

Numerical and Statistical Methods for Bioengineering

This is the first MATLAB-based numerical methods textbook for bioengineers that uniquely integrates modeling concepts with statistical analysis, while maintaining a focus on enabling the user to report the error or uncertainty in their result. Between traditional numerical method topics of linear modeling concepts, non-linear root finding, and numerical integration, chapters on hypothesis testing, data regression, and probability are interweaved. A unique feature of the book is the inclusion of examples from clinical trials and bioinformatics, which are not found in other numerical methods textbooks for engineers. With a wealth of biomedical engineering examples, case studies on topical biomedical research, and the inclusion of end of chapter problems, this is a perfect core text for a one-semester undergraduate course.

Michael R. King is an Associate Professor of Biomedical Engineering at Cornell University. He is an expert on the receptor-mediated adhesion of circulating cells, and has developed new computational and in vitro models to study the function of leukocytes, platelets, stem, and cancer cells under flow. He has co-authored two books and received numerous awards, including the 2008 ICNMM Outstanding Researcher Award from the American Society of Mechanical Engineers, and received the 2009 Outstanding Contribution for a Publication in the International Journal *Clinical Chemistry*.

Nipa A. Mody is currently a postdoctoral research associate at Cornell University in the Department of Biomedical Engineering. She received her Ph.D. in Chemical Engineering from the University of Rochester in 2008 and has received a number of awards including a Ruth L. Kirschstein National Research Service Award (NRSA) from the NIH in 2005 and the Edward Peck Curtis Award for Excellence in Teaching from University of Rochester in 2004.

CAMBRIDGE TEXTS IN BIOMEDICAL ENGINEERING

Series Editors

W. MARK SALTZMAN, Yale University
SHU CHIEN, University of California, San Diego

Series Advisors

WILLIAM HENDEE, Medical College of Wisconsin
ROGER KAMM, Massachusetts Institute of Technology
ROBERT MALKIN, Duke University
ALISON NOBLE, Oxford University
BERNHARD PALSSON, University of California, San Diego
NICHOLAS PEPPAS, University of Texas at Austin
MICHAEL SEFTON, University of Toronto
GEORGE TRUSKEY, Duke University
CHENG ZHU, Georgia Institute of Technology

Cambridge Texts in Biomedical Engineering provides a forum for high-quality accessible textbooks targeted at undergraduate and graduate courses in biomedical engineering. It covers a broad range of biomedical engineering topics from introductory texts to advanced topics including, but not limited to, biomechanics, physiology, biomedical instrumentation, imaging, signals and systems, cell engineering, and bioinformatics. The series blends theory and practice, aimed primarily at biomedical engineering students, it also suits broader courses in engineering, the life sciences and medicine.

Numerical and Statistical Methods for Bioengineering

Applications in MATLAB

Michael R. King and Nipa A. Mody
Cornell University



CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

Cambridge University Press is part of the University of Cambridge.

It further the University's mission by disseminating knowledge in the pursuit of education, learning and research at the highest international levels of excellence.

www.cambridge.org

Information on this title: www.cambridge.org/9780521871587

© M. King and N. Mody 2011

This publication is in copyright. Subject to statutory exception and to the provisions of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published 2011

Reprinted 2015

Printed in the United Kingdom at the University Press, Cambridge

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

Library of Congress Cataloging in Publication data

King, Michael R.

Numerical and statistical methods for bioengineering : applications in MATLAB / Michael R. King, Nipa A. Mody.

p. cm. — (Cambridge texts in biomedical engineering)

ISBN 978-0-521-87158-7 (hardback)

1. Biomedical engineering — Statistical methods. 2. Biomedical engineering — Mathematics.

3. MATLAB. I. Mody, Nipa A. II. Title. III. Series.

R857.M34K56 2010

610.28—dc22

2010025717

ISBN 978-0-521-87158-7 Hardback

Additional resources for this publication at www.cambridge.org/9780521871587

Cambridge University Press has no responsibility for the persistence or accuracy of URLs for external or third-party internet websites referred to in this publication, and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

MATLAB® is a registered trademark of The MathWorks, Inc.

