a tutorial on principal components analysis

A Tutorial on Principal Component Analysis: Unveiling the Power of Dimensionality Reduction

Author: Dr. Evelyn Reed, PhD in Statistics, Associate Professor of Data Science at the University of California, Berkeley. Dr. Reed has over 15 years of experience in statistical modeling and machine learning, with a focus on high-dimensional data analysis.

Publisher: Springer Nature, a leading global scientific publisher known for its rigorous peer-review process and high-quality publications in mathematics, statistics, and computer science.

Editor: Dr. Michael Chen, PhD in Applied Mathematics, Senior Editor at Springer Nature with extensive experience in editing statistical and data science publications.

Keywords: principal component analysis, PCA, dimensionality reduction, data analysis, feature extraction, machine learning, tutorial, explained, step-by-step, python, R, algorithm

Summary: This comprehensive tutorial on principal component analysis (PCA) provides a thorough understanding of this powerful dimensionality reduction technique. Starting with foundational concepts, the tutorial progressively builds towards advanced applications. It covers the mathematical underpinnings of PCA, including eigenvectors and eigenvalues, and explains how PCA transforms high-dimensional data into a lower-dimensional representation while preserving maximum variance. The tutorial includes step-by-step explanations, illustrative examples using Python and R, and practical applications across diverse fields. It addresses common challenges and misconceptions associated with PCA and concludes with a discussion of its limitations and alternatives. This resource aims to equip readers with the knowledge and skills to effectively apply PCA in their own data analysis projects.

1. Introduction: What is Principal Component Analysis (PCA)?

This tutorial on principal component analysis will guide you through the intricacies of this vital technique used in data analysis and machine learning. PCA is a powerful dimensionality reduction method that transforms a large dataset with many variables into a smaller dataset with fewer variables, called principal components. These principal components are linear combinations of the original variables, capturing the maximum possible variance in the data. The key advantage is that PCA helps reduce noise, improve model performance, and visualize high-dimensional data effectively. This tutorial on principal component analysis will demystify the process, making it accessible to anyone with a basic understanding of linear algebra and statistics.

2. Mathematical Foundations: Eigenvectors and Eigenvalues

A crucial understanding of this tutorial on principal component analysis lies in grasping the concepts of eigenvectors and eigenvalues. Eigenvectors are vectors that, when multiplied by a matrix (our covariance matrix in PCA), only change in scale, not direction. The scaling factor is the eigenvalue, indicating the magnitude of the variance along that eigenvector. In PCA, we seek the eigenvectors of the data's covariance matrix. These eigenvectors represent the principal components, and their corresponding eigenvalues signify the amount of variance explained by each component. This tutorial on principal component analysis will delve deeper into the calculations and interpretations.

3. The PCA Algorithm: A Step-by-Step Guide

This section of our tutorial on principal component analysis provides a step-by-step guide to performing PCA:

1. Data Standardization: Center and scale the data to ensure that variables with larger values don't disproportionately influence the analysis.

2. Covariance Matrix Calculation: Compute the covariance matrix of the standardized data, reflecting the relationships between variables.

3. Eigenvalue Decomposition: Find the eigenvalues and eigenvectors of the covariance matrix. Eigenvectors are the principal components, and eigenvalues represent the variance explained.

4. Component Selection: Select the principal components with the highest eigenvalues, retaining the ones that explain a sufficient portion of the total variance (e.g., 95%).

5. Dimensionality Reduction: Project the original data onto the selected principal components to obtain the reduced-dimensionality representation.

This tutorial on principal component analysis will provide clear examples using Python and R to illustrate each step.

4. Implementing PCA using Python and R

This tutorial on principal component analysis will include practical coding examples. We'll showcase how to implement PCA using both Python (with libraries like scikit-learn and NumPy) and R (with base R functions or packages like `prcomp`). The examples will cover data loading, preprocessing, PCA application, and visualization of results. This hands-on approach will solidify your understanding of the theoretical concepts discussed earlier.

5. Interpreting PCA Results: Visualizing and Understanding Principal Components

After performing PCA, interpreting the results is critical. This tutorial on principal component analysis will guide you through visualizing the principal components using biplots (showing both variables and data points in the reduced space) and scree plots (displaying the eigenvalues to assess variance explained). We'll discuss how to interpret the loadings (contributions of original variables to each principal component) to understand the underlying structure of the data.

6. Applications of PCA: Real-World Examples

PCA finds extensive use across various fields:

Image Compression: Reducing the size of images while preserving essential features.
Gene Expression Analysis: Identifying patterns in gene expression data to understand biological processes.
Financial Modeling: Reducing the dimensionality of financial market data for risk assessment and portfolio optimization.
Anomaly Detection: Identifying outliers in datasets by projecting them onto the principal components.
Feature Extraction for Machine Learning: Improving the performance of machine learning models by using principal components as input features.

This tutorial on principal component analysis will illustrate these applications with examples.

7. Limitations and Alternatives to PCA

While PCA is a powerful tool, it has limitations:

Linearity Assumption: PCA assumes linear relationships between variables, which might not always hold true.
Sensitivity to Scaling: The results can be affected by the scales of the original variables.
Interpretability Challenges: Interpreting higher-order principal components can sometimes be difficult.

Alternatives like t-SNE (t-distributed Stochastic Neighbor Embedding) and UMAP (Uniform Manifold Approximation and Projection) offer non-linear dimensionality reduction techniques. This tutorial on principal component analysis will briefly introduce these alternatives and discuss their suitability compared to PCA.

8. Advanced Topics in PCA

This tutorial on principal component analysis will briefly touch upon advanced concepts including:

Kernel PCA: Extending PCA to handle non-linear relationships using kernel functions.

Sparse PCA: Finding sparse principal components, which can be more interpretable.
Robust PCA: Handling outliers and noisy data more effectively.

9. Conclusion

This tutorial on principal component analysis has provided a comprehensive overview of this fundamental dimensionality reduction technique. By understanding its mathematical foundations, implementation details, and interpretation strategies, you can effectively leverage PCA in your data analysis workflows. Remember to carefully consider its limitations and explore alternative methods when necessary. The power of PCA lies in its ability to simplify complex datasets, revealing hidden patterns and improving the performance of various analytical tasks.

FAQs:

1. What is the difference between PCA and Factor Analysis? While both reduce dimensionality, PCA focuses on variance maximization, while factor analysis aims to identify latent variables explaining the correlations between observed variables.

1. How do I choose the optimal number of principal components? Common methods include the scree plot, explained variance ratio, and Kaiser criterion.

1. Can PCA handle missing data? Yes, imputation techniques can be used to handle missing values before applying PCA.

1. Is PCA suitable for categorical data? No, PCA is primarily designed for continuous data. For categorical data, consider techniques like correspondence analysis.

1. How does PCA handle high-dimensionality? The curse of dimensionality is mitigated by reducing the number of variables while preserving important information.

