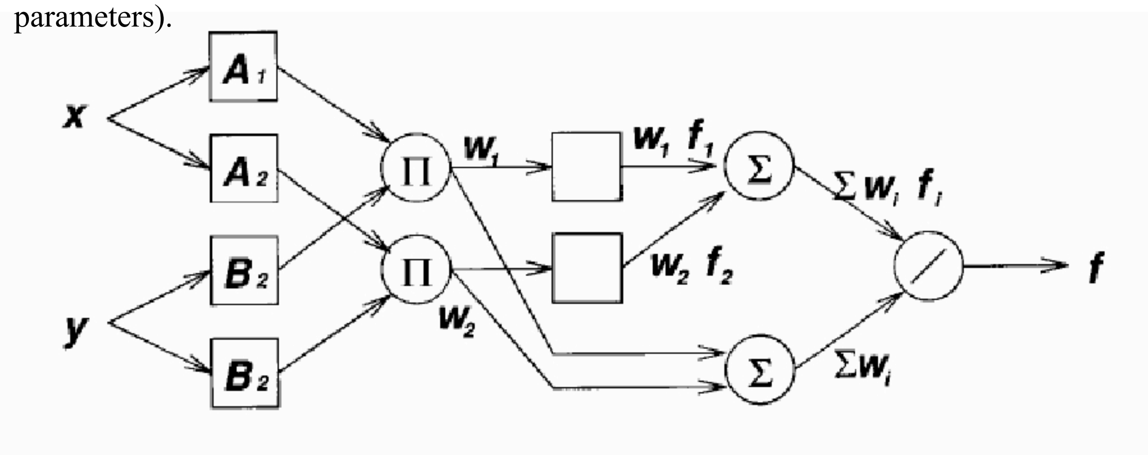
$$O_{3,i} = \overline{w_i} = \frac{w_i}{w_1 + w_2 + w_3 + w_4 + w_5 + w_6 + w_7 + w_8}, \qquad \mathbf{i=1,2,3,4,5,6,7,8.}$$

Για ευκολία, οι έξοδοι αυτού του επιπέδου ονομάζονται κανονικοποιημένοι βαθμοί ενεργοποίησης (normalized firing strengths).

**Επίπεδο 4:** Κάθε κόμβος *i* σε αυτό το επίπεδο είναι ένας προσαρμόσιμος κόμβος με μια συνάρτηση κόμβου.

$$O_{4,i} = \overline{w_i} * f_i = \overline{w_i} * (p_i * x + q_i * y + s_i * z + r_i)$$

όπου: $w_i$ - η έξοδος του επιπέδου 3, $\{p_i, q_i, s_i, r_i\}$- το σύνολο των παραμέτρων. Οι παράμετροι σε αυτό το επίπεδο αναφέρονται ως *επακόλουθοι* παράμετροι (consequent parameters).

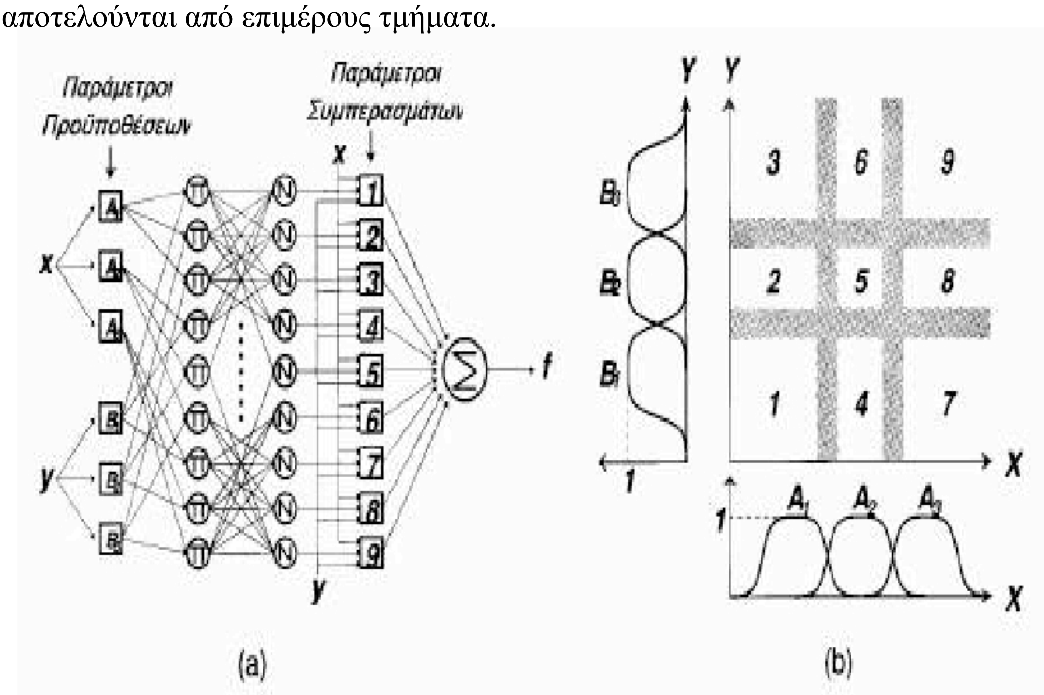



Σχήμα 5.3 Η αρχιτεκτονική του ANFIS για το ασαφές σύνολο Sugeno, όπου η κανονικοποίηση βαρών πραγματοποιείται στο τελευταίο επίπεδο.

**Επίπεδο 5:** Ο μοναδικός κόμβος σε αυτό το επίπεδο είναι ένας σταθερός κόμβος Σ που υπολογίζει τη συνολική έξοδο σαν το ολικό άθροισμα όλων των εισερχόμενων σημάτων:

$$\textbf{\textit{overall output}} = O_{5,i} = \sum_i \overline{w_i} * f_i = \frac{\sum_i w_i * f_i}{\sum_i w_i}$$

Αυτό το προσαρμοστικό δίκτυο είναι λειτουργικά ισοδύναμο με το ασαφές μοντέλο Sugeno. Μπορούμε να συνδυάσουμε τα επίπεδα 3 και 4 για να αποκτήσουμε ισοδύναμο δίκτυο με τέσσερα μόνο επίπεδα. Με το ίδιο δείγμα μπορούμε να πραγματοποιήσουμε την κανονικοποίηση των βαρών στο τελευταίο επίπεδο. Το σχήμα 5.3 απεικονίζει ένα ANFIS αυτού του τύπου. Στην ακραία περίπτωση μπορούμε να συρρικνώσουμε ακόμα και όλο το δίκτυο, σε έναν μόνο προσαρμοστικό κόμβο με το ίδιο σύνολο παραμέτρων. Η ανάθεση συναρτήσεων κόμβων και η σύνθεση του δικτύου είναι αυθαίρετες, εφόσον κάθε κόμβος και κάθε επίπεδο πραγματοποιούν λειτουργίες που είναι σημαντικές και έχουν την δυνατότητα να αποτελούνται από επιμέρους τμήματα.




Σχήμα 5.4 (α) Η αρχιτεκτονική του ANFIS για το ασαφές μοντέλο Sugeno με 2 εισόδους και εννέα κανόνες (β) ο χώρος εισόδου, χωρισμένος σε εννέα ασαφής περιοχές.

Εκτός από το μηχανισμό συμπεράσματος του Sugeno, μπορούμε να κατασκευάσουμε ANFIS και με το μοντέλο Mandami. Το Sugeno χρησιμοποιείται πιο συχνά, μιας και η διακρίνεται για τη διαφάνεια και την αποτελεσματικότητα του. Στο σχήμα 5.4 (α) βλέπουμε μια αρχιτεκτονική ANFIS που είναι ισοδύναμη με ένα ασαφές μοντέλο Sugeno πρώτου βαθμού δυο εισόδων και έντεκα κανόνων, όπου σε κάθε είσοδο θεωρούμε ότι αντιστοιχούν τρεις συναρτήσεις συμμετοχής. Το σχήμα 5.4 (β) απεικονίζει πως δυο διαστάσεων χώρος εισόδου είναι χωρισμένος σε εννέα υπέρθετες (overlapping) ασαφείς περιοχές όπου κάθε μια ελέγχεται από έναν κανόνα if – then. Αυτό σημαίνει ότι το μέρος των προϋποθέσεων ενός κανόνα προσδιορίζει μια ασαφή περιοχή, ενώ το μέρος των συμπερασμάτων προσδιορίζει την έξοδο μέσα στην περιοχή.

## *5.2 ΥΒΡΙΔΙΚΟΣ ΑΛΓΟΡΙΘΜΟΣ ΕΚΠΑΙΔΕΥΣΗΣ*

Το ANFIS χρησιμοποιεί έναν υβριδικό αλγόριθμο εκμάθησης για να προσδιορίσει τις παραμέτρους του ασαφούς συστήματος τύπου Sugeno. Εφαρμόζει έναν συνδυασμό της μεθόδου των ελαχίστων τετραγώνων (least-squares) και της μεθόδου οπισθόδρομης βαθμωτής ελαχιστοποίησης (back propagation gradient descent) για την εκπαίδευση παραμέτρων των συναρτήσεων συμμετοχής του FIS ώστε να μιμηθεί ένα δοσμένο σύνολο δεδομένων εκπαίδευσης.

### 5.2.1 ΕΚΤΙΜΗΤΗΣ ΕΛΑΧΙΣΤΩΝ ΤΕΤΡΑΓΩΝΩΝ (Least-squares estimator)

Στο γενικό πρόβλημα ελαχίστων τετραγώνων, η έξοδος του γραμμικού μοντέλου *y* δίνεται από τη γραμμικά παραμετροποιημένη έκφραση:

$$y = \theta_1 f_1(u) + \theta_2 f_2(u) + \dots \theta_n f_n(u) \qquad 5.1$$

όπου $\boldsymbol{u} = [\, u_1, \dots\dots u_p]^T$ είναι το διάνυσμα εισόδου του μοντέλου $f_1, \dots, f_n$ είναι γνωστές συναρτήσεις του $\boldsymbol{u}$ και $\theta_1, \dots, \theta_n$ είναι γνωστές παράμετροι που θα υπολογιστούν. Η εξίσωση (5.1) καλείται **συνάρτηση παλινδρόμησης**, και τα $\theta_\iota$ ονομάζονται **συντελεστές παλινδρόμησης**.

