IWMS-2019

The 27th International Workshop
in Matrices and Statistics

Workshop Booklet

6-9 June 2019 | Shanghai-China

Warm welcome to the 27th International Workshop on Matrices and Statistics (IWMS): Shanghai, China (6–9 June 2019)

As Honorary Chair of the International Organizing Committee (IOC) of the 27th International Workshop on Matrices and Statistics (IWMS) in Shanghai, China (6–9 June 2019), I am very pleased to extend a warm welcome to all the participants. Special thanks go to IOC Chair Jeffrey J. Hunter and to Yonghui Liu, Chair of the Local Organizing Committee (LOC) and their teams. I also thank Ka Lok Chu and Simo Puntanen for their help with the 4th international mini-symposium on "Magic squares, prime numbers and postage stamps" (IWMS-27/2019-M4) and, in particular, for our joint Poster 1: "An introduction to some magic squares by Paul Daniels and by Steve Martin, and to the Kostabi/Leigh Bereshit bara Elohim drawing Nova Ratio for Pope Emeritus Benedict XVI, all illustrated philatelically". This poster is part of our ongoing McGill University Dept. of Mathematics and Statistics Report 2019-01: "Some magic squares, magic hexagons and some prime-friendly wines with special emphasis on the prime number 19 illustrated philatelically", by Ka Lok Chu, Simo Puntanen & George P. H. Styan.

As observed online at `https://www.sis.uta.fi/tilasto/iwms/IWMS-history.pdf` in "A short history of the International Workshop on Matrices and Statistics (IWMS)" by Simo Puntanen and George P. H. Styan, the first IWMS took place almost 30 years ago at the University of Tampere in Tampere, Finland (6–8 August 1990) and the most recent IWMS was held almost exactly a year ago in the Multimedia Centre at Dawson College, Westmount/Montréal (Québec), Canada (5–7 June 2018). The next IWMS will almost certainly be held in the Manipal Academy of Higher Education (formerly known as Manipal University), Karnataka State, South India, in December 2020.

We have established an open-access website online at `http://www.sis.uta.fi/tilasto/iwms/` for the IWMS at the University of Tampere, where our aim is to put all associated reports and photographs of the IWMS series from 1990 onwards, including those published in *Image: The Bulletin of the International Linear Algebra Society*. Complete videos, prepared by Jarmo Niemelä and Reijo Sund, of the talks at two pre-IWMS Tampere conferences in statistics in 1987 and 1990 are on YouTube.

A very warm welcome to you all!

George P. H. Styan, Honorary Chair IWMS-IOC
Professor Emeritus of Msthematics and Statistics, McGill University, Montréal (Québec), Canada.

The 27th International Workshop on Matrices and Statistics

Workshop Booklet

Sponsor: Shanghai University of International Business and Economics,
Shanghai, China

International Organizing Committee (IOC)

- Jeffrey J. Hunter (New Zealand) (Chair)
- Dietrich von Rosen (Sweden)(Vice-Chair)
- George P. H. Styan (Canada) (Honorary Chair)
- S. Ejaz Ahmed (Canada)
- Francisco Carvalho (Portugal)
- Katarzyna Filipiak (Poland)
- Daniel Klein (Slovakia)
- Augustyn Markiewicz (Poland)
- Simo Puntanen (Finland)
- Julia Volaufova (USA)
- Hans Joachim Werner (Germany)

Local Organizing Committee (LOC)

- Yonghui Liu (Chair)
- Hui Liu (Vice-Chair)
- Chengcheng Hao
- Cihai Sun

Contents

Preface of Abstract and Program Booklet 5
Expression of Thanks 6
Workshop Venue 7
Schedule of IWMS-2019 9
    Thursday, 6 June 2019 9
    Friday, 7 June 2019 10
    Saturday, 8 June 2019 12
    Sunday, 9 June 2019 13
Abstracts 15
List of Participants 61
Introduction to Shanghai University of International Business and Economics 64
Introduction to the School of Statistics and Information 65

Preface of Abstract and Program Booklet

On behalf of the International Organizing Committee (IOC) I have much pleasure in welcoming you to the 27th International Workshop on Matrices and Statistics, IWMS-2019. The IOC is grateful for the offer by Professor Yonghui Liu to host this event in Shanghai over the period 6-9 June 2019 at Shanghai University of International Business and Economics. We appreciate all the efforts that his team have expended to ensure that your participation is worthwhile.

