SIAM Review Vol. 62, Issue 4 (December 2020)

Book Reviews

BOOK REVIEWS

The first and featured review is on the book *Introduction to Numerical Methods for Variational Problems*, by Hans Petter Langtangen and Kent-Andre Mardal. It is one of the last books co-authored by Hans Petter Langtangen, who was a very influential, enthusiastic researcher and teacher, and nevertheless a very kind person. I met him several times during my time on the editorial board of SISC, where he used to be the editor-in-chief. The scientific community in the field of scientific computing owes a lot to him. The reviewer of the book, Akil Narayan, has written a very careful and diligent review on this book pointing out its nonstandard features, which let it stand out from the crowd and make it very worthwhile reading.

The featured review is followed by Daniele Boffi's review on the book *Convection-Diffusion Problems. An Introduction to Their Analysis and Numerical Solution*, by Martin Stynes and David Stynes. Daniele recommends the book as a "concise and well-organized introduction to the approximation of convection-diffusion problems." Among our reviews, we have three more books related to data science: *Neural Networks and Statistical Learning*, by Ke-Lin Du and M. N. S. Swamy, which is reviewed by Jan Pablo Burgard with the comment that the book "can be seen as a central reference point for the mathematical understanding and implementation of the core ideas of neuronal networks and statistical learning techniques"; John Harlim's *Data-Driven Computational Methods: Parameter and Operator Estimations*, very positively reviewed by Nikolas Kantas; and the book *Mathematics for Machine Learning*, by Marc Peter Deisenroth, A. Aldo Faisal, and Cheng Soon Og.

Finally, I would like to mention the book by Doug Arnold on *Finite Element Exterior Calculus*, which is reviewed by the expert Ralf Hiptmair, who praises the book as "a work bridging the divide sometimes separating what is labeled pure and applied mathematics."

Volker H. Schulz
Section Editor
sirev.br@gmail.com

SIAM REVIEW
Vol. 62, No. 4, pp. 985–994

Book Reviews

Edited by Volker H. Schulz

Featured Review: Introduction to Numerical Methods for Variational Problems. *By Hans Petter Langtangen and Kent-Andre Mardal.* Springer, Cham, 2019. $79.99. xvi+395 pp., hardcover. ISBN 978-3-030-23787-5.

Computational methods are enormously useful tools in many disciplines, but mastery of the mathematical and computational nuances of various aspects of these tools can be a daunting task. The authors of *Numerical Methods for Variational Problems* recognize that this lack of accessibility can dissuade practitioners from using these tools, hence the motivation for writing a practical text that is understandable for nonexperts who wish both to gain competence in algorithms for solving scientific computing problems and to understand the overall mathematical ideas without becoming overwhelmed with technicalities.

The result is an impressively broad tour of the anatomy of algorithms for function approximation, numerical solutions to partial differential equations (PDEs), nonlinear optimization, and iterative methods for linear systems of equations. The main focus is on finite element methods (FEMs) for linear stationary PDEs, which the text uses as a vehicle for introducing numerous concepts in scientific computing and functional and numerical analysis. The eponymous variational problems theme pervades most chapters in the book, and the authors pay particular attention to crafting several of the discussed algorithms from variational forms or ideas.

The appropriate audience for this textbook is scientists or scientists-in-training who are looking to gain a survey-like understanding of scientific computing. The main practical outcome for these readers will be the ability to use modern software tools to compute numerical solutions to fairly nontrivial multidimensional PDEs using FEMs. Some knowledge of basic numerical analysis and differential equations is expected. For example, the authors point out that some fundamental knowledge of finite-difference methods is assumed in order to communicate some ideas and frame the narrative of FEMs. Advanced undergraduates and graduate students in the physical sciences and engineering will find this text particularly helpful. With heavy focus on explicit computation and algorithmic implementation, there is very little "theorem-proof" style mathematics or presentation. In several places the authors highlight the lack of rigor in their presentation in order to provide enough keywords and breadcrumbs for interested readers to delve deeper by consulting other references. The authors point out that this text is intended to replace neither formal mathematical treatment of FEMs, such as [2], nor more algorithmic hands-on approaches such as [4]. Instead, the book positions itself as a practical introduction to numerical methods that contains some discussion of mathematical formalities.

