

Parallel Computing For Data Science With Examples In R C

Parallel Computing: Accelerating Computational Science and Engineering (CSE)
 Real-Time Parallel Computing
 Data Analysis in the Distributed World
 Introduction to HPC with MPI for Data Science
 Applied Parallel Computing
 Handbook of Parallel Computing
 Vector Models for Data-parallel Computing
 Big Data and High Performance Computing
 IoT Big Data Stream Processing
 With Examples in R, C++ and CUDA
 Mastering Functional Programming
 Concepts and Practice
 Parallel Scientific Computation
 Nature Inspired Computing for Data Science
 Big Data Analysis with Python
 Perform data collection, data processing, wrangling, visualization, and model building using Python
 Parallel and High Performance Computing
 Parallel Algorithms in Computational Science and Engineering
 Parallel Processing for Scientific Computing
 Python Data Analysis
 16th International Conference, PaCT 2021, Kaliningrad, Russia, September 13–18, 2021, Proceedings
 Handbook of Parallel Computing and Statistics
 The Future of Data Science and Parallel Computing
 Scientific Parallel Computing
 Julia Cookbook
 Combine Spark and Python to unlock the powers of parallel computing and machine learning
 A Structured Approach Using BSP
 Parallel R
 Programming Massively Parallel Processors
 Data Science with Python and Dask
 Algorithms and Parallel Computing
 Deep Learning and Parallel Computing Environment for Bioengineering Systems
 Parallel Computing: Fundamentals, Applications and New Directions
 Learn R and Python in Parallel
 Data Intensive Computing Applications for Big Data
 Models, Algorithms and Applications
 Image Analysis
 Parallel Computing Technologies
 Advances in Computing and Data Sciences

*Parallel Computing For
 Data Science With
 Examples In R C*

Downloaded from
archive.imba.com by guest

SAUNDERS CHAMBERS

**Parallel Computing: Accelerating
 Computational Science and
 Engineering (CSE)** Mit Press

This book is concerned with the aspects of real-time, parallel computing which are specific to the analysis of digitized images including both the symbolic and semantic data derived from such images. The subjects covered encompass processing, storing, and transmitting images and image data. A variety of techniques and algorithms for the analysis and manipulation of images are explored both theoretically and in terms of implementation in hardware and software.

The book is organized into four topic areas: (1) the retical development, (2) languages for image processing, (3) new computer techniques, and (4) implementation in special purpose real-time digital systems. Computer utilization, methodology, and design for image analysis presents special and unusual problems. One author (Nagao)* points out that, "Human perception of a scene is very complex. It has not been made clear how perception functions, what one sees in a picture, and how one understands the whole picture. It is almost certain that one carries out a very quick trial-and-error process, starting from the detection of gross prominent features and then analyzing details, using one's knowledge of the world. " Another author (Duff)

makes the observation that, "It is therefore more difficult to write computer programs which deal with images than those which deal with numbers, human thinking about arithmetic being a largely conscious activity.

Real-Time Parallel Computing Elsevier
 Get to grips with processing large volumes of data and presenting it as engaging, interactive insights using Spark and Python. Key Features Get a hands-on, fast-paced introduction to the Python data science stack Explore ways to create useful metrics and statistics from large datasets Create detailed analysis reports with real-world data Book Description Processing big data in real time is challenging due to scalability, information inconsistency, and fault tolerance. Big

