CHAPTER 1

Introduction to Neural Networks

What You Will Learn

- Scope of neural networks and MATLAB.
- How neural network is used to learn patterns and relationships in data.
- The aim of neural networks.
- About fuzzy logic .
- Use of MATLAB to develop the applications based on neural networks.

1.1 Neural Processing

Artificial neural networks are the result of academic investigations that use mathematical formulations to model nervous system operations. The resulting techniques are being successfully applied in a variety of everyday business applications.

Neural networks (NNs) represent a meaningfully different approach to using computers in the workplace. A neural network is used to learn patterns and relationships in data. The data may be the results of a market research effort, a production process given varying operational conditions, or the decisions of a loan officer given a set of loan applications. Regardless of the specifics involved, applying a neural network is substantially different from traditional approaches.

Traditionally a programmer or an analyst specifically ‘codes’ for every facet of the problem for the computer to ‘understand’ the situation. Neural networks do not require explicit coding of the problems. For example, to generate a model that performs a sales forecast, a neural network needs to be given only raw data related to the problem. The raw data might consist of history of past sales, prices, competitors’ prices and other economic variables. The neural network sorts through this information and produces an understanding of the factors impacting sales. The model can then be called upon to provide a prediction of future sales given a forecast of the key factors.

These advancements are due to the creation of neural network learning rules, which are the algorithms used to ‘learn’ the relationships in the data. The learning rules enable the network to ‘gain knowledge’ from available data and apply that knowledge to assist a manager in making key decisions.

What are the Capabilities of Neural Networks?

In principle, NNs can compute any computable function, i.e. they can do everything a normal digital computer can do. Especially anything that can be represented as a mapping between vector spaces can be approximated to arbitrary precision by feedforward NNs (which is the most often used type).

In practice, NNs are especially useful for mapping problems, which are tolerant of some errors, have lots of example data available, but to which hard and fast rules cannot easily be applied. However, NNs are, as of now, difficult to apply successfully to problems that concern manipulation of symbols and memory.

Who is Concerned with Neural Networks?

Neural Networks are of interest to quite a lot of people from different fields:

- Computer scientists want to find out about the properties of non-symbolic information processing with neural networks and about learning systems in general.
- Engineers of many kinds want to exploit the capabilities of neural networks in many areas (e.g. signal processing) to solve their application problems.
- Cognitive scientists view neural networks as a possible apparatus to describe models of thinking and conscience (High-level brain function).
- Neuro-physiologists use neural networks to describe and explore medium-level brain function (e.g. memory, sensory system).
- Physicists use neural networks to model phenomena in statistical mechanics and for a lot of other tasks.
- Biologists use Neural Networks to interpret nucleotide sequences.
- Philosophers and some other people may also be interested in Neural Networks to gain knowledge about the human systems namely behavior, conduct, character, intelligence, brilliance and other psychological feelings. Environmental nature and related functioning, marketing business as well as designing of any such systems can be implemented via Neural networks.

1.2 Neural Networks — An Overview

The development of Artificial Neural Network started 50 years ago. Artificial neural networks (ANNs) are gross simplifications of real (biological) networks of neurons. The paradigm of neural networks, which began during the 1940s, promises to be a very important tool for studying the structure-function relationship of the human brain. Due to the complexity and incomplete understanding of biological neurons, various architectures of artificial neural networks have been reported in the literature. Most of the ANN structures used commonly for many applications often consider the behavior of a single neuron as the basic computing unit for describing neural information processing operations. Each computing unit, i.e. the artificial neuron in the neural network is based on the concept of an ideal neuron. An ideal neuron is assumed to respond optimally to the applied inputs. However, experimental studies in neuro-physiology show that the response of a biological neuron appears random and only by averaging many observations it is possible to obtain predictable results. Inspired by this observation, some researchers have developed neural structures based on the concept of neural populations.

In common with biological neural networks, ANNs can accommodate many inputs in parallel and encode the information in a distributed fashion. Typically the information that is stored in a neural net is shared by many of its processing units. This type of coding is in sharp contrast to traditional memory schemes, where a particular piece of information is stored in only one memory location. The recall process is time consuming and generalization is usually absent. The distributed storage scheme provides many advantages, most important of them being the redundancy in information representation. Thus, an ANN can undergo partial destruction of its structure and still be able to function well. Although redundancy can also be built into other types of systems, ANN has a natural way of implementing this. The result is a natural fault-tolerant system which is very similar to biological systems.