The purpose of these Workshops is "to stimulate research and, in an informal setting, to foster the interaction of researchers in the interface between statistics and matrix theory. The Workshop will provide a forum through which statisticians may be better informed of the latest developments and newest techniques in linear algebra and matrix theory and may exchange ideas with researchers from a wide variety of countries"

We have structured the workshop along the lines that we introduced at IWMS-2015 by having a number of Plenary Speakers, leaders in a variety of fields that comprise the themes of the workshop, together with a number of Mini-symposia that bring together a number of researchers working in related cognate areas. The aim of this is to foster interaction, not only between the speakers, but also between the Workshop participants. We have also encouraged participants to offer contributed talks in order to give them the opportunity to advise the statistical and matrix theory communities of current research being undertaken in these areas. We have striven to ensure that Workshop format does not, where possible, have overlapping presentations so that everyone has the opportunity to share in all the talks.

I do hope that you find the Workshop a valuable opportunity to interact with others working at the cutting edge of research in their research fields. I encourage you all meet with others and share the experiences with each other. Above all, I wish you all an enjoyable time together.

Jeffrey J Hunter

Chair of the International Organizing Committee
Professor Emeritus of Mathematics, Auckland University of Technology, New Zealand
Professor Emeritus of Statistics, Massey University, New Zealand
Email addresses:
jeffreyjhunter@gmail.com, jeffrey.hunter@aut.ac.nz, j.hunter@massey.ac.nz

Expression of Thanks

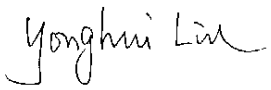
As the Chair of Local Organizing Committee (LOC) of the "27th International Workshop on Matrices and Statistics" (IWMS-2019), I wish to extend warmest welcome and best wishes to all the mathematicians, statisticians and data scientists for participating in this workshop.

It has been nine years since IWMS-2010 has held in Shanghai. Shanghai economy has changed a lot over the last ten years, so does our subject. It is always a pleasure to meet old and new friends and exchange latest developments in matrix theory and statistics from all over the world. We thank members of IOC to provide us such a chance. Special thanks go to Prof. Jeffrey J. Hunter, chair of IOC, Prof. Dietrich von Rosen, vice chair of IOC and Dr. Simo Puntanen for supporting us during preparing this workshop. We also appreciate Prof. Kai-Tai Fang, Prof. Julia Volaufova, Dr. Simo Puntanen, Prof. Dietrich von Rosen, Prof. Shuangzhe Liu for their great efforts to organize mini-symposia and invite 27 outstanding invited speakers in specific areas. We thank all plenary speakers and chairs of sessions for their attendance and fully supports. We thank Prof. George Styan for his preparation of excellent posters in Magic Squares and Postage Stamps. We thank the leaders of Shanghai University of International Business and Economics for the fund support to prepare this workshop smoothly. Also I would like to thank my colleagues of School of Statistics and Information, Dr. Chengcheng Hao, Dr. Hui Liu, Mr. Cihai Sun, Dr. Jialin Chen, Dr. Rui Li, Mr. Baoxingbang Xiao, Ms. Jie Chen, Ms. Chen Yang, Ms. Yue Zheng etc. for their efforts of preparing this workshop!

Sincerely hope that all participants have a nice stay in Shanghai!

Thank you!

Yonghui Liu



Chair of the Local Organizing Committee of IWMS-2019
School of Statistics and Information
Shanghai University of International Business and Economics
liuyh@lsec.cc.ac.cn

Workshop Venue

The Workshop will take place at the Room 317 Zonghe Building of Shanghai University of International Business and Economics Gubei Campus (上海对外经贸大学-古北校区-综合楼317), located in the center of Hongqiao business area.