Data Analysis with Python teaches you how to use tools that can control this data avalanche for you. With this book, you'll learn practical techniques to aggregate data into useful dimensions for posterior analysis, extract statistical measurements, and transform datasets into features for other systems. The book begins with an introduction to data manipulation in Python using pandas. You'll then get familiar with statistical analysis and plotting techniques. With multiple hands-on activities in store, you'll be able to analyze data that is distributed on several computers by using Dask. As you progress, you'll study how to aggregate data for plots when the entire data cannot be accommodated in memory. You'll also explore Hadoop (HDFS and YARN), which will help you tackle larger datasets. The book also covers Spark and explains how it interacts with other tools. By the end of this book, you'll be able to bootstrap your own Python environment, process large files, and manipulate data to generate statistics, metrics, and graphs. What you will learn Use Python to read and transform data into different formats Generate basic statistics and metrics using data on disk Work with computing tasks distributed over a cluster Convert data from various sources into storage or querying formats Prepare data for statistical analysis, visualization, and machine learning Present data in the form of effective visuals Who this book is for Big Data Analysis with Python is designed for Python developers, data analysts, and data scientists who want to get hands-on with methods to control data and transform it into impactful insights. Basic knowledge of statistical measurements and relational databases will help you to understand various concepts explained in this book.

Data Analysis in the Distributed World

Packt Publishing Ltd

Parallel Computing Architectures and APIs: IoT Big Data Stream Processing commences from the point high-performance uniprocessors were becoming increasingly complex, expensive, and power-hungry. A basic trade-off exists between the use of one or a small number of such complex processors, at one extreme, and a moderate to very large number of simpler processors, at the other. When combined with a high-bandwidth, interprocessor communication facility leads to significant simplification of the design process. However, two major roadblocks prevent the widespread adoption of such moderately to massively parallel architectures: the interprocessor

communication bottleneck, and the difficulty and high cost of algorithm/software development. One of the most important reasons for studying parallel computing architectures is to learn how to extract the best performance from parallel systems. Specifically, you must understand its architectures so that you will be able to exploit those architectures during programming via the standardized APIs. This book would be useful for analysts, designers and developers of high-throughput computing systems essential for big data stream processing emanating from IoT-driven cyber-physical systems (CPS). This pragmatic book: Devolves uniprocessors in terms of a ladder of abstractions to ascertain (say) performance characteristics at a particular level of abstraction Explains limitations of uniprocessor high performance because of Moore's Law Introduces basics of processors, networks and distributed systems Explains characteristics of parallel systems, parallel computing models and parallel algorithms Explains the three primary categorical representatives of parallel computing architectures, namely, shared memory, message passing and stream processing Introduces the three primary categorical representatives of parallel programming APIs, namely, OpenMP, MPI and CUDA Provides an overview of Internet of Things (IoT), wireless sensor networks (WSN), sensor data processing, Big Data and stream processing Provides introduction to 5G communications, Edge and Fog computing Parallel Computing Architectures and APIs: IoT Big Data Stream Processing discusses stream processing that enables the gathering, processing and analysis of high-volume, heterogeneous, continuous Internet of Things (IoT) big data streams, to extract insights and actionable results in real time. Application domains requiring data stream management include military, homeland security, sensor networks, financial applications, network management, web site performance tracking, real-time credit card fraud detection, etc.

Introduction to HPC with MPI for Data Science "O'Reilly Media, Inc."

The book provides a practical guide to computational scientists and engineers to help advance their research by exploiting the superpower of supercomputers with many processors and complex networks. This book focuses on the design and analysis of basic parallel algorithms, the key components for composing larger packages for a wide range of applications. Applied Parallel Computing Princeton University Press