1. What are the computational costs of PCA? The computational complexity is largely determined by the eigenvalue decomposition step, which can be demanding for very large datasets.

1. How can I interpret the principal components? By examining the loadings, which show the contribution of each original variable to each principal component.

1. What are some software packages for PCA? Many statistical software packages offer PCA functionality, including R, Python (scikit-learn), MATLAB, and SPSS.

1. When should I NOT use PCA? PCA may not be suitable if the data is highly non-linear or if the goal is to find clusters rather than reduce dimensionality.

Related Articles:

1. "A Practical Guide to PCA in R": This article provides a detailed walkthrough of implementing PCA in R, including data preprocessing and visualization techniques.
2. "Understanding Eigenvalues and Eigenvectors for PCA": A focused tutorial on the linear algebra underpinnings of PCA, making the mathematical concepts more accessible.
3. "PCA for Image Compression: A Case Study": Demonstrates the application of PCA in compressing images, highlighting its practical use in image processing.
4. "Comparing PCA and t-SNE for Dimensionality Reduction": A comparative analysis of PCA and t-SNE, discussing their strengths and weaknesses in different scenarios.
5. "PCA in Python with scikit-learn: A Step-by-Step Tutorial": A detailed tutorial on using the scikit-learn library in Python for PCA implementation.
6. "Interpreting PCA Results: A Guide to Biplots and Scree Plots": Focuses on interpreting PCA outputs, including visualizing and understanding principal components.
7. "Robust PCA for Noisy Data: Handling Outliers Effectively": Addresses the issue of outliers in data and presents robust PCA methods to deal with them.
8. "Kernel PCA for Non-linear Dimensionality Reduction": Explores kernel PCA, an extension of standard PCA for handling non-linear relationships in data.
9. "Applications of PCA in Bioinformatics: Gene Expression Analysis": Shows the application of PCA in bioinformatics, focusing on gene expression data analysis.

a tutorial on principal components analysis: *Principal Component Analysis* I.T. Jolliffe, 2013-03-09 Principal component analysis is probably the oldest and best known of the It was first introduced by Pearson (1901), techniques ofmultivariate analysis. and developed independently by Hotelling (1933). Like many multivariate methods, it was not widely used until the advent of electronic computers, but it is now weIl entrenched in virtually every statistical computer package. The central idea of principal component analysis is to reduce the dimen sionality of a data set in which there are a large number of interrelated variables, while retaining as much as possible of the variation present in the data set. This reduction is achieved by transforming to a new set of variables, the principal components, which are uncorrelated, and which are ordered so that the first few retain most of the variation present in all of the original variables. Computation of the principal components reduces to the solution of an eigenvalue-eigenvector problem for a positive-semidefinite symmetrie matrix. Thus, the definition and computation of principal components are straightforward but, as will be seen, this apparently simple technique has a wide variety of different applications, as weIl as a number of different deri vations. Any feelings that principal component analysis is a narrow subject should soon be dispelled by the present book; indeed some quite broad topics which are related to principal component analysis receive no more than a brief mention in the final two chapters.

a tutorial on principal components analysis: Practical Guide To Principal Component Methods in R Alboukadel KASSAMBARA, 2017-08-23 Although there are several good books on principal component methods (PCMs) and related topics, we felt that many of them are either too theoretical or too advanced. This book provides a solid practical guidance to summarize, visualize and interpret the most important information in a large multivariate data sets, using principal component methods in R. The visualization is based on the factoextra R package that we developed for creating easily beautiful ggplot2-based graphs from the output of PCMs. This book contains 4 parts. Part I provides a quick introduction to R and presents the key features of FactoMineR and factoextra. Part II describes classical principal component methods to analyze data sets containing, predominantly, either continuous or categorical variables. These methods include: Principal Component Analysis (PCA, for continuous variables), simple correspondence analysis (CA, for large contingency tables formed by two categorical variables) and Multiple CA (MCA, for a data set with more than 2 categorical variables). In Part III, you'll learn advanced methods for analyzing a data set containing a mix of variables (continuous and categorical) structured or not into groups: Factor Analysis of Mixed Data (FAMD) and Multiple Factor Analysis (MFA). Part IV covers hierarchical clustering on principal components (HCPC), which is useful for performing clustering with a data set containing only categorical variables or with a mixed data of categorical and continuous variables.

a tutorial on principal components analysis: *Data-Driven Science and Engineering* Steven L. Brunton, J. Nathan Kutz, 2022-05-05 A textbook covering data-science and machine learning methods for modelling and control in engineering and science, with Python and MATLAB®.

a tutorial on principal components analysis: Places Rated Almanac David Savageau, 1993 This sometimes controversial bestseller, completely updated with all new statistics, is packed with timely facts and unbiased information on more than 300 metropolitan areas in the U.S. and Canada. Each city is ranked according to costs of living, crime rates, cultural life, and environmental factors.

a tutorial on principal components analysis: An Introduction to Statistical Learning Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani, Jonathan Taylor, 2023-08-01 An Introduction to Statistical Learning provides an accessible overview of the field of statistical learning, an essential toolset for making sense of the vast and complex data sets that have emerged in fields ranging from biology to finance, marketing, and astrophysics in the past twenty years. This book presents some of the most important modeling and prediction techniques, along with relevant applications. Topics include linear regression, classification, resampling methods, shrinkage approaches, tree-based methods, support vector machines, clustering, deep learning, survival analysis, multiple testing, and more. Color graphics and real-world examples are used to illustrate the methods presented. This book is targeted at statisticians and non-statisticians alike, who wish to use cutting-edge statistical learning techniques to analyze their data. Four of the authors co-wrote An Introduction to Statistical Learning, With Applications in R (ISLR), which has become a mainstay of undergraduate and graduate classrooms worldwide, as well as an important reference book for data scientists. One of the keys to its success was that each chapter contains a tutorial on implementing the analyses and methods presented in the R scientific computing environment. However, in recent years Python has become a popular language for data science, and there has been increasing demand for a Python-based alternative to ISLR. Hence, this book (ISLP) covers the same materials as ISLR but with labs implemented in Python. These labs will be useful both for Python novices, as well as experienced users.

