Linear Algebra for Data Science, Machine Learning, and Signal Processing

Maximize student engagement and understanding of matrix methods in data-driven applications with this modern teaching package. Students are introduced to matrices in two preliminary chapters, before progressing to advanced topics such as the nuclear norm, proximal operators, and convex optimization. Highlighted applications include low-rank approximation, matrix completion, subspace learning, logistic regression for binary classification, robust principal component analysis, dimensionality reduction, and Procrustes problems.

Extensively tested in the classroom, the book includes over 200 multiple-choice questions suitable for in-class interactive learning or quizzes, as well as homework exercises (with solutions available for instructors). It encourages active learning with engaging "explore" questions, with answers at the back of each chapter, and Julia code examples to demonstrate how the math is actually used in practice. A suite of computational notebooks offers a hands-on learning experience for students. This is a perfect textbook for upper-level undergraduates and first-year graduate students who have taken a prior course in linear algebra basics.

Jeffrey A. Fessler is the William L. Root Professor of Electrical Engineering and Computer Science at the University of Michigan. He received the Edward Hoffman Medical Imaging Scientist Award in 2013, and an IEEE EMBS Technical Achievement Award in 2016. He received the 2023 Steven S. Attwood Award, the highest honor awarded to a faculty member by the College of Engineering at the University of Michigan. He is a fellow of the IEEE.

Raj Rao Nadakuditi is an Associate Professor of Electrical Engineering and Computer Science at the University of Michigan. He received the Jon R. and Beverly S. Holt Award for Excellence in Teaching in 2018 and the Ernest and Bettine Kuh Distinguished Faculty Award in 2021.

"The authors provide a comprehensive contemporary presentation of linear algebra, demonstrating its foundational and intrinsic value to modern subjects, such as machine/deep learning, data science, and signal processing. The presentation is fun, exciting, topic-diverse, classroom tested, and addresses practical implementation in ways that jump start students' use."

Christ D. Richmond, *Duke University*

"This is an excellent and timely text that addresses the specific needs of data science (DS), machine learning (ML), and signal processing (SP). Its nicely crafted coverage is designed to prepare students in the areas of DS/ML/SP, in particular, by drawing thoughtful examples from these fields. With increasing demands from data-based sciences, there is a pressing need for a book in 'the new linear algebra,' and this text fills this gap."

Yousef Saad, *University of Minnesota*

"With the emergence of Graphics Processing Units (GPUs), the importance of linear algebra for machine learning cannot be overstated. This is a thoughtful and timely work on the topic of linear algebra for machine learning, which I anticipate will be one of the definitive textbooks in this field."

Vahid Tarokh, *Duke University*

"To see the spirit of this book, just look at pages 1 and 2. A painting is deblurred by linear algebra. Great ideas and how to use them in real time – all on display!"

Gilbert Strang, *Massachusetts Institute of Technology*

Linear Algebra for Data Science, Machine Learning, and Signal Processing

JEFFREY A. FESSLER
University of Michigan, Ann Arbor

RAJ RAO NADAKUDITI
University of Michigan, Ann Arbor

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

Shaftesbury Road, Cambridge CB2 8EA, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre, New Delhi – 110025, India

103 Penang Road, #05–06/07, Visioncrest Commercial, Singapore 238467

Cambridge University Press is part of Cambridge University Press & Assessment, a department of the University of Cambridge.

We share the University's mission to contribute to society through the pursuit of education, learning and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/highereducation/isbn/9781009418140

DOI: 10.1017/9781009418164

First published 2024

A catalogue record for this publication is available from the British Library

Library of Congress Cataloging-in-Publication Data
Names: Fessler, Jeffrey A., 1964- author. | Nadakuditi, Raj Rao, 1977- author.
Title: Linear algebra for data science, machine learning, and signal processing / Jeffrey A. Fessler, Raj Rao Nadakuditi.
Description: New York, NY : Cambridge University Press, 2024. | Includes bibliographical references and index.
Identifiers: LCCN 2023050878 (print) | LCCN 2023050879 (ebook) | ISBN 9781009418140 (hardback) | ISBN 9781009418164 (epub)
Subjects: LCSH: Algebras, Linear–Textbooks. | Matrices–Textbooks. | Machine learning–Mathematics–Textbooks. | Signal processing–Mathematics–Textbooks. | Julia (Computer program language)–Textbooks.
Classification: LCC QA184.2 .F47 2024 (print) | LCC QA184.2 (ebook) | DDC 512/.5–dc23/eng/20231117
LC record available at https://lccn.loc.gov/2023050878
LC ebook record available at https://lccn.loc.gov/2023050879