Understand data analysis pipelines using machine learning algorithms and techniques with this practical guide Key Features Prepare and clean your data to use it for exploratory analysis, data manipulation, and data wrangling Discover supervised, unsupervised, probabilistic, and Bayesian machine learning methods Get to grips with graph processing and sentiment analysis Book Description Data analysis enables you to generate value from small and big data by discovering new patterns and trends, and Python is one of the most popular tools for analyzing a wide variety of data. With this book, you'll get up and running using Python for data analysis by exploring the different phases and methodologies used in data analysis and learning how to use modern libraries from the Python ecosystem to create efficient data pipelines. Starting with the essential statistical and data analysis fundamentals using Python, you'll perform complex data analysis and modeling, data manipulation, data cleaning, and data visualization using easy-to-follow examples. You'll then understand how to conduct time series analysis and signal processing using ARMA models. As you advance, you'll get to grips with smart processing and data analytics using machine learning algorithms such as regression, classification, Principal Component Analysis (PCA), and clustering. In the concluding chapters, you'll work on real-world examples to analyze textual and image data using natural language processing (NLP) and image analytics techniques, respectively. Finally, the book will demonstrate parallel computing using Dask. By the end of this data analysis book, you'll be equipped with the skills you need to prepare data for analysis and create meaningful data visualizations for forecasting values from data. What you will learn Explore data science and its various process models Perform data manipulation using NumPy and pandas for aggregating, cleaning, and handling missing values Create interactive visualizations using Matplotlib, Seaborn, and Bokeh Retrieve, process, and store data in a wide range of formats Understand data preprocessing and feature engineering using pandas and scikit-learn Perform time series analysis and signal processing using sunspot cycle data Analyze textual data and image data to perform advanced analysis Get up to speed with parallel computing using Dask Who this book is for This book is for data analysts, business analysts, statisticians, and data scientists looking to learn how to use Python for data analysis. Students and academic faculties will also find this book

useful for learning and teaching Python data analysis using a hands-on approach. A basic understanding of math and working knowledge of the Python programming language will help you get started with this book.

Handbook of Parallel Computing Packt Publishing Ltd

An overview of the most prominent contemporary parallel processing programming models, written in a unique tutorial style. With the coming of the parallel computing era, computer scientists have turned their attention to designing programming models that are suited for high-performance parallel computing and supercomputing systems. Programming parallel systems is complicated by the fact that multiple processing units are simultaneously computing and moving data. This book offers an overview of some of the most prominent parallel programming models used in high-performance computing and supercomputing systems today. The chapters describe the programming models in a unique tutorial style rather than using the formal approach taken in the research literature. The aim is to cover a wide range of parallel programming models, enabling the reader to understand what each has to offer. The book begins with a description of the Message Passing Interface (MPI), the most common parallel programming model for distributed memory computing. It goes on to cover one-sided communication models, ranging from low-level runtime libraries (GASNet, OpenSHMEM) to high-level programming models (UPC, GA, Chapel); task-oriented programming models (Charm++, ADLB, Scioto, Swift, CnC) that allow users to describe their computation and data units as tasks so that the runtime system can manage computation and data movement as necessary; and parallel programming models intended for on-node parallelism in the context of multicore architecture or attached accelerators (OpenMP, Cilk Plus, TBB, CUDA, OpenCL). The book will be a valuable resource for graduate students, researchers, and any scientist who works with data sets and large computations. Contributors Timothy Armstrong, Michael G. Burke, Ralph Butler, Bradford L. Chamberlain, Sunita Chandrasekaran, Barbara Chapman, Jeff Daily, James Dinan, Deepak Eachempati, Ian T. Foster, William D. Gropp, Paul Hargrove, Wen-mei Hwu, Nikhil Jain, Laxmikant Kale, David Kirk, Kath Knobe, Ariram Krishnamoorthy, Jeffery A. Kuehn, Alexey Kukanov, Charles E. Leiserson, Jonathan Lifflander, Ewing Lusk, Tim Mattson, Bruce Palmer, Steven C. Pieper, Stephen W. Poole, Arch D.