a tutorial on principal components analysis: A User's Guide to Principal Components J. Edward Jackson, 2005-01-21 WILEY-INTERSCIENCE PAPERBACK SERIES The Wiley-Interscience Paperback Series consists of selected books that have been made more accessible to consumers in an effort to increase global appeal and general circulation. With these new unabridged softcover volumes, Wiley hopes to extend the lives of these works by making them available to future generations of statisticians, mathematicians, and scientists. From the Reviews of A User’s Guide to Principal Components The book is aptly and correctly named–A User’s Guide. It is the kind of book that a user at any level, novice or skilled practitioner, would want to have at hand for autotutorial, for refresher, or as a general-purpose guide through the maze of modern PCA. –Technometrics I recommend A User’s Guide to Principal Components to anyone who is running multivariate analyses, or who contemplates performing such analyses. Those who write their own software will find the book helpful in designing better programs. Those who use off-the-shelf software will find it invaluable in interpreting the results. –Mathematical Geology

a tutorial on principal components analysis: *Python Data Science Handbook* Jake VanderPlas, 2016-11-21 For many researchers, Python is a first-class tool mainly because of its libraries for storing, manipulating, and gaining insight from data. Several resources exist for individual pieces of this data science stack, but only with the Python Data Science Handbook do you get them all—IPython, NumPy, Pandas, Matplotlib, Scikit-Learn, and other related tools. Working scientists and data crunchers familiar with reading and writing Python code will find this comprehensive desk reference ideal for tackling day-to-day issues: manipulating, transforming, and cleaning data; visualizing different types of data; and using data to build statistical or machine learning models. Quite simply, this is the must-have reference for scientific computing in Python. With this handbook, you’ll learn how to use: IPython and Jupyter: provide computational environments for data scientists using Python NumPy: includes the ndarray for efficient storage and manipulation of dense data arrays in Python Pandas: features the DataFrame for efficient storage and manipulation of labeled/columnar data in Python Matplotlib: includes capabilities for a flexible range of data visualizations in Python Scikit-Learn: for efficient and clean Python implementations of the most important and established machine learning algorithms

a tutorial on principal components analysis: Independent Component Analysis James V. Stone, 2004 A tutorial-style introduction to a class of methods for extracting independent signals from a mixture of signals originating from different physical sources; includes MatLab computer code examples. Independent component analysis (ICA) is becoming an increasingly important tool for analyzing large data sets. In essence, ICA separates an observed set of signal mixtures into a set of statistically independent component signals, or source signals. In so doing, this powerful method can extract the relatively small amount of useful information typically found in large data sets. The applications for ICA range from speech processing, brain imaging, and electrical brain signals to telecommunications and stock predictions. In Independent Component Analysis, Jim Stone presents the essentials of ICA and related techniques (projection pursuit and complexity pursuit) in a tutorial style, using intuitive examples described in simple geometric terms. The treatment fills the need for a basic primer on ICA that can be used by readers of varying levels of mathematical sophistication, including engineers, cognitive scientists, and neuroscientists who need to know the essentials of this evolving method. An overview establishes the strategy implicit in ICA in terms of its essentially physical underpinnings and describes how ICA is based on the key observations that different physical processes generate outputs that are statistically independent of each other. The book then describes what Stone calls the mathematical nuts and bolts of how ICA works. Presenting only essential mathematical proofs, Stone guides the reader through an exploration of the fundamental characteristics of ICA. Topics covered include the geometry of mixing and unmixing; methods for blind source separation; and applications of ICA, including voice mixtures, EEG, fMRI, and fetal heart monitoring. The appendixes provide a vector matrix tutorial, plus basic demonstration computer code that allows the reader to see how each mathematical method described in the text translates into working Matlab computer code.

a tutorial on principal components analysis: An Introduction to Applied Multivariate Analysis with R Brian Everitt, Torsten Hothorn, 2011-04-23 The majority of data sets collected by researchers in all disciplines are multivariate, meaning that several measurements, observations, or recordings are taken on each of the units in the data set. These units might be human subjects, archaeological artifacts, countries, or a vast variety of other things. In a few cases, it may be sensible to isolate each variable and study it separately, but in most instances all the variables need to be examined simultaneously in order to fully grasp the structure and key features of the data. For this purpose, one or another method of multivariate analysis might be helpful, and it is with such methods that this book is largely concerned. Multivariate analysis includes methods both for describing and exploring such data and for making formal inferences about them. The aim of all the techniques is, in general sense, to display or extract the signal in the data in the presence of noise and to find out what the data show us in the midst of their apparent chaos. An Introduction to Applied Multivariate Analysis with R explores the correct application of these methods so as to extract as much information as possible from the data at hand, particularly as some type of graphical representation, via the R software. Throughout the book, the authors give many examples of R code used to apply the multivariate techniques to multivariate data.

a tutorial on principal components analysis: *Python Machine Learning* Sebastian Raschka, 2015-09-23 Unlock deeper insights into Machine Leaning with this vital guide to cutting-edge predictive analytics About This Book Leverage Python's most powerful open-source libraries for deep learning, data wrangling, and data visualization Learn effective strategies and best practices to improve and optimize machine learning systems and algorithms Ask – and answer – tough questions of your data with robust statistical models, built for a range of datasets Who This Book Is For If you want to find out how to use Python to start answering critical questions of your data, pick up Python Machine Learning – whether you want to get started from scratch or want to extend your data science knowledge, this is an essential and unmissable resource. What You Will Learn Explore how to use different machine learning models to ask different questions of your data Learn how to build neural networks using Keras and Theano Find out how to write clean and elegant Python code that will optimize the strength of your algorithms Discover how to embed your machine learning model in a web application for increased accessibility Predict continuous target outcomes using regression analysis Uncover hidden patterns and structures in data with clustering Organize data using effective pre-processing techniques Get to grips with sentiment analysis to delve deeper into textual and social media data In Detail Machine learning and predictive analytics are transforming the way businesses and other organizations operate. Being able to understand trends and patterns in complex data is critical to success, becoming one of the key strategies for unlocking growth in a challenging contemporary marketplace. Python can help you deliver key insights into your data – its unique capabilities as a language let you build sophisticated algorithms and statistical models that can reveal new perspectives and answer key questions that are vital for success. Python Machine Learning gives you access to the world of predictive analytics and demonstrates why Python is one of the world's leading data science languages. If you want to ask better questions of data, or need to improve and extend the capabilities of your machine learning systems, this practical data science book is invaluable. Covering a wide range of powerful Python libraries, including scikit-learn, Theano, and Keras, and featuring guidance and tips on everything from sentiment analysis to neural networks, you'll soon be able to answer some of the most important questions facing you and your organization. Style and approach Python Machine Learning connects the fundamental theoretical principles behind machine learning to their practical application in a way that focuses you on asking and answering the right questions. It walks you through the key elements of Python and its powerful machine learning libraries, while demonstrating how to get to grips with a range of statistical models.

