Matrix Methods in Data Mining and Pattern Recognition. *By Lars Elden.* SIAM, Philadelphia, PA, 2007. $69.00. x+224 pp., softcover. ISBN 978-0-898716-26-9.

The author indicates that the book is intended as an undergraduate text for an introduction to data mining for students with some background in scientific computing or numerical analysis. Graduate students from other disciplines (engineering and the physical sciences, for example) will find the background material on linear algebra especially useful. The author seeks to demonstrate that linear algebra is a key player in the development of problem solving techniques in data mining and pattern recognition. One could easily use this book as a text for a second (semester) course in applied linear algebra.

The first nine chapters of the book are devoted to fundamental concepts of linear algebra and matrix decompositions. Topics of this first part of the book include matrix multiplication, matrix norms, rank, linear systems and least squares problems, orthogonality, QR decomposition, singular value decomposition, tensor decomposition, k-means, and nonnegative matrix factorization. Several MATLAB code examples are used to demonstrate the definition, solution, or factorization described. Examples drawn from text or web mining and image analysis are used throughout this section of the book.

The second part of the book spans five chapters and is primarily dedicated to applications of data mining. Topics in handwriting classification, text mining, web page ranking, word and sentence extraction, and face recognition are covered. The effects of different models and/or parameter choices are wonderfully illustrated and help the reader grasp the strengths and weaknesses of competing approaches. In many cases, the author has provided the small data matrix and MATLAB code so that the reader can reproduce the results illustrated. Students will find this feature of the book very useful.

The third and final part of the book is a subject-packed chapter on the algorithms for computing the various matrix decompositions used in the application section. MATLAB code examples and references to available software are provided. Particular attention to the differences in computing dense and sparse matrix decompositions is given.

This is undoubtedly an application-oriented book—only a limited number of mathematical proofs are included. The author intended to reveal the existence and properties of fundamental methods in linear algebra that are used in common data mining tasks. He has succeeded and I wager that both students and faculty will want to keep a copy of this book for their reference shelf long after the semester ends.

Michael W. Berry
University of Tennessee