The aim of neural networks is to mimic the human ability to adapt to changing circumstances and the current environment. This depends heavily on being able to learn from events that have happened in the past and to be able to apply this to future situations. For example the decisions made by doctors are rarely based on a single symptom because of the complexity of the human body; since one symptom could signify any number of problems. An experienced doctor is far more likely to make a sound decision than a trainee, because from his past experience he knows what to look out for and what to ask, and may have etched on his mind a past mistake, which he will not repeat. Thus the senior doctor is in a superior position than the trainee. Similarly it would be beneficial if machines, too, could use past events as part of the criteria on which their decisions are based, and this is the role that artificial neural networks seek to fill.

Artificial neural networks consist of many nodes, i.e. processing units analogous to neurons in the brain. Each node has a node function, associated with it which along with a set of local parameters determines the output of the node, given an input. Modifying the local parameters may alter the node function. Artificial Neural Networks thus is an information-processing system. In this information-processing system, the elements called **neurons**, process the information. The signals are transmitted by means of connection links. The links possess an associated weight, which is multiplied along with the incoming signal (net input) for any typical neural net. The output signal is obtained by applying activations to the net input.

The neural net can generally be a single layer or a multilayer net. The structure of the simple artificial neural net is shown in Fig. 1.1.

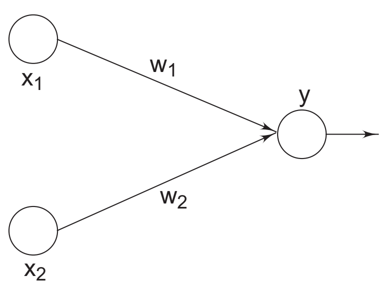

Fig. 1.1 *A Simple Artificial Neural Net*

Figure 1.1 shows a simple artificial neural net with two input neurons (x_1, x_2) and one output neuron (y). The interconnected weights are given by w_1 and w_2. In a single layer net there is a single layer of weighted interconnections.

A typical multi-layer artificial neural network, abbreviated as MNN, comprises an input layer, output layer and hidden (intermediate) layer of neurons. MNNs are often called layered networks. They can implement arbitrary complex input/output mappings or decision surfaces separating different patterns. A three-layer MNN is shown in Fig. 1.2, and a simplified block diagram representation in Fig. 1.3. In a MNN, a layer of input units is connected to a layer of hidden units, which is connected to the layer of output units. The activity of neurons in the input layer represents the raw information that is fed into the network. The activity of neurons in the hidden layer is determined by the activities of the input neurons and the connecting weights between the input and hidden units. Similarly, the behavior of the output units depends on the activity of the neurons in the hidden layer and the connecting weights between the hidden and the output layers. This simple neural structure is interesting because neurons in the hidden layers are free to construct their own representation of the input.

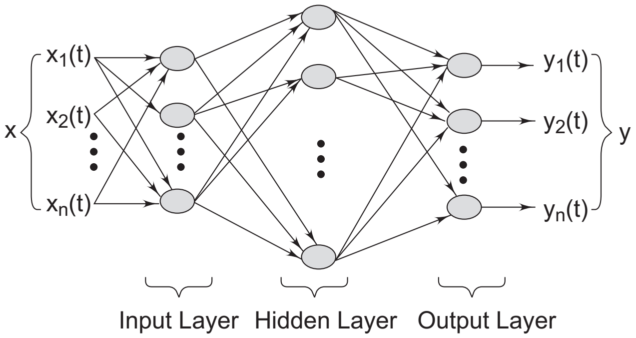

Fig. 1.2 *A Densely Interconnected Three-layered Static Neural Network. Each Shaded Circle, or Node, Represents an Artificial Neuron*

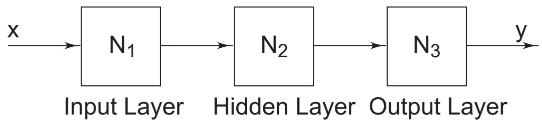

Fig. 1.3 *A Block Diagram Representation of a Three-layered MNN*

MNNs provide an increase in computational power over a single-layer neural network unless there is a nonlinear activation function between layers. Many capabilities of neural networks, such as nonlinear functional approximation, learning, generalization, etc are in fact performed due to the nonlinear activation function of each neuron.

