Preface

What do you picture when you hear the term neural networks with high-dimensional parameters? Naturally, all neural networks can receive and send two or more signals, process high-dimensional data, and have a large number of parameters. Why are they specified to have high-dimensional parameters? This book describes neural networks with parameters (weights and threshold values) that are high-dimensional, such as complex numbers, quaternions, and $N$-dimensional vectors. The neural network with high-dimensional parameters is high-dimensional in this sense. But what is the significance of creating neural networks with high-dimensional parameters? The answer can be found in (Hirose, 2003, 2006; Nitta, 2008), and in every chapter of this book together with descriptions of possible developments. I had the idea of creating neural networks with high-dimensional parameters in April 1990, when I started to work at the Electrotechnical Laboratory (now National Institute of Advanced Industrial Science and Technology (AIST)) and was looking for a theme to study. During a seminar for newcomers, I learned about a study on a complex autoregressive model by Dr. Otsu (presently an AIST Fellow) (Sekita, Kurita, & Otsu, 1992), mentioning that complex regression coefficients form rotation invariant for two-dimensional figures, and was inspired by his comment that it went well when it was extended to complex numbers. The complex autoregressive models were later extended to quaternionic models, and so forth. (Tanaka, 1996). The section to which I was assigned was working on neural networks, and I started to study complex-valued neural networks under the supervision of Dr. Furuya (presently Professor of Toho University) (Nitta, & Furuya, 1991). At that time in 1990, I assumed I was the first in the world to extend neural networks to complex-valued models but I was wrong, as described below.

Up to the early 1980s, there were many studies on symbolic processing. I was also engaged in R&D of the Expert shell. Along with the development of Von Neumann-type computers, studies started on information processing different from symbolic processing but similar to that in the human brain. The neural network is one of such processes. Neural network is a network composed of artificial neurons and can be trained to find nonlinear relationships in data. Because there are many introductory references (for example, (Rojas, 1996)), only typical examples are outlined in this Preface. The first study on neural networks was reported by MacCulloch and Pitts in 1943. Stimulated by the results of anatomical and physiological studies, they proposed a network model consisting of a small number of very simple neurons and showed that the model could be used for logical calculations, and so forth. The original model of the present neural network was proposed by Rosenblatt in 1958 and was called Perceptron. In 1969, Minsky and Perpert showed mathematically that Perceptron cannot solve linearly non-separable problems (Minsky, & Papert, 1969), which requires identifying multi-dimensional data that are mutually interrelated. Perceptron was shown to be applicable only to simple problems. A feedforward neural network is a network in which signals are transmitted only in one direction. Rumelhart, Hinton, and Williams (1986) proposed a learning algorithm called back-propagation (BP), which was applicable to multilayer feedforward neural networks (Multilayer Perceptron). Multilayer feedforward neural networks with BP attracted attention as they could solve linearly non-separable problems, which could not be solved by Perceptron. By learning, the multilayer feedforward neural network acquires the ability to generalize. With this helpful ability, the network can output a sort of answer to unlearned patterns. This generalization ability is very useful when using the network in various fields. Hopfield proposed a kind of fully connected recurrent neural network (Hopfield, 1984; Hopfield & Tank, 1985). In fully connected recurrent neural networks, signals are transmitted not only in one direction but also in the op-
posite direction, and the signals can pass through the same neurons not only once but many times. The operation of the network is very simple. First, an arbitrary neuron is selected, and a simple computation is performed. Then the result of the computation is transmitted to all neurons in the network. Hopfield defined an index for showing the behavior of a fully connected recurrent neural network as a whole and called it energy function. He proved mathematically that the energy function decreases monotone with time and showed that combinational optimization problems could be quickly solved by approximation by using the monotone decreasing characteristic of the energy function. Kohonen (1995) showed that concept formation can be achieved using a neural network. Concept formation means automatic classification of a large amount of data. Two two-dimensional planes installed with two or more neurons are used. One plane is for receiving the input pattern (the input layer), and the other outputs the results (the output layer). The neurons on the input and output layers are weighted with weight parameters. The weight value is modified by Hebbian learning, which involves changing the value of the weight parameter according to the activity of the neuron, and a larger change is made at a higher activity.