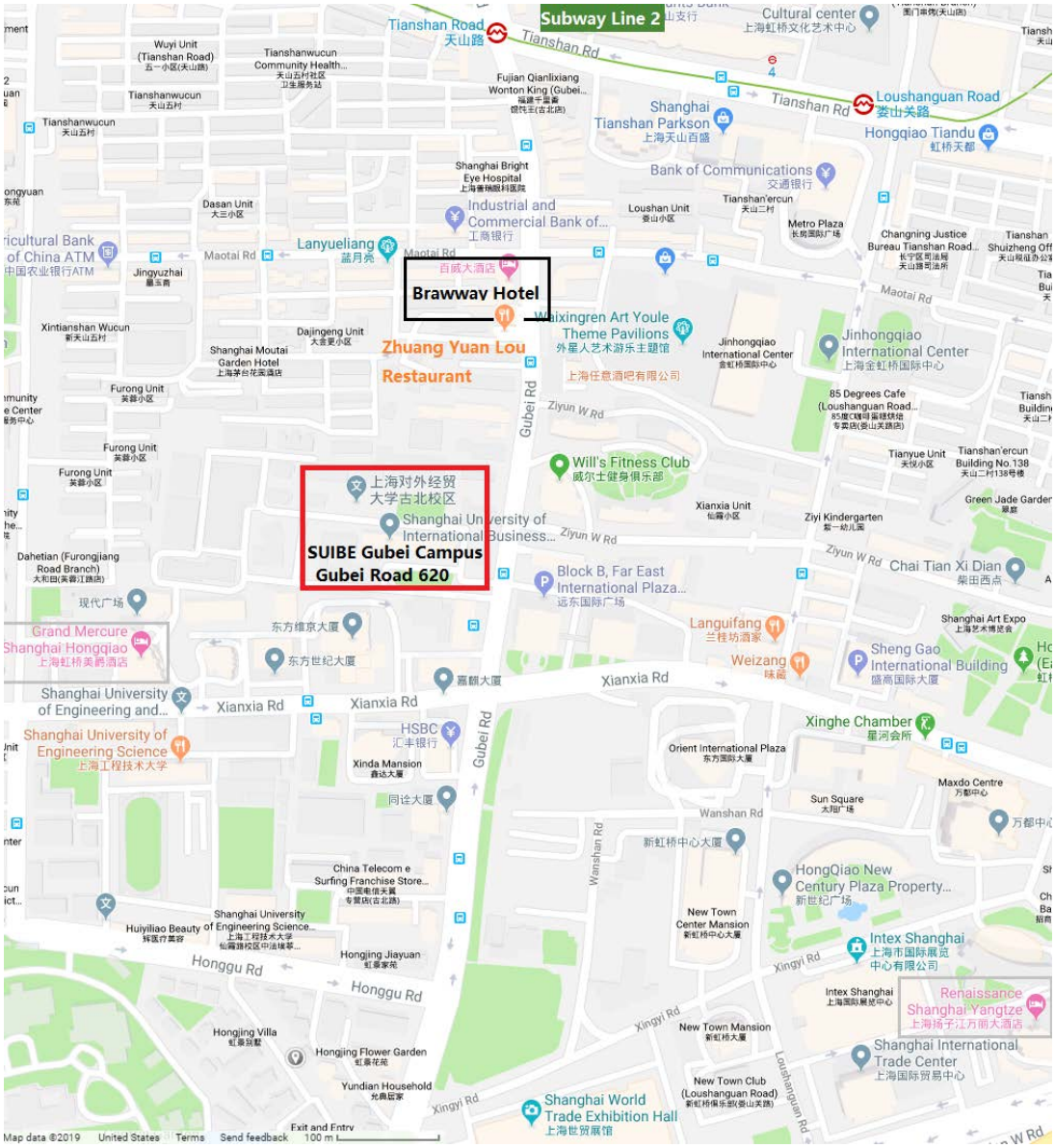
It's 10 minutes' drive from Hongqiao Airport, Hongqiao railway station, 45 minutes from Pudong International Airport, and near metro network.

Workshop Venue

Reporting room 317, Zonghe Building

Dinning

Dinning room on the 2nd floor of the Zonghe Building


Subway Line 2
Tianshan Road
天山路