Για να προσδιοριστούν οι άγνωστες παράμετροι $\theta_\iota$, συνήθως πρέπει να εκτελεστούν πειράματα για να βρεθεί ένα σύνολο δεδομένων εκπαίδευσης που αποτελείται από ζευγάρια δεδομένων $\{(\boldsymbol{u}_i, y_i), i = 1 \dots m\}$ τα οποία αντιπροσωπεύουν τα επιθυμητά ζευγάρια εισόδου – εξόδου του συστήματος που θα μοντελοποιηθεί. Η αντικατάσταση κάθε ζεύγους στοιχείων στην εξίσωση (5.1) παράγει ένα σύνολο γραμμικών εξισώσεων *m*:

$$
\begin{aligned}
&f_1(u_1)\theta_1 + f_2(u_1)\theta_2 + ... + f_n(u_1)\theta_n = y_1\\
&f_1(u_1)\theta_1 + f_2(u_2)\theta_2 + ... + f_n(u_2)\theta_n = y_2\\
&...\\
&f_1(u_1)\theta_1 + f_2(u_n)\theta_2 + ... + f_n(u_n)\theta_n = y_n
\end{aligned}
\qquad 5.2
$$

Σε μορφή πινάκων, οι προηγούμενες εξισώσεις μπορούν να γραφτούν σε συνοπτική μορφή:

**Αθ = y** 5.3

όπου **A** είναι ένας $m \times n$ πίνακας (**πίνακας σχεδιασμού**) :

$$\boldsymbol{A} = \begin{pmatrix} f_1(u_1) & \dots & f_n(u_1) \\ \vdots & \ddots & \vdots \\ f_1(u_m) & \cdots & f_n(u_m) \end{pmatrix}$$

**θ** είναι ένα $n \times 1$ διάνυσμα άγνωστων παραμέτρων:

$$\boldsymbol{\theta} = \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_n \end{bmatrix}$$

και **y** είναι ένα $m \times 1$ διάνυσμα εξόδου:

$$\boldsymbol{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

Η *i-οστή* σειρά του ενωμένου πίνακα δεδομένων $\left[\boldsymbol{A} \vdots \boldsymbol{y}\right]$, που δηλώνεται με $\left[\boldsymbol{a}_i^T \vdots \boldsymbol{y}_i\right]$, σχετίζεται με το *i-οστό* ζευγάρι δεδομένων $\left(\boldsymbol{u}_i ; y_i\right)$ μέσω της

$$\boldsymbol{a}_i^T = \left[f_1(\boldsymbol{u}_i), \dots, f_n(\boldsymbol{u}_i)\right]$$

Το μεγαλύτερο μέρος των υπολογισμών βασίζεται στους πίνακες **A** και **y**, μερικές φορές γίνεται αναφορά στο $\left(\boldsymbol{a}_i^T ; y_i\right)$ σαν το *i-οστό* ζευγάρι δεδομένων του συνόλου δεδομένων εκπαίδευσης.

Για να προσδιοριστεί μεμονωμένα το άγνωστο διάνυσμα **θ**, είναι απαραίτητο να ισχύει $m \geq n$. Αν ο **A** είναι τετραγωνικός (*m*=*n*) και αντιστρέψιμος, τότε η εξίσωση (5.3) μπορεί να λυθεί ως προς τον άγνωστο **x** και γίνεται :

$$\theta = A^{-1} y \qquad 5.4$$

Ο *m* είναι συνήθως μεγαλύτερος από τον *n*, που σημαίνει ότι υπάρχουν περισσότερα ζευγάρια στοιχείων από τις παραμέτρους. Σε αυτήν τη περίπτωση, μια ακριβής λύση που να ικανοποιεί όλες τις *m* εξισώσεις δεν είναι πάντα δυνατή, δεδομένου ότι τα στοιχεία μπορεί_να ΄΄ μολυνθούν ΄΄ από θόρυβο, ή το μοντέλο μπορεί να μην είναι κατάλληλο για τη περιγραφή του συστήματος στόχου. Κατά συνέπεια η εξίσωση (5.3) πρέπει να τροποποιηθεί με την ενσωμάτωση ενός διανύσματος λάθους **e** για να αποτελέσει το τυχαίο λάθος θορύβου ή το τυχαίο λάθος διαμόρφωσης ως εξής:

$$A\theta + e = y \qquad 5.5$$

Τώρα, αντί της εύρεσης της ακριβούς λύσης στην εξίσωση (5.3), πρέπει να βρεθεί το $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$ που ελαχιστοποιεί το άθροισμα του τετραγωνικού σφάλματος και ορίζεται ως:

$$E(\theta) = \sum_{i=1}^{m} (y_i - a_i^r \theta)^2 = e^r e = (y - A\theta)^r (y - A\theta) \qquad 5.6$$

όπου **e = y - Αθ** είναι το διάνυσμα λάθους που παράγεται από μια συγκεκριμένη επιλογή του **θ**. πρέπει να σημειωθεί ότι το $E(\theta)$ είναι σε τετραγωνική μορφή και έχει ένα μοναδικό ελάχιστο $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$. Το ακόλουθο θεώρημα δηλώνει έναν απαραίτητο όρο που ικανοποιείται από τον εκτιμητή ελαχίστων τετραγώνων $\widehat{\boldsymbol{\theta}}$.

***Θεώρημα 5.2.1.1*** Εκτιμητής ελαχίστων τετραγώνων

Το τετραγωνικό σφάλμα στην εξίσωση (5.6) ελαχιστοποιείται όταν $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$, ο οποίος καλείται εκτιμητής ελαχίστων τετραγώνων (LSE) και ο οποίος ικανοποιεί την κανονική εξίσωση

$$A^r A\hat{\theta} = A^r y \qquad 5.7$$

Αν ο $A^T A$ είναι αντιστρέψιμος, ο $\widehat{\boldsymbol{\theta}}$ είναι μοναδικός και δίνεται από την εξίσωση

$$\hat{\theta} = \left(A^T A\right)^{-1} A^{-T} y \qquad 5.8$$

## 5.2.2 ΟΠΙΣΘΟΔΡΟΜΗΣΗ ΓΙΑ ΠΡΩΣΟ-ΤΡΟΦΟΔΟΤΟΥΜΕΝΑ ΔΙΚΤΥΑ

Η ενότητα αυτή παρουσιάζει ένα βασικό κανόνα εκμάθησης α προσαρμόσιμα δίκτυα, που είναι στην ουσία η πιο απλή μέθοδος βαθμωτής ελαχιστοποίησης. Το κεντρικό μέρος αυτού του κανόνα εκμάθησης αφορά στο πώς να επιλεγεί επαναληπτικά ένα διάνυσμα κλίση στο οποίο κάθε στοιχείο ορίζεται ως η παράγωγος ενός μέτρου σφάλματος ως προς μια παράμετρο. Αυτό γίνεται με τη βοήθεια του κανόνα αλυσίδας, ενός βασικού τύπου για το διαφορισμό σύνθετων συναρτήσεων ο οποίος αναλύεται σε κάθε εγχειρίδιο μαθηματικών. Η διαδικασία εύρεσης ενός διανύσματος κλίσης σε ένα δίκτυο αναφέρεται γενικά σαν **οπισθοδρόμηση** (back propagation) επειδή το διάνυσμα κλίσης υπολογίζεται σε κατεύθυνση αντίθετη από τη ροή της εξίσου κάθε κόμβου. Μόλις επιλεγεί η κλίση, διάφορες τεχνικές βελτιστοποίησης και παλινδρόμησης βασισμένες στις παραγώγους είναι διαθέσιμες για την ενημέρωση των παραμέτρων. Ειδικότερα, αν χρησιμοποιούμε το διάνυσμα κλίσης σε μια απλή μέθοδο βαθμωτής ελαχιστοποίησης, το προκύπτον παράδειγμα εκμάθησης αναφέρεται συχνά σαν **κανόνας οπισθόδρομης εκμάθησης**.

Έστω ότι δεδομένο πρωσοτροφοδοτούμενο προσαρμοστικό δίκτυο στην αναπαράσταση του σε επίπεδα έχει *L* επίπεδα και το επίπεδο *I* (όπου *l*= 0,1…,L;*l*=0 αντιπροσωπεύει το επίπεδο εισαγωγής) έχει *N(l)*. Τότε η έξοδος και η συνάρτηση του κόμβου *i [i*=1,…, *N(l)*] στο επίπεδο *l* μπορούν να αναπαρασταθούν σαν $x_{l,i}$ και $f_{l,i}$, αντίστοιχα. Ας υποτεθεί ότι δεν υπάρχει καμία σύνδεση μεταξύ μη συνεχόμενων στρωμάτων. Δεδομένου ότι η έξοδος ενός κόμβου εξαρτάται από τα συνεχόμενα σήματα και το σύνολο των παραμέτρων του κόμβου, προκύπτει η ακόλουθη γενική έκφραση για τη συνάρτηση των κόμβων $f_{l,i}$:

$$x_{l,i} = f_{l,i}\left(x_{l-1,1},\ldots,x_{l,N(l-1)},\alpha,\beta,\gamma,\ldots\right) \qquad 5.9$$

Όπου α, β, γ κλπ. Είναι οι παράμετροι αυτού του κόμβου.

Υποθέτοντας ότι το δοσμένο σύνολο δεδομένων εκπαίδευσης έχει *P* καταχωρήσεις, μπορεί να οριστεί ένα μέτρο σφάλματος για την *p-οστή* (1≤*p*≤*P*) καταχώρηση των δεδομένων εκπαίδευσης σαν το άθροισμα των τετραγωνικών σφαλμάτων:

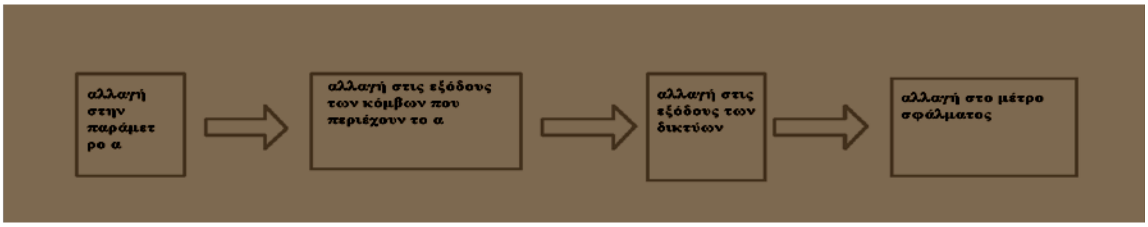
$$E_p = \sum_{k=1}^{N(l)}\left(d_k - x_{L,k}\right)^2 \qquad 5.10$$

Όπου $d_k$ είναι το *k-οστό* συστατικό του πραγματικού διανύσματος εξόδου και $x_{L,k}$ είναι το *k-οστό* συστατικό του πραγματικού διανύσματος εξίσου που παράγεται με την παρουσίαση του *p-οστού* διανύσματος εισόδου στο δίκτυο. (Για σημειογραφική απλότητα, παραλείπεται ο δείκτης *p* και από το $d_k$ και από το $x_{L,k}$). Προφανώς, όταν το $E_p$ είναι ίσο με το μηδέν, το δίκτυο είναι ικανό να αναπαράγει ακριβώς το επιθυμητό διάνυσμα εξόδου στο *p-στο* ζευγάρι δεδομένων εκπαίδευσης. Ο στόχος εδώ είναι να ελαχιστοποιηθεί ένα συνολικό μέτρο σφάλματος, που ορίζεται ως:

$$E = \sum\nolimits_{p=1}^{P} E_p$$

Επιπλέον, ας υποτεθεί ότι το $E_p$ εξαρτάται μόνο από τους κόμβους εξόδου.

Για να χρησιμοποιηθεί η βαθμωτή ελαχιστοποίηση ώστε να ελαχιστοποιηθεί το μέτρο σφάλματος, πρέπει να βρεθεί το διάνυσμα κλίσης. Πριν υπολογιστεί το διάνυσμα κλίσης, πρέπει να παρατηρηθούν οι ακόλουθες αιτιώδεις σχέσεις:




όπου τα βέλη ⇒ δείχνουν τις αιτιώδεις σχέσεις. Δηλαδή, μια μικρή αλλαγή σε μια παράμετρο ***α*** θα επηρεάσει την έξοδο του κόμβου που περιέχει το ***α***. Αυτό με τη σειρά του θα επηρεάσει την έξοδο του τελευταίου επιπέδου και συνεπώς το μέτρο του σφάλματος.

Η βασική αρχή στον υπολογισμό του διανύσματος κλίσης είναι να περαστεί μια σειρά από πληροφορίες παραγώγων ξεκινώντας από το επίπεδο εξόδου και πηγαίνοντας ανάποδα από επίπεδο σε επίπεδο έως ότου καταλήξει η διαδικασία στο επίπεδο εισόδου.

