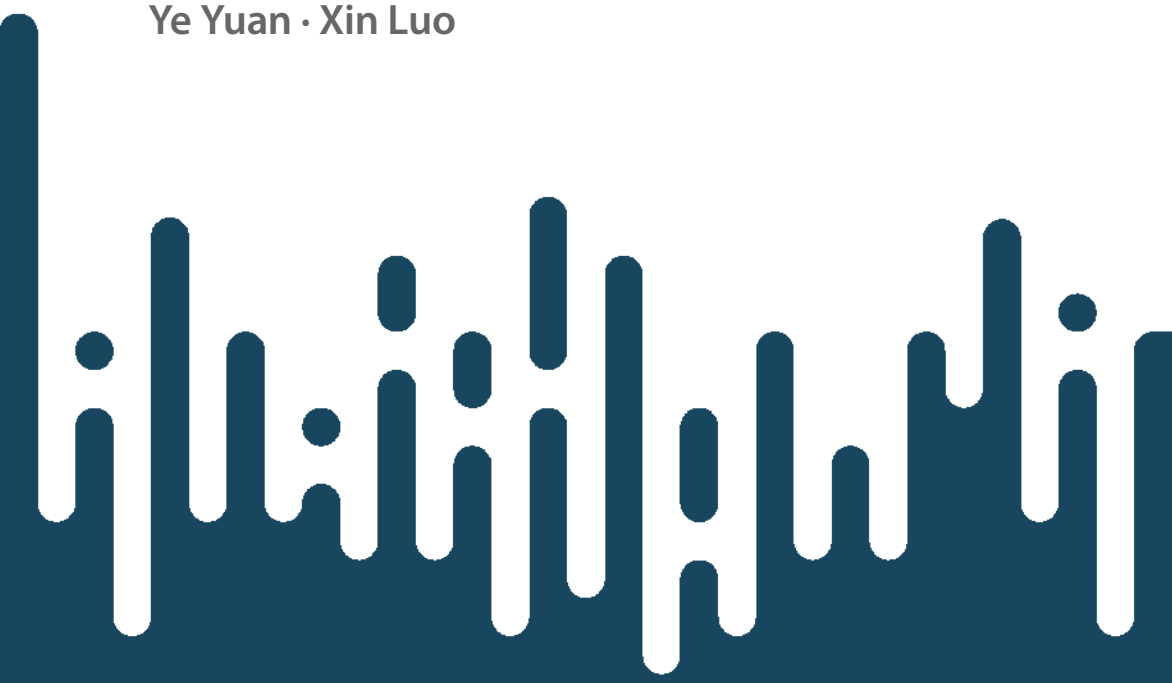
SpringerBriefs in Computer Science

Ye Yuan · Xin Luo



Latent Factor Analysis for High-dimensional and Sparse Matrices

A particle swarm optimization-based approach

Springer

SpringerBriefs in Computer Science

Series Editors

Stan Zdonik, Brown University, Providence, RI, USA

Shashi Shekhar, University of Minnesota, Minneapolis, MN, USA

Xindong Wu, University of Vermont, Burlington, VT, USA

Lakhmi C. Jain, University of South Australia, Adelaide, SA, Australia

David Padua, University of Illinois Urbana-Champaign, Urbana, IL, USA

Xuemin Sherman Shen, University of Waterloo, Waterloo, ON, Canada

Borko Furht, Florida Atlantic University, Boca Raton, FL, USA

V. S. Subrahmanian, University of Maryland, College Park, MD, USA

Martial Hebert, Carnegie Mellon University, Pittsburgh, PA, USA

Katsushi Ikeuchi, University of Tokyo, Tokyo, Japan

Bruno Siciliano, Università di Napoli Federico II, Napoli, Italy

Sushil Jajodia, George Mason University, Fairfax, VA, USA

Newton Lee, Institute for Education, Research and Scholarships, Los Angeles, CA, USA

SpringerBriefs present concise summaries of cutting-edge research and practical applications across a wide spectrum of fields. Featuring compact volumes of 50 to 125 pages, the series covers a range of content from professional to academic.