One of the major strengths of the text is its focus on software implementation. This is apparent both in the main text, where software code is included in the text

Publishers are invited to send books for review to Book Reviews Editor, SIAM, 3600 *Market St.,* 6^{th} *Floor, Philadelphia, PA* 19104-2688.

for demonstration, and in the end-of-chapter exercises that practice implementation aspects of algorithms. Of course, implementing FEM solvers itself requires substantial coding, but the authors mitigate this learning curve by leaning on the existing and well-documented FEniCS software suite [1]. In particular, the authors utilize the freely available programming language Python as the scripting tool for all demonstrations. The code in the book uses high-level Python interfaces provided by the FEniCS software. Using high-level FEniCS routines runs the risk of obfuscating some important details from the reader, but the authors ensure that nearly all of the technical issues required in FEM solvers are presented in the text for at least simple one-dimensional PDEs. Thus, the reader is left with both a feeling of ownership of software results that they have produced as well as a strong technical knowledge of the inner workings of the software.

Another pedagogical focus of the book is on explicit symbolic computation. FEM formulations often use bilinear forms to describe the numerical scheme. However, readers without a strong mathematical background and understanding of variational methods will find this confusing at first. To demystify this and other mathematical notation, the authors place an emphasis on explicit symbolic computation for special cases. As the authors state, this exercise, sometimes tedious for those with a deeper understanding, provides a necessary crutch for readers who may be unfamiliar with notational technicalities. This theme of explicit computation is followed throughout the book, with explicit computations of least squares approximations of simple functions and of entries in some FEM mass and stiffness matrices. The authors also recognize that such algebraic manipulations can become cumbersome to accomplish by hand, so they often appeal to symbolic computing tools. With their choice of language being Python, the authors utilize the SymPy package [6] as their workhorse for demonstration of analytical computations. All code for in-text demonstrations and end-of-chapter exercises is included in an online, freely available repository [5].

The technical contents of the book are well summarized in Chapter 1, where a distillation of the book's goals is presented and motivated through a whirlwind introduction to the FEM. The authors then narrow their focus in Chapters 2 and 3, where the basics of function approximation with numerical method is introduced. Here the theme of variational methods begins with a discussion and comparison of Galerkin and least squares formulations. Lessons from numerical analysis, such as the Runge Phenomenon, errors due to floating-point arithmetic, and computational complexity, are introduced and discussed on an as-needed basis. Even in these early chapters, the authors call attention to mathematical and computational extensions to multivariate approximation problems.

The technical core of the book begins in Chapters 4 and 5, where variational methods for solving differential equations are formulated as weighted residual methods. To their credit, the authors assume very little prior scientific computing experience from the reader, and take effort to explain various differential equation models and identify the desiderata that practitioners place on numerical methods for solving PDEs. These chapters introduce foundational concepts familiar to experienced FEM practitioners: test and trial spaces, Galerkin and collocation methods, and natural and essential boundary conditions. The authors use this jargon sparingly, sometimes simply in passing to inform readers; the focus is on numerous examples, the actual computations, and relevance to solving PDEs. With a singularly perturbed convection-diffusion problem, the authors even employ the demonstration of failure of the presented numerical schemes as a pedagogical tool. Chapter 5 in particular delves into the computational details of one-dimensional FEM solvers, with a discussion about sparse matrices and the computational savings they deliver, and a fairly detailed discretized formulation for several kinds of boundary conditions. The chapter ends with demonstrations and exercises for two-dimensional FEM simulations, immediately communicating to the reader the generality of the intuition they have gained in one dimension.

Chapters 6–9 take the reader through more advanced topics for FEMs: time-dependent problems, systems of PDEs, penalty methods for boundary conditions, and nonlinear PDEs. Throughout, the theme of examples and explicit computation, with accompanying code, is practiced and emphasized. This approach reveals the difficulties of these problems to readers and discusses potential solutions without onerous notational burden. Chapter 10 ends the technical narrative of the book with a discussion of iterative methods for linear systems from a variational formulation point of view. Such a presentation is not standard in many numerical linear algebra texts, but its articulation here provides yet another demonstration to the audience of the utility of variational methods in numerical computing.

This book was published after the passing of the first author, Hans Petter Langtangen, who is a well-known and respected figure in computational science and applied mathematics [3]. Through his dedication to seeing this text reach completion, Kent-Andre Mardal has ensured that Hans Petter Langtangen's spirit and passion for computational science, education, and interdisciplinary collaboration continues to manifest through a text that is lively, conversational, informative, approachable, and interactive.