Το σήμα σφάλματος $\in_{l,i}$ ορίζεται σαν τη παράγωγο του μέτρου σφάλματος $E_p$ ως προς την έξοδο του κόμβου *i* στο επίπεδο *l*, λαμβάνοντας υπόψη και τις έμμεσες πορείες.

$$\in_{i,i} = \frac{\partial^+ E_p}{\partial x_{i,i}} \qquad 5.11$$

Η έκφραση αυτή ονομάστηκε διατεταγμένη παράγωγος Werbos.

Το σήμα σφάλματος για τον *i-οστό* κόμβο εξόδου (στο επίπεδο *L*) μπορεί να υπολογιστεί απ’ ευθείας:

$$\in_{1,i} = \frac{\partial^+ E_p}{\partial x_{1,i}} = \frac{\partial E_p}{\partial x_{1,i}} \qquad 5.12$$

Αυτό ισούται με $\in_{L,i} = -2\left(d_i - x_{L,i}\right)$αν το $E_p$ ορίζεται όπως στην εξίσωση (5.10). για τον εσωτερικό κόμβο στην *i-οστή* θέση του επιπέδου *l*, το σήμα σφάλματος μπορεί να βρεθεί από τον κανόνα αλυσίδας:

$$\in_{l,i} = \underbrace{\frac{\partial^+ E_p}{\partial x_{l,i}}}_{\substack{error\ signal \\ at\ layer\ l}} = \sum_{m=1}^{N(l+1)} \underbrace{\frac{\partial^+ E_p}{\partial x_{l+1,m}}}_{\substack{error\ signal \\ at\ layer\ l+1}} \frac{\partial^+ f_{l+1,m}}{\partial x_{1,i}} = \sum_{m=1}^{N(l+1)} \in_{l+1,m} \frac{\partial f_{l+1,m}}{\partial x_{l,i}} \qquad 5.13$$

όπου $0 \le l \le L-1$. Δηλαδή το σήμα σφάλματος ενός εσωτερικού κόμβου στο στρώμα *l* μπορεί να εκφραστεί σαν γραμμικός συνδυασμός των κόμβων στο στρώμα *l*+1. Επομένως, για οποιαδήποτε *l* και *i* [και $1 \le i \le N(l)$], μπορούν να βρεθούν τα:

$$\in_{l,i} = \frac{\partial^+ E_p}{\partial x_{l,i}}$$

εφαρμόζοντας πρώτα την εξίσωση (5.12) μια φορά για να υπολογιστούν τα σήματα σφάλματος στο επίπεδο εξόδου, και έπειτα εφαρμόζοντας την εξίσωση (5.13) επαναληπτικά έως ότου καταλήξει η διαδικασία στο επιθυμητό επίπεδο *l*. Η διαδικασία αυτή καλείται οπισθοδρόμηση δεδομένου ότι τα σήματα σφάλματος λαμβάνονται διαδοχικά από το επίπεδο εξόδου προς το επίπεδο εισόδου.

Το διάνυσμα κλίσης ορίζεται ως η παράγωγος του μέτρου σφάλματος ως προς κάθε παράμετρο, έτσι πρέπει να εφαρμοστεί ο κανόνας της αλυσίδας ξανά για να βρεθεί το διάνυσμα κλίσης. Εάν ***a*** είναι μια παράμετρος του *i-στού* κόμβου στο επίπεδο *l*, τότε ισχύει:

$$\frac{\partial^+ E_p}{\partial a} = \frac{\partial^+ E_p}{\partial x_{l,i}} \frac{\partial f_{l,i}}{\partial a} = \in_{l,i} \frac{\partial f_{l,i}}{\partial a} \qquad 5.14$$

Ας σημειωθεί ότι αν επιτραπεί στην παράμετρο **α** να μοιράζεται μεταξύ διαφορετικών κόμβων, τότε η εξίσωση (5.14) πρέπει να αλλαχθεί σε μια πιο γενική μορφή:

$$\frac{\partial^+ E_p}{\partial a} = \sum_{x^* \in S} \frac{\partial^+ E_p}{\partial x^*} \frac{\partial f^*}{\partial a} \qquad 5.15$$

όπου S είναι το σύνολο των κόμβων που περιέχουν το $a$ σαν παράμετρο ενώ $x^*$ και $f^*$ είναι η έξοδος και η συνάρτηση, αντίστοιχα, ενός γενικού κόμβου στο S. Η παράγωγος του γενικού μέτρου σφάλματος $E$ ως προς το $a$ είναι:

$$\frac{\partial^+ E_p}{\partial a} = \sum_{p=1}^{P} \frac{\partial^+ E_p}{\partial a} \qquad 5.16$$

Συνεπώς, για την απλούστερη βαθμωτή ελαχιστοποίηση χωρίς ελαχιστοποίηση γραμμών, ο τύπος για τη γενική παράμετρο $a$ είναι:

$$\Delta\alpha = -\eta \frac{\partial^+ E_p}{\partial a} \qquad 5.17$$

όπου το n είναι ο βαθμός εκμάθησης, το οποίο μπορεί να εκφραστεί περεταίρω ως

$$\eta = \frac{\kappa}{\sqrt{\sum_\alpha \left(\frac{\partial E}{\partial a}\right)^2}} \qquad 5.18$$

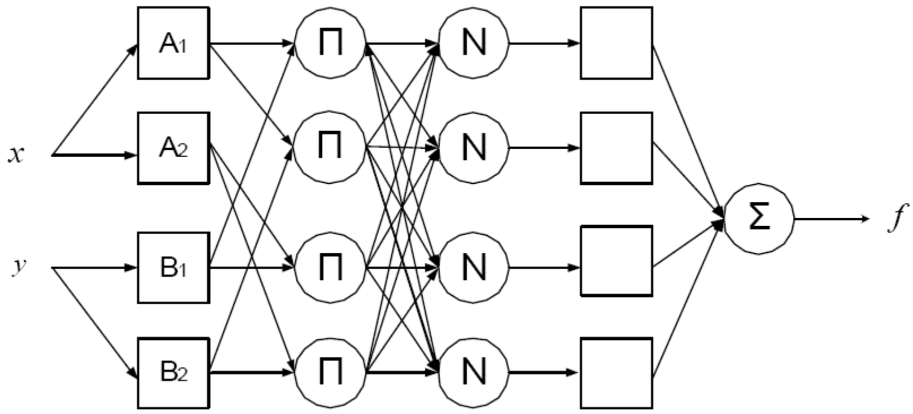
όπου το $\kappa$ είναι το μέγεθος βήματος, το μήκος δηλαδή κάθε μετάβασης κατά μήκος της κατεύθυνσης κλίσης στο διάστημα παραμέτρου. Συνήθως το μέγεθος του βήματος μπορεί να αλλαχθεί για να μεταβληθεί η ταχύτητα σύγκλισης.

Όταν ένα πρωσοτροφοδοτούμενο δίκτυο n-κόμβων αναπαριστάται στην τοπολογική του διάταξη, μπορεί να υπολογίζει το μέτρο σφάλματος $E_p$ σαν την έξοδο ενός επιπλέον κόμβου με ένδειξη $n$+1, του οποίου η συνάρτηση κόμβου $f_{n+1}$ μπορεί να οριστεί από τις εξόδους κάθε κόμβου με μικρότερη ένδειξη. (Επομένως αν το $E_p$ μπορεί να εξαρτάται άμεσα από οποιουσδήποτε κόμβους ). Εφαρμόζοντας πάλι τον κανόνα αλυσίδας του σήματος σφάλματος $\in_i = {\partial E_p}/{\partial x_i}$

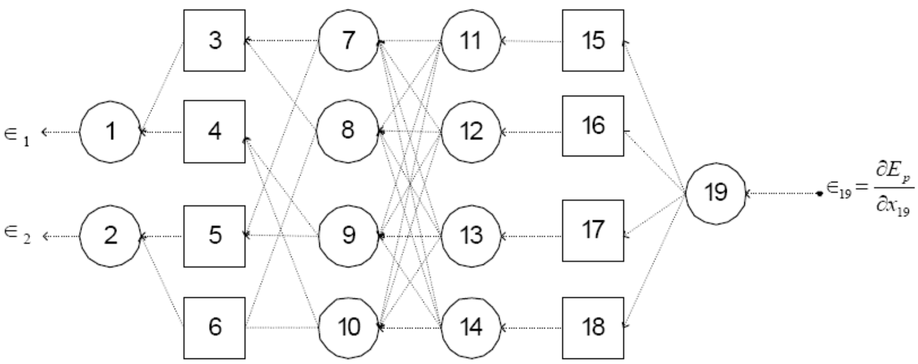
$$\in_i = \frac{\partial f_{n+1}}{\partial x_i} + \sum_{i<j\le n} \in_j \frac{\partial f_j}{\partial x_i} \qquad 5.19$$

όπου ο πρώτος όρος δείχνει μια άμεση επίδραση του $x_i$ στο $E_p$ μέσω της άμεσης διαδρομής από τον κόμβο *i* στον κόμβο *n*+1 και κάθε όρος παραγώγου στο άθροισμα δείχνει την έμμεση επίδραση του $x_i$ στο $E_p$. Μόλις βρεθεί το σήμα σφάλματος για κάθε κόμβο, τότε το διάνυσμα κλίσης για τις παραμέτρους παράγεται όπως πριν.

Ένας άλλος συστηματικός τρόπος να υπολογιστούν τα σήματα σφάλματος είναι μέσω της αναπαράστασης του δικτύου διάδοσης σφάλματος. Το ακόλουθο σχήμα δείχνει το μοντέλο ANFIS και το δίκτυο διάδοσης σφάλματος του:



Σχήμα 5.5 (a): Το προσαρμόσιμο δίκτυο



Σχήμα 5.5 (b): Το δίκτυο διάδοσης σφάλματος

Τώρα θα υπολογιστούν τα σήματα σφάλματος στους εσωτερικούς κόμβους. Χρησιμοποιούνται τα $f_i$ και $x_i$ για το συμβολισμό της συνάρτησης και της εξόδου του κόμβου *i*. Η έξοδος του κόμβου *i* είναι το σήμα σφάλματος αυτού του κόμβου στο πραγματικό προσαρμόσιμο δίκτυο. Σε σύμβολα, αν επιλεχθεί το τετραγωνικό μέτρο σφάλματος για $E_p$, τότε ισχύει το ακόλουθο: $\in_{19} = -2\left(d_{19} - x_{19}\right)$

Αυτό επειδή ο κόμβος 19 είναι μόνο ένας κόμβος προσωρινής αποθήκευσης στο δίκτυο διάδοσης σφάλματος. Για τους κόμβους 15, 16, 17 και 18 ισχύουν:

$$\in_{18} = \frac{\partial^+ E_p}{\partial x_{18}} = \frac{\partial^+ E_p}{\partial x_{19}} \frac{\partial f_{19}}{\partial x_{18}} = \in_{19} \frac{\partial f_{19}}{\partial x_{18}}$$

$$\in_{17} = \frac{\partial^+ E_p}{\partial x_{17}} = \frac{\partial^+ E_p}{\partial x_{19}} \frac{\partial f_{19}}{\partial x_{17}} = \in_{19} \frac{\partial f_{19}}{\partial x_{17}}$$

$$\in_{16} = \frac{\partial^+ E_p}{\partial x_{16}} = \frac{\partial^+ E_p}{\partial x_{19}} \frac{\partial f_{19}}{\partial x_{16}} = \in_{19} \frac{\partial f_{19}}{\partial x_{16}}$$

$$\in_{15} = \frac{\partial^+ E_p}{\partial x_{15}} = \frac{\partial^+ E_p}{\partial x_{19}} \frac{\partial f_{19}}{\partial x_{15}} = \in_{19} \frac{\partial f_{19}}{\partial x_{15}}$$