ISBN 978-1-009-41814-0 Hardback

Additional resources for this publication at www.cambridge.org/FesslerNadakuditi

To Sue

To Tata, Bama, Dad, Reizo and Ms. Gracie

Contents

Preface

Overview

Modern methods in data science, machine learning, and signal processing (DS–ML–SP) all build extensively on matrix methods and linear algebra. Often students who are interested in DS–ML–SP are advised to "go take a linear algebra course" with the promise that the material learned there will be useful later in more advanced courses. The content in this book is designed to teach important linear algebra ideas in an integrated way with computational methods in the context of DS–ML–SP applications. The focus here is on using matrix methods to pose and solve DS–ML–SP problems, rather than to provide rigorous proofs of linear algebra theorems. Traditional linear algebra and numerical linear algebra courses spend considerable time focusing on solving $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$. Solving systems of equations is essential for physics-based applications described by partial differential equations, whereas modern DS–ML–SP applications are data-driven and rarely reduce to solving $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$. Thus, this book treats the topic of solving linear systems only very briefly so that we can get to "the good stuff" like regularized least squares regression (Chapter 5), multidimensional scaling (Chapter 6), low-rank matrix approximation and Procrustes analysis (Chapter 7), Markov chains and the PageRank method (Chapter 8), logistic regression for binary classification, (Chapter 9), and matrix completion (Chapter 10). Put another way, after establishing a foundation in Chapters 1–3, every chapter that follows has further mathematical methods and models, along with one or more compelling applications to motivate and illustrate them. The goal of this book is to provide mathematical foundations for subsequent DS–ML–SP courses while also introducing matrix-based DS–ML–SP methods and applications that are useful in their own right.

Software

Nearly every mathematical concept in this book has corresponding operations in software, and this book describes those operations using the JULIA language [1]. This relatively new language has many benefits. JULIA is designed in a way that allows the code to look very similar to the math, facilitating the translation of algorithm ideas to working software. JULIA uses dynamic typing so it is suitable for interactive and educational use, yet it is very fast because it is compiled. JULIA is open source and its git-based package manager greatly facilitates reproducible research. Readers do need to know JULIA to begin using this book; the language borrows a lot of ideas from MATLAB and Python (among others), so readers familiar with those tools will be able to follow the examples easily. Readers can view the JULIA code examples as pseudocode even if they prefer to use other languages, learning some JULIA along the way. There are many tutorials online as well as other books based on JULIA that provide useful references, for example, [2]. Every figure in this book was generated using JULIA.

Textbook Use

The content in this book has been used in two first-year graduate courses (EECS 505 and EECS 551) at the University of Michigan since at least 2016, taken by several thousand students over that time. Since 2017, those courses have used JULIA as the primary (505) or only (551) language for illustrating and implementing the ideas. Senior-level undergraduates with mathematical maturity have also taken these courses. The courses include weekly discussion sections where students apply the techniques to real data (like handwritten digit classification) using Jupyter notebooks. (The Ju in Jupyter is for JULIA.)

Several methods in the book are illustrated in JULIA demos at the website https://github.com/JeffFessler/book-la-demo. These demos were created using the convenient `Literate.jl` and `Documenter.jl` tools in JULIA. Those tools generate HTML output that is easily viewed in a browser, as well as Jupyter notebooks that students can use and modify.

A prior undergraduate-level course in linear algebra is likely to be helpful as background for getting the most out of this book. A prior undergraduate-level course in digital signal processing is helpful for understanding a few of the examples, but is not essential for most of the book.