REFERENCES

[1] M. Alnæs, J. Blechta, J. Hake, A. Johansson, B. Kehlet, A. Logg, C. Richardson, J. Ring, M. E. Rognes, and G. N. Wells, *The FEniCS Project Version* 1.5, Arch. Numer. Software, 3 (2015), no. 100.

[2] S. Brenner and R. Scott, *The Mathematical Theory of Finite Element Methods*, 3rd ed., Springer, 2007.

[3] X. Cai, J. S. Hesthaven, M. Peters, M. E. Rognes, and A. Tveito, *Obituaries: Hans Petter Langtangen*, SIAM News, March 2017.

[4] H. P. Langtangen, *Computational Partial Differential Equations: Numerical Methods and Diffpack Programming*, 2nd ed., Springer-Verlag, 2003.

[5] H. P. Langtangen and K.-A. Mardal, *Resources for the book Introduction to Numerical Methods for Variational Problems*, http://hplgit.github.io/fem-book/doc/web/.

[6] A. Meurer et al., *SymPy: Symbolic computing in Python*, PeerJ Comput. Sci., 3 (2017), e103.

Akil Narayan
University of Utah

Convection-Diffusion Problems. An Introduction to Their Analysis and Numerical Solution. *By M. Stynes and D. Stynes.* American Mathematical Society, Providence, RI, 2018. $38.00. viii+156 pp., hardcover. ISBN 978-1-470-44868-4.

Have you ever looked for a concise introduction to the theory and the approximation of convection-diffusion problems? This 156-page book (including preface and index) might be precisely what you need.

The book is an extended and updated version of [1]: exercises and other material have been added in order to attract numerical analysts who have not worked on this subject before. Advanced readers will find the contents too basic, as this book is not intended for them (although they might find its introductory and tutorial style useful for the preparation of a course on this subject); if you are looking for a classical reference you are redirected by the authors to [2, 3].

The text is a well-organized version of the lecture notes for a course taught by Martin Stynes during Summer 2015. In this sense, the material is very well balanced and the covered topics are mostly self-contained.

The mission of the book is clearly described in the preface: "For many years I felt that an easier, more introductory book was needed to encourage new people to enter our fascinating research area." I found the contents and the level perfectly suited for this goal.

After a motivating example presented in Chapter 1, where some fundamental concepts are introduced, including the notion of boundary layer, maximum principle, and barrier functions, Chapter 2 deals with convection-diffusion problems in one dimension. The model equation is the boundary value problem

$$\begin{cases} -\varepsilon u''(x) + a(x)u'(x) + b(x)u(x) = f(x) \\ \qquad\qquad \text{for } x \in (0,1), \\ u(0) = g_0, \\ u(1) = g_1, \end{cases}$$

where ε is a positive constant not larger than 1, the convection coefficient $a(x)$ is bounded below by a positive constant, and the reaction coefficient $b(x)$ is nonnegative.

The finite difference approximation of the one-dimensional model is presented in Chapter 3, which is the main part of the book. Here the authors present the basic ideas related to upwinding, artificial diffusion, and Shishkin meshes.

The extension to two dimensions is presented in Chapter 4, where the following problem is studied:

$$\begin{cases} -\varepsilon \Delta u(x,y) + \mathbf{a}(x,y) \cdot \operatorname{grad} u(x,y) \\ \quad + b(x,y)u(x,y) = f(x,y) \text{ for } x \in \Omega \subset \mathbb{R}^2, \\ u(x,y) = g(x,y) \text{ on } \partial\Omega. \end{cases}$$

Some interesting examples are discussed, together with a survey of various a priori estimates. The section "General comments on numerical methods" describes the essence of stabilization for beginners.

The final chapter extends the discussion to the finite element method. Without pretending to be exhaustive, the text covers the Petrov–Galerkin approach, the SUPG stabilization, higher-order methods, Shishkin meshes, discontinuous Galerkin methods, and the use of adaptive meshes. While some basic analysis is presented, the aim of this section is to provide the reader with pointers to possible choices for the approximation of the problem under consideration.