Αυτό είναι επειδή όλοι αυτοί οι κόμβοι εξαρτώνται από τον κόμβο 19. Αυτό ισχύει και για τους κόμβους 11, 12, 13 και 14 με τη διαφορά ότι κάθε ένας από αυτούς εξαρτάται από διαφορετικό κόμβο:

$$\in_{14} = \frac{\partial^+ E_p}{\partial x_{14}} = \frac{\partial^+ E_p}{\partial x_{18}} \frac{\partial f_{18}}{\partial x_{14}} = \in_{18} \frac{\partial f_{18}}{\partial x_{14}}$$

$$\in_{13} = \frac{\partial^+ E_p}{\partial x_{13}} = \frac{\partial^+ E_p}{\partial x_{17}} \frac{\partial f_{17}}{\partial x_{13}} = \in_{17} \frac{\partial f_{17}}{\partial x_{13}}$$

$$\in_{12} = \frac{\partial^+ E_p}{\partial x_{12}} = \frac{\partial^+ E_p}{\partial x_{16}} \frac{\partial f_{16}}{\partial x_{12}} = \in_{16} \frac{\partial f_{16}}{\partial x_{12}}$$

$$\in_{11} = \frac{\partial^+ E_p}{\partial x_{11}} = \frac{\partial^+ E_p}{\partial x_{15}} \frac{\partial f_{15}}{\partial x_{11}} = \in_{15} \frac{\partial f_{15}}{\partial x_{11}}$$

Αντιθέτως, οι κόμβοι 7, 8, 9 και 10 εξαρτώνται από τέσσερις διαφορετικούς κόμβους όπως φαίνεται στο σχήμα 5.5. Έτσι ισχύει:

$$\in_{10} = \frac{\partial^+ E_p}{\partial x_{10}} = \in_{14} \frac{\partial f_{14}}{\partial x_{10}} + \in_{13} \frac{\partial f_{13}}{\partial x_{10}} + \in_{12} \frac{\partial f_{12}}{\partial x_{10}} + \in_{11} \frac{\partial f_{11}}{\partial x_{10}}$$

$$\in_9 = \frac{\partial^+ E_p}{\partial x_9} = \in_{14} \frac{\partial f_{14}}{\partial x_9} + \in_{13} \frac{\partial f_{13}}{\partial x_9} + \in_{12} \frac{\partial f_{12}}{\partial x_9} + \in_{11} \frac{\partial f_{11}}{\partial x_9}$$

$$\in_8 = \frac{\partial^+ E_p}{\partial x_8} = \in_{14} \frac{\partial f_{14}}{\partial x_8} + \in_{13} \frac{\partial f_{13}}{\partial x_8} + \in_{12} \frac{\partial f_{12}}{\partial x_8} + \in_{11} \frac{\partial f_{11}}{\partial x_8}$$

$$\in_7 = \frac{\partial^+ E_p}{\partial x_7} = \in_{14} \frac{\partial f_{14}}{\partial x_7} + \in_{13} \frac{\partial f_{13}}{\partial x_7} + \in_{12} \frac{\partial f_{12}}{\partial x_7} + \in_{11} \frac{\partial f_{11}}{\partial x_7}$$

Οι κόμβοι 3, 4, 5 και 6 εξαρτώνται από δυο διαφορετικούς κόμβους και σαν αποτέλεσμα προκύπτουν οι ακόλουθες σχέσεις:

$$\in_6 = \frac{\partial^+ E_p}{\partial x_6} = \frac{\partial^+ E_p}{\partial x_{10}} \frac{\partial f_{10}}{\partial x_6} + \frac{\partial^+ E_p}{\partial x_8} \frac{\partial f_8}{\partial x_6} = \in_{10} \frac{\partial f_{10}}{\partial x_6} + \in_8 \frac{\partial f_8}{\partial x_6}$$

$$\in_5 = \frac{\partial^+ E_p}{\partial x_5} = \frac{\partial^+ E_p}{\partial x_9} \frac{\partial f_9}{\partial x_5} + \frac{\partial^+ E_p}{\partial x_7} \frac{\partial f_7}{\partial x_5} = \in_9 \frac{\partial f_9}{\partial x_5} + \in_7 \frac{\partial f_7}{\partial x_5}$$

$$\in_4 = \frac{\partial^+ E_p}{\partial x_4} = \frac{\partial^+ E_p}{\partial x_{10}} \frac{\partial f_{10}}{\partial x_4} + \frac{\partial^+ E_p}{\partial x_9} \frac{\partial f_9}{\partial x_4} = \in_{10} \frac{\partial f_{10}}{\partial x_4} + \in_9 \frac{\partial f_9}{\partial x_4}$$

$$\in_3 = \frac{\partial^+ E_p}{\partial x_3} = \frac{\partial^+ E_p}{\partial x_8} \frac{\partial f_8}{\partial x_3} + \frac{\partial^+ E_p}{\partial x_7} \frac{\partial f_7}{\partial x_3} = \in_8 \frac{\partial f_8}{\partial x_3} + \in_7 \frac{\partial f_7}{\partial x_3}$$

Τέλος, οι κόμβοι 1 και 2 από δυο κόμβους και έτσι ισχύει:

$$\in_2 = \frac{\partial^+ E_p}{\partial x_2} = \frac{\partial^+ E_p}{\partial x_6} \frac{\partial f_6}{\partial x_2} + \frac{\partial^+ E_p}{\partial x_5} \frac{\partial f_5}{\partial x_2} = \in_6 \frac{\partial f_2}{\partial x_2} + \in_5 \frac{\partial f_5}{\partial x_2}$$

$$\in_1 = \frac{\partial^+ E_p}{\partial x_1} = \frac{\partial^+ E_p}{\partial x_4} \frac{\partial f_4}{\partial x_1} + \frac{\partial^+ E_p}{\partial x_3} \frac{\partial f_3}{\partial x_1} = \in_4 \frac{\partial f_4}{\partial x_1} + \in_3 \frac{\partial f_3}{\partial x_1}$$

## *5.3. ΣΥΝΔΙΑΣΜΟΣ ΒΑΘΜΩΤΗΣ ΕΛΑΧΙΣΤΟΠΟΙΗΣΗΣ ΚΑΙ ΕΚΤΙΜΗΤΗ ΕΛΑΧΙΣΤΩΝ ΤΕΤΡΑΓΩΝΩΝ*

Η έξοδος ενός προσαρμόσιμου δικτύου είναι γραμμική σε μερικές από τις παραμέτρους του δικτύου. Έτσι μπορούν να προσδιοριστούν αυτές οι γραμμικές παράμετροι με τη γραμμική μέθοδο ελαχίστων τετραγώνων που περιγράφτηκε παραπάνω. Αυτή η προσέγγιση οδηγεί σε έναν υβριδικό κανόνα εκμάθησης που συνδυάζει τη βαθμωτή ελαχιστοποίηση (SD) και τον εκτιμητή ελαχίστων τετραγώνων (LSE) για γρήγορο προσδιορισμό παραμέτρων.

Από την αρχιτεκτονική δομή του ANFIS παρατηρούμε ότι οι τιμές των παραμέτρων από τις προϋποθέσεις είναι προκαθορισμένες (αμετάβλητες), η ολική έξοδος μπορεί να εκφραστεί σαν ο γραμμικός συνδυασμός των παραμέτρων από τα συμπεράσματα. Σε σύμβολα, η έξοδος $f$ στο σχήμα 5.4(β) ξαναγράφεται:

$$
\begin{aligned}
f &= \frac{w_1}{w_1+w_2} f_1 + \frac{w_2}{w_1+w_2} f_1 \\
&= \overline{w_1}\left(p_1 x + q_1 y + r_1\right) + \overline{w_2}\left(p_2 x + q_2 y + r_2\right) \\
&= \left(\overline{w_1} x\right) p_1 + \left(\overline{w_1} y\right) q_1 + \left(\overline{w_1}\right) r_1 + \left(\overline{w_2} x\right) p_2 + \left(\overline{w_2} y\right) q_1 + \left(\overline{w_2}\right) r_2
\end{aligned}
$$

η οποία είναι γραμμική στις παραμέτρους των συμπερασμάτων $p_1$, $q_1$, $r_1$, $p_2$, $q_1$ και $r_2$.

Έστω ότι:

$S$ = σύνολο όλων των παραμέτρων

$S_1$ = σύνολο των μη (μη γραμμικών) παραμέτρων από τις προϋποθέσεις

$S_2$ = σύνολο (γραμμικών) παραμέτρων από τα συμπεράσματα

Το προσαρμόσιμο δίκτυο έχει μια έξοδο που αναπαριστάται ως:

$o = F(i, S)$ 5.20

όπου i είναι το διάνυσμα των μεταβλητών εισόδου, S είναι το σύνολο των παραμέτρων και $F$ είναι η συνολική συνάρτηση που εφαρμόζεται από το προσαρμόσιμο δίκτυο

στην εξίσωση $S = S_1 \oplus S_2$ 5.21

και H(.) και F(.) είναι η ταυτοτική συνάρτηση και η συνάρτηση του FIS αντιστοίχως στην εξίσωση $H \circ F$ είναι γραμμική στα στοιχεία του $S_2$, τότε εφαρμόζοντας την $H$ στην εξίσωση (5.20), προκύπτει ότι $H\left(o\right) = H \circ F\left(B_i, S\right)$.

Το $\oplus$ αντιπροσωπεύει την άμεση άθροιση (direct sum).

Η H(.) είναι ταυτοτική συνάρτηση και η F(.,.) είναι η συνάρτηση του συστήματος ασαφούς συμπεράσματος αντίστοιχα. Κατά συνέπεια, ο υβριδικός αλγόριθμος μάθησης που αναπτύχθηκε παραπάνω μπορεί να εφαρμοστεί άμεσα.

Αναλυτικότερα, όπως συμβαίνει και στα νευρωνικά δίκτυα, θα ξεχωρίσουν δύο φάσεις στη διαδικασία εκπαίδευσης:

1. Πέρασμα προς τα εμπρός (forward pass): Στη φάση αυτή το σήμα εισόδου διαδίδεται από το επίπεδο 1 μέχρι το επίπεδο 4 και οι παράμετροι ($p_i, q_i, r_i$) i=1,2 εκτιμώνται με τη μέθοδο ελαχίστων τετραγώνων.
2. Πέρασμα προς τα πίσω (backward pass): Στη φάση αυτή έχοντας μια ένδειξη του λάθους, πραγματοποιείται μια διόρθωση κατά μια ποσότητα που ορίζεται

από τη μέθοδο της βαθμωτής κατάβασης (Gradient Descent), των μεταβλητών $a_i$, $b_i$, $c_i$

Το ANFIS διασπά το σύνολο των παραμέτρων του σε δυο υποσύνολα, εκ των οποίων το ένα αποτελεί γραμμικό σύνολο παραμέτρων και, επομένως, είναι δυνατό να εκπαιδευτεί με γραμμικούς αλγορίθμους, όπως η μέθοδος των ελαχίστων τετραγώνων (least squared). Οι γραμμικοί αλγόριθμοι εκμάθησης επιτυγχάνουν ´´ συνολικά ελάχιστα ´´ (global minimums) της συνάρτησης κόστους στο χώρο των παραμέτρων τους και είναι αποδοτικοί από πλευράς απαιτούμενου υπολογιστικού χρόνου. Το δεύτερο υποσύνολο παραμέτρων εκπαιδεύεται με αλγόριθμους που μπορούν να δημιουργήσουν μη-γραμμικές απεικονίσεις, όπως η βαθμωτή κατάβαση (gradient descent). Οι αλγόριθμοι αυτοί είναι απαιτητικοί από πλευράς υπολογιστικού χρόνου εκπαίδευσης και δεν υπάρχει εγγύηση για την πραγματοποίηση του συνολικού ελαχίστου της συνάρτησης κόστους στο χώρο τον παραμέτρων τους. Ο αλγόριθμος εκμάθησης του ANFIS συνδυάζει περάσματα προς τα εμπρός και προς τα πίσω. Στο πέρασμα εμπρός γίνεται η μάθηση των γραμμικών παραμέτρων και στο πέρασμα προς τα πίσω γίνεται η προσαρμογή των μη-γραμμικών αντίστοιχα. Στον παρακάτω πίνακα 5.3.1συνοψίζονται οι δραστηριότητες του κάθε περάσματος.