a tutorial on principal components analysis: *The Elements of Statistical Learning* Trevor Hastie, Robert Tibshirani, Jerome Friedman, 2013-11-11 During the past decade there has been an explosion in computation and information technology. With it have come vast amounts of data in a variety of fields such as medicine, biology, finance, and marketing. The challenge of understanding these data has led to the development of new tools in the field of statistics, and spawned new areas such as data mining, machine learning, and bioinformatics. Many of these tools have common underpinnings but are often expressed with different terminology. This book describes the important ideas in these areas in a common conceptual framework. While the approach is statistical, the emphasis is on concepts rather than mathematics. Many examples are given, with a liberal use of color graphics. It should be a valuable resource for statisticians and anyone interested in data mining in science or industry. The book’s coverage is broad, from supervised learning (prediction) to unsupervised learning. The many topics include neural networks, support vector machines, classification trees and boosting---the first comprehensive treatment of this topic in any book. This major new edition features many topics not covered in the original, including graphical models, random forests, ensemble methods, least angle regression & path algorithms for the lasso, non-negative matrix factorization, and spectral clustering. There is also a chapter on methods for “wide” data (p bigger than n), including multiple testing and false discovery rates. Trevor Hastie, Robert Tibshirani, and Jerome Friedman are professors of statistics at Stanford University. They are prominent researchers in this area: Hastie and Tibshirani developed generalized additive models and wrote a popular book of that title. Hastie co-developed much of the statistical modeling software and environment in R/S-PLUS and invented principal curves and surfaces. Tibshirani proposed the lasso and is co-author of the very successful An Introduction to the Bootstrap. Friedman is the co-inventor of many data-mining tools including CART, MARS, projection pursuit and gradient boosting.

a tutorial on principal components analysis: Independent Component Analysis Aapo Hyvärinen, Juha Karhunen, Erkki Oja, 2004-04-05 A comprehensive introduction to ICA for students and practitioners Independent Component Analysis (ICA) is one of the most exciting new topics in fields such as neural networks, advanced statistics, and signal processing. This is the first book to provide a comprehensive introduction to this new technique complete with the fundamental mathematical background needed to understand and utilize it. It offers a general overview of the basics of ICA, important solutions and algorithms, and in-depth coverage of new applications in image processing, telecommunications, audio signal processing, and more. Independent Component Analysis is divided into four sections that cover: * General mathematical concepts utilized in the book * The basic ICA model and its solution * Various extensions of the basic ICA model * Real-world applications for ICA models Authors Hyvarinen, Karhunen, and Oja are well known for their contributions to the development of ICA and here cover all the relevant theory, new algorithms, and applications in various fields. Researchers, students, and practitioners from a variety of disciplines will find this accessible volume both helpful and informative.

a tutorial on principal components analysis: Grokking Machine Learning Luis Serrano, 2021-12-14 Grokking Machine Learning presents machine learning algorithms and techniques in a way that anyone can understand. This book skips the confused academic jargon and offers clear explanations that require only basic algebra. As you go, you'll build interesting projects with Python, including models for spam detection and image recognition. You'll also pick up practical skills for cleaning and preparing data.

a tutorial on principal components analysis: Principles of Data Mining Max Bramer, 2016-11-09 This book explains and explores the principal techniques of Data Mining, the automatic extraction of implicit and potentially useful information from data, which is increasingly used in commercial, scientific and other application areas. It focuses on classification, association rule mining and clustering. Each topic is clearly explained, with a focus on algorithms not mathematical formalism, and is illustrated by detailed worked examples. The book is written for readers without a strong background in mathematics or statistics and any formulae used are explained in detail. It can be used as a textbook to support courses at undergraduate or postgraduate levels in a wide range of subjects including Computer Science, Business Studies, Marketing, Artificial Intelligence, Bioinformatics and Forensic Science. As an aid to self study, this book aims to help general readers develop the necessary understanding of what is inside the 'black box' so they can use commercial data mining packages discriminatingly, as well as enabling advanced readers or academic researchers to understand or contribute to future technical advances in the field. Each chapter has practical exercises to enable readers to check their progress. A full glossary of technical terms used is included. This expanded third edition includes detailed descriptions of algorithms for classifying streaming data, both stationary data, where the underlying model is fixed, and data that is time-dependent, where the underlying model changes from time to time - a phenomenon known as concept drift.

a tutorial on principal components analysis: Applications and Innovations in Intelligent Systems XIII Ann Macintosh, Richard Ellis, Tony Allen, 2007-10-27 The papers in this volume are the refereed application papers presented at AI-2005, the Twenty-fifth SGAI International Conference on Innovative Techniques and Applications of Artificial Intelligence, held in Cambridge in December 2005. The papers present new and innovative developments in the field, divided into sections on Synthesis and Prediction, Scheduling and Search, Diagnosis and Monitoring, Classification and Design, and Analysis and Evaluation. This is the thirteenth volume in the Applications and Innovations series. The series serves as a key reference on the use of AI Technology to enable organisations to solve complex problems and gain significant business benefits. The Technical Stream papers are published as a companion volume under the title Research and Development in Intelligent Systems XXII.

a tutorial on principal components analysis: Biplots in Practice Michael J. Greenacre, 2010 Este libro explica las aplicaciones específicas y las interpretaciones del biplot en muchas áreas del análisis multivariante. regresión, modelos lineales generalizados, análisis de componentes principales, análisis de correspondencias y análisis discriminante.

a tutorial on principal components analysis: Introduction to Environmental Forensics Brian L. Murphy, Robert D. Morrison, 2014-07-30 The third edition of Introduction to Environmental Forensics is a state-of-the-art reference for the practicing environmental forensics consultant, regulator, student, academic, and scientist, with topics including compound-specific isotope analysis (CSIA), advanced multivariate statistical techniques, surrogate approaches for contaminant source identification and age dating, dendroecology, hydrofracking, releases from underground storage tanks and piping, and contaminant-transport modeling for forensic applications. Recognized international forensic scientists were selected to author chapters in their specific areas of expertise and case studies are included to illustrate the application of these methods in actual environmental forensic investigations. This edition provides updates on advances in various techniques and introduces several new topics. - Provides a comprehensive review of all aspects of environmental forensics - Coverage ranges from emerging statistical methods to state-of-the-art analytical techniques, such as gas chromatography-combustion-isotope ratio mass spectrometry and polytopic vector analysis - Numerous examples and case studies are provided to illustrate the application of these forensic techniques in environmental investigations

a tutorial on principal components analysis: Principal Manifolds for Data Visualization and Dimension Reduction Alexander N. Gorban, Balázs Kégl, Donald C. Wunsch, Andrei Zinovyev, 2007-09-11 The book starts with the quote of the classical Pearson definition of PCA and includes reviews of various methods: NLPCA, ICA, MDS, embedding and clustering algorithms, principal manifolds and SOM. New approaches to NLPCA, principal manifolds, branching principal components and topology preserving mappings are described. Presentation of algorithms is supplemented by case studies. The volume ends with a tutorial PCA deciphers genome.