ANNs have become a technical folk legend. The market is flooded with new, increasingly technical software and hardware products, and many more are sure to come. Among the most popular hardware implementations are Hopfield, Multilayer Perceptron, Self-organizing Feature Map, Learning Vector Quantization, Radial Basis Function, Cellular Neural, and Adaptive Resonance Theory (ART) networks, Counter Propagation networks, Back Propagation networks, Neo-cognitron, etc. As a result of the existence of all these networks, the application of the neural network is increasing tremendously.

Thus artificial neural network represents the major extension to computation. They perform the operations similar to that of the human brain. Hence it is reasonable to expect a rapid increase in our understanding of artificial neural networks leading to improved network paradigms and a host of application opportunities.

1.3 The Rise of Neurocomputing

A majority of information processing today is carried out by digital computers. This has led to the widely held misperception that information processing is dependent on digital computers. However, if we look at cybernetics and the other disciplines that form the basis of information science, we see that information processing originates with living creatures in their struggle to survive in their environments, and that the information being processed by computers today accounts for only a small part — the automated portion — of this. Viewed in this light, we can begin to consider the possibility of information processing devices that differ from conventional computers. In fact, research aimed at realizing a variety of different types of information processing devices is already being carried out, albeit in the shadows of the major successes achieved in the realm of digital computers. One direction that this research is taking is toward the development of an information processing device that mimics the structures and operating principles found in the information processing systems possessed by humans and other living creatures.

Digital computers developed rapidly in and after the late 1940's and after originally being applied to the field of mathematical computations, have found expanded applications in a variety of areas, like text (word), symbol, image and voice processing, i.e. pattern information processing, robotic control and artificial intelligence. However, the fundamental structure of digital computers is based on the principle of sequential (serial) processing, which has little if anything in common with the human nervous system.

The human nervous system, it is now known, consists of an extremely large number of nerve cells, or neurons, which operate in parallel to process various types of information. By taking a hint from the structure of the human nervous system, we should be able to build a new type of advanced parallel information processing device.

In addition to the increasingly large volumes of data that we must process as a result of recent developments in sensor technology and the progress of information technology, there is also a growing requirement to simultaneously gather and process huge amounts of data from multiple sensors and other sources. This situation is creating a need in various fields to switch from conventional computers that process information sequentially, to parallel computers equipped with multiple processing elements aligned to operate in parallel to process information.

Besides the social requirements just cited, a number of other factors have been at work during the 1980's to prompt research on new forms of information processing devices. For instance, recent neurophysiological experiments have shed considerable light on the structure of the brain, and even in fields such as cognitive science, which study human information processing processes at the macro level, we are beginning to see proposals for models that call for multiple processing elements aligned to operate in parallel. Research in the fields of mathematical science and physics is also concentrating more on the mathematical analysis of systems comprising multiple elements that interact in complex ways. These factors gave birth to a major research trend aimed at clarifying the structures and operating principles inherent in the information processing systems of human beings and other animals, and constructing an information processing device based on these structures and operating principles. The term **'neurocomputing'** is used to refer to the information engineering aspects of this research.

1.4 MATLAB – An Overview

Dr. Cleve Moler, chief scientist at MathWorks, Inc., originally wrote MATLAB to provide easy access to matrix software developed in the LINPACK and EISPACK projects. The first version was written in the late 1970s for use in courses in matrix theory, linear algebra, and numerical analysis. MATLAB is therefore built upon a foundation of sophisticated matrix software, in which the basic data element is a matrix that does not require predimensioning.

MATLAB is a product of The MathWorks, Inc. and is an advanced interactive software package specially designed for scientific and engineering computation. The MATLAB environment integrates graphical illustrations with precise numerical calculations, and is a powerful, easy-to-use, and comprehensive tool for performing all kinds of computations and scientific data visualization. MATLAB has proven to be a very flexible and useful tool for solving problems in many areas. MATLAB is a high-performance language for technical computing. It integrates computation, visualization and programming in an easy-to-use environment where problems and solutions are expressed in familiar mathematical notation. Typical areas of application of MATLAB include:

- Math and computation
- Algorithm development
- Modeling, simulation and prototyping
- Data analysis, exploration, and visualization
- Scientific and engineering graphics
- Application development, including graphical user interface building

MATLAB is an interactive system whose basic element is an array that does not require dimensioning. This helps in solving many computing problems, especially those with matrix and vector formulations, in a fraction of the time it would take to write a program in a scalar non-interactive language such as C or FORTRAN. Mathematics is the common language of science and engineering. Matrices, differential equations, arrays of data, plots and graphs are the basic building blocks of both applied mathematics and MATLAB. It is the underlying mathematical base that makes MATLAB accessible and powerful. MATLAB allows expressing the entire algorithm in a few dozen lines, to compute the solution with great accuracy in about a second. Therefore it is especially helpful for technical analysis, algorithm prototyping and application development.