Πίνακας 5.3.1: Δραστηριότητες κατά την εκπαίδευση

| | **εμπροσθόδρομο πέρασμα** | **οπισθόδρομο πέρασμα** |
|---|---|---|
| παράμετροι των προϋποθέσεων | προκαθορισμένες (αμετάβλητες) | βαθμωτή φθίνουσα μέθοδος |
| παράμετροι των συμπερασμάτων | εκτιμήτρια ελαχίστων τετραγώνων | προκαθορισμένες (αμετάβλητες) |
| σήματα | έξοδοι των κόμβων | σήματα σφάλματος |

Το πλεονέκτημα της μεθόδου έγκειται στο γεγονός της ύπαρξης τόσο γραμμικών όσο και μη γραμμικών μεθόδων, που την καθιστά ταχύτερη από τα κλασσικά νευρωνικά δίκτυα. Οι παράμετροι των συμπερασμάτων που ευρίσκονται με αυτόν τον τρόπο, είναι βέλτιστες υπό την προϋπόθεση ότι οι παράμετροι των προϋποθέσεων είναι προκαθορισμένες (αμετάβλητες). Αναλόγως, η υβριδική προσέγγιση συγκλίνει πολύ γρηγορότερα μιας και μειώνει τις διαστάσεις του διαστήματος αναζήτησης της αρχικής καθαρής μεθόδου οπισθόδρομης διάδοσης.

Μέχρι τώρα χρησιμοποιούσαμε προκαθορισμένες και αυθαίρετα επιλεγμένες συναρτήσεις συμμετοχής. Πολλές φορές έχουμε μια συλλογή από δεδομένα εισόδου / εξόδου και δεν μπορούμε να χρησιμοποιήσουμε ένα προκαθορισμένο μοντέλο βασισμένο στα στοιχεία αυτά ώστε να ξέρουμε ποιες συναρτήσεις συμμετοχής και με ποιες παραμέτρους να χρησιμοποιήσουμε. Τότε μπορούμε να βοηθηθούμε από το ANFIS, ώστε να βρούμε τις καταλληλότερες συναρτήσεις συμμετοχής.

Οι νεύρο-προσαρμοστικές τεχνικές είναι αρκετά απλές. Έτσι παρέχεται μια μέθοδος ώστε το ασαφές μοντέλο να εκπαιδευτεί με τις πληροφορίες που του δίνει το σύνολο των δεδομένων, ώστε να υπολογιστούν οι παράμετροι των συναρτήσεων συμμετοχής με σκοπό την εύρεση του καλύτερου τρόπου που θα επιτρέπει στο σύστημα εξαγωγής συμπερασμάτων να ανιχνεύει τα δεδομένα εισόδου / εξόδου. Η μέθοδος αυτή είναι παρόμοια με αυτή των νευρωνικών δικτύων.

Η προσαρμογή των παραμέτρων των συναρτήσεων συμμετοχής από ένα ANFIS γίνεται με αλγόριθμους εκπαίδευσης είτε μόνο με back propagation είτε με έναν υβριδικό αλγόριθμο, συνδυασμό back propagation και ελαχίστων τετραγώνων.

Ο υπολογισμός των παραμέτρων αυτών, που αλλάζουν κατά τη διάρκεια της διαδικασίας εκμάθησης, διευκολύνεται από ένα βαθμωτό διάνυσμα που μας δείχνει πόσο καλά το FIS μοντελοποιεί τα δεδομένα εισόδου / εξόδου για το εν λόγω σύνολο παραμέτρων. Μόλις αποκτήσουμε το βαθμωτό διάνυσμα μπορούμε να εφαρμόσουμε κάποιον από τους πολλούς αλγόριθμους βελτιστοποίησης, για να βελτιώσουμε αυτές τις παραμέτρους, ούτως ώστε να μειώσουμε το μέγεθος του σφάλματος (συνήθως το άθροισμα των τετραγώνων των διαφορών μεταξύ πραγματικών και επιθυμητών τιμών).

## *5.4 ΠΕΡΙΟΡΙΣΜΟΙ ΤΟΥ ANFIS*

Οι κυριότεροι περιορισμοί του ANFIS που αναφέρονται στην διεθνή βιβλιογραφία είναι:

- Βασίζεται σε ένα FIS τύπου Sugeno.
- Έχει μια έξοδο, που λαμβάνεται με τη μέθοδο της αποασαφοποίησης σταθμισμένου μέσου. Όλες οι συναρτήσεις συμμετοχής εξίσου πρέπει να είναι του ίδιου τύπου, είτε γραμμικές είτε σταθερές.
- Δεν μπορεί να γίνεται κοινή χρήση κανόνων. Διαφορετικοί κανόνες δεν μπορούν να έχουν την ίδια συνάρτηση συμμετοχής εξόδου.
- Πρέπει να υπάρχουν βάρη σε κάθε κανόνα
- Δεν μπορεί να χρησιμοποιηθεί οποιαδήποτε συνάρτηση συμμετοχής, παρά μόνο προκαθορισμένες που επιβάλλουν οι ορισμοί του ANFIS.

# ΚΕΦΑΛΑΙΟ 6: ΑΠΟΤΕΛΕΣΜΑΤΑ ΑΠΟ ΤΗ ΧΡΗΣΗ ΤΟΥ ANFIS

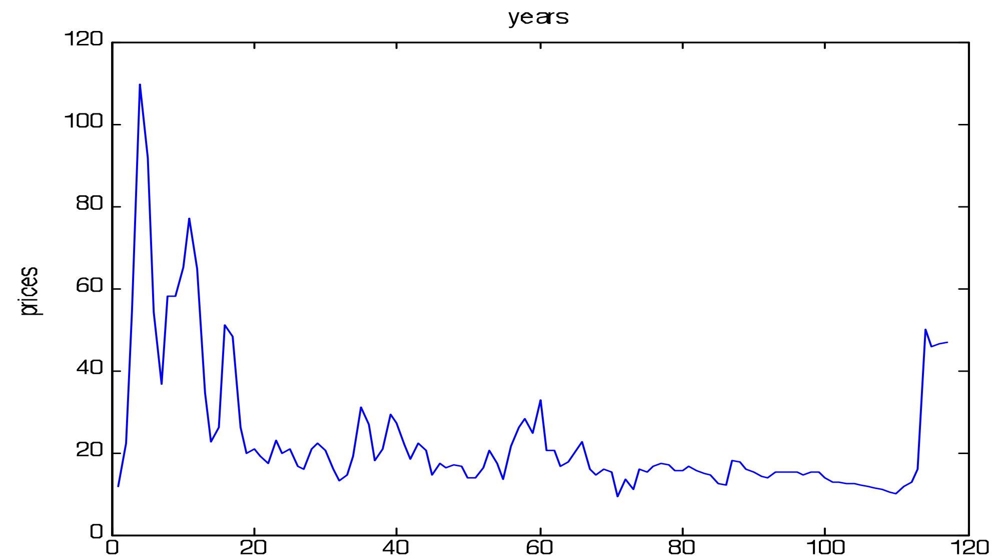
## *6.1. ΠΕΡΙΓΡΑΦΗ ΤΟΥ ΠΡΟΓΡΑΜΜΑΤΟΣ ΠΡΟΒΛΕΨΗΣ*

Στο πρακτικό κομμάτι της εργασίας κατασκευάστηκε αλγόριθμος σε Matlab, με τη χρήση του Fuzzy Logic Toolbox, το οποίο περιέχει όλα τα απαραίτητα εργαλεία για την ολοκληρωμένη ανάπτυξη και αναλυτική χρήση ενός Προσαρμοστικού Νεύρο-Ασαφούς Συστήματος Συμπεράσματος (Adaptive Neuro Fuzzy Inference System, ANFIS) για τη πρόβλεψη της τιμής του πετρελαίου ($/ βαρέλι).

Ο αλγόριθμος δέχεται δεδομένα από μια χρονοσειρά που περιέχει ετήσιες τιμές του αργού πετρελαίου για την περίοδο 1861-2008. Με βάση αυτές τις τιμές προβλέπει τις αναμενόμενες τιμές για το μέλλον. Η χρονοσειρά που χρησιμοποιούμε, περιέχει 147 παρατηρήσεις (2008-1861).

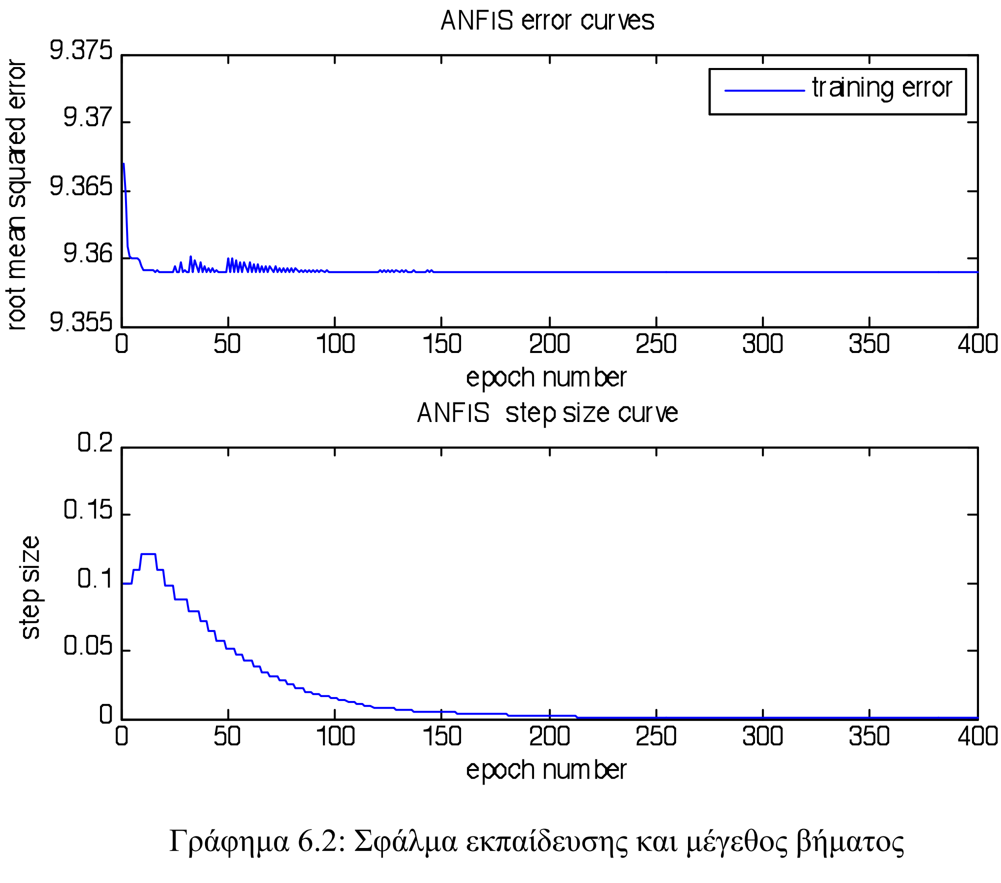
Στην περίπτωση μας χρησιμοποιήσαμε τα πρώτα 117 δεδομένα για την εκπαίδευση του συστήματος δημιουργώντας ένα πίνακα 147×2, όπου τα τελευταία 30 δεδομένα τα χρησιμοποιήσαμε για την αξιολόγηση της προβλεπτικής ικανότητας του συστήματος.