In conclusion, this book is easy to read and represents a concise and well-organized introduction to the approximation of convection-diffusion problems. It can certainly be used as a textbook for a short course on the subject; the presence of several examples and exercises is definitely an added value that makes the book appealing for students and researchers who have not yet worked on singularly perturbed problems.

REFERENCES

[1] M. Stynes, *Steady-state convection-diffusion problems*, Acta Numer., 14 (2005), pp. 445–508.

[2] H.-G. Roos, M. Stynes, and L. Tobiska, *Numerical Methods for Singularly Perturbed Differential Equations: Convection-Diffusion and Flow Problems*, Springer Ser. Comput. Math. 24, Springer-Verlag, Berlin, 1996.

[3] H.-G. Roos, M. Stynes, and L. Tobiska, *Robust Numerical Methods for Singularly Perturbed Differential Equations: Convection-Diffusion-Reaction and Flow Problems*, 2nd ed., Springer Ser. Comput. Math. 24, Springer-Verlag, Berlin, 2008.

Daniele Boffi
King Abdullah University of Science and Technology, and
University of Pavia

Neural Networks and Statistical Learning. Second Edition. *By Ke-Lin Du and M. N. S. Swamy.* Springer, London, 2019. \$159.99. xxx+988 pp., hardcover. ISBN 978-1-4471-7451-6.

Since the 1990s the use of neuronal networks and statistical learning techniques has grown extraordinarily. The research field around these techniques is vivid and continuously evolving, making a thorough overview within one book nearly impossible. However, Ke-Lin Du and M. N. S. Swamy, in their book *Neural Networks and Statistical Learning*, discuss a very broad and substantial subset of this vast field.

This second edition extends the first edition by six chapters, yielding in total 31 chapters and two appendices. Each chapter ends with some well-chosen problems that readers can use to self-assess their understanding of the material. Further, all references used within a given chapter are given directly at the chapter's end. This makes the chapters useful as standalone reading material.

The first three chapters cover basic vocabulary and concepts for the understanding of the tackled problem classes. These are very helpful to laying the foundation needed for the subsequent chapters for students and researchers coming from other research fields. Then, step by step, the theory behind neuronal networks is developed.

In Chapter 4 the most basic concepts—the one-neuron perceptron and the single-layer perceptron—are introduced. Subsequently, in Chapters 5 and 6, different kinds of multilayer perceptrons and their parameter estimation are presented. In particular, the very important back-propagation algorithm is described in detail, and two major approaches—incremental learning and batch learning—are compared. Further, in Chapter 24 the concept of deep learning is introduced. In deep learning a multilayer network with hidden layers is constructed, resembling the hierarchical structure of the human neocortex. As a method to incorporate human experience into data analysis, fuzzy sets are introduced in Chapter 26.

Chapters 7 and 8 deal with associative memory networks, with a major focus on the different extensions of the Hopfield models. It is also discussed how Hopfield networks may be used for optimization purposes themselves in different optimization fields. Different recurrent neural networks are provided in Chapter 12 which are generalizations of the Hopfield models. Additionally, in Chapter 23 the Boltzmann machine is introduced briefly as a further generalization. All these networks are useful when trying to approximate dynamical systems.

The two major chapters on clustering, Chapters 9 and 10, encompass a great variety of clustering models and algorithms such as self-organizing maps, the nearest neighbor algorithm, and C-means clustering. It is also discussed how to assess the number of clusters in a data set, as in most cases this has to be determined by the user. Besides these two chapters, further chapters also treat methods that are able to perform clustering. For example, in Chapter 21 support vector machines (SVMs), which are quite popular methods, especially in big data applications, are discussed. In SVMs the kernel method is very much used and thus discussed in the previous Chapter 20 in detail. Also radial basis function (RBF) networks are discussed in Chapter 11. Besides clustering, these last two methods are also explained in the context of approximating functions in a high-dimensional setting.

In Chapters 13, 15, and 16 typical methods used in multivariate statistics are proposed. Besides the well-known principal component analysis (PCA), independent component analysis and discriminant analysis are also discussed. While PCA describes the data in a lower rank dimensional space, the matrix completion method discussed in Chapter 19 uses the lower rank approximation to fill in gaps in a higher rank matrix. Perhaps most prominent problem that can be solved by this method is the Netflix problem.

The methods introduced so far can be classified into supervised (the output is labeled) and unsupervised (the output is not labeled) learning. A third basic learning method is reinforcement learning, which is introduced in Chapter 17. This chapter provides a good understanding of the central learning concepts.