Robison, Frank Schlimbach, Rajeev Thakur, Abhinav Vishnu, Justin M. Wozniak, Michael Wilde, Kathy Yelick, Yili Zheng
Vector Models for Data-parallel Computing Springer

Parallel computing has been the enabling technology of high-end machines for many years. Now, it has finally become the ubiquitous key to the efficient use of any kind of multi-processor computer architecture, from smart phones, tablets, embedded systems and cloud computing up to exascale computers. This book presents the proceedings of ParCo2013 – the latest edition of the biennial International Conference on Parallel Computing – held from 10 to 13 September 2013, in Garching, Germany. The conference focused on several key parallel computing areas. Themes included parallel programming models for multi- and manycore CPUs, GPUs, FPGAs and heterogeneous platforms, the performance engineering processes that must be adapted to efficiently use these new and innovative platforms, novel numerical algorithms and approaches to large-scale simulations of problems in science and engineering. The conference programme also included twelve mini-symposia (including an industry session and a special PhD Symposium), which comprehensively represented and intensified the discussion of current hot topics in high performance and parallel computing. These special sessions covered large-scale supercomputing, novel challenges arising from parallel architectures (multi-/manycore, heterogeneous platforms, FPGAs), multi-level algorithms as well as multi-scale, multi-physics and multi-dimensional problems. It is clear that parallel computing – including the processing of large data sets (“Big Data”) – will remain a persistent driver of research in all fields of innovative computing, which makes this book relevant to all those with an interest in this field.

Big Data and High Performance Computing Createspace Independent Publishing Platform

This volume gives an overview of the state-of-the-art with respect to the development of all types of parallel computers and their application to a wide range of problem areas. The international conference on parallel computing ParCo97 (Parallel Computing 97) was held in Bonn, Germany from 19 to 22 September 1997. The first conference in this biannual series was held in 1983 in Berlin. Further conferences were held in Leiden (The Netherlands), London (UK), Grenoble (France) and Gent (Belgium). From the

outset the aim with the ParCo (Parallel Computing) conferences was to promote the application of parallel computers to solve real life problems. In the case of ParCo97 a new milestone was reached in that more than half of the papers and posters presented were concerned with application aspects. This fact reflects the coming of age of parallel computing. Some 200 papers were submitted to the Program Committee by authors from all over the world. The final programme consisted of four invited papers, 71 contributed scientific/industrial papers and 45 posters. In addition a panel discussion on Parallel Computing and the Evolution of Cyberspace was held. During and after the conference all final contributions were refereed. Only those papers and posters accepted during this final screening process are included in this volume. The practical emphasis of the conference was accentuated by an industrial exhibition where companies demonstrated the newest developments in parallel processing equipment and software. Speakers from participating companies presented papers in industrial sessions in which new developments in parallel computing were reported.

IoT Big Data Stream Processing Simon and Schuster

In large projects, programmers tend to get overwhelmed by their complexity. It can be hard to keep track of all the interdependencies in the code-base and how its state changes on runtime. The solution to these problems is Functional Programming, a paradigm specifically designed to deal with the complexity of software development. Mastering ...

With Examples in R, C++ and CUDA

Springer Science & Business Media
This two-volume set (CCIS 905 and CCIS 906) constitutes the refereed proceedings of the Second International Conference on Advances in Computing and Data Sciences, ICACDS 2018, held in Dehradun, India, in April 2018. The 110 full papers were carefully reviewed and selected from 598 submissions. The papers are centered around topics like advanced computing, data sciences, distributed systems organizing principles, development frameworks and environments, software verification and validation, computational complexity and cryptography, machine learning theory, database theory, probabilistic representations.

Mastering Functional Programming IOS Press

Deep Learning and Parallel Computing Environment for Bioengineering Systems delivers a significant forum for the technical advancement of deep learning in

parallel computing environment across bio-engineering diversified domains and its applications. Pursuing an interdisciplinary approach, it focuses on methods used to identify and acquire valid, potentially useful knowledge sources. Managing the gathered knowledge and applying it to multiple domains including health care, social networks, mining, recommendation systems, image processing, pattern recognition and predictions using deep learning paradigms is the major strength of this book. This book integrates the core ideas of deep learning and its applications in bio engineering application domains, to be accessible to all scholars and academicians. The proposed techniques and concepts in this book can be extended in future to accommodate changing business organizations' needs as well as practitioners' innovative ideas. Presents novel, in-depth research contributions from a methodological/application perspective in understanding the fusion of deep machine learning paradigms and their capabilities in solving a diverse range of problems Illustrates the state-of-the-art and recent developments in the new theories and applications of deep learning approaches applied to parallel computing environment in bioengineering systems Provides concepts and technologies that are successfully used in the implementation of today's intelligent data-centric critical systems and multi-media Cloud-Big data