Έπειτα δημιουργήσαμε ένα ANFIS, με τη βοήθεια του fuzzy logic toolbox, το οποίο υπολογίζει τις καταλληλότερες παραμέτρους των συναρτήσεων συμμετοχής (MF parameters). Για λόγους βοηθητικούς (για να έχουμε καλύτερη οπτική παρατήρηση), παρουσιάζουμε στο γράφημα 6.1 τα δεδομένα εισόδου συναρτήσει του χρόνου και στη συνέχεια εξάγουμε στο γράφημα 6.2 που περιέχει δυο ξεχωριστές γραφικές παραστάσεις, η 1$^η$ με το όνομα ΄΄error curves΄΄ και η δεύτερη με το όνομα ΄΄step sizes΄΄. Στην 1$^η$ γραφική παράσταση απεικονίζονται τα σφάλματα εκπαίδευσης και ελέγχου (ως τη ρίζα του μέσου τετραγωνικού σφάλματος-RMSE) συναρτήσει του αριθμού των επαναλήψεων (epochs).




Γράφημα 6.1: Δεδομένα Εκπαίδευσης

Στη δεύτερη γραφική παράσταση για λόγους αναφοράς σχεδιάζουμε την καμπύλη του μεγέθους βήματος συναρτήσει των εποχών. Αυτό μας επιτρέπει να προσαρμόσουμε το αρχικό μέγεθος βήματος στο βέλτιστο, βλέποντας τους αντίστοιχους ρυθμούς αύξησης και μείωσης πάνω στο γράφημα.

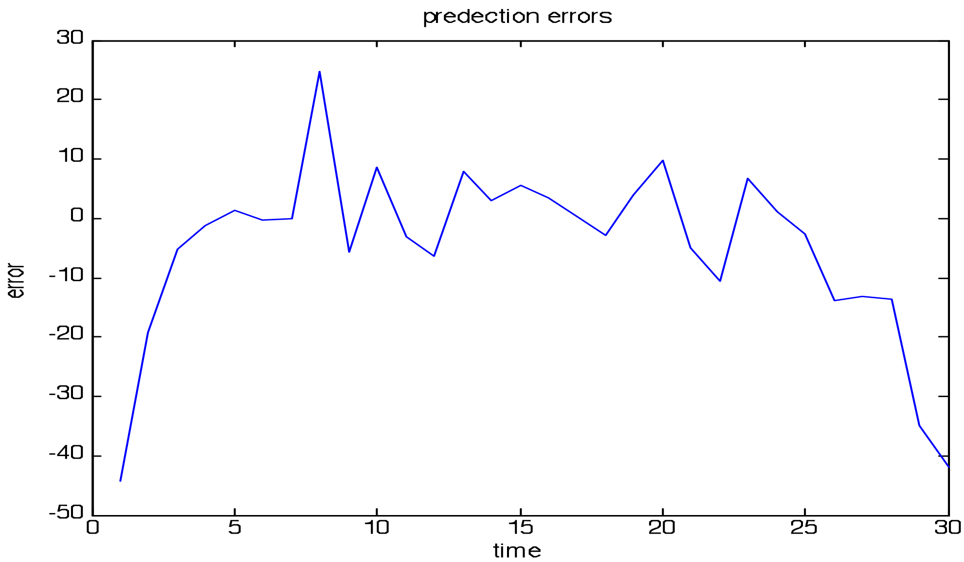



Γράφημα 6.2: Σφάλμα εκπαίδευσης και μέγεθος βήματος

Εν συνεχεία αξιολογούμε το ANFIS συγκρίνοντας τα αποτελέσματα που μας κοινή τους παρουσίαση στο γράφημα 6.3.

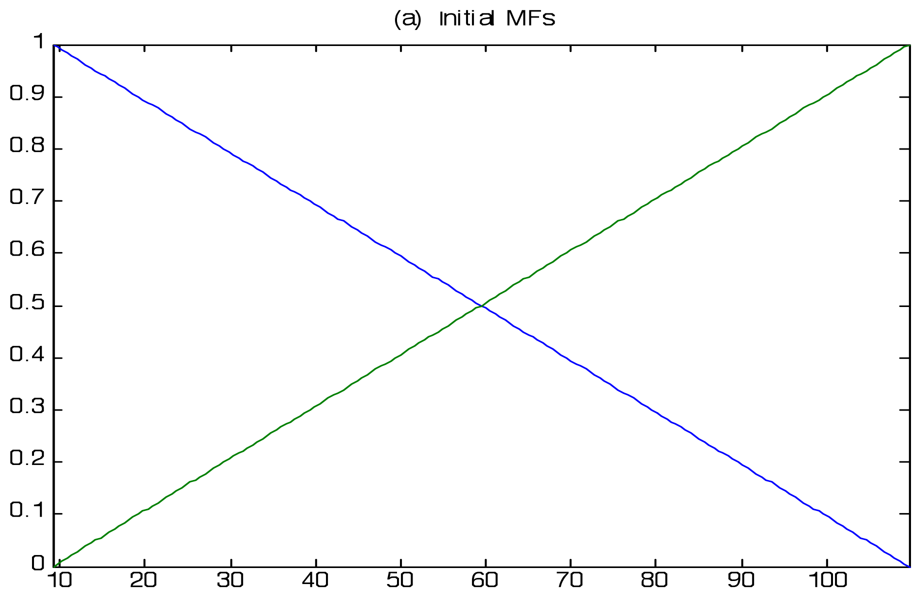
Γράφημα 6.3: Σύγκριση προβλέψεων του ANFIS έναντι των επιθυμητών για τη χρονική περίοδο 1978-2008.

Παίρνοντας τη διαφορά (δηλαδή το σφάλμα πρόβλεψης) των δύο προηγούμενων αποτελεσμάτων (πραγματικά-ANFIS) δημιουργούμε το γράφημα 6.4 το οποίο απεικονίζει τα σφάλματα συναρτήσει των παρατηρήσεων.

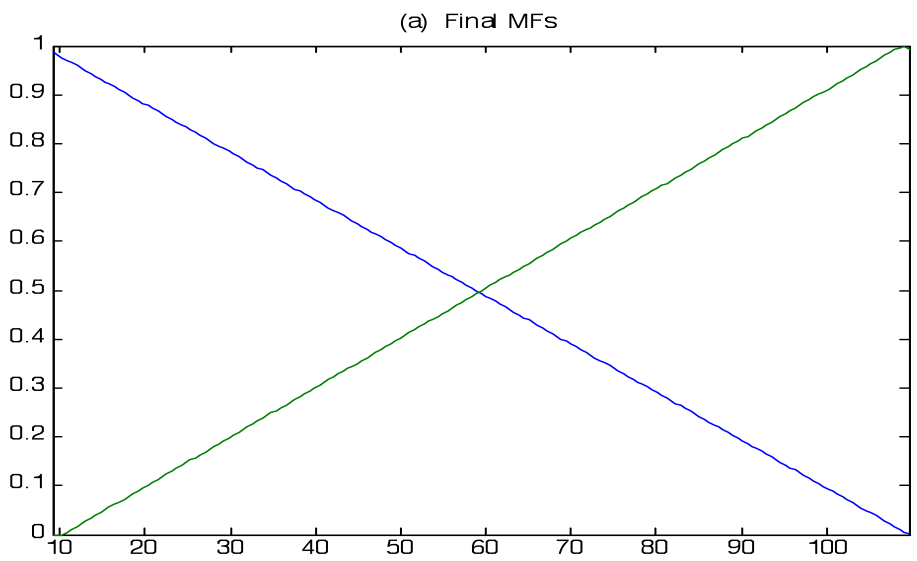



Γράφημα 6.4: Σφάλματα πρόβλεψης του ANFIS για τη περίοδο 1978-2008

Τις αρχικές συναρτήσεις συμμετοχής (initial MFs) που προέκυψαν από την προσαρμογή των παραμέτρων τους με τη βοήθεια του ANFIS τις παρουσιάζουμε στο γράφημα 6.5 και τις τελικές συναρτήσεις συμμετοχής (final MFs) που προέκυψαν μετά την εκπαίδευση του ANFIS τις παρουσιάζουμε στο γράφημα 6.6.