Typical topics might include:

- A timely report of state-of-the art analytical techniques
- A bridge between new research results, as published in journal articles, and a contextual literature review
- A snapshot of a hot or emerging topic
- An in-depth case study or clinical example
- A presentation of core concepts that students must understand in order to make independent contributions

Briefs allow authors to present their ideas and readers to absorb them with minimal time investment. Briefs will be published as part of Springer's eBook collection, with millions of users worldwide. In addition, Briefs will be available for individual print and electronic purchase. Briefs are characterized by fast, global electronic dissemination, standard publishing contracts, easy-to-use manuscript preparation and formatting guidelines, and expedited production schedules. We aim for publication 8–12 weeks after acceptance. Both solicited and unsolicited manuscripts are considered for publication in this series.

**Indexing: This series is indexed in Scopus, Ei-Compendex, and zbMATH **

Ye Yuan • Xin Luo

Latent Factor Analysis for High-dimensional and Sparse Matrices

A particle swarm optimization-based approach

Springer

Ye Yuan
Computer and Information Science
Southwest University
Chongqing, China

Xin Luo
Computer and Information Science
Southwest University
Chongqing, China

ISSN 2191-5768 ISSN 2191-5776 (electronic)
SpringerBriefs in Computer Science
ISBN 978-981-19-6702-3 ISBN 978-981-19-6703-0 (eBook)
https://doi.org/10.1007/978-981-19-6703-0

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022
This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Preface

High-dimensional and sparse (HiDS) matrices are commonly seen in many big data-related industrial applications like electronic commerce, cloud services, social networks, and wireless sensor networks. Despite its extreme sparse, an HiDS matrix contains rich knowledge regarding desired patterns like users' potential favorites, item clusters, and topological neighbors. Hence, how to efficiently and accurately extract desired knowledge from it for various data analysis tasks becomes a highly interesting issue.

Latent factor analysis (LFA) has proven to be highly efficient for addressing an HiDS matrix owing to its high scalability and efficiency. However, an LFA model's performance relies heavily on its hyper-parameters, which should be chosen with care. The common strategy is to employ grid-search to tune these hyper-parameters. However, it requires expensive computational cost to search the whole candidate space, especially when building an LFA model on an HiDS matrix. Hence, how to implement a hyper-parameter-free LFA model becomes a significant issue.

In this book, we incorporate the principle of particle swarm optimization (PSO) into latent factor analysis, thereby achieving four effective hyper-parameter-free latent factor analysis models. The first model is a Learning rate-free LFA (L^2FA) model, which utilizes a standard PSO algorithm into the learning process by building a swarm of learning rates applied to the same group. The second model is a *L*earning rate and *R*egularization coefficient-free *LFA* (LRLFA) model, which build a swarm by taking the learning rate and regularization coefficient of every single LFA-based model as particles, and then apply particle swarm optimization to make them adaptation according to a predefined fitness function. The third model is a *G*eneralized and *A*daptive *LFA* (GALFA) model, which implement self-adaptation of the regularization coefficient and momentum coefficient for excellent practicability via PSO. The last model is an *A*dvanced *L*earning rate-free *LFA* (AL^{2}FA) model. Before building this model, we first propose a novel position-transitional particle swarm optimization (P^{2}SO) algorithm by incorporate more dynamic information into the particle's evolution for preventing premature convergence. And then, a P^{2}SO algorithm is utilized into the training process to make the learning rate adaptation without

accuracy loss. Since hyper-parameter adaptation is done within only one full training process, thereby greatly reducing computational cost. Hence, it fits the need of real applications with high scalability and efficiency.