Two chapters encompass methods to improve the previously proposed techniques. In Chapter 25 different types of ensemble learning are described. In ensemble learning multiple weak learners are combined into a strong learner. In Chapter 28 software- and hardware-related implementation issues for neural networks are discussed.

The last big block of chapters encompasses different fields of applications. Besides pattern recognition, data mining and big data applications are also considered.

All in all the book finds a good balance in giving a broad overview of the vast research field of neuronal networks and statistical learning techniques and providing detailed fundamental descriptions of the central methods and algorithms.

With the rapid increase in demand for skilled data scientists, statistical learning and especially neuronal networks are becoming more and more relevant. This can be seen, e.g., by the immense amount of master programs oftentimes entitled "Data Science." A quick search on https://www.mastersportal.com yielded over 3400 M.Sc. Programs entitled or including the string "Data Science" in the description. Most of them are situated in North America and Europe. *Neural Networks and Statistical Learning* by Ke-Lin Du and M. N. S. Swamy can be seen as a central reference point for the mathematical understanding and implementation of the core ideas of neuronal networks and statistical learning techniques.

JAN PABLO BURGARD
Trier University

Finite Element Exterior Calculus. *By Douglas N. Arnold.* SIAM, Philadelphia, PA, 2018. xii+120 pp. CBMS-NSF Regional Conference Series in Applied Mathematics. Vol. 93. ISBN 978-1-611975-53-6. https://doi.org/10.1137/1.9781611975543.

Open a classical textbook on the design and mathematical analysis of finite-element methods and you get the impression that the central and almost sole concern is the estimation of approximation errors, of consistency errors, and of norms in function spaces in general. The strong emphasis put on these aspects is partly due to focus in these texts on those simple scalar second-order elliptic boundary value problems.

If finite-element theory had its roots in the field of computational electromagnetics, I guess most textbooks would resemble D. N. Arnold's book *Finite Element Exterior Calculus*. The reason is that the laws of electromagnetism as expressed by Maxwell's famous equations

$$\mathbf{curl}\,\mathbf{E} = -\partial_t(\boldsymbol{\mu}\mathbf{H}), \quad \mathbf{curl}\,\mathbf{H} = \partial_t(\boldsymbol{\epsilon}\mathbf{E}) + \mathbf{j}$$

reflect profound algebraic structure alongside variational principles. The algebraic structure manifests itself in constraints like $\partial_t \operatorname{div} \mathbf{B} = 0$ and the possibility to express the fields through potentials. All this we owe to

(1)
$$C^\infty(\Omega) \xrightarrow{\text{grad}} (C^\infty(\Omega))^3 \xrightarrow{\mathbf{curl}} (C^\infty(\Omega))^3 \xrightarrow{\text{div}} C^\infty(\Omega)$$

being an *exact sequence* provided that the domain $\Omega \subset \mathbb{R}^3$ has vanishing higher *Betti numbers*. "Exact sequence," "Betti number"—these are core concepts of *algebra* and *topology* and, suddenly, they have become relevant to understanding Maxwell's equations.

More precisely, (1) is the so-called vector-proxy representation of one specimen of a deRham complex. In differential geometry, (1) is written as a relationship between spaces $C^\infty\Lambda^\ell(\Omega)$ of smooth differential ℓ-forms on Ω:

(2)
$$C^\infty\Lambda^0(\Omega) \xrightarrow{d} C^\infty\Lambda^1(\Omega) \xrightarrow{d} C^\infty\Lambda^2(\Omega) \xrightarrow{d} C^\infty\Lambda^3(\Omega)\,.$$

This is (1) in the language of *exterior calculus*, the calculus of differential forms, and in this framework (1) can be generalized to manifolds of any dimension.

The elegance and power of exterior calculus in electromagnetic modeling had long been realized in theoretical physics. Conversely, in computational electromagnetics this perspective was adopted only sluggishly, starting with the groundbreaking works of A. Bossavit. From there, it seeped into numerical analysis, spurred by efforts to understand and remedy the mystery of spurious solutions haunting conventional finite-element methods. The discovery or invention—however one may prefer to view it—of finite-element exterior calculus (FEEC) by D. N. Arnold and his long-time collaborators R. Falk and R. Winther marks the culmination of this development. It also marks another success of the fundamental principle that the preservation of essential algebraic structure in the process of discretization is necessary for achieving stability, accuracy, and robustness of the resulting discrete model. In my opinion, this sentence also wraps up the chief philosophical message of Arnold's book.