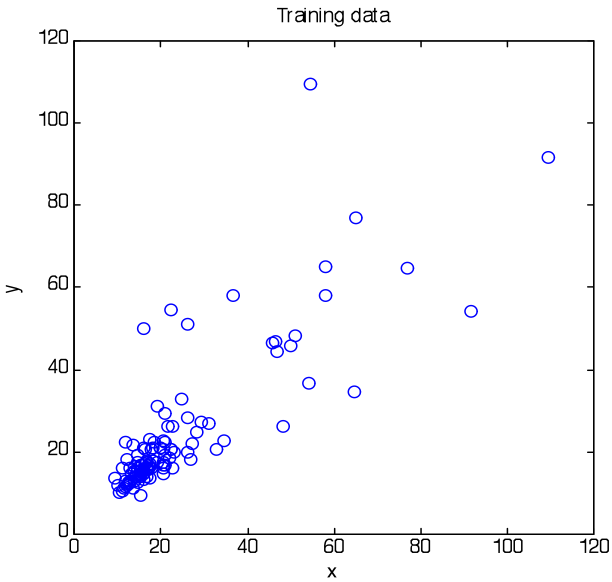



Γράφημα 6.5: Αρχικές συναρτήσεις συμμετοχής

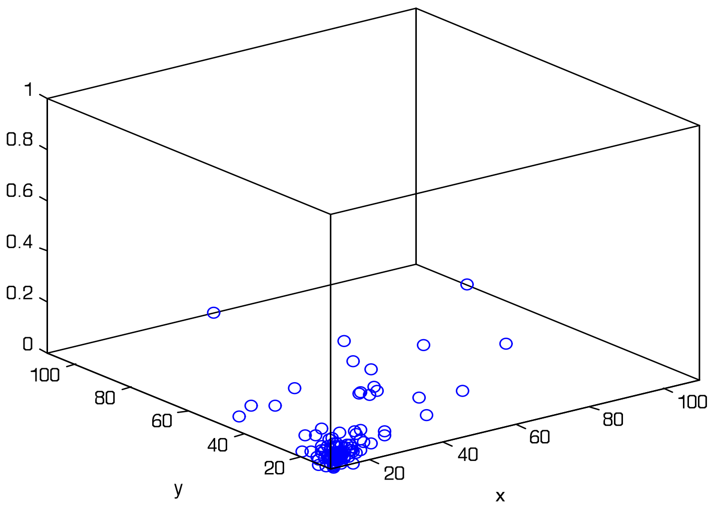



Γράφημα 6.6: Τελικές συναρτήσεις συμμετοχής

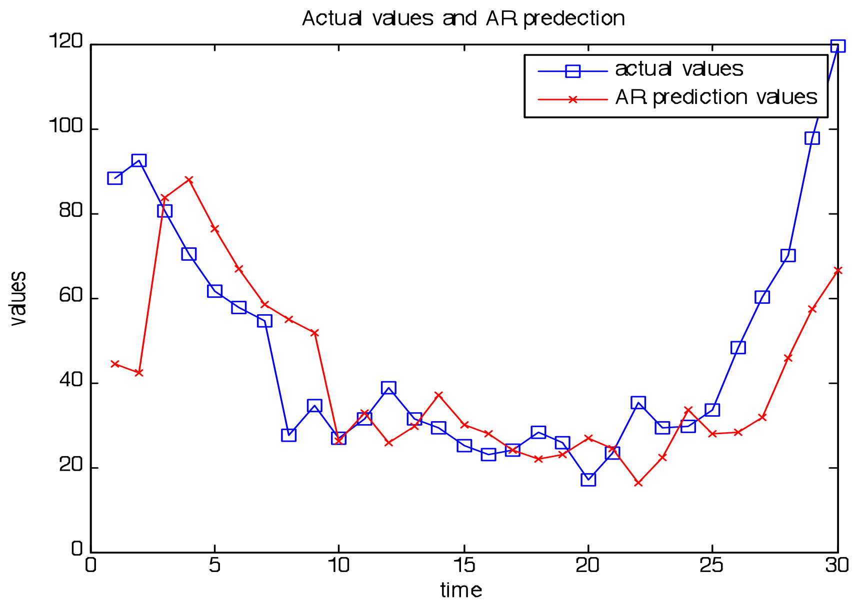
Στο γράφημα 6.7 παρουσιάζεται η διασπορά των δεδομένων εκπαίδευσης σε δυσδιάστατο επίπεδο και στο 6.8 σε τρισδιάστατο επίπεδο.



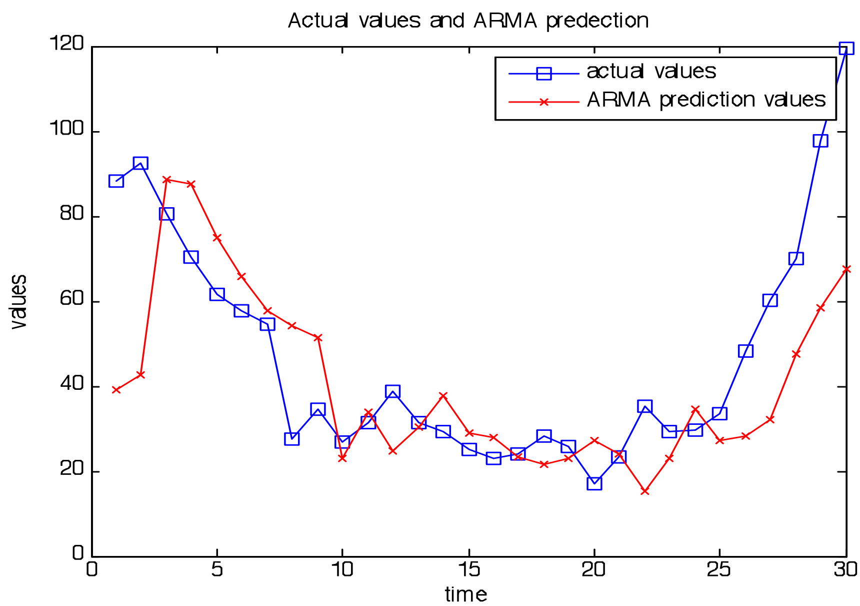

Γράφημα 6.7: Διασπορά των δεδομένων εκπαίδευσης 2D



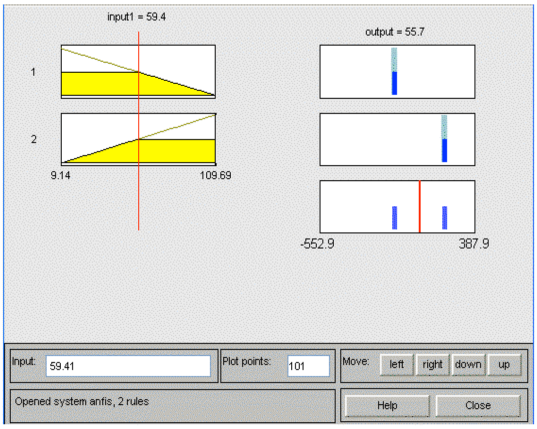

Γράφημα 6.8: Διασπορά των δεδομένων εκπαίδευσης 3D




Σχήμα 6.9: Σύγκριση προβλέψεων του μοντέλου AR έναντι των επιθυμητών τιμών για τη χρονική περίοδο 1978-2008.



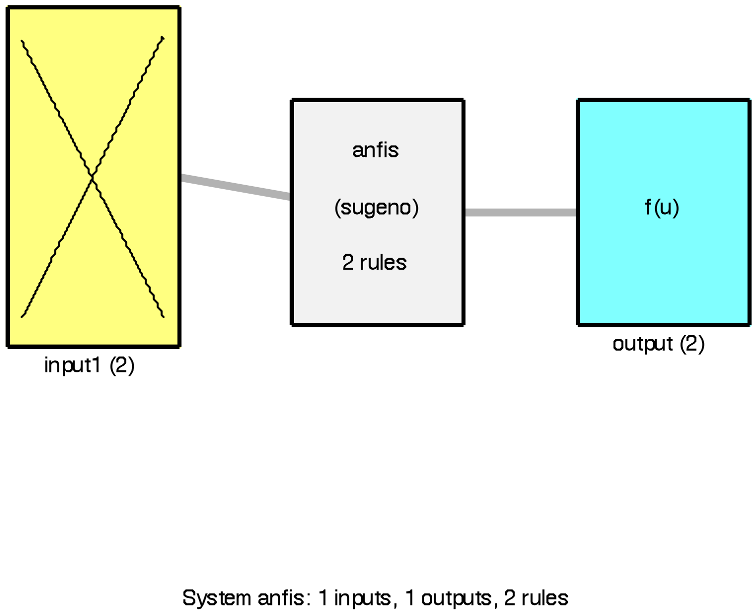

Σχήμα 6.10: Σύγκριση προβλέψεων του μοντέλου ARMA έναντι των επιθυμητών τιμών για τη χρονική περίοδο 1978-2008




Σχήμα 6.11: Οι υπολογισμοί του ANFIS

Στο σχήμα 6.11 απεικονίζεται γραφικά ο τρόπος υπολογισμού της εξόδου με τη βοήθεια των κανόνων.

Στο γράφημα 6.12 παρουσιάζεται η δομή του ANFIS που χρησιμοποιήσαμε. Διακρίνονται η είσοδος με τη τριγωνική συνάρτηση συμμετοχής, ο μηχανισμός συμπερασμού με τους 2 κανόνες και η έξοδος του συστήματος.




Σχήμα 6.12: Παρουσίαση της τελικής δομής του ANFIS.

Οι κανόνες του ANFIS που χρησιμοποιήσαμε είναι οι εξής:

1. If (input1 is in1mf1) then (output is out1mf1) (1)
2. If (input1 is in1mf2) then (output is out1mf2) (1)

Στον πίνακα 6.1 παρουσιάζονται συγκεντρωμένα τα χαρακτηριστικά του ANFIS.

Πίνακας 6.1: Χαρακτηριστικά του ANFIS

| | |
|---|---|
| Τύπος ασαφούς μηχανισμού συμπερασμού | Sugeno |
| Αριθμός εισόδων | 1 |
| Ονομασία εισόδου | Input1 |
| Ονομασία εξόδου | Output1 |
| Αριθμός συναρτήσεων συμμετοχής | 2 τριγωνικές |
| Αριθμός κανόνων | 2 |
| And method | Product |
| Or method | Max |
| Imp. method | Product |
| Agg. method | Max |
| Defuzz. method | Wtaver |
| Αριθμός κόμβων | 12 |
| Αριθμός γραμμικών παραμέτρων | 4 |

| | |
|---|---|
| Αριθμός μη γραμμικών παραμέτρων | 6 |
| Αριθμός ζευγών εκπαίδευσης (training data pairs) | 117 |

Ακολούθησε η εξαγωγή τεσσάρων τύπων σφαλμάτων ώστε να μπορέσει να γίνει η σύγκριση μεταξύ των μοντέλων.

Τα σφάλματα είναι:

- **Το μέσο τετραγωνικό σφάλμα (MSE)**

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

- **Η ρίζα του μέσου τετραγωνικού σφάλματος (RMSE)**

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

- **Το μέσο απόλυτο σφάλμα (MAE)**

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

- **Το μέσο απόλυτο ποσοστιαίο σφάλμα (MAPE)**

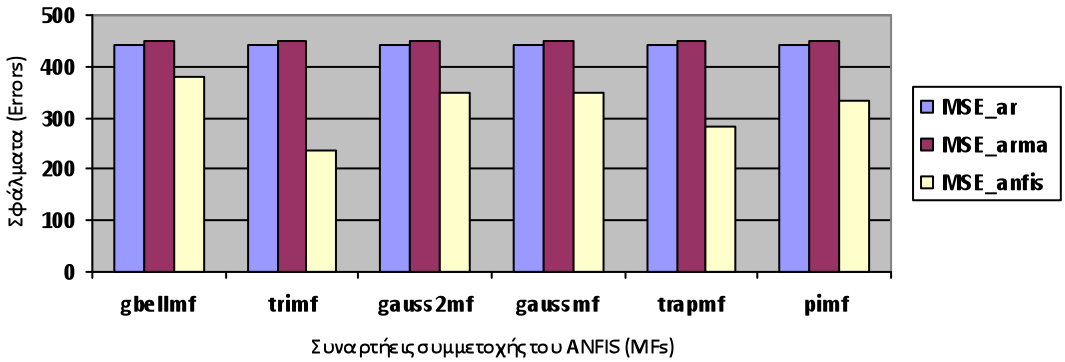
$$MAPE = \frac{100}{N}\sum_{t=1}^{N} |e_t|$$ , όπου

$$e_t = \frac{F_t - A_t}{A_t}$$ όπου $F_t$ και $A_t$ είναι πραγματικές τιμές.

Στη συνέχεια τρέξαμε το ANFIS για κάθε είδους συνάρτησης συμμετοχής ξεχωριστά για να δούμε ποιο είδος συναρτήσεως μας δίνει τα μικρότερα σφάλματα και τα συγκρίναμε με αυτά που μας δίνουν τα μοντέλα AR και ARMA, μέσω του System Identification Toolbox της Matlab. Παρακάτω παρουσιάζονται ο σχετικός πίνακας αξιολόγησης του ANFIS και το γράφημα με τις συναρτήσεις συμμετοχής και τα σφάλματα. Τα μικρότερα σφάλματα τα είχαμε με τη τριγωνική συνάρτηση συμμετοχής, ενώ το ANFIS παρουσίασε μικρότερα σφάλματα από τα άλλα δυο μοντέλα για κάθε είδους συνάρτησης συμμετοχής.

