and makes a short detour into the idea of noncomputable numbers.

The second classical topic deals with computational efficiency and with ways of measuring the difficulty of actually implementing an algorithm. The concepts of a difficulty function and equivalence classes of such functions are introduced, and it is shown that in this formulation there doesn't appear to be a way of defining an intrinsic difficulty of a problem. To facilitate a comparable argument when quantum mechanics is added, the author defines a "language for efficiency" rather different from the Turing machine context used in discussing computability.

After a section on probabilistic computing and a short section on finite-dimensional quantum mechanics, Geroch has the tools to discuss the impact of quantum mechanics on computability and computational efficiency. He begins with a detailed discussion of Grover's algorithm and indulges his original audience of graduate students by spending a little time discussing the theoretical feasibility of implementing the several steps of that algorithm. This is a nice summary of the algorithm, and it is presented in a way that permits its use as an example in the last five chapters that deal with quantum-assisted computability and quantum-assisted efficiency. It is in those five chapters that he discusses what is probably his motivating issue: while quantum-assisted computation appears to enhance computational efficiency, a formal proof is still lacking.

Geroch's book is an accessible introduction to some of the results and open questions involved in computability and computational efficiency, both with and without quantum computation. If one of his goals was to produce a treatise that would encourage physics graduate students to read further into the subject of quantum-assisted computation, he has succeeded admirably.

ARTHUR PITTENGER
University of Maryland

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

This is the fourth in the growing SIAM book series *Fundamentals of Algorithms* edited by Nick Higham. These books are short and fairly narrowly focused. Each presents algorithms for solving a few specific problems, together with the background material necessary for understanding the algorithms.

Eldén's book discusses five application areas in data mining and pattern recognition that are amenable to treatment by matrix methods. He states in the preface that his primary audience is undergraduate students who have already had a course in scientific computing or numerical analysis. I would add that such a prerequisite course had better have a substantial and pretty serious component on matrix computations.

The book has three parts, but the heart of the book is Part II, which consists of five chapters (10–14) and about 65 pages. Each chapter discusses one application.

Chapter 10 is on machine classification of handwritten digits, an important problem for postal services around the world. Each digit is represented as a grayscale image of 16×16 pixels. Thus, it is a vector in $\mathbb{R}^{256}$. Sets of training digits determine which regions of $\mathbb{R}^{256}$ are occupied by which digits. The singular value decomposition (SVD) is used as a tool to create a low-rank approximation and a low-dimensional subspace that captures the important information about each digit.

The effects of simple transformations, such as rotation, stretching, or thickening of a digit, are discussed. The notion of *tangent distance*, a method of measuring the distance between digits that is insensitive to small transformations, is introduced.

Chapter 11 is on text mining. Given a large set of documents, how does one determine which documents are relevant to a given query? To this end, a term-document matrix A is built with one row for each term and one column for each document. The (i, j) element is positive if and only if the ith term appears in the jth document. The value of a_{ij} can be weighted by the importance of the term and/or the document. Thus, each document is represented by a vector, a column of A. Queries are also documents and can be represented as vectors in the same way. A simple measure

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

of the relevance of a document to a given query is the cosine of the angle between the query vector and the document vector.

This measure of relevance is only modestly successful. Chapter 11 consists mainly of ways of improving on it. Each of these methods replaces the term-document matrix by a low-rank approximation in an attempt to capture the important information and discard the irrelevant details. The first is *latent semantic indexing*, which uses the SVD to create the low-rank approximation. Viable alternatives are based on clustering and on a nonnegative matrix factorization. Lanczos bidiagonalization, which builds a Krylov subspace starting from the query vector, is also considered. This method can give good results after a very small number of steps (very low-rank approximation), but becomes less effective if too many steps are taken.

Chapter 12, which is about ranking webpages, consists mainly of a description of Google's PageRank algorithm. For a Web search engine it does not suffice to identify all webpages that have been judged relevant to a given query. It must also make a decision about which of the many relevant pages are most important and therefore worthy of being recommended to the user. Google uses PageRank to help make this decision. The idea of PageRank is that a page is important if there are many links to it from other important pages. This circular definition of importance leads to a gigantic eigenvalue problem,

$$r = Qr,$$

where Q is the *Google matrix*, a positive, column-stochastic matrix determined by the link structure of the Web. Its dimension is $n \times n$, where n is the number of pages in the Web. Perron–Frobenius theory guarantees that the equation $r = Qr$ has an essentially unique positive solution r, which can be computed by the power method. The importance of the ith webpage is given by r_i, the ith component of r.

Chapter 13 is on automatic key word and key sentence extraction. Given a document, how can we decide what are the most important terms and the most important sentences in that document? Instead of a term-document matrix, we can build a term-sentence matrix A, in which each sentence is treated as a separate document. Each term and each sentence is given a *saliency score*, which is a measure of its importance. A sentence is considered important if it contains many important terms, and a term is considered important if it is contained in many important sentences. This circular definition leads to another eigenvalue problem, which is actually a singular value problem for the term-sentence matrix. The saliency values for the terms and sentences are the entries of the dominant left and right singular vectors of A, respectively.

Just as in Chapter 11, we can improve performance here by replacing the matrix A by a low-rank approximation, which can be done using the SVD, a clustering method, or a nonnegative matrix factorization.

Chapter 14 considers the question of facial recognition. It is difficult for a machine to match two photographs of the same face because they can be taken from different angles, in different lighting, and with different facial expressions. The method described in the book stores several images of each of several people as a third-order tensor. A tensor SVD (HOSVD) is used to analyze the data. It is safe to say that this application is the least well developed of the five applications discussed in the book. This would also make it the ripest for future development.

So far we have been talking about Part II of the book. Part I, which is about 110 pages long, is a minicourse in matrix computations, including such standard topics as linear systems and least squares problems, QR decomposition, and SVD. Examples pertaining to data mining and pattern recognition are included. Some nonstandard topics that are included are tensor decompositions, clustering, and nonnegative matrix factorization. The presentation is too compressed to be readable by a neophyte, but it could serve as a useful review and reinforcement for a reader who has already had a course in matrix computations.

Part III, which consists of a single chapter about 30 pages long, gives a brief overview of methods for computing matrix decompositions. Topics discussed include perturbation theory, the power method, reduc-

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

tion to tridiagonal form, the QR algorithm, SVD computation, and Arnoldi and Lanczos methods. Clearly the treatment is very condensed.

The book has an accompanying website that has theory questions (exercises), computer assignments, and more.

If you are planning to teach a course on data mining or applied matrix computations, consider using this book as a text or a supplement. Or pick it up and read it for your own edification, as I did.

DAVID S. WATKINS
Washington State University

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.