a tutorial on principal components analysis: Statistical Parametric Mapping: The Analysis of Functional Brain Images William D. Penny, Karl J. Friston, John T. Ashburner, Stefan J. Kiebel, Thomas E. Nichols, 2011-04-28 In an age where the amount of data collected from brain imaging is increasing constantly, it is of critical importance to analyse those data within an accepted framework to ensure proper integration and comparison of the information collected. This book describes the ideas and procedures that underlie the analysis of signals produced by the brain. The aim is to understand how the brain works, in terms of its functional architecture and dynamics. This book provides the background and methodology for the analysis of all types of brain imaging data, from functional magnetic resonance imaging to magnetoencephalography. Critically, Statistical Parametric Mapping provides a widely accepted conceptual framework which allows treatment of all these different modalities. This rests on an understanding of the brain's functional anatomy and the way that measured signals are caused experimentally. The book takes the reader from the basic concepts underlying the analysis of neuroimaging data to cutting edge approaches that would be difficult to find in any other source. Critically, the material is presented in an incremental way so that the reader can understand the precedents for each new development. This book will be particularly useful to neuroscientists engaged in any form of brain mapping; who have to contend with the real-world problems of data analysis and understanding the techniques they are using. It is primarily a scientific treatment and a didactic introduction to the analysis of brain imaging data. It can be used as both a textbook for students and scientists starting to use the techniques, as well as a reference for practicing neuroscientists. The book also serves as a companion to the software packages that have been developed for brain imaging data analysis. - An essential reference and companion for users of the SPM software - Provides a complete description of the concepts and procedures entailed by the analysis of brain images - Offers full didactic treatment of the basic mathematics behind the analysis of brain imaging data - Stands as a compendium of all the advances in neuroimaging data analysis over the past decade - Adopts an easy to understand and incremental approach that takes the reader from basic statistics to state of the art approaches such as Variational Bayes - Structured treatment of data analysis issues that links different modalities and models - Includes a series of appendices and tutorial-style chapters that makes even the most sophisticated approaches accessible

a tutorial on principal components analysis: Generalized Principal Component Analysis René Vidal, Yi Ma, Shankar Sastry, 2016-04-11 This book provides a comprehensive introduction to the latest advances in the mathematical theory and computational tools for modeling high-dimensional data drawn from one or multiple low-dimensional subspaces (or manifolds) and potentially corrupted by noise, gross errors, or outliers. This challenging task requires the development of new algebraic, geometric, statistical, and computational methods for efficient and robust estimation and segmentation of one or multiple subspaces. The book also presents interesting real-world applications of these new methods in image processing, image and video segmentation, face recognition and clustering, and hybrid system identification etc. This book is intended to serve as a textbook for graduate students and beginning researchers in data science, machine learning, computer vision, image and signal processing, and systems theory. It contains ample illustrations, examples, and exercises and is made largely self-contained with three Appendices which survey basic concepts and principles from statistics, optimization, and algebraic-geometry used in this book. René Vidal is a Professor of Biomedical Engineering and Director of the Vision Dynamics and Learning Lab at The Johns Hopkins University. Yi Ma is Executive Dean and Professor at the School of Information Science and Technology at ShanghaiTech University. S. Shankar Sastry is Dean of the College of Engineering, Professor of Electrical Engineering and Computer Science and Professor of Bioengineering at the University of California, Berkeley.

a tutorial on principal components analysis: Multivariate Statistics for Wildlife and Ecology

Research Kevin McGarigal, Samuel A. Cushman, Susan Stafford, 2013-12-01 With its focus on the practical application of the techniques of multivariate statistics, this book shapes the powerful tools of statistics for the specific needs of ecologists and makes statistics more applicable to their course of study. It gives readers a solid conceptual understanding of the role of multivariate statistics in ecological applications and the relationships among various techniques, while avoiding detailed mathematics and the underlying theory. More importantly, the reader will gain insight into the type of research questions best handled by each technique and the important considerations in applying them. Whether used as a textbook for specialised courses or as a supplement to general statistics texts, the book emphasises those techniques that students of ecology and natural resources most need to understand and employ in their research. While targeted for upper-division and graduate students in wildlife biology, forestry, and ecology, and for professional wildlife scientists and natural resource managers, this book will also be valuable to researchers in any of the biological sciences.

a tutorial on principal components analysis: Hands-On Machine Learning with R Brad Boehmke, Brandon M. Greenwell, 2019-11-07 Hands-on Machine Learning with R provides a practical and applied approach to learning and developing intuition into today's most popular machine learning methods. This book serves as a practitioner's guide to the machine learning process and is meant to help the reader learn to apply the machine learning stack within R, which includes using various R packages such as glmnet, h2o, ranger, xgboost, keras, and others to effectively model and gain insight from their data. The book favors a hands-on approach, providing an intuitive understanding of machine learning concepts through concrete examples and just a little bit of theory. Throughout this book, the reader will be exposed to the entire machine learning process including feature engineering, resampling, hyperparameter tuning, model evaluation, and interpretation. The reader will be exposed to powerful algorithms such as regularized regression, random forests, gradient boosting machines, deep learning, generalized low rank models, and more! By favoring a hands-on approach and using real word data, the reader will gain an intuitive understanding of the architectures and engines that drive these algorithms and packages, understand when and how to tune the various hyperparameters, and be able to interpret model results. By the end of this book, the reader should have a firm grasp of R's machine learning stack and be able to implement a systematic approach for producing high quality modeling results. Features: · Offers a practical and applied introduction to the most popular machine learning methods. · Topics covered include feature engineering, resampling, deep learning and more. · Uses a hands-on approach and real world data.

a tutorial on principal components analysis: *Mathematics for Machine Learning* Marc Peter Deisenroth, A. Aldo Faisal, Cheng Soon Ong, 2020-04-23 Distills key concepts from linear algebra, geometry, matrices, calculus, optimization, probability and statistics that are used in machine learning.

a tutorial on principal components analysis: Forecasting: principles and practice Rob J Hyndman, George Athanasopoulos, 2018-05-08 Forecasting is required in many situations. Stocking an inventory may require forecasts of demand months in advance. Telecommunication routing requires traffic forecasts a few minutes ahead. Whatever the circumstances or time horizons involved, forecasting is an important aid in effective and efficient planning. This textbook provides a comprehensive introduction to forecasting methods and presents enough information about each method for readers to use them sensibly.

a tutorial on principal components analysis: *Probabilistic Machine Learning* Kevin P. Murphy, 2022-03-01 A detailed and up-to-date introduction to machine learning, presented through the unifying lens of probabilistic modeling and Bayesian decision theory. This book offers a detailed and up-to-date introduction to machine learning (including deep learning) through the unifying lens of probabilistic modeling and Bayesian decision theory. The book covers mathematical background (including linear algebra and optimization), basic supervised learning (including linear and logistic regression and deep neural networks), as well as more advanced topics (including transfer learning and unsupervised learning). End-of-chapter exercises allow students to apply what they have learned, and an appendix covers notation. Probabilistic Machine Learning grew out of the author's 2012 book, Machine Learning: A Probabilistic Perspective. More than just a simple update, this is a completely new book that reflects the dramatic developments in the field since 2012, most notably deep learning. In addition, the new book is accompanied by online Python code, using libraries such as scikit-learn, JAX, PyTorch, and Tensorflow, which can be used to reproduce nearly all the figures; this code can be run inside a web browser using cloud-based notebooks, and provides a practical complement to the theoretical topics discussed in the book. This introductory text will be followed by a sequel that covers more advanced topics, taking the same probabilistic approach.

