Automation and Control Engineering Series

Subspace Learning of Neural Networks

Jian Cheng Lv
Zhang Yi
Jiliu Zhou

CRC Press
Taylor & Francis Group

Subspace Learning of Neural Networks

AUTOMATION AND CONTROL ENGINEERING
A Series of Reference Books and Textbooks

Series Editors

FRANK L. LEWIS, Ph.D.,
Fellow IEEE, Fellow IFAC
Professor
Automation and Robotics Research Institute
The University of Texas at Arlington

SHUZHI SAM GE, Ph.D.,
Fellow IEEE
Professor
Interactive Digital Media Institute
The National University of Singapore

Subspace Learning of Neural Networks, *Jian Cheng Lv, Zhang Yi, and Jiliu Zhou*

Reliable Control and Filtering of Linear Systems with Adaptive Mechanisms, *Guang-Hong Yang and Dan Ye*

Reinforcement Learning and Dynamic Programming Using Function Approximators, *Lucian Buşoniu, Robert Babuška, Bart De Schutter, and Damien Ernst*

Modeling and Control of Vibration in Mechanical Systems, *Chunling Du and Lihua Xie*

Analysis and Synthesis of Fuzzy Control Systems: A Model-Based Approach, *Gang Feng*

Lyapunov-Based Control of Robotic Systems, *Aman Behal, Warren Dixon, Darren M. Dawson, and Bin Xian*

System Modeling and Control with Resource-Oriented Petri Nets, *Naiqi Wu and MengChu Zhou*

Sliding Mode Control in Electro-Mechanical Systems, Second Edition, *Vadim Utkin, Jürgen Guldner, and Jingxin Shi*

Optimal Control: Weakly Coupled Systems and Applications, *Zoran Gajić, Myo-Taeg Lim, Dobrila Skatarić, Wu-Chung Su, and Vojislav Kecman*

Intelligent Systems: Modeling, Optimization, and Control, *Yung C. Shin and Chengying Xu*

Optimal and Robust Estimation: With an Introduction to Stochastic Control Theory, Second Edition, *Frank L. Lewis, Lihua Xie, and Dan Popa*

Feedback Control of Dynamic Bipedal Robot Locomotion, *Eric R. Westervelt, Jessy W. Grizzle, Christine Chevallereau, Jun Ho Choi, and Benjamin Morris*

Intelligent Freight Transportation, *edited by Petros A. Ioannou*

Modeling and Control of Complex Systems, *edited by Petros A. Ioannou and Andreas Pitsillides*

Wireless Ad Hoc and Sensor Networks: Protocols, Performance, and Control, *Jagannathan Sarangapani*

Stochastic Hybrid Systems, *edited by Christos G. Cassandras and John Lygeros*

Hard Disk Drive: Mechatronics and Control, *Abdullah Al Mamun, Guo Xiao Guo, and Chao Bi*

Autonomous Mobile Robots: Sensing, Control, Decision Making and Applications, *edited by Shuzhi Sam Ge and Frank L. Lewis*

Automation and Control Engineering Series

Subspace Learning of Neural Networks

Jian Cheng Lv

Sichuan University
Chengdu, People's Republic of China

Zhang Yi

Sichuan University
Chengdu, People's Republic of China

Jiliu Zhou

Sichuan University
Chengdu, People's Republic of China

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

Printed in the United States of America on acid-free paper
10 9 8 7 6 5 4 3 2 1

International Standard Book Number-13: 978-1-4398-1536-6 (Ebook-PDF)

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

Dedication

To all of our loved ones

Preface

Principal component analysis (PCA) neural networks, minor component analysis (MCA) neural networks and independent component analysis (ICA) neural networks can approximate a subspace of input data by learning. These networks inspired by biology and psychology provide a novel way for parallel online computation of a subspace. An input of these neural networks can be used at once so that they can enable fast adaptation in a nonstationary environment. Although these networks are almost linear neural models, they have found many applications, including applications relating to signal and image processing, video analysis, data mining, and pattern recognition.

The learning algorithms of these neural networks play a vital role in subspace learning. These subspace learning algorithms make these networks learn low-dimensional linear and multilinear models in a high-dimensional space, wherein specific statistical properties can be well preserved. The book will be mainly focused on the convergence analysis of these subspace learning algorithms and the ways to extend the use of these networks to fields such as biomedical signal processing, biomedical image processing, and surface fitting to name just a few.