Concepts and Practice SIAM

This book constitutes the proceedings of the 16th International Conference on Parallel Computing Technologies, PaCT 2021, which was held during September 13-18, 2021. The conference was planned to take place in Kaliningrad, Russia, but changed to an online event due to the COVID-19 pandemic. The 24 full and 12 short papers included in this book were carefully reviewed and selected from 62 submissions. They were organized in topical sections as follows: parallel programming methods and tools; applications; memory-efficient data structures; experimental studies; job management; essential algorithms; computing services; and cellular automata.

Parallel Scientific Computation

CreateSpace

Parallel processing has been an enabling technology in scientific computing for more than 20 years. This book is the first in-depth discussion of parallel computing in 10 years; it reflects the mix of topics that mathematicians, computer scientists, and computational scientists focus on to

make parallel processing effective for scientific problems. Presently, the impact of parallel processing on scientific computing varies greatly across disciplines, but it plays a vital role in most problem domains and is absolutely essential in many of them. Parallel Processing for Scientific Computing is divided into four parts: The first concerns performance modeling, analysis, and optimization; the second focuses on parallel algorithms and software for an array of problems common to many modeling and simulation applications; the third emphasizes tools and environments that can ease and enhance the process of application development; and the fourth provides a sampling of applications that require parallel computing for scaling to solve larger and realistic models that can advance science and engineering.

Nature Inspired Computing for Data Science Packt Publishing Ltd

Amazon's #1 Bestseller Dr. Ganapathi Pulipaka is a Chief Data Scientist and SAP Technical Lead for one of the largest firms in the world. He is also a PostDoc Research Scholar in Computer Science Engineering in Big Data Analytics, Machine Learning, Robotics, IoT, Artificial Intelligence as part of Doctor of Computer Science program from Colorado Technical University, CO with another PhD in Data Analytics, Information Systems, and Enterprise Resource Management, California University, Irvine. A technology leader in artificial intelligence, SAP development, and solution architecture. A project/program manager for application development of SAP systems, machine learning, deep learning systems, application development management, basis, infrastructure, and consulting delivery services offering expertise in delivery execution and executive interaction. Experienced implementing ASAP, Agile, Agile ASAP 8, HANA ASAP 8, Activate, Prince2, SCRUM, and Waterfall SDLC project methodologies. Implemented multiple SAP programs/projects managing a team size of 60+ members, managing budget more than \$5M to \$10M with SAP backend databases of Oracle, IBM DB2, Sybase, Informix, MS SQL server on Mac OS and Linux environments. His background is in Computer Science with a professional skillset and two decades of management and hands-on development experience in Machine Learning in TensorFlow, Python, and R, Deep Learning in TensorFlow, Python, and R, SAP ABAP S/4 HANA 1609, SAP S/4 HANA 1710, SAP IBP on SAP Cloud Platform, Big Data, IaaS, IoT, Data Science, Apache Hadoop, Apache Kafka, Apache Spark, Apache