a tutorial on principal components analysis: An Easy Guide to Factor Analysis Paul Kline, 2014-02-25 Factor analysis is a statistical technique widely used in psychology and the social sciences. With the advent of powerful computers, factor analysis and other multivariate methods are now available to many more people. An Easy Guide to Factor Analysis presents and explains factor analysis as clearly and simply as possible. The author, Paul Kline, carefully defines all statistical terms and demonstrates step-by-step how to work out a simple example of principal components analysis and rotation. He further explains other methods of factor analysis, including confirmatory and path analysis, and concludes with a discussion of the use of the technique with various examples. An Easy Guide to Factor Analysis is the clearest, most comprehensible introduction to factor analysis for students. All those who need to use statistics in psychology and the social sciences will find it invaluable. Paul Kline is Professor of Psychometrics at the University of Exeter. He has been using and teaching factor analysis for thirty years. His previous books include Intelligence: the psychometric view (Routledge 1990) and The Handbook of Psychological Testing (Routledge 1992).

a tutorial on principal components analysis: *ADKAR* Jeff Hiatt, 2006 In his first complete text on the ADKAR model, Jeff Hiatt explains the origin of the model and explores what drives each building block of ADKAR. Learn how to build awareness, create desire, develop knowledge, foster ability and reinforce changes in your organization. The ADKAR Model is changing how we think about managing the people side of change, and provides a powerful foundation to help you succeed at change.

a tutorial on principal components analysis: Macroeconomic Forecasting in the Era of Big Data Peter Fuleky, 2019-11-28 This book surveys big data tools used in macroeconomic forecasting and addresses related econometric issues, including how to capture dynamic relationships among variables; how to select parsimonious models; how to deal with model uncertainty, instability, non-stationarity, and mixed frequency data; and how to evaluate forecasts, among others. Each chapter is self-contained with references, and provides solid background information, while also reviewing the latest advances in the field. Accordingly, the book offers a valuable resource for researchers, professional forecasters, and students of quantitative economics.

a tutorial on principal components analysis: *Statistics for Marketing and Consumer Research* Mario Mazzocchi, 2008-05-22 Balancing simplicity with technical rigour, this practical guide to the statistical techniques essential to research in marketing and related fields, describes each method as well as showing how they are applied. The book is accompanied by two real data sets to replicate examples and with exercises to solve, as well as detailed guidance on the use of appropriate software including: - 750 powerpoint slides with lecture notes and step-by-step guides to run analyses in SPSS (also includes screenshots) - 136 multiple choice questions for tests This is augmented by in-depth discussion of topics including: - Sampling - Data management and statistical packages - Hypothesis testing - Cluster analysis - Structural equation modelling

a tutorial on principal components analysis: MATLAB® Recipes for Earth Sciences Martin H. Trauth, Robin Gebbers, Norbert Marwan, 2007 Introduces methods of data analysis in geosciences using MATLAB such as basic statistics for univariate, bivariate and multivariate datasets, jackknife and bootstrap resampling schemes, processing of digital elevation models, gridding and contouring, geostatistics and kriging, processing and georeferencing of satellite images, digitizing from the screen, linear and nonlinear time-series analysis and the application of linear time-invariant and adaptive filters. Includes a brief description of each method and numerous examples demonstrating how MATLAB can be used on data sets from earth sciences.

a tutorial on principal components analysis: Hands-On Unsupervised Learning Using Python Ankur A. Patel, 2019-02-21 Many industry experts consider unsupervised learning the next frontier in artificial intelligence, one that may hold the key to general artificial intelligence. Since the majority of the world's data is unlabeled, conventional supervised learning cannot be applied. Unsupervised learning, on the other hand, can be applied to unlabeled datasets to discover meaningful patterns buried deep in the data, patterns that may be near impossible for humans to uncover. Author Ankur Patel shows you how to apply unsupervised learning using two simple, production-ready Python frameworks: Scikit-learn and TensorFlow using Keras. With code and hands-on examples, data scientists will identify difficult-to-find patterns in data and gain deeper business insight, detect anomalies, perform automatic feature engineering and selection, and generate synthetic datasets. All you need is programming and some machine learning experience to get started. Compare the strengths and weaknesses of the different machine learning approaches: supervised, unsupervised, and reinforcement learning Set up and manage machine learning projects end-to-end Build an anomaly detection system to catch credit card fraud Clusters users into distinct and homogeneous groups Perform semisupervised learning Develop movie recommender systems using restricted Boltzmann machines Generate synthetic images using generative adversarial networks

a tutorial on principal components analysis: Modern Statistics for Modern Biology SUSAN. HUBER HOLMES (WOLFGANG.), Wolfgang Huber, 2018

a tutorial on principal components analysis: R and Data Mining Yanchang Zhao, 2012-12-31 R and Data Mining introduces researchers, post-graduate students, and analysts to data mining using R, a free software environment for statistical computing and graphics. The book provides practical methods for using R in applications from academia to industry to extract knowledge from vast amounts of data. Readers will find this book a valuable guide to the use of R in tasks such as classification and prediction, clustering, outlier detection, association rules, sequence analysis, text mining, social network analysis, sentiment analysis, and more.Data mining techniques are growing in popularity in a broad range of areas, from banking to insurance, retail, telecom, medicine, research, and government. This book focuses on the modeling phase of the data mining process, also addressing data exploration and model evaluation.With three in-depth case studies, a quick reference guide, bibliography, and links to a wealth of online resources, R and Data Mining is a valuable, practical guide to a powerful method of analysis. - Presents an introduction into using R for data mining applications, covering most popular data mining techniques - Provides code examples and data so that readers can easily learn the techniques - Features case studies in real-world applications to help readers apply the techniques in their work

a tutorial on principal components analysis: Modern Psychometrics with R Patrick Mair, 2018-09-20 This textbook describes the broadening methodology spectrum of psychological measurement in order to meet the statistical needs of a modern psychologist. The way statistics is used, and maybe even perceived, in psychology has drastically changed over the last few years; computationally as well as methodologically. R has taken the field of psychology by storm, to the point that it can now safely be considered the lingua franca for statistical data analysis in psychology. The goal of this book is to give the reader a starting point when analyzing data using a particular method, including advanced versions, and to hopefully motivate him or her to delve deeper into additional literature on the method. Beginning with one of the oldest psychometric model formulations, the true score model, Mair devotes the early chapters to exploring confirmatory factor analysis, modern test theory, and a sequence of multivariate exploratory method. Subsequent chapters present special techniques useful for modern psychological applications including correlation networks, sophisticated parametric clustering techniques, longitudinal measurements on a single participant, and functional magnetic resonance imaging (fMRI) data. In addition to using real-life data sets to demonstrate each method, the book also reports each method in three parts--

first describing when and why to apply it, then how to compute the method in R, and finally how to present, visualize, and interpret the results. Requiring a basic knowledge of statistical methods and R software, but written in a casual tone, this text is ideal for graduate students in psychology. Relevant courses include methods of scaling, latent variable modeling, psychometrics for graduate students in Psychology, and multivariate methods in the social sciences.