The usual real-valued neural networks have been applied to various fields such as telecommunications, robotics, bioinformatics, image processing, and speech recognition, in which complex numbers (2 dimensions) are often used with the Fourier transformation. This indicates that complex-valued neural networks, whose parameters (weights and threshold values) are all complex numbers, are useful. In addition, in the human brain, an action potential may have different pulse patterns, and the distance between pulses may be different. This suggests that it is appropriate to introduce complex numbers representing phase and amplitude into neural networks. Furthermore, it is obvious that vectors with more than 2 dimensions are used in the real world to represent a cluster of something, for example, a 4-dimensional vector consisting of height, width, depth and time, and an \( N \)-dimensional vector consisting of \( N \) particles and so on. Thus, a model neuron that can deal with \( N \) signals as a cluster, is useful.

Aizenberg, Ivaskiv, Pospelov, and Hudiakov (1971) (former Soviet Union) proposed a complex-valued neuron model for the first time, and although it was only available in the Russian literature at the time, their work can now be read in English (Aizenberg, Aizenberg & Vandewalle, 2000). Prior to that time, most researchers other than Russians had assumed that the first persons to propose a complex-valued neuron were Widrow, McCool, and Ball (1975). Interest in the field of neural networks started to grow around 1990, and various types of complex-valued neural network models were subsequently proposed. Since then, their characteristics have been researched, making it possible to solve some problems which could not be solved with the real-valued neuron, and to solve many complicated problems more simply and efficiently. From 2001, several special sessions on complex-valued neural networks have been organized in several international conferences (KES, 2001, 2002, 2003; ICONIP, 2002, 2004; ICANN/ICONIP, 2003; IJCNN, 2006, 2008; ICANN, 2007).

There appear to be several approaches for extending the real-valued neural network to higher dimensions. One approach is to extend the number field, that is, from real numbers \( x \) (1 dimension), to complex numbers \( z = x + iy \) (2 dimensions), to quaternions \( q = a + ib + jc + kd \) (4 dimensions; see Chapter XVI), to octonions (8 dimensions), to sedenions (16 dimensions), and so forth (Weyl, 1946; Nitta, 1995; Arena, Fortuna, Muscato, & Xibilia, 1998; Pearson, 2003; Nitta, 2005; Buchholz, & Sommer, 2008). In this approach, the dimension of the input signal fed into the neural network is restricted to the form of \( 2n, n = 1, 2, \ldots \), that is, 1, 2, 4, 8, 16, ... Another approach is to extend the dimension of the threshold values and weights from 1 dimension to \( N \) dimensions using \( N \)-dimensional real-value vectors. In this approach, the dimension of the input signal fed into the neural network takes a natural number, that is, \( N = 1, 2, 3, 4, \ldots \). Moreover, there are two types of the latter approach: (a) weights are \( N \)-dimensional matrices (Nitta, & Garis, 1992; Nitta, 2006), or (b) weights are \( N \)-dimensional vectors (Nitta, 1993, 2007; Kobayashi, 2004). Also, there is an approach using hyperbolic numbers (2 dimensions) (Buchholz, & Sommer, 2000; Nitta, & Buchholz, 2008). Hyperbolic numbers, which are closely related to the popular complex numbers, are numbers of the form \( z = x + uy \) where \( x, y \) are real numbers and \( u \) is called unipotent which has the algebraic property that \( u \neq \pm 1 \) but \( u^2 = 1 \) (Sobczyk, 1995). Quantum neural networks can be viewed as one type of complex-valued neural network (see Chapters XIII-XV).
This book describes the latest developments in the theories and applications of neural networks with high-dimensional parameters which have been progressing in recent years. Graduate students and researchers will easily acquire the fundamental knowledge needed to be at the forefront of research, while practitioners will readily absorb the information required for applications. This book also provides a snapshot of current research and thus serves as a workbench for further developments in neural networks with high-dimensional parameters.

The following four books related to neural networks with high-dimensional parameters have been published: (Arena et al., 1998; Aizenberg et al., 2000; Hirose, 2003; Hirose, 2006). (Arena et al., 1998) was the first monograph on neural networks with high-dimensional parameters, and described the results of research on complex-valued neural networks, vectorial neural networks, and quaternion neural networks. The results of research up to 1997 are well organized in the monograph. The detailed descriptions of the function approximation capabilities of complex-valued neural networks with an analytic activation function and networks with a non-analytic activation function are excellent. (Aizenberg et al., 2000) is a comprehensive book on the complex-valued neuron models proposed by the authors, and is well organized from theories to applications. (Hirose, 2003) is an edited book which contains 14 chapters on complex-valued neural networks written by various authors. The chapter on the Clifford neural network written by Pearson is of interest to the complex-valued neural network community. (Hirose, 2006) is a translation of a book in Japanese (Hirose, 2005) that systematically describes complex-valued neural networks in the first half, and application examples obtained by the author’s laboratory in the second half.