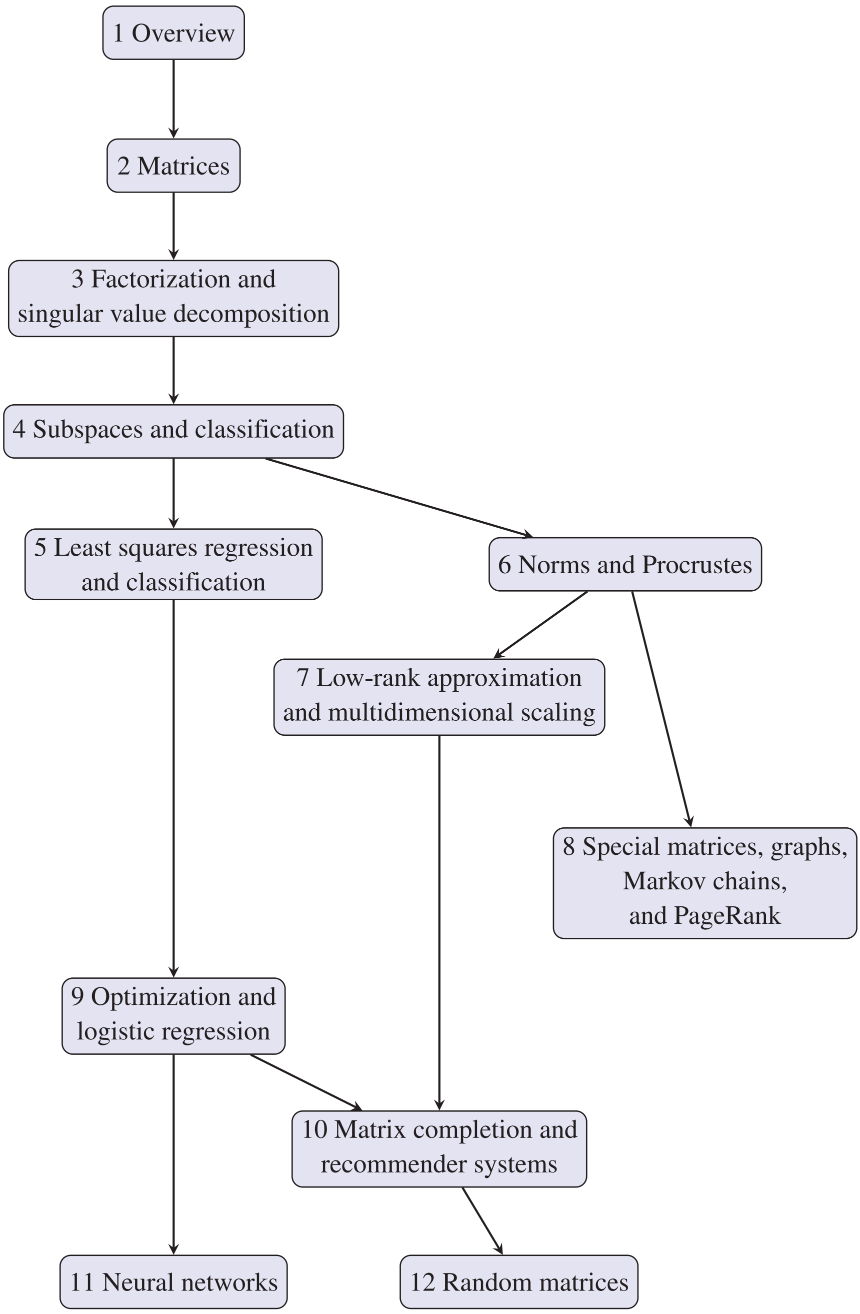
Instructor Resources

Embedded in the chapters are over 200 multiple-choice questions that instructors can use for in-class active learning exercises, or for self study by readers.

There are over 150 exercises at the ends of the chapters. Typeset solutions to these problems are available for instructors on the book's web page https://doi.org/10.1017/9781009418164. Also available there are slides of the material for classroom use. One version of the slides is a skeleton format with key equations omitted that an instructor can complete interactively during a lecture. (This is how the first author teaches this material.)

Organization

The following diagram illustrates how the book chapters are related. The first few chapters provide a foundation that should be read in sequence. There is more flexibility in the ordering of the subsequent chapters.



Related Books

Books we used as references when preparing this material include [3], [4], [5], and [6]. None of those books use Julia to illustrate the ideas.

Other books that provide useful linear algebra background, also using Julia, are [2], [7], and [8]. Those books have less depth in DS–ML–SP applications. [9] describes Julia and uses it for some ML applications with less matrix fundamentals. Other books using Julia for related topics include [10], [11], and [12].

There are many graduate-level books on DS–ML–SP topics that give a brief review of linear algebra concepts before delving into more advanced material. The material in this book will provide the reader with a more thorough foundation in preparation for more advanced DS–ML–SP courses.

Acknowledgments

The authors donate a portion of their royalties to organizations that empower groups that have been historically disadvantaged in science, technology, engineering, and mathematics fields, including ostem.org.

From JF

Thanks to Prof. Zhongming Liu, Prof. Yong Long, graduate student instructors Caroline Crockett, David Hong, Steven Whitaker, Haowei Xiang, and postdocs Rodrigo Lobos, Greg Ongie, and Dan Weller, and numerous past students, including Yongli He, Winston Wang, Emma Shedden, Zicheng Jin, Ege Taga and Yixuan (Isaac) Jia, for many corrections and suggestions. Special thanks to Matt Raymond for particularly detailed and insightful suggestions that refined and clarified the content of each chapter. Thanks to Raj for starting me on this journey with his handwritten notes. The best way to learn is to teach from a colleague's insightful lecture notes!

From RN

Special thanks to Jonas Kersulis and Brian Moore for helping create, edit, and test the many computationally centered homework problems.

This book would not be possible without them and the various graduate student instructors (Arvind Prasadan, David Hong, Hao Wu, Rishi Sonthalia, Dipak Narayan, Yash Sanjay Bhalgat, and Raj Tejas Suryaprakash) who helped edit, test, and refine the many homework problems, often right before they were about to go live to hundreds of students. Thanks to Simon Danisch for helping start this journey in 2017 by porting my MATLAB demos to JULIA.

Thanks in particular to Gil Strang for his encouragement, feedback, and support, and for his inspiration during the very special semester of Spring 2017 when we launched and taught 18.065 at MIT using some of the material in this book.

Multiple thanks to Alan Edelman for years of encouragement and inspiration, and for teaching me so much (including JULIA). A reader might sometimes recognize Alan and Gil's style in the way the math and code are presented. That's no accident. This book is infused with their DNA and years of me soaking in their thoughts and ideas on so many matters, particularly on how elegant math produces elegant code and vice versa. All they taught me about how to see math and linear algebra makes me love it, and to want to share it with you in this book, even more.