MATLAB's two-and three-dimensional graphics are object oriented. MATLAB is thus both an environment and a matrix/vector-oriented programming language, which enables the user to build own reusable tools. The user can create his own customized functions and programs (known as M-files) in MATLAB code. The Toolbox is a specialized collection of M-files for working on particular classes of problems. MATLAB Documentation Set has been written, expanded and put online for ease of use. The set includes online help, as well as hypertext-based and printed manuals. The commands in MATLAB are expressed in a notation close to that used in mathematics and engineering. There is a very large set of these commands and functions, known as MATLAB M-files. As a result, solving problems through MATLAB is faster than the other traditional programming. It is easy to modify the functions since most of the M-files can be opened and modified. For ensuring high performance, the MATLAB software has been written in optimized C and coded in assembly language.

The main features of MATLAB can be summarized as:

- Advance algorithms for high-performance numerical computations, especially in the field of matrix algebra.
- A large collection of predefined mathematical functions and the ability to define one's own functions.
- Two- and three-dimensional graphics for plotting and displaying data.
- A complete online help system.
- Powerful, matrix/vector-oriented, high-level programming language for individual applications.

- Ability to cooperate with programs written in other languages and for importing and exporting formatted data.
- Toolboxes available for solving advanced problems in several application areas.

Figure 1.4 shows the main features and capabilities of MATLAB.