Fittingly, the book starts by elaborating the theory of chain complexes and homology, the algebraic tools for describing the structure underlying (2). Then it switches to functional analysis, the study of unbounded operators, before moving on to the ultimate connection of both worlds in the theory of Hilbert complexes, which puts (2) in a framework of Sobolev spaces and offers the right setting for numerical analysis. As a Hilbert complex of vector proxies (1) reads

(3)
$$H^1(\Omega) \xrightarrow{\text{grad}} \boldsymbol{H}(\mathbf{curl},\Omega) \xrightarrow{\mathbf{curl}} \boldsymbol{H}(\mathrm{div},\Omega) \xrightarrow{\mathrm{div}} L^2(\Omega)\,.$$

In that setting variational formulations of boundary value problems are natural, paving the way for an abstract treatment of stability and convergence of Galerkin methods based on the key concept of bounded cochain projections π_h^ℓ. Concretely, for the deRham complex in three-dimensional Euclidean space, we want

(4)
$$\begin{array}{ccccccc}
H^1(\Omega) & \xrightarrow{\text{grad}} & \boldsymbol{H}(\mathbf{curl},\Omega) & \xrightarrow{\mathbf{curl}} & \boldsymbol{H}(\mathrm{div},\Omega) & \xrightarrow{\mathrm{div}} & L^2(\Omega) \\
\downarrow \pi_h^0 & & \downarrow \pi_h^1 & & \downarrow \pi_h^2 & & \downarrow \pi_h^3 \\
V_h^0 & \xrightarrow{\text{grad}} & V_h^1 & \xrightarrow{\mathbf{curl}} & V_h^2 & \xrightarrow{\mathrm{div}} & V_h^3
\end{array}$$

to commute, the π_h^ℓ to supply continuous projections onto subspaces V_h^ℓ of the respective function spaces. Afterwards, the constructions of spaces V_h^ℓ of piecewise polynomial discrete differential forms on simplicial meshes is discussed in detail. Again, this relies on the judicious use of techniques from (linear) algebra. The book closes with a review of special extensions of FEEC and of developments beyond exterior calculus in a narrow sense with a focus on linear elasticity.

The book is a shining example of fine mathematical writing, of keeping a perfect balance between explaining ideas and concepts and their inevitable technical aspects, between generality and detail, between conciseness and redundancy, and between lucidity and rigor. With remarkable didactic skill, the author has structured and linked the broad variety of topics addressed in the book. All this renders it suitable as both a "bedtime reading" for mathematicians and a primer on the fascinating topic of FEEC. In particular, this book is an excellent foundation for a graduate course intended to whet the appetite for modern numerical analysis of students from areas of pure mathematics. Hardly surprising, because it grew out of a series of lectures the author gave to a broad audience.

Of course, with only limited time available for the lectures, the author had to omit certain aspects of FEEC. Well, he should have added them in the final book. In my opinion, it would have been appropriate to devote a separate chapter to cover eigenvalue problems and discrete compactness [1, 2] and to cover the recent breakthroughs in the construction of bounded cochain projections [3]. A chapter addressing ideas connected with discrete Hodge operators [4, 5] could also have been included.

This does not compromise the great value of the book as a work bridging the divide sometimes separating what is labeled "pure" and "applied" mathematics. I hope it will also have lasting impact as an impressive demonstration that the theory of finite-element methods is by no means a "mindless exercise in endless estimates," but rather riveting mathematics full of deep ideas.

REFERENCES

[1] D. Boffi, *Finite element approximation of eigenvalue problems*, Acta Numer., 19 (2010), pp. 1–120.

[2] D. Boffi, M. Costabel, M. Dauge, L. Demkowicz, and R. Hiptmair, *Discrete compactness for the p-version of discrete differential forms*, SIAM J. Numer. Anal., 49 (2011), pp. 135–158, https://doi.org/10.1137/090772629.

[3] R. S. Falk and R. Winther, *Local bounded cochain projections*, Math. Comp., 83 (2014), pp. 2631–2656.

[4] R. Hiptmair, *Discrete Hodge operators*, Numer. Math., 90 (2001), pp. 265–289.