A crucial issue of concern in a practical application is the convergence of the subspace learning algorithms of these neural networks. The convergence of these algorithms determines whether these applications can be successful. The book will analyze the convergence of these learning algorithms by mainly using discrete deterministic time (DDT) method. To guarantee their nondivergence, invariant sets of some algorithms will be obtained and global boundedness of some algorithms is studied. Then, the convergence conditions of these algorithms will be derived. Cauchy convergence principle and inequalities analysis method, and so on, will be used rigorously to prove the convergence. Furthermore, the book establishes a relationship between an SDT algorithm and the corresponding DDT algorithm by using block algorithms. This not only can overcome the shortcomings of DDT method, but also can get a good convergence and accuracy in practice. Finally, the chaotic and robust properties of some algorithms will also be studied. These results obtained lay the sound theoretical foundation of these networks and guarantee the successful applications of these algorithms in practice.

The book not only benefits the researcher of subspace learning algorithms, but also improves the quality of data mining, image processing, and signal processing. Besides its research contributions and applications, the book could also serve as a good example for pushing the latest technologies in neural networks to some application community.

Scope and Contents of This Book

This book provides an analysis framework for convergence analysis of subspace learning algorithms of neural networks. The emphasis is on the analysis method, which can be generalized to the study of other learning algorithms. Our work builds a theoretical understanding of the convergence behavior of some subspace learning algorithms through the analysis framework. In addition, this book uses real-life examples to illustrate the performance of learning algorithms and instructs readers on how to apply them to practical applications. The book is organized as follows.

Chapter 1 provides a brief introduction to linear neural networks and subspace learning algorithms of neural networks. Some frequently used notations and preliminaries are given. Basic discussions on the methods for convergence analysis are presented which should lay the foundation for subsequent chapters.

In the following chapters, convergence of subspace learning algorithms is analyzed to lay the theoretical foundation for successful applications of these networks. In Chapter 2, the convergence of Oja's and Xu's algorithms with constant learning rates is studied in detail. The global convergence of Oja's algorithm with the adaptive learning rate is analyzed in Chapter 3. In Chapter 4, the convergence of Generalized Hebbian Algorithm (GHA) with adaptive learning rates is studied. MCA learning algorithms and the Hyvärinen-Oja's ICA learning algorithm are analyzed in Chapters 5 and 6, respectively. In Chapter 7, chaotic behaviors of subspace learning algorithms are presented.

Some problems concerning a practical application are discussed in chapters 8, 9, 10, 11, and some real-life examples are given to illustrate the performance of these subspace learning algorithms.

The contents of this book are mainly based on our research publications on this subject, which over the years have accumulated into a complete and unified coverage of the topic. It will serve as an interesting reference for postgraduates, researchers, and engineers who may be keen to use these neural networks in applications. Undoubtedly, there are other excellent works in this area, which we hope to have included in the references for the readers. We should also like to point out that at the time of this writing, many problems relating to subspace learning remained unresolved, and the book may contain personal views and conjecture of the authors that may not appeal to all sectors of readers. To this end, readers are encouraged to send us criticisms and suggestions, and we look forward to discussion and collaboration on the topic.

Acknowledgments

This book was supported in part by the National Science Foundation of China under grants 60971109 and 60970013.

Jian Cheng Lv
Zhang Yi
Jiliu Zhou

January 2010

List of Figures

List of Tables

Contents

1

Introduction

1.1 Introduction

Subspace leaning neural networks provide a parallel online automated learning of low-dimensional models in a nonstationary environment. It is commonly known that automated learning of low-dimensional linear or multilinear models from training data has become a standard paradigm in computer vision [54, 55]. Thus, these neural networks used to learning a low-dimensional model have many important applications in computer vision, such as structure from motion, motion estimation, layer extraction, objection recognition, and object tracking [18, 19, 50, 56, 55].

Subspace learning algorithms used to update the weights of these networks pay a vital important role in applications. This book will focus on the subspace learning algorithms of principal component analysis (PCA) neural networks, minor component analysis (MCA) neural networks, and independent component analysis (ICA) neural networks. Generally, they are linear neural networks.

1.1.1 Linear Neural Networks

Inspired by biological neural networks, a simple neural model is designed to mimic its biological counterpart, the neuron. The model accepted the weighted set of input x responds with an output y, as shown in Figure 1.1. The vital effect of synapse between neurons is presented by the weights w. The mathematical model of a linear single neuron is as follows:

$$y(k) = w^T(k)x(k), (k = 0, 1, 2, \ldots), \tag{1.1}$$

where $y(k)$ is the network output, the input sequence

$$\{x(k)\,|x(k) \in R^n(k = 0, 1, 2, \ldots)\} \tag{1.2}$$

is a zero-mean stochastic process, each $w(k) \in R^n(k = 0, 1, 2, \ldots)$ is a weight vector.

