Inference and Learning from Data

Volume II

This extraordinary three-volume work, written in an engaging and rigorous style by a world authority in the field, provides an accessible, comprehensive introduction to the full spectrum of mathematical and statistical techniques underpinning contemporary methods in data-driven learning and inference.

This second volume, *Inference*, builds on the foundational topics established in Volume I to introduce students to techniques for inferring unknown variables and quantities, including Bayesian inference, Markov chain Monte Carlo methods, maximum likelihood, variational inference, hidden Markov models, Bayesian networks, and reinforcement learning.

A consistent structure and pedagogy is employed throughout this volume to reinforce student understanding, with over 350 end-of-chapter problems (including solutions for instructors), 180 solved examples, almost 200 figures, datasets, and downloadable Matlab code. Supported by sister volumes *Foundations* and *Learning*, and unique in its scale and depth, this textbook sequence is ideal for early-career researchers and graduate students across many courses in signal processing, machine learning, statistical analysis, data science, and inference.

Ali H. Sayed is Professor and Dean of Engineering at École Polytechnique Fédérale de Lausanne (EPFL), Switzerland. He has also served as Distinguished Professor and Chairman of Electrical Engineering at the University of California, Los Angeles (UCLA), USA, and as President of the IEEE Signal Processing Society. He is a member of the US National Academy of Engineering (NAE) and The World Academy of Sciences (TWAS), and a recipient of several awards, including the 2022 IEEE Fourier Award and the 2020 IEEE Norbert Wiener Society Award. He is a Fellow of the IEEE, EURASIP, and AAAS.

Inference and Learning from Data

Volume II: Inference

ALI H. SAYED
École Polytechnique Fédérale de Lausanne
University of California at Los Angeles

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/9781009218269

DOI: 10.1017/9781009218245

First published 2023

Printed in the United Kingdom by Bell and Bain Ltd

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

ISBN - 3 Volume Set 978-1-009-21810-8 Hardback
ISBN - Volume I 978-1-009-21812-2 Hardback
ISBN - Volume II 978-1-009-21826-9 Hardback
ISBN - Volume III 978-1-009-21828-3 Hardback

Additional resources for this publication at www.cambridge.org/sayed-vol2.

In loving memory of my parents

Contents

Cambridge University Press & Assessment
978-1-009-21826-9 — Inference and Learning from Data
Volume 2
Ali H. Sayed
Frontmatter
More Information

978-1-009-21826-9 — Inference and Learning from Data
Volume 2
Ali H. Sayed
Frontmatter
More Information

Cambridge University Press & Assessment
978-1-009-21826-9 — Inference and Learning from Data
Volume 2
Ali H. Sayed
Frontmatter
More Information

Preface

Learning directly from data is critical to a host of disciplines in engineering and the physical, social, and life sciences. Modern society is literally driven by an interconnected web of data exchanges at rates unseen before, and it relies heavily on decisions inferred from patterns in data. There is nothing fundamentally wrong with this approach, except that the inference and learning methodologies need to be anchored on solid foundations, be fair and reliable in their conclusions, and be robust to unwarranted imperfections and malicious interference.

P.1 EMPHASIS ON FOUNDATIONS

Given the explosive interest in data-driven learning methods, it is not uncommon to encounter claims of superior designs in the literature that are substantiated mainly by sporadic simulations and the potential for "life-changing" applications rather than by an approach that is founded on the well-tested scientific principle to inquiry. For this reason, one of the main objectives of this text is to highlight, in a unified and formal manner, the firm mathematical and statistical pillars that underlie many popular data-driven learning and inference methods. This is a nontrivial task given the wide scope of techniques that exist, and which have often been motivated independently of each other. It is nevertheless important for practitioners and researchers alike to remain cognizant of the common foundational threads that run across these methods. It is also imperative that progress in the domain remains grounded on firm theory. As the aphorism often attributed to Lewin (1945) states, "*there is nothing more practical than a good theory.*" According to Bedeian (2016), this saying has an even older history.

Rigorous data analysis, and conclusions derived from experimentation and theory, have been driving science since time immemorial. As reported by Heath (1912), the Greek scientist Archimedes of Syracuse devised the now famous Archimedes' Principle about the volume displaced by an immersed object from observing how the level of water in a tub rose when he sat in it. In the account by Hall (1970), Gauss' formulation of the least-squares problem was driven by his desire to predict the future location of the planetoid Ceres from observations of its location over 41 prior days. There are numerous similar examples by notable scientists where experimentation led to hypotheses and from there