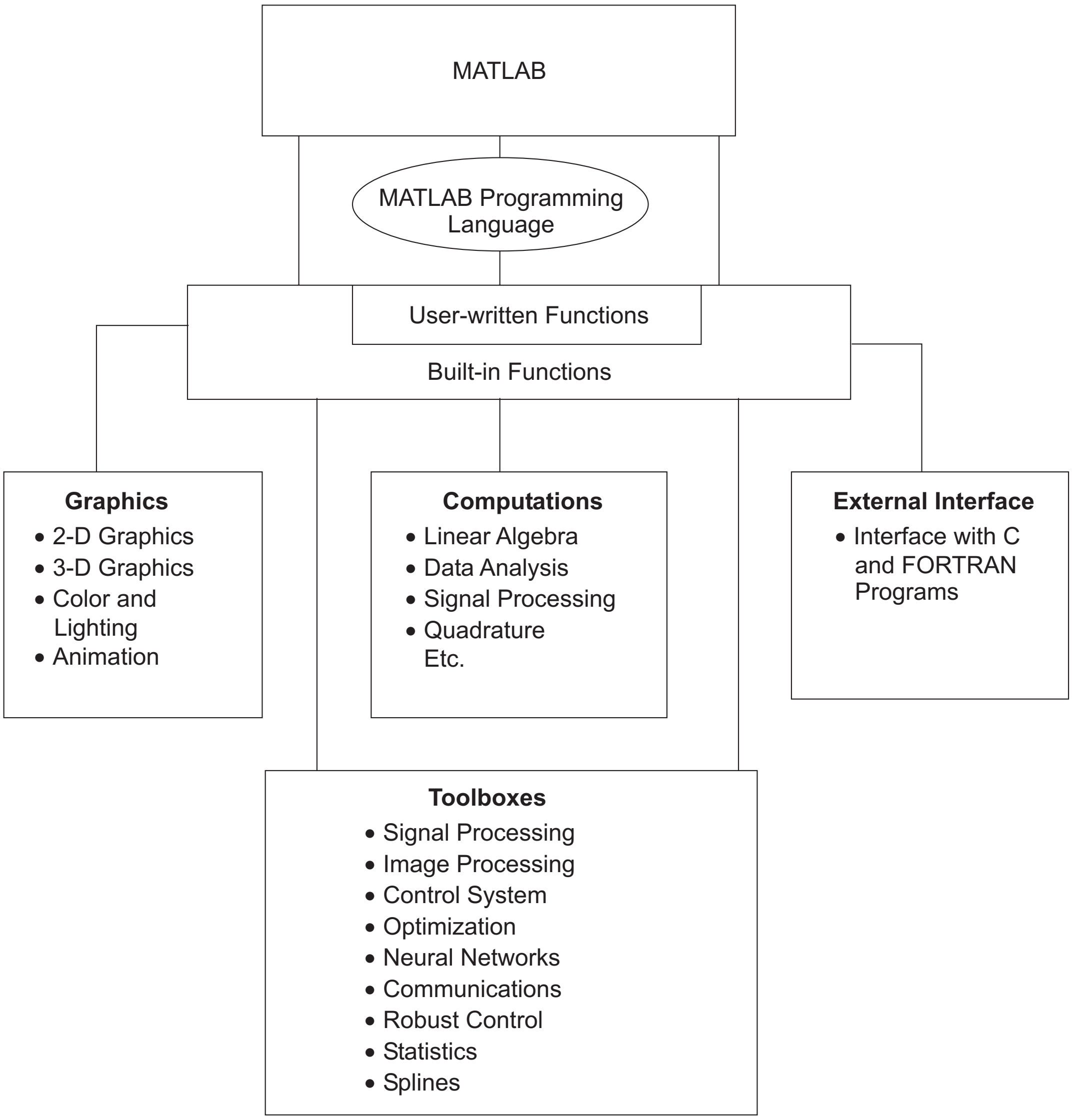

Fig. 1.4 *Features and Capabilities of MATLAB*

An optional extension of the core of MATLAB called SIMULINK is also available. SIMULINK means SIMUlating and LINKing the environment. SIMULINK is an environment for simulating linear and non-linear dynamic systems, by constructing block diagram models with an easy to use graphical user interface.

SIMULINK is a MATLAB toolbox designed for the dynamic simulation of linear and non-linear systems as well as continuous and discrete-time systems. It can also display information graphically. MATLAB is an interactive package for numerical analysis, matrix computation, control system design, and linear system analysis and design available on most CAEN (Computer Aided Engineering Network) platforms (Macintosh, PCs, Sun, and Hewlett-Packard). In addition to the standard functions provided by MATLAB, there exist a large set of Toolboxes, or collections of functions and procedures, available as part of the MATLAB package. Toolboxes are libraries of MATLAB functions used to customize MATLAB for solving particular class of problems. Toolboxes are a result of some of the world's top researches in specialized fields. They are equivalent to prepackaged 'off-the-shelf' software solution for a particular class of problem or technique. It is a collection of special files called M-files that extend the functionality of the base program. The various Toolboxes available are:

- *Control System:* Provides several features for advanced control system design and analysis.
- *Communications:* Provides functions to model the components of a communication system's physical layer.
- *Signal Processing:* Contains functions to design analog and digital filters and apply these filters to data and analyze the results.
- *System Identification:* Provides features to build mathematical models of dynamical systems based on observed system data.
- *Robust Control:* Allows users to create robust multivariable feedback control system designs based on the concept of the singular-value Bode plot.
- *Simulink:* Allows you to model dynamic systems graphically.
- *Neural Network:* Allows you to simulate neural networks.
- *Fuzzy Logic:* Allows for manipulation of fuzzy systems and membership functions.
- *Image Processing:* Provides access to a wide variety of functions for reading, writing, and filtering images of various kinds in different ways.
- *Analysis:* Includes a wide variety of system analysis tools for varying matrices.
- *Optimization:* Contains basic tools for use in constrained and unconstrained minimization problems.
- *Spline:* Can be used to find approximate functional representations of data sets.
- *Symbolic:* Allows for symbolic (rather than purely numeric) manipulation of functions.
- *User Interface Utilities:* Includes tools for creating dialog boxes, menu utilities, and other user interactions for script files.

MATLAB has been used as an efficient tool, all over this text to develop the applications based on Neural Nets.

Review Questions

1.1 How did neurocomputing originate?

1.2 What is a multilayer net? Describe with a neat sketch.

1.3 State some of the popular neural networks.
1.4 Briefly discuss the key characteristics of MATLAB.
1.5 List the basic arithmetic operations that can be performed in MATLAB.
1.6 What is the necessity of SIMULINK package available in MATLAB?
1.7 Discuss in brief about the GUI toolbox feature of MATLAB.
1.8 What is meant by toolbox and list some of the toolboxes available for MATLAB?