a tutorial on principal components analysis: XAFS for Everyone Scott Calvin, 2013-05-20 XAFS for Everyone provides a practical, thorough guide to x-ray absorption fine-structure (XAFS) spectroscopy for both novices and seasoned practitioners from a range of disciplines. The text is enhanced with more than 200 figures as well as cartoon characters who offer informative commentary on the different approaches used in XAFS spectroscopy. The book covers sample preparation, data reduction, tips and tricks for data collection, fingerprinting, linear combination analysis, principal component analysis, and modeling using theoretical standards. It describes both near-edge (XANES) and extended (EXAFS) applications in detail. Examples throughout the text are drawn from diverse areas, including materials science, environmental science, structural biology, catalysis, nanoscience, chemistry, art, and archaeology. In addition, five case studies from the literature demonstrate the use of XAFS principles and analysis in practice. The text includes derivations and sample calculations to foster a deeper comprehension of the results. Whether you are encountering this technique for the first time or looking to hone your craft, this innovative and engaging book gives you insight on implementing XAFS spectroscopy and interpreting XAFS experiments and results. It helps you understand real-world trade-offs and the reasons behind common rules of thumb.

a tutorial on principal components analysis: Natural Language Processing with Python Steven Bird, Ewan Klein, Edward Loper, 2009-06-12 This book offers a highly accessible introduction to natural language processing, the field that supports a variety of language technologies, from predictive text and email filtering to automatic summarization and translation. With it, you'll learn how to write Python programs that work with large collections of unstructured text. You'll access richly annotated datasets using a comprehensive range of linguistic data structures, and you'll understand the main algorithms for analyzing the content and structure of written communication. Packed with examples and exercises, Natural Language Processing with Python will help you: Extract information from unstructured text, either to guess the topic or identify named entities Analyze linguistic structure in text, including parsing and semantic analysis Access popular linguistic databases, including WordNet and treebanks Integrate techniques drawn from fields as diverse as linguistics and artificial intelligence This book will help you gain practical skills in natural language processing using the Python programming language and the Natural Language Toolkit (NLTK) open source library. If you're interested in developing web applications, analyzing multilingual news sources, or documenting endangered languages -- or if you're simply curious to have a programmer's perspective on how human language works -- you'll find Natural Language Processing with Python both fascinating and immensely useful.

a tutorial on principal components analysis: Linear Algebra and Its Applications David C. Lay, 2013-07-29 NOTE: This edition features the same content as the traditional text in a convenient, three-hole-punched, loose-leaf version. Books a la Carte also offer a great value--this format costs significantly less than a new textbook. Before purchasing, check with your instructor or review your course syllabus to ensure that you select the correct ISBN. Several versions of Pearson's MyLab & Mastering products exist for each title, including customized versions for individual schools, and registrations are not transferable. In addition, you may need a CourseID, provided by your instructor, to register for and use Pearson's MyLab & Mastering products. xxxxxxxxxxxxxxx For courses in linear algebra.This package includes MyMathLab(R). With traditional linear algebra texts, the course is relatively easy for students during the early stages as material is presented in a familiar, concrete setting. However, when abstract concepts are introduced, students often hit a wall. Instructors seem to agree that certain concepts (such as linear independence, spanning, subspace, vector space, and linear transformations) are not easily understood and require time to assimilate. These concepts are fundamental to the study of linear algebra, so students' understanding of them is vital to mastering the subject. This text makes these concepts more accessible by introducing them early in a familiar, concrete Rn setting, developing them gradually, and returning to them throughout the text so that when they are discussed in the abstract, students are readily able to understand. Personalize learning with MyMathLabMyMathLab is an online homework, tutorial, and assessment program designed to work with this text to engage students and improve results. MyMathLab includes assignable algorithmic exercises, the complete eBook, interactive figures, tools to personalize learning, and more.

a tutorial on principal components analysis: Nature of Computation and Communication Phan Cong Vinh, Nguyen Huu Nhan, 2022-01-04 This book constitutes the refereed post-conference proceedings of the 7th International Conference on Nature of Computation and Communication, ICTCC 2021, held in October 2021. Due to COVID-19 pandemic the conference was held virtually. The 17 revised full papers presented were carefully selected from 43 submissions. The papers of ICTCC 2021 cover formal methods for self-adaptive systems and discuss natural approaches and techniques for natural computing systems and their applications.

a tutorial on principal components analysis: Data Mining Florin Gorunescu, 2011-03-10 The knowledge discovery process is as old as Homo sapiens. Until some time ago this process was solely based on the ‘natural personal' computer provided by Mother Nature. Fortunately, in recent decades the problem has begun to be solved based on the development of the Data mining technology, aided by the huge computational power of the 'artificial' computers. Digging intelligently in different large databases, data mining aims to extract implicit, previously unknown and potentially useful information from data, since “knowledge is power”. The goal of this book is to provide, in a friendly way, both theoretical concepts and, especially, practical techniques of this exciting field, ready to be applied in real-world situations. Accordingly, it is meant for all those who wish to learn how to explore and analysis of large quantities of data in order to discover the hidden nugget of information.

a tutorial on principal components analysis: Advances in Computer Graphics Tomoyuki Nishita, Qunsheng Peng, 2006-06-22 This is the refereed proceedings of the 24th Computer Graphics International Conference, CGI 2006. The 38 revised full papers and 37 revised short papers presented were carefully reviewed. The papers are organized in topical sections on rendering and texture, efficient modeling and deformation, digital geometry processing, shape matching and shape analysis, face, virtual reality, motion and image, as well as CAGD.

Additional Resources

A Tutorial on Principal Component Analysis
This tutorial focuses on building a solid intuition for how and why principal component analysis works; furthermore, it crystallizes this knowledge by deriving from simple intuitions, the …

Lecture Notes on Principal Component Analysis
The task of principal component analysis (PCA) is to reduce the dimensionality of some high-dimensional data points by linearly projecting them onto a lower-dimensional space in such a …

A tutorial on Principal Components Analysis - Otago
This tutorial is designed to give the reader an understanding of Principal Components Analysis (PCA). PCA is a useful statistical technique that has found application in fields such as face …

A tutorial on principal component analysis - compute.dtu.dk
Principal Component Analysis (PCA) learning objectives Describe the concept

of principal component analysis Explain why principal component analysis can be beneficial when there is …

Principal Component Analysis (PCA) - Stony Brook University
Principal Component Analysis (PCA) takes a data matrix of n objects by p variables, which may be correlated, and summarizes it by uncorrelated axes (principal components or principal …

PRINCIPAL COMPONENTS ANALYSIS (PCA)
The formal name for this approach of rotating data such that each successive axis displays a decreasing among of variance is known as Principal Components Analysis, or PCA. PCA …