Πίνακας 6.2: Σφάλματα ANFIS,AR,ARMA για κάθε είδους συνάρτησης συμμετοχής.

| Membership Function (MF) / Errors (Arma,Ar,Anfis) | gbellmf | trimf | gauss2mf | gaussmf | trapmf | pimf |
|---|---|---|---|---|---|---|
| MSE_ar | 441,8338 | 441,8338 | 441,8338 | 441,8338 | 441,8338 | 441,8338 |
| MSE_arma | 449,9177 | 449,9177 | 449,9177 | 449,9177 | 449,9177 | 449,9177 |
| MSE_anfis | 379,4917 | **235,4245** | 347,3808 | 347,9985 | 281,4473 | 334,7059 |
| RMSE_ar | 21,0198 | 21,0198 | 21,0198 | 21,0198 | 21,0198 | 21,0198 |
| RMSE_arma | 21,2113 | 21,2113 | 21,2113 | 21,2113 | 21,2113 | 21,2113 |
| RMSE_anfis | 19,4805 | **15,3435** | 18,6382 | 18,6547 | 16,7764 | 18,2950 |
| MAE_ar | 14,7827 | 14,7827 | 14,7827 | 14,7827 | 14,7827 | 14,7827 |
| MAE_arma | 15,0728 | 15,0728 | 15,0728 | 15,0728 | 15,0728 | 15,0728 |
| MAE_anfis | 12,5231 | **10,0170** | 12,2237 | 12,2206 | 11,3795 | 12,0876 |
| MAPE_ar | 28,4463 | 28,4463 | 28,4463 | 28,4463 | 28,4463 | 28,4463 |
| MAPE_arma | 29,1857 | 29,1857 | 29,1857 | 29,1857 | 29,1857 | 29,1857 |
| MAPE_anfis | 23,7699 | **20,9808** | 23,6008 | 23,5930 | 23,0723 | 23,6414 |




Σχήμα 6.13: Σφάλματα ANFIS,AR,ARMA για κάθε είδους συνάρτησης συμμετοχής

Πίνακας 6.3: Σφάλματα του ANFIS σε σύγκριση με τα μοντέλα AR και ARMA για διαφορετικό αριθμό γκαουσιανών συναρτήσεων συμμετοχής.

| Number of Membership Functions (gaussmf) | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|---|
| MSE_ar | 441,8338 | 441,8338 | 441,8338 | 441,8338 | 441,8338 | 441,8338 | 441,8338 | 441,8338 | 441,8338 | 441,8338 |
| MSE_arma | 449,9177 | 449,9177 | 449,9177 | 449,9177 | 449,9177 | 449,9177 | 449,9177 | 449,9177 | 449,9177 | 449,9177 |
| MSE_anfis | 347,9985 | 382,4038 | 510,0619 | 544,568 | 6689,797 | 4423,744 | 2954,704 | 19964841,179 | 183634289,37 | 3248879,97 |
| RMSE_ar | 21,0198 | 21,0198 | 21,0198 | 21,0198 | 21,0198 | 21,0198 | 21,0198 | 21,0198 | 21,0198 | 21,0198 |
| RMSE_arma | 21,2113 | 21,2113 | 21,2113 | 21,2113 | 21,2113 | 21,2113 | 21,2113 | 21,2113 | 21,2113 | 21,2113 |
| RMSE_anfis | 18,6547 | 19,5551 | 22,5846 | 23,3360 | 81,791 | 66,5112 | 171,892 | 4468,203 | 13551,172 | 1802,465 |
| MAE_ar | 14,7827 | 14,7827 | 14,7827 | 14,7827 | 14,7827 | 14,7827 | 14,7827 | 14,7827 | 14,7827 | 14,7827 |
| MAE_arma | 15,0728 | 15,0728 | 15,0728 | 15,0728 | 15,0728 | 15,0728 | 15,0728 | 15,0728 | 15,0728 | 15,0728 |
| MAE_anfis | 12,2206 | 13,1083 | 15,0814 | 14,3805 | 42,6312 | 34,9089 | 707,4118 | 1859,818 | 4545,668 | 503,367 |
| MAPE_ar | 28,4463 | 28,4463 | 28,4463 | 28,4463 | 28,4463 | 28,4463 | 28,4463 | 28,4463 | 28,4463 | 28,4463 |
| MAPE_arma | 29,1857 | 29,1857 | 29,1857 | 29,1857 | 29,1857 | 29,1857 | 29,1857 | 29,1857 | 29,1857 | 29,1857 |
| MAPE_anfis | 23,5930 | 25,5379 | 32,2343 | 31,0837 | 93,4242 | 62,5689 | 111,2276 | 2291,938 | 4816,002 | 528,574 |

Στον παραπάνω πίνακα παρουσιάζονται τα σφάλματα που μας δίνει το ANFIS για κάθε αριθμό από 2-11 γκαουσιανών συναρτήσεων συμμετοχής. Παρατηρούμε ότι όσο αυξάνουμε τον αριθμό των συναρτήσεων συμμετοχής τόσο αυξάνονται τα σφάλματα του ANFIS.

## *ΕΠΙΛΟΓΟΣ*

Ένα ANFIS για να μπορέσει να μελετηθεί απαιτεί από το χρήστη να έχει ένα ανεπτυγμένο γνωστικό υπόβαθρο, το οποίο θα περιλαμβάνει τις στατιστικές μεθόδους, την αριθμητική ανάλυση, τα αλγεβρικά συστήματα και τους πίνακες. Επίσης να γνωρίζει τι είναι τα νευρωνικά δίκτυα και πως μπορούν να εκπαιδευτούν και να γίνουν προσαρμοστικά καθώς και τη θεωρία της ασαφούς λογικής.

Έτσι μπορούμε να πούμε ότι αυτή η πολυπλοκότητα του δίνει τη δύναμη να μπορεί να προσαρμόζεται εύκολα στις διάφορες απαιτήσεις των όποιων προβλημάτων. Δίνει έτσι μικρότερα σφάλματα, αρά και καλύτερα αποτελέσματα. Η εκπαίδευση ενός τέτοιου συστήματος μπορεί να γίνει εύκολα, καθώς μπορούμε να αντιληφθούμε άμεσα τυχόν προβλήματα μέσα από τους κανόνες και τα γραφήματα που εξάγουμε.

Αντίθετα, ενώ η πολυπλοκότητα του σημαίνει και καλύτερες δυνατότητες προσαρμογής, αυτόματα σημαίνει και μεγαλύτερες ανάγκες για επεξεργαστική ισχύ. Δηλ. χρειάζεται περισσότερο χρόνο για να ολοκληρώσει την εξαγωγή αποτελεσμάτων σε σχέση με ένα απλούστερο παραδοσιακό μοντέλο καθώς επίσης απαιτεί μεγαλύτερες επενδύσεις σε τεχνολογικό εξοπλισμό.

Ένα ακόμα μειονέκτημα αποτελεί και το γεγονός ότι χρειαζόμαστε πολλά δεδομένα για την εκπαίδευση του συστήματος. Σε αντίθετη περίπτωση τα αποτελέσματα που παίρνουμε οδηγούν στη μη γενίκευση του μοντέλου.

Συμπερασματικά μπορούμε να πούμε ότι το ANFIS παρουσίασε καλύτερα αποτελέσματα από τα μοντέλα AR και ARMA. Επίσης το AR παρουσίασε καλύτερα αποτελέσματα από το ARMA.

## *ΒΙΒΛΙΟΓΡΑΦΙΑ*


1. Sam Mirmirani and Hsi Cheng Li (2004). A comparison of var and neural networks with genetic algorithm in forecasting price of oil.
2. Ying Fan, Qiang Liang, Yi-Ming Wei (2006). A generalized pattern matching approach for multistep prediction of crude oil price.
3. Mohammad Reza Amin-Naseri, Ehsan Ahmadi Gharacheh. A hybrid artificial intelligence approach to monthly forecasting of crude oil price time series.
4. Michael Ye, John Zyren, Joanne Shore (2005). A monthly crude oil spot price forecasting model using relative inventories.
5. Xun Zhang, K.K Lai, Shoy-Yang Wang (2008). A new approach for crude oil price analysis based on Empirical Mode Decomposition.
6. Jinlan Liu, Yin Bai, Bin Li. A New Approach to Forecast Crude Oil Price Based on Fuzzy Neural Network.
7. Wen Xie, Lean Yu, Shanying Xu, and Shouyang Wang (2006). A new method for crude oil price forecasting based on support vector machines.
8. Shouyang Wang, Lean Yu, K.K. Lai (2004). A novel hybrid AI system framework for crude oil price forecasting.
9. Claudio Morana (2001). A semiparametric approach to short-termoil price forecasting.
10. Akihiro Amano (1987). A Small Forecasting Model of the World Oil Market.

11. Aistis Raudys (2005). Accuracy of MLP Based Data VisualizationUsed in Oil Prices Forecasting Task.
12. Mehrzad Zamani (2004). An Econometrics Forecasting Model of Short Term Oil Spot Price.
13. Thalia Chantziara , George Skiadopoulos (2007). Can the dynamics of the term structure of petroleum_futures be forecasted? Evidence from major markets.
14. M. A. Kaboudan (2001). Compumetric Forecasting of Crude Oil Prices.
15. B. Zimberg. Crude oil price forecasting with anfis.
16. Wang Shouyang, YU Lean, K. K. Lai (2005). Crude oil price forecasting with TEI@I methodology.
17. Yejing Bao , Xun Zhang , Lean Yu , and Shouyang Wang (2007). Crude Oil Price Prediction Based On Multi-scale_Decomposition.
18. Chiara Longo, Matteo Manera, Anil Markandya and Elisa Scarpa (2007). Evaluating the Empirical Performance of Alternative Econometric Models for Oil Price Forecasting.
19. Komsan Suriya (2006). Forecasting Crude Oil Price Using Neural Networks.
20. Michael Ye, John Zyren, Joanne Shore (2002). Forecasting Crude Oil Spot Price Using OECD Petroleum Inventory Levels.
21. Michael Ye, John Zyren, Joanne Shore (2005). Forecasting short-run crude oil price using high- and low-inventory variables.
22. Thomas A. Knetsch (2006). Forecasting the price of crude oil via convenience yield predictions.
23. Jesus Crespo Cuaresma, Adusei Jumah and Sohbet Karbuz (2002). Modelling and Forecasting Oil Prices: The Role of Asymmetric Cycles.
24. Lean Yu, Kin Keung Lai, Shouyang Wang, Kaijian He (2007). Oil Price Forecasting with an EMD-Based Multiscale Neural Network Learning Paradigm.
25. Salah Abosedraa, Hamid Baghestani (2003). On the predictive accuracy of crude oil futures prices.
26. Bruce Abramson a'*, Anthony Finizza (1995). Probabilistic forecasts from probabilistic models" a case study in the oil market.
27. E.S. Dokuchaev, A.M. Rogacheva, E.V. Evtushenko (2006). The forecasting of the world oil price by summing up linear trend and periodic functions.
28. Bruce Abramson, Anthony Finizza (1991). Using belief networks to forecast oil prices.
29. Shahriar Yousefi , Ilona Weinreich , Dominik Reinarz (2005). Wavelet-based prediction of oil prices.
30. Χρυσή Μινουδάκη (2007). Πρόβλεψη κατανάλωσης νερού στο Ν.Χανίων χρησιμοποιώντας το νεύρο-ασαφές σύστημα (ANFIS).
31. Εμμανουήλ Μ. Βουμβουλάκης (2003). Βραχυπρόθεσμη πρόβλεψη φορτίου με χρήση νευρωνικών δικτύων και ασαφούς λογικής.

## Internet Sites

1. http://www.ioga.com/Special/crudeoil_Hist.htm

2. http://www.inflationdata.com/inflation/Inflation_Rate/Historical_Oil_Prices_Table.asp
3. http://paul.kedrosky.com/archives/2008/06/18/crude_oil_price.html
4. http://www.mathworks.com/access/helpdesk/help/toolbox/fuzzy/fp715dup12.html