Storm, Apache Flink, SQL, NoSQL, Tableau, PowerBI, Mathematics, Data Mining, Statistical Framework, SIEM, SAP, SAP ERP/ECC 6.0 NetWeaver Portals, SAP PLM, cProjects, R/3, BW, SRM 5.0, CRM 7.4, 7.3, 7.2, 7.1, 7.0, Java, C, C++, VC++, SAP CRM-IPM, SAP CRM- Service management, SAP CRM-Banking, SAP PLM Web UI 7.47, xRPM, SCM 7.1 APO, DP, SNP, SNC, FSCM, FSCD, SCEM, EDI. CRM ABAP/OO, ABAP, CRM Web UI/BOL/GENIL/ABAP Objects, SAP Netweaver Gateway (OData), SAP Mobility, SAP Fiori, Information Security, CyberSecurity, Governance, Risk Controls, and Compliance, SAP Fiori HANA, ABAP Webdynpros, BSPs, EDI/ALE, CRM Middleware, CRM Workflow, JavaScript, SAP KW 7.3 SAP Content server, SAP TREX Server, SAP KPro, SAP PI (PO), SAP BPC, Script logics, Azure, SAP BPM, SAP UI5, SAP BRM, Unix, Linux, macOS, and always looking for patterns in data and performing extractions to provide new meanings and insights through algorithms and analytics.

Big Data Analysis with Python Springer

This contributed volume highlights two areas of fundamental interest in high-performance computing: core algorithms for important kernels and computationally demanding applications. The first few chapters explore algorithms, numerical techniques, and their parallel formulations for a variety of kernels that arise in applications. The rest of the volume focuses on state-of-the-art applications from diverse domains. By structuring the volume around these two areas, it presents a comprehensive view of the application landscape for high-performance computing, while also enabling readers to develop new applications using the kernels. Readers will learn how to choose the most suitable parallel algorithms for any given application, ensuring that theory and practicality are clearly connected. Applications using these techniques are illustrated in detail, including: Computational materials science and engineering Computational cardiovascular analysis Multiscale analysis of wind turbines and turbomachinery Weather forecasting Machine learning techniques Parallel Algorithms in Computational Science and Engineering will be an ideal reference for applied mathematicians, engineers, computer scientists, and other researchers who utilize high-performance computing in their work.

Perform data collection, data processing, wrangling, visualization, and model building using Python Packt Publishing Ltd
Parallel Computing for Data Science: With Examples in R, C++ and CUDA is one of

the first parallel computing books to concentrate exclusively on parallel data structures, algorithms, software tools, and applications in data science. It includes examples not only from the classic "n observations, p variables" matrix format but also from time series,

Parallel and High Performance Computing Newnes

Building upon the wide-ranging success of the first edition, *Parallel Scientific Computation* presents a single unified approach to using a range of parallel computers, from a small desktop computer to a massively parallel computer. The author explains how to use the bulk synchronous parallel (BSP) model to design and implement parallel algorithms in the areas of scientific computing and big data, and provides a full treatment of core problems in these areas, starting from a high-level problem description, via a sequential solution algorithm to a parallel solution algorithm and an actual parallel program written in BSPlib. Every chapter of the book contains a theoretical section and a practical section presenting a parallel program and numerical experiments on a modern parallel computer to put the theoretical predictions and cost analysis to the test. Every chapter also presents extensive bibliographical notes with additional discussions and pointers to relevant literature, and numerous exercises which are suitable as graduate student projects. The second edition provides new material relevant for big-data science such as sorting and graph algorithms, and it provides a BSP approach towards new hardware developments such as hierarchical architectures with both shared and distributed memory. A single, simple hybrid BSP system suffices to handle both types of parallelism efficiently, and there is no need to master two systems, as often happens in alternative approaches. Furthermore, the second edition brings all algorithms used up to date, and it includes new material on high-performance linear system solving by LU decomposition, and improved data partitioning for sparse matrix computations. The book is accompanied by a software package BSPedupack, freely available online from the author's homepage, which contains all programs of the book and a set of test driver programs. This package written in C can be run using modern BSPlib implementations such as MulticoreBSP for C or BSPonMPI.