A Tutorial on Principal Component Analysis
Principal component analysis (PCA) is a standard tool in mod-ern data analysis - in diverse fields from neuroscience to com-puter graphics - because it is a simple, non-parametric method for …

Lecture 15: Principal Component Analysis - Duke University
Principal Component Analysis, or simply PCA, is a statistical procedure concerned with elucidating the covari-ance structure of a set of variables. In particular it allows us to identify …

Principal Components Analysis (PCA) in Matlab
Each column corresponds to a principal component. Variance explained is used when deciding how many PCs to keep. Statistic measuring how far each observation is from the "center" of …

Lecture 14 Principal Component Analysis
Principal component analysis (PCA) is one of the most valuable results of applied linear algebra. It is widely used { from neuroscience to computer graphics { because it is an easy way to …

PRINCIPAL COMPONENTS ANALYSIS (PCA)
The formal name for this approach of rotating data such that each successive axis displays a decreasing among of variance is known as Principal Components Analysis, or PCA. PCA …

Tutorial on Principal Component Analysis
2.12 Interpretation Principal component analysis models X as a linear combination of uncorrelated hidden sources, which are called the principal components. If our goal is to decompose X into …

A Tutorial on Principal Component Analysis - Nottingham
Principal component analysis (PCA) is a standard tool in mod-ern data analysis - in diverse fields from neuroscience to com-puter graphics - because it is a simple, non-parametric method for …

Principal Components Analysis with Spatial Data - BioMedware
This tutorial will undertake a Principal Components Analysis (PCA) of geographically distributed data in SpaceStat. The data are homeownership and socioeconomic data for the state of …

BEGINNER'S GUIDE TO PRINCIPAL COMPONENT ANALYS
te the impact of individual variables on subjects. PCA is a statistical technique that transforms the original correlated variables into a new set of variables called the principal components, which …

Principal Component Analysis - Duke University

Principal Component Analysis (PCA) is the general name for a technique which uses sophis-ticated underlying mathematical principles to transforms a number of possibly correlated …

A TUTORIAL ON PRINCIPAL COMPONENT ANALYSIS
This tutorial focuses on building a solid intuition for how and why principal component analysis works; furthermore, it crystallizes this knowledge by deriving from first prin-cipals, the …

A Tutorial on Principal Component Analysis
This manuscript focuses on building a solid intuition for how and why principal component analysis works. This manuscript crystallizes this knowledge by deriving from simple intuitions, …

Principal Component Analysis and Optimization: A Tutorial
Principal component analysis (PCA) is one of the most widely used multivariate techniques in statistics. It is commonly used to reduce the dimensionality of data in order to examine its …

Principal Component Analysis (PCA) Tutorial
PCA finds new variables, called principal components, that are linear combinations of the original variables, capturing the directions of maximum variance in the data. This technique is widely …

A Tutorial on Principal Component Analysis
This tutorial focuses on building a solid intuition for how and why principal component analysis works; furthermore, it crystallizes this knowledge by deriving from simple intuitions, the …

Lecture Notes on Principal Component Analysis
The task of principal component analysis (PCA) is to reduce the dimensionality of some high-dimensional data points by linearly projecting them onto a lower-dimensional space in such a …

A tutorial on Principal Components Analysis - Otago
This tutorial is designed to give the reader an understanding of Principal Components Analysis (PCA). PCA is a useful statistical technique that has found application in fields such as face …

A tutorial on principal component analysis - compute.dtu.dk
Principal Component Analysis (PCA) learning objectives Describe the concept of principal component analysis Explain why principal component analysis can be beneficial when there is …

Principal Component Analysis (PCA) - Stony Brook University
Principal Component Analysis (PCA) takes a data matrix of n objects by p variables, which may be correlated, and summarizes it by uncorrelated axes (principal components or principal …

PRINCIPAL COMPONENTS ANALYSIS (PCA)
The formal name for this approach of rotating data such that each successive axis displays a decreasing among of variance is known as Principal Components Analysis, or PCA. PCA …

A Tutorial on Principal Component Analysis
Principal component analysis (PCA) is a standard tool in mod-ern data analysis - in diverse fields from neuroscience to com-puter graphics - because it is a simple, non-parametric method for …

Lecture 15: Principal Component Analysis - Duke University
Principal Component Analysis, or simply PCA, is a statistical procedure concerned with elucidating the covari-ance structure of a set of variables. In particular it allows us to identify …

Principal Components Analysis (PCA) in Matlab
Each column corresponds to a principal component. Variance explained is used when deciding how many PCs to keep. Statistic measuring how far each observation is from the “center” of …

Lecture 14 Principal Component Analysis
Principal component analysis (PCA) is one of the most valuable results of applied linear algebra. It is widely used { from neuroscience to computer graphics { because it is an easy way to …

PRINCIPAL COMPONENTS ANALYSIS (PCA)
The formal name for this approach of rotating data such that each successive axis displays a decreasing among of variance is known as Principal Components Analysis, or PCA. PCA …

Tutorial on Principal Component Analysis
2.12 Interpretation Principal component analysis models X as a linear combination of uncorrelated hidden sources, which are called the principal components. If our goal is to decompose X into …

A Tutorial on Principal Component Analysis - Nottingham
Principal component analysis (PCA) is a standard tool in mod-ern data analysis - in diverse fields from neuroscience to com-puter graphics - because it is a simple, non-parametric method for …

Principal Components Analysis with Spatial Data
This tutorial will undertake a Principal Components Analysis (PCA) of geographically distributed data in SpaceStat. The data are homeownership and socioeconomic data for the state of …

BEGINNER’S GUIDE TO PRINCIPAL COMPONENT ANALYS
te the impact of individual variables on subjects. PCA is a statistical technique that transforms the original correlated variables into a new set of variables called the principal components, which …

Principal Component Analysis - Duke University
Principal Component Analysis (PCA) is the general name for a technique which uses sophis-ticated underlying mathematical principles to transforms a number of possibly correlated …

A TUTORIAL ON PRINCIPAL COMPONENT ANALYSIS
This tutorial focuses on building a solid intuition for how and why principal component analysis works; furthermore, it crystallizes this knowledge by deriving from first prin-cipals, the …

A Tutorial on Principal Component Analysis
This manuscript focuses on building a solid intuition for how and why principal component analysis works. This manuscript crystallizes this knowledge by deriving from simple intuitions, …

Principal Component Analysis and Optimization: A Tutorial
Principal component analysis (PCA) is one of the most widely used multivariate techniques in statistics. It is commonly used to reduce the dimensionality of data in order to examine its …

Principal Component Analysis (PCA) Tutorial
PCA finds new variables, called principal components, that are linear combinations of the original variables, capturing the directions of maximum variance in the data. This technique is widely …

A Tutorial On Principal Components Analysis

A Tutorial On Principal Components Analysis

Related Articles

- a theoretical analysis of contrastive unsupervised representation learning
- a regional bank implemented an automated solution to streamline brainly
- a union b intersection c venn diagram

Back to Home