This is the first book about how to incorporate particle swarm optimization into latent factor analysis for implementing effective hyper-parameter adaptation. It is intended for professionals and graduate students involved in LFA on HiDS data. It is assumed that the reader has a basic knowledge of mathematics, as well as a certain background in data mining. The reader can get an overview on how to implement efficient hyper-parameter adaptation in an LFA model. We hope this monograph will be a useful reference for students, researchers, and professionals to understand the basic methodologies of hyper-parameter adaptation via PSO. The readers can immediately conduct extensive researches and experiments on the real applications HiDS data involved in this book.

Chongqing, China Ye Yuan
June 2022 Xin Luo

Contents

1 Introduction 1
1.1 Background 1
1.2 Preliminaries 2
1.2.1 Symbol and Abbreviation Appointment 2
1.2.2 An LFA Model 3
1.2.3 Particle Swarm Optimization 6
1.3 Book Organization 6
References 7

2 Learning Rate-Free Latent Factor Analysis via PSO 11
2.1 Overview 11
2.2 An LFA Model with SGD Algorithm 12
2.3 The Proposed L^{2}FA Model 14
2.3.1 Learning Rate Adaptation via PSO 14
2.3.2 Algorithm Design and Analysis 15
2.4 Experimental Results and Analysis 17
2.4.1 General Settings 17
2.4.2 Performance Comparison 18
2.4.3 Effect of Swarm Size 20
2.4.4 Summary 22
2.5 Conclusions 22
References 25

3 Learning Rate and Regularization Coefficient-Free Latent Factor Analysis via PSO 29
3.1 Overview 29
3.2 An SGD-Based LFA Model 31
3.3 The Proposed LRLFA Model 32
3.3.1 Learning Rate and Regularization Coefficient Adaptation via PSO 32
3.3.2 Linearly Decreasing Inertia Weight Incorporation 35

3.3.3 Algorithm Design and Analysis . . . 36
3.4 Experimental Results and Analysis . . . 37
3.4.1 General Settings . . . 37
3.4.2 Effect of Swarm Size . . . 39
3.4.3 Performance Comparison . . . 41
3.4.4 Summary . . . 45
3.5 Conclusions . . . 45
References . . . 45

4 Regularization and Momentum Coefficient-Free Non-negative Latent Factor Analysis via PSO . . . 51
4.1 Overview . . . 51
4.2 An SLF-NMU-Based NLFA Model . . . 53
4.3 The Proposed GALFA Model . . . 54
4.3.1 A Generalized Learning Objective . . . 54
4.3.2 A Generalized SLF-NMU-Based Learning Rule . . . 55
4.3.3 Generalized-Momentum Incorporation . . . 56
4.3.4 Regularization and Momentum Coefficient Adaptation via PSO . . . 58
4.3.5 Algorithm Design and Analysis . . . 60
4.4 Experimental Results and Analysis . . . 63
4.4.1 General Settings . . . 63
4.4.2 Parameter Sensitivity . . . 64
4.4.3 Comparison with State-of-the-Art Models . . . 68
4.4.4 Summary . . . 71
4.5 Conclusions . . . 71
References . . . 72

5 Advanced Learning Rate-Free Latent Factor Analysis via P^2SO . . . 75
5.1 Overview . . . 75
5.2 An SGD-Based LFA Model . . . 76
5.3 The Proposed AL^2FA Model . . . 78
5.3.1 A P^2SO Algorithm . . . 78
5.3.2 Learning Rate Adaptation via P^2SO . . . 80
5.3.3 Algorithm Design and Analysis . . . 81
5.4 Experimental Results and Analysis . . . 83
5.4.1 General Settings . . . 83
5.4.2 Effect of ρ . . . 84
5.4.3 Comparison Results . . . 85
5.4.4 Summary . . . 87
5.5 Conclusions . . . 87
References . . . 88

6 Conclusion and Future Directions . . . 91
6.1 Conclusion . . . 91
6.2 Discussion . . . 92
References . . . 92