Consider input output relation

$$y(k) = W^T x(k), \tag{1.3}$$

where

$$y(k) = [y_1(k), y_2(k), \ldots, y_m(k)]^T$$

and

$$W = \begin{bmatrix} w_{11} & w_{12} & \ldots & w_{1m} \\ w_{21} & w_{22} & \ldots & w_{2m} \\ \ldots & \ldots & \ldots & \ldots \\ w_{n1} & w_{n2} & \ldots & w_{nm} \end{bmatrix}.$$

This is a multineuron linear model, as shown in Figure 1.2.

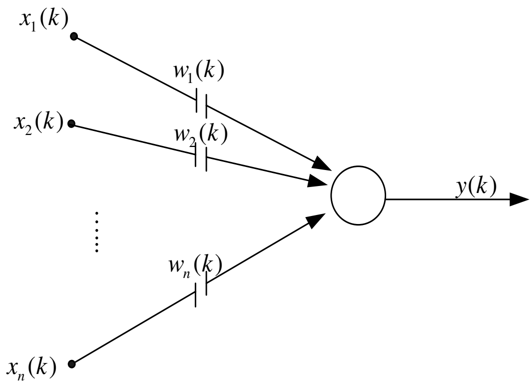

FIGURE 1.1
A single neuron model.

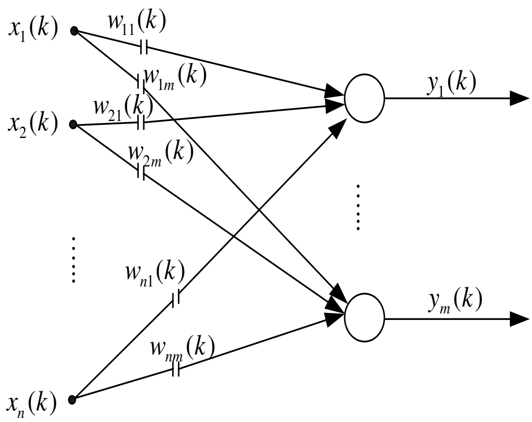

FIGURE 1.2
A multineuron model.

These linear networks have been widely studied [25, 28, 70] and used for many fields involving signal processing [66], prediction [80], associative memory [167], function approximation [180], power system [15, 153], chemistry [196], and so on. The weight of these neural networks, which represents the strength of the connection between neurons, pays a very important role in the different applications. A variety of learning algorithms are used to update the weight so that the networks have the different applications with the corresponding weight. For instance, the adaptive linear neuron (ADALINE) network is one of the most widely used neural networks in practical applications, which was introduced by Widrow and M. Hoff in 1960 [181]. The least mean square (LMS) algorithm is used to update the weight so that ADALINE can be used as a adaptive filter [75].

In this book, these linear networks are used to extract the principal components, or minor components, or independent components from input data. It is required that these networks must approximate a low-dimensional model.

1.1.2 Subspace Learning Algorithms

Subspace learning algorithms are used to update the weights of these linear networks so that these networks intend to approximate a low-dimensional linear model. Generally, an original stochastic discrete time (SDT) algorithm is formulated as

$$w(k+1) = w(k) \pm \eta(k) \triangle w(k), \tag{1.4}$$

where $\eta(k)$ is learning rate and $\triangle w(k)$ determines the change at time k.

This book will mainly discuss the following subspace learning algorithms: PCA learning algorithms, MCA learning algorithms, and ICA learning algorithms.

1.1.2.1 PCA Learning Algorithms

Principal component analysis (PCA) is a traditional statistical technique in multivariate analysis, stemming from the early work of Pearson [141]. It is closely related to Karhunen-Loève (KL) transform, or the Hotelling transform [82]. The purpose of PCA is to reduce the dimensionality of a given data set, while retaining as much as possible of the information present in the data set.

Definition 1.1 *A vector is called the first principal component direction if the vector is along the eigenvector associated with the largest eigenvalue of the covariance matrix of a given data set, and a vector is called the second principal component direction if the vector is along the eigenvector associated with the second largest eigenvalue of the covariance matrix of a given data set, and so on.*

Definition 1.2 *Principal components are the variances that are obtained by projecting the given data onto the principal component directions.*