It took a long time for mathematicians to accept complex numbers (Ebbinghaus, et al., 1988). During the Renaissance, when complex numbers were first discovered, they were called *quantitates impossibiles*. They were carefully calculated but were not recognized in mathematics. In the mid 19th century mathematicians finally recognized the real power of complex numbers. Today, physicists do not hesitate to speak of complex numbers as physical targets. Complex numbers appear in Schrödinger’s equation of quantum mechanics and are used in electrical engineering quite naturally. Complex numbers, which were once called *quantitates impossibiles*, are now firmly established in all fields of natural science and engineering, and scientists and engineers do not hesitate to use them in calculations. Unlike complex numbers, complex-valued neural networks were easily accepted in general. In my experience, I received only several negative comments in 1991 when I first proposed a complex-valued neural network. This quick recognition was likely because studies have focused on the engineering usefulness of complex-valued neural networks. Actually, most studies on complex-valued neural networks have been on engineering applications (usefulness) and were independent from those on the brain. It will be interesting to understand the actual relationships with the neural network of the brain. Neural networks are frequently grouped in soft computing together with evolutionary computation and fuzzy computation. Hybridizing neural networks with high-dimensional parameters with evolutionary computation or fuzzy computation looks promising (Chapter XV describes an example) for extending the potential of neural networks with high-dimensional parameters. In practice, there are many more study results and fields of application than are described in this book. Initially, this book was to contain 26 chapters, but the number was reduced to 16 for various reasons. It is a pity that we could not include the study on chaos by Nemoto and Saito (2002) and the studies on fractals by Miura and Aiyoshi (2003). For the special issue on complex-valued neural networks of the *International Journal of Neural Systems* (Rao, Nitta, & Murthy, 2008), for which I served as guest editor, 24 papers were submitted. Special sessions on complex-valued neural networks are also held in many international conferences as described above. We hope that readers all over the world will find this book both useful and enjoyable.

**ORGANIZATION OF THE BOOK**

The book is divided into three main sections: Complex-Valued Neural Network Models and Their Analysis (Chapters I-VI), Applications of Complex-Valued Neural Networks (Chapters VII-XII), and Models with High-Dimensional Parameters (Chapters XIII-XVI).
A brief description of each of the chapters follows.

**Chapter I** applies information geometry to complex-valued Boltzmann machines. The author of this chapter constructs the complex-valued Boltzmann machines, and investigates the structure of the complex-valued Boltzmann manifold. The author also derives an effective learning algorithm, called an *em* algorithm, for complex-valued Boltzmann machines with hidden neurons. Some important notions of information geometry, exponential families, mixture families, Kullback-Leibler divergence, connections, geodesics, Fisher metrics, potential functions and so on are explained for readers who are unfamiliar with information geometry.

**Chapter II** introduces the complex-valued network inversion method to solve inverse problems with complex numbers. The original network inversion is applied to usual multilayer neural networks with real-valued inputs and outputs, which solves inverse problems to estimate causes from results using a multilayer neural network. Regularization for the complex-valued network inversion is explained, which solves difficulties attributable to the ill-posedness of inverse problems.

**Chapter III** attempts to extend the clustering ensemble method and the Kolmogorov’s Spline Network to complex numbers, in the context of adaptive dynamic modeling of time-variant multidimensional data. The chapter is intended to provide an introduction to these subjects and to stimulate the participation of both young and experienced researchers in solving challenging and important problems in theory and practice related to this area.

**Chapter IV** describes a complex-variable version of the Hopfield neural network (CHNN), which can exist in both fixed point and oscillatory modes. In the fixed-point mode, CHNN is similar to a continuous-time Hopfield network. In the oscillatory mode, when multiple patterns are stored, the network wanders chaotically among patterns. It is shown that adaptive connections can be used to control chaos and increase memory capacity. Electronic implementation of the network in oscillatory dynamics, with fixed and adaptive connections, shows an interesting tradeoff between energy expenditure and retrieval performance. Some interesting applications are presented.

**Chapter V** presents global stability conditions for discrete-time and continuous-time complex-valued recurrent neural networks, which are regarded as nonlinear dynamical systems. Global asymptotic stability conditions for these networks are derived by suitably choosing activation functions. According to these stability conditions, there are classes of discrete-time and continuous-time complex-valued recurrent neural networks whose equilibrium point is globally asymptotically stable.