[5] T. Tarhasaari, L. Kettunen, and A. Bossavit, *Some realizations of a*

discrete Hodge: A reinterpretation of finite element techniques, IEEE Trans. Mag., 35 (1999), pp. 1494–1497.

RALF HIPTMAIR
ETH Zurich

Data-Driven Computational Methods: Parameter and Operator Estimations. *By John Harlim.* Cambridge University Press, Cambridge, UK, 2018. $69.99. xii+158 pp., softcover. ISBN 978-1-108-47247-0.

Working with data and models is extremely important in most modern scientific disciplines, including engineering, economics, and finance. This book by John Harlim is an interesting contribution that provides a survey of different techniques and approaches related to data analysis and uncertainty quantification using stochastic differential equations. The book covers some basics on Monte Carlo, ensemble Kalman filters, spectral decomposition methods, and kernel methods for stochastic differential equations. The choice of topics and presentation style is strongly aligned with the research work of the author, and the book provides a nice introduction to each of the aforementioned topics. I found the writing style fairly clear and overall enjoyed reading through it.

The book is organized such that each chapter can be read independently. Chapter 1 provides the main setup, an outline of what follows, and motivating examples. Chapter 2 contains an overview of Markov chains, Markov chain Monte Carlo, and some ideas on how it can be applied for complex or high-dimensional dynamical system, e.g., using surrogate models. Chapter 3 discusses the problem of filtering or estimation of hidden dynamical states and static parameters for hidden Markov models. The main tool presented here is the ensemble Kalman filter, which is a common choice for high-dimensional models. The chapter concludes with ideas on parameter estimation and model reduction, which are both very useful and often overlooked in other books or papers. Chapters 4 and 5 look at spectral methods and how to decompose dynamics for effective model reduction and uncertainty quantification. An overview of polynomial chaos, Galerkin and collocation methods is presented in Chapter 4, and Chapter 5 presents Karhunen–Loève expansions and principal components analysis. Finally, Chapter 6 combines some ideas from Riemannian geometry, kernel methods, and nonparametric estimation for stochastic differential equations. The book also contains some appendices with some background results and definitions, on probability, stochastic processes, and geometry.

I would classify the book in the broader topic of data assimilation, which is an area that combines methods and ideas from statistics and applied mathematics and aims to address challenging high-dimensional applications in weather prediction, geosciences, and many more. The book is quite short (about 150 pages), so I would put it in the handbook category. The size and variety of the material makes it very convenient when used as a reference for courses and student projects. I have used it for both undergraduate and MSc-level projects and found that the concise and comprehensive exposition encouraged students to dive faster into implementation and quickly gain hands-on experience with the methods. The variety of ideas and methods also had a positive effect in encouraging students to try different approaches and methods.

The book is certainly timely, and it is not surprising that other books have appeared recently in this area; see, e.g., [1, 2, 3, 4]. These are longer (especially [2, 3]) and tend to go into more depth in various topics. In my opinion this book is distinct as a short introductory treatment, and I would recommend it to uninitiated readers from applied mathematics or engineering as a first point of reference before looking to other sources that provide more details. The size and comprehensive style of the book is in this sense a strength, but at the same time this means that researchers or practitioners wishing to understand well one particular topic need to also use more advanced references.

Regarding omissions (or, better, suggestions worth considering for a second edition), given its scope and usefulness in teaching, some exercises involving some derivations or computing would be useful.

Also in each chapter a neat add-on could be a summary of links to other parts of the literature. That would be useful for novice readers to expand their reading, while the present book maintains brevity and clarity. Such a discussion could include pointing to papers and books for more details in different topics. One could mention references close in spirit to this book such as those mentioned here as well as relevant parts of the literature from other topics such as machine learning. A critical discussion of the literature could also be combined with several points such as historical context, relevance for high-dimensional problems, and interesting open questions of theoretical or practical interest.

To conclude, I believe this book is useful for students or researchers entering in the topic of data assimilation or interested in statistical and computational methods for stochastic differential equations. It complements nicely other recent books in the field and gives a concise overview of some recent research activity in a very comprehensive style.

REFERENCES

[1] K. Law, A. Stuart, and K. Zygalakis, *Data Assimilation*, Springer, 2015.