MATLAB® is a trademark of The MathWorks, Inc. and is used with permission. The MathWorks does not warrant the accuracy of the text or exercises in this book. This book's use or discussion of MATLAB® software or related products does not constitute endorsement or sponsorship by The MathWorks of a particular pedagogical approach or particular use of the MATLAB® software.

Contents

<i>Preface</i>	<i>page ix</i>
1 Types and sources of numerical error	1
1.1 Introduction	1
1.2 Representation of floating-point numbers	4
1.2.1 How computers store numbers	7
1.2.2 Binary to decimal system	7
1.2.3 Decimal to binary system	9
1.2.4 Binary representation of floating-point numbers	10
1.3 Methods used to measure error	16
1.4 Significant digits	18
1.5 Round-off errors generated by floating-point operations	20
1.6 Taylor series and truncation error	26
1.6.1 Order of magnitude estimation of truncation error	28
1.6.2 Convergence of a series	32
1.6.3 Finite difference formulas for numerical differentiation	33
1.7 Criteria for convergence	39
1.8 End of Chapter 1: key points to consider	40
1.9 Problems	40
References	46
2 Systems of linear equations	47
2.1 Introduction	47
2.2 Fundamentals of linear algebra	53
2.2.1 Vectors and matrices	53
2.2.2 Matrix operations	56
2.2.3 Vector and matrix norms	64
2.2.4 Linear combinations of vectors	66
2.2.5 Vector spaces and basis vectors	69
2.2.6 Rank, determinant, and inverse of matrices	71
2.3 Matrix representation of a system of linear equations	75
2.4 Gaussian elimination with backward substitution	76
2.4.1 Gaussian elimination without pivoting	76
2.4.2 Gaussian elimination with pivoting	84
2.5 LU factorization	87
2.5.1 LU factorization without pivoting	88
2.5.2 LU factorization with pivoting	93
2.5.3 The MATLAB lu function	95
2.6 The MATLAB backslash (\) operator	96
2.7 III-conditioned problems and the condition number	97
2.8 Linear regression	101

2.9	Curve fitting using linear least-squares approximation	107
2.9.1	The normal equations	109
2.9.2	Coefficient of determination and quality of fit	115
2.10	Linear least-squares approximation of transformed equations	118
2.11	Multivariable linear least-squares regression	123
2.12	The MATLAB function <code>polyfit</code>	124
2.13	End of Chapter 2: key points to consider	125
2.14	Problems	127
	References	139
3	Probability and statistics	141
3.1	Introduction	141
3.2	Characterizing a population: descriptive statistics	144
3.2.1	Measures of central tendency	145
3.2.2	Measures of dispersion	146
3.3	Concepts from probability	147
3.3.1	Random sampling and probability	149
3.3.2	Combinatorics: permutations and combinations	154
3.4	Discrete probability distributions	157
3.4.1	Binomial distribution	159
3.4.2	Poisson distribution	163
3.5	Normal distribution	166
3.5.1	Continuous probability distributions	167
3.5.2	Normal probability density	169
3.5.3	Expectations of sample-derived statistics	171
3.5.4	Standard normal distribution and the z statistic	175
3.5.5	Confidence intervals using the z statistic and the t statistic	177
3.5.6	Non-normal samples and the central-limit theorem	183
3.6	Propagation of error	186
3.6.1	Addition/subtraction of random variables	187
3.6.2	Multiplication/division of random variables	188
3.6.3	General functional relationship between two random variables	190
3.7	Linear regression error	191
3.7.1	Error in model parameters	193
3.7.2	Error in model predictions	196
3.8	End of Chapter 3: key points to consider	199
3.9	Problems	202
	References	208
4	Hypothesis testing	209
4.1	Introduction	209
4.2	Formulating a hypothesis	210
4.2.1	Designing a scientific study	211
4.2.2	Null and alternate hypotheses	217
4.3	Testing a hypothesis	219
4.3.1	The p value and assessing statistical significance	220
4.3.2	Type I and type II errors	226
4.3.3	Types of variables	228
4.3.4	Choosing a hypothesis test	230