**Chapter VI** presents models of fully connected complex-valued neural networks which are complex-valued extensions of Hopfield-type neural networks and discusses methods of studying their dynamics. In particular, the author investigates existence conditions of energy functions for complex-valued Hopfield-type neural networks. As an application of the energy function, a qualitative analysis of the network by utilizing the energy function is shown and a synthesis method of complex-valued associative memories is discussed.

**Chapter VII** addresses a grey-box approach to complex-valued RBF modeling and develops a complex-valued symmetric RBF (SRBF) network model. The application of this SRBF network is demonstrated using nonlinear beamforming assisted detection for multiple-antenna aided wireless systems that employ complex-valued modulation schemes. Two training algorithms for this complex-valued SRBF network are proposed. The effectiveness of the proposed complex-valued SRBF network and the efficiency of the two training algorithms in a nonlinear beamforming application are demonstrated.

**Chapter VIII** illustrates the application of various types of complex-valued neural networks such as radial basis function networks (RBFN), multilayer feedforward networks and recurrent neural networks for training sequence-based as well as blind equalization of communication channels. The structures and algorithms for these equalizers are presented and performances based on simulation studies are analyzed, highlighting their advantages and the important issues involved.

**Chapter IX** presents the complex backpropagation (BP) algorithm for complex backpropagation neural networks (BPN) consisting of suitable node activation functions having multi-saturated output regions. The complex BPN is used as a nonlinear adaptive equalizer that can deal with both quadrature amplitude modulation (QAM) and phase shift key (PSK) signals of constellations of any size. In addition, four nonlinear blind equalization schemes using complex BPN for M-ary QAM signals are described and their learning algorithms are presented.
Chapter X presents new design methods for the complex-valued multistate Hopfield associative memories (CVHAMs). The stability of the presented CVHAM is analyzed by using the energy function approach which shows that in synchronous update mode a CVHAM is guaranteed to converge to a fixed point from any given initial state. Next, a generalized intraconnected bidirectional associative memory (GIBAM) is introduced, which is a complex generalization of the intraconnected BAM (IBAM).

Chapter XI proposes a method for automatically estimating nuclear magnetic resonance (NMR) spectra of metabolites in the living body by magnetic resonance spectroscopy (MRS) without human intervention or complicated calculations. In the method, the problem of NMR spectrum estimation is transformed into the estimation of the parameters of a mathematical model of the NMR signal. To estimate these parameters, the author designed a complex-valued Hopfield neural network, noting that NMR signals are essentially complex-valued.

Chapter XII introduces an independent component analysis (ICA) approach to the separation of linear and nonlinear mixtures in the complex domain. Source separation is performed by an extension of the INFOMAX approach to the complex environment. The neural network approach is based on an adaptive activation function, whose shape is properly modified during learning. A simple adaptation algorithm is derived and several experimental results are shown to demonstrate the effectiveness of the proposed method.

Chapter XIII introduces the authors’ qubit neural network, which is a multilayered neural network composed of quantum bit neurons. In this description, it is indispensable to use the complex-valued representation, which is based on the concept of quantum bits (qubits). The authors clarify that this model outperforms the conventional neural networks via computer simulations such as a benchmark test.

Chapter XIV shows the effectiveness of incorporating quantum dynamics and then proposes a neuromorphic adiabatic quantum computation algorithm based on the adiabatic change of Hamiltonian. The proposed method can be viewed as a complex-valued neural network because a qubit operates like a neuron. Next, the performance of the proposed algorithm is studied by applying it to a combinatorial optimization problem. Finally, the authors discuss learning ability and hardware implementation.

Chapter XV studies neural structures with weights that follow the model of the quantum harmonic oscillator. The proposed neural networks have stochastic weights which are calculated from the solution of Schrödinger’s equation under the assumption of a parabolic (harmonic) potential. The learning of the stochastic weights is analyzed. In the case of associative memories the proposed neural model results in an exponential increase of pattern storage capacity (number of attractors).

Chapter XVI describes two types of quaternionic neural network model. One type is a multilayer perceptron based on 3D geometrical affine transformations by quaternions. The operations that can be performed in this network are translation, dilatation, and spatial rotation in three-dimensional space. The other type is a Hopfield-type recurrent network whose parameters are directly encoded into quaternions. The fundamental properties of these networks are presented.

REFERENCES