Parallel Algorithms in Computational Science and Engineering World Scientific
The book 'Data Intensive Computing Applications for Big Data' discusses the

technical concepts of big data, data intensive computing through machine learning, soft computing and parallel computing paradigms. It brings together researchers to report their latest results or progress in the development of the above mentioned areas. Since there are few books on this specific subject, the editors aim to provide a common platform for researchers working in this area to exhibit their novel findings. The book is intended as a reference work for advanced undergraduates and graduate students, as well as multidisciplinary, interdisciplinary and transdisciplinary research workers and scientists on the subjects of big data and cloud/parallel and distributed computing, and explains didactically many of the core concepts of these approaches for practical applications. It is organized into 24 chapters providing a comprehensive overview of big data analysis using parallel computing and addresses the complete data science workflow in the cloud, as well as dealing with privacy issues and the challenges faced in a data-intensive cloud computing environment. The book explores both fundamental and high-level concepts, and will serve as a manual for those in the industry, while also helping beginners to understand the basic and advanced aspects of big data and cloud computing.

Parallel Processing for Scientific Computing CRC Press

Programming Massively Parallel Processors: A Hands-on Approach, Second Edition, teaches students how to program massively parallel processors. It offers a detailed discussion of various techniques for constructing parallel programs. Case studies are used to demonstrate the development process, which begins with computational thinking and ends with effective and efficient parallel programs. This guide shows both student and professional alike the basic concepts of parallel programming and GPU architecture. Topics of performance, floating-point format, parallel patterns, and dynamic parallelism are covered in depth. This revised edition contains more parallel programming examples, commonly-used libraries such as Thrust, and explanations of the latest tools. It also provides new coverage of CUDA 5.0, improved performance, enhanced development tools, increased hardware support, and more; increased coverage of related technology, OpenCL and new material on algorithm patterns, GPU clusters, host programming, and data parallelism; and two new case studies (on MRI reconstruction and molecular visualization) that explore the latest

applications of CUDA and GPUs for scientific research and high-performance computing. This book should be a valuable resource for advanced students, software engineers, programmers, and hardware engineers. New coverage of CUDA 5.0, improved performance, enhanced development tools, increased hardware support, and more. Increased coverage of related technology, OpenCL and new material on algorithm patterns, GPU clusters, host programming, and data parallelism. Two new case studies (on MRI reconstruction and molecular visualization) explore the latest applications of CUDA and GPUs for scientific research and high-performance computing.

Python Data Analysis MIT Press

Over 40 recipes to get you up and running with programming using Julia
About This Book Follow a practical approach to learn Julia programming the easy way
Get an extensive coverage of Julia's packages for statistical analysis
This recipe-based approach will help you get familiar with the key concepts in Julia
Who This Book Is For This book is for data scientists and data analysts who are familiar with the basics of the Julia language. Prior experience of working with high-level languages such as MATLAB, Python, R, or Ruby is expected.
What You Will Learn
Extract and handle your data with Julia
Uncover the concepts of metaprogramming in Julia
Conduct statistical analysis with StatsBase.jl and Distributions.jl
Build your data science models
Find out how to visualize your data with Gadfly
Explore big data concepts in Julia
In Detail Want to handle everything that Julia can throw at you and get the most of it every day? This practical guide to programming with Julia for performing numerical computation will make you more productive and able work with data more efficiently. The book starts with the main features of Julia to help you quickly refresh your knowledge of functions, modules, and arrays. We'll also show you how to utilize the Julia language to identify, retrieve, and transform data sets so you can perform data analysis and data manipulation. Later on, you'll see how to optimize data science programs with parallel computing and memory allocation. You'll get familiar with the concepts of package development and networking to solve numerical problems using the Julia platform. This book includes recipes on identifying and classifying data science problems, data modelling, data analysis, data manipulation, meta-programming, multidimensional arrays, and parallel computing. By the end of the book, you

will acquire the skills to work more effectively with your data. Style and

approach This book has a recipe-based

approach to help you grasp the concepts of Julia programming.

Related with Parallel Computing For Data Science With Examples In R C:

- Present For A Language Teacher Crossword : [click here](#)