vii	Contents	
	4.4 Parametric tests and assessing normality	231
	4.5 The z test	235
	4.5.1 One-sample z test	235
	4.5.2 Two-sample z test	241
	4.6 The t test	244
	4.6.1 One-sample and paired sample t tests	244
	4.6.2 Independent two-sample t test	249
	4.7 Hypothesis testing for population proportions	251
	4.7.1 Hypothesis testing for a single population proportion	256
	4.7.2 Hypothesis testing for two population proportions	257
	4.8 One-way ANOVA	260
	4.9 Chi-square tests for nominal scale data	274
	4.9.1 Goodness-of-fit test	276
	4.9.2 Test of independence	281
	4.9.3 Test of homogeneity	285
	4.10 More on non-parametric (distribution-free) tests	288
	4.10.1 Sign test	289
	4.10.2 Wilcoxon signed-rank test	292
	4.10.3 Wilcoxon rank-sum test	296
	4.11 End of Chapter 4: key points to consider	299
	4.12 Problems	299
	References	308
5	Root-finding techniques for nonlinear equations	310
	5.1 Introduction	310
	5.2 Bisection method	312
	5.3 Regula-falsi method	319
	5.4 Fixed-point iteration	320
	5.5 Newton's method	327
	5.5.1 Convergence issues	329
	5.6 Secant method	336
	5.7 Solving systems of nonlinear equations	338
	5.8 MATLAB function <code>fzero</code>	346
	5.9 End of Chapter 5: key points to consider	348
	5.10 Problems	349
	References	353
6	Numerical quadrature	354
	6.1 Introduction	354
	6.2 Polynomial interpolation	361
	6.3 Newton–Cotes formulas	371
	6.3.1 Trapezoidal rule	372
	6.3.2 Simpson's 1/3 rule	380
	6.3.3 Simpson's 3/8 rule	384
	6.4 Richardson's extrapolation and Romberg integration	387
	6.5 Gaussian quadrature	391
	6.6 End of Chapter 6: key points to consider	402
	6.7 Problems	403
	References	408

7	Numerical integration of ordinary differential equations	409
7.1	Introduction	409
7.2	Euler’s methods	416
7.2.1	Euler’s forward method	417
7.2.2	Euler’s backward method	428
7.2.3	Modified Euler’s method	431
7.3	Runge–Kutta (RK) methods	434
7.3.1	Second-order RK methods	434
7.3.2	Fourth-order RK methods	438
7.4	Adaptive step size methods	440
7.5	Multistep ODE solvers	451
7.5.1	Adams methods	452
7.5.2	Predictor–corrector methods	454
7.6	Stability and stiff equations	456
7.7	Shooting method for boundary-value problems	461
7.7.1	Linear ODEs	463
7.7.2	Nonlinear ODEs	464
7.8	End of Chapter 7: key points to consider	472
7.9	Problems	473
	References	478
8	Nonlinear model regression and optimization	480
8.1	Introduction	480
8.2	Unconstrained single-variable optimization	487
8.2.1	Newton’s method	488
8.2.2	Successive parabolic interpolation	492
8.2.3	Golden section search method	495
8.3	Unconstrained multivariable optimization	500
8.3.1	Steepest descent or gradient method	502
8.3.2	Multidimensional Newton’s method	509
8.3.3	Simplex method	513
8.4	Constrained nonlinear optimization	523
8.5	Nonlinear error analysis	530
8.6	End of Chapter 8: key points to consider	533
8.7	Problems	534
	References	538
9	Basic algorithms of bioinformatics	539
9.1	Introduction	539
9.2	Sequence alignment and database searches	540
9.3	Phylogenetic trees using distance-based methods	554
9.4	End of Chapter 9: key points to consider	557
9.5	Problems	558
	References	558
	<i>Appendix A Introduction to MATLAB</i>	560
	<i>Appendix B Location of nodes for Gauss–Legendre quadrature</i>	576
	<i>Index for MATLAB commands</i>	578
	<i>Index</i>	579

Preface

Biomedical engineering programs have exploded in popularity and number over the past 20 years. In many programs, the fundamentals of engineering science are taught from textbooks borrowed from other, more traditional, engineering fields: statics, transport phenomena, circuits. Other courses in the biomedical engineering curriculum are so multidisciplinary (think of tissue engineering, Introduction to BME) that this approach does not apply; fortunately, excellent new textbooks have recently emerged on these topics. On the surface, numerical and statistical methods would seem to fall into this first category, and likely explains why biomedical engineers have not yet contributed textbooks on this subject. I mean ... math is math, right? Well, not exactly.

There exist some unique aspects of biomedical engineering relevant to numerical analysis. Graduate research in biomedical engineering is more often *hypothesis driven*, compared to research in other engineering disciplines. Similarly, biomedical engineers in industry design, test, and produce medical devices, instruments, and drugs, and thus must concern themselves with human clinical trials and gaining approval from regulatory agencies such as the US Food & Drug Administration. As a result, statistics and hypothesis testing play a bigger role in biomedical engineering and must be taught at the curricular level. This increased emphasis on statistical analysis is reflected in special “program criteria” established for biomedical engineering degree programs by the Accreditation Board for Engineering and Technology (ABET) in the USA.

There are many general textbooks on numerical methods available for undergraduate and graduate students in engineering; some of these use MATLAB as the teaching platform. A good undergraduate text along these lines is *Numerical Methods with Matlab* by G. Recktenwald, and a good graduate-level reference on numerical methods is the well-known *Numerical Recipes* by W. H. Press *et al.* These texts do a good job of covering topics such as programming basics, nonlinear root finding, systems of linear equations, least-squares curve fitting, and numerical integration, but tend not to devote much space to statistics and hypothesis testing. Certainly, topics such as genomic data and design of clinical trials are not covered. But beyond the basic numerical algorithms that may be common to engineering and the physical sciences, one thing an instructor learns is that *biomedical engineering students want to work on biomedical problems!* This requires a biomedical engineering instructor to supplement a general numerical methods textbook with a gathering of relevant lecture examples, homework, and exam problems, a labor-intensive task to be sure and one that may leave students confused and unsatisfied with their textbook investment. This book is designed to fill an unmet need, by providing a complete numerical and statistical methods textbook, tailored to the unique requirements of the modern BME curriculum and implemented in MATLAB, which is inundated with examples drawn from across the spectrum of biomedical science.

This book is designed to serve as the primary textbook for a one-semester course in numerical and statistical methods for biomedical engineering students. The level of the book is appropriate for sophomore year through first year of graduate studies, depending on the pace of the course and the number of advanced, optional topics that are covered. A course based on this book, together with later opportunities for implementation in a laboratory course or senior design project, is intended to fulfil the statistics and hypothesis testing requirements of the program criteria established by ABET, and served this purpose at the University of Rochester. The material within this book formed the basis for the required junior-level course “Biomedical computation,” offered at the University of Rochester from 2002 to 2008. As of Fall 2009, an accelerated version of the “Biomedical computation” course is now offered at the masters level at Cornell University. It is recommended that students have previously taken calculus and an introductory programming course; a semester of linear algebra is helpful but not required. It is our hope that this book will also serve as a valuable reference for bioengineering practitioners and other researchers working in quantitative branches of the life sciences such as biophysics and physiology.

Format

As with most textbooks, the chapters have been organized so that concepts are progressively built upon as the reader advances from one chapter to the next. Chapters 1 and 2 develop basic concepts, such as types of errors, linear algebra concepts, linear problems, and linear regression, that are referred to in later chapters. Chapters 3 (Probability and statistics) and 5 (Nonlinear root-finding techniques) draw upon the material covered in Chapters 1 and 2. Chapter 4 (Hypothesis testing) exclusively draws upon the material covered in Chapter 3, and can be covered at any point after Chapter 3 (Sections 3.1 to 3.5) is completed. The material on linear regression error in Chapter 3 should precede the coverage of Chapter 8 (Nonlinear model regression and optimization). The following chapter order is strongly recommended to provide a seamless transition from one topic to the next:

Chapter 1 → Chapter 2 → Chapter 3 → Chapter 5 → Chapter 6 → Chapter 8.

Chapter 4 can be covered at any time once the first three chapters are completed, while Chapter 7 can be covered at any time after working through Chapters 1, 2, 3, and 5. Chapter 9 covers an elective topic that can be taken up at any time during a course of study.

The examples provided in each chapter are of two types: Examples and Boxes. The problems presented in the Examples are more straightforward and the equations simpler. Examples either illustrate concepts already introduced in earlier sections or are used to present new concepts. They are relatively quick to work through compared to the Boxes since little additional background information is needed to understand the example problems. The Boxes discuss biomedical research, clinical, or industrial problems and include an explanation of relevant biology or engineering concepts to present the nature of the problem. In a majority of the Boxes, the equations to be solved numerically are derived from first principles to provide a more complete understanding of the problem. The problems covered in Boxes can be more challenging and require more involvement by

the reader. While the Examples are critical in mastering the text material, the choice of which boxed problems to focus on is left to the instructor or reader.

As a recurring theme of this book, we illustrate the implementation of numerical methods through programming with the technical software package MATLAB. Previous experience with MATLAB is not necessary to follow and understand this book, although some prior programming knowledge is recommended. The best way to learn how to program in a new language is to jump right into coding when a need presents itself. Sophistication of the code is increased gradually in successive chapters. New commands and programming concepts are introduced on a need-to-know basis. Readers who are unfamiliar with MATLAB should first study Appendix A, Introduction to MATLAB, to orient themselves with the MATLAB programming environment and to learn the basic commands and programming terminology. Examples and Boxed problems are accompanied by the MATLAB code containing the numerical algorithm to solve the numerical or statistical problem. The MATLAB programs presented throughout the book illustrate code writing practice. We show two ways to use MATLAB as a tool for solving numerical problems: (1) by developing a program (m-file) that contains the numerical algorithm, and (2) using built-in functions supplied by MATLAB to solve a problem numerically. While self-written numerical algorithms coded in MATLAB are instructive for teaching, MATLAB built-in functions that compute the numerical solution can be more efficient and robust for use in practice. The reader is taught to integrate MATLAB functions into their written code to solve a specific problem (e.g. the backslash operator).

The book has its own website hosted by Cambridge University Press at www.cambridge.org/kingmody. All of the m-files and numerical data sets within this book can be found at this website, along with additional MATLAB programs relevant to the topics and problems covered in the text.

Acknowledgements

First and foremost, M. R. K. owes a debt of gratitude to Professor David Leighton at the University of Notre Dame. The “Biomedical computation” course that led to this book was closely inspired by Leighton’s “Computer methods for chemical engineers” course at Notre Dame. I (Michael King) had the good fortune of serving as a teaching assistant and later as a graduate instructor for that course, and it shaped the format and style of my own teaching on this subject. I would also like to thank former students, teaching assistants, and faculty colleagues at the University of Rochester and Cornell University; over the years, their valuable input has helped continually to improve the material that comprises this book. I thank my wife and colleague Cindy Reinhart-King for her constant support. Finally, I thank my co-author, friend, and former student Nipa Mody: without her tireless efforts this book would not exist.

N.A.M. would like to acknowledge the timely and helpful advice on creating bioengineering examples from the following business and medical professionals: Khyati Desai, Shimoni Shah, Shital Modi, and Pinky Shah. I (Nipa Mody) also thank Banu Sankaran and Ajay Sadarangani for their general feedback on parts of the book. I am indebted to the support provided by the following faculty and staff members of the Biomedical Engineering Department of the University of Rochester: Professor Richard E. Waugh, Donna Porcelli, Nancy Gronska, Mary

Gilmore, and Gayle Hurlbutt, in this endeavor. I very much appreciate the valiant support of my husband, Anand Mody, while I embarked on the formidable task of book writing.

We thank those people who have critically read versions of this manuscript, in particular, Ayotunde Oluwakorede Ositelu, Aram Chung, and Bryce Allio, and also to our reviewers for their excellent recommendations. We express many thanks to Michelle Carey, Sarah Matthews, Christopher Miller, and Irene Pizzie at Cambridge University Press for their help in the planning and execution of this book project and for their patience.

If readers wish to suggest additional topics or comments, please write to us. We welcome all comments and criticisms as this book (and the field of biomedical engineering) continue to evolve.