[2] S. Reich and C. Cotter, *Probabilistic Forecasting and Bayesian Data Assimilation*, Cambridge University Press, 2015.

[3] S. J. Fletcher, *Data Assimilation for the Geosciences: From Theory to Application*, Elsevier, 2017.

[4] P. J. Van Leeuwen, Y. Cheng, and S. Reich, *Nonlinear Data Assimilation*, Springer, 2015.

Nikolas Kantas
Imperial College London

Mathematics for Machine Learning. *By Marc Peter Deisenroth, A. Aldo Faisal, and Cheng Soon Og.* Cambridge University Press, Cambridge, UK, 2020. $46.99. xviii+411 pp., softcover. ISBN 978-1-108-45514-5.

In this current age of machine learning and data science, more and more introductory books are entering the market that emphasize the importance of mathematics in these fields. It is no wonder that Gilbert Strang, being a mathematician himself, is the author of one such book, the excellent introductory text [1]. The book under review, by Marc Deisenroth, Aldo Faisal, and Cheng Soon Og, is different in so far as it is written not by members of math departments but rather computer scientists on the forefront of machine learning. Over and over again, this book emphasizes the importance of solid mathematical foundations for understanding advanced concepts of machine learning. We as a mathematical community should understand this as an important challenge and invitation to delve further into the deep foundations of machine learning and data science.

In terms of content, the book starts immediately after high school, literally: "The book assumes the reader to have mathematical knowledge commonly covered in high school mathematics and physics." In seven chapters of mathematics, which mainly contain undergraduate math (linear algebra, vector calculus, probabilities and only the basics of optimization), it then leads the reader from page 249 onward to the ML concepts of regression, PCA, Gaussian mixture models, and support vector machines. Even though I didn't attend an American high school myself (just a German "gymnasium"), this should be quite a ride if you rely exclusively on the book. Didactically, it is difficult in some places, especially if there are many definitions, but relatively few precise theorems and even fewer formal proofs.

Very nice is that the reader is constantly motivated by numerous cross-references to ML to work through the somewhat drier mathematical aspects. Really successful are the numerous explanatory illustrations, which help to explain even difficult concepts in a catchy way. Each chapter concludes with many instructive exercises. An outstanding feature of this book is the additional material presented on the website https://mml-book.com. In addition to a free PDF version of the book, complementary Jupyter notebooks are provided along with a constantly updated list of errata. Unfortunately Jupyter notebooks are available only for the first part of the book on mathematical foundations and not (yet?) for the second part, which deals with selected aspects of machine learning. I found the application of PCA to the MNIST dataset very interesting in section 10.3.3, where you can get a very nice insight into this dataset. Before this, I knew the MNIST dataset only as a test example for artificial neural networks. By the way, it is surprising that neural networks play a very small role in this book. They do not at all appear in Part II about central machine learning problems, and in Part I they are only briefly mentioned in section 5.6 about backpropagation and automatic differentiation, and indirectly in section 7.1.3 about stochastic gradient descent.

However, there is something to criticize about this book, which is by and large very comprehensive: as a numerical optimizer, it pains me to see the terms Jacobian and gradient defined as synonyms in section 5.2. Where gradients are used algorithmically, e.g., in section 7.1 about optimization using gradient descent, the gradient previously defined as a row vector, because it is synonymous with the Jacobian, is then transposed seemingly arbitrarily. But everything becomes less arbitrary and ultimately much simpler if, on the other hand, the consensus of the optimization community is followed and the derivative is regarded as a linearized mapping, the Jacobian as its matrix representation and therefore a row vector, and the gradient as its (Riesz) representation with the help of an inner product that still has to be chosen, and thus a column vector.

The book may make a good companion to a respective course on mathematical foundations of machine learning and data science, as you have them in many data science curricula nowadays. We have this as a first course in our local Master curriculum Data Science. And there is the right place, students get a refresher of everything they are supposed to have learned in their respective bachelor curriculum, which may be in diverse disciplines. It competes among others with the book [1] that I already mentioned and which I prefer because of its mathematical rigor, but this is probably a matter of taste. When looking for a book as a companion for a course, the book of Marc Deisenroth, Aldo Faisal and Cheng Soon Og should be considered a good choice.

REFERENCE

[1] G. Strang, *Linear Algebra and Learning from Data*, Wellesley Cambridge Press, 2019.

Volker H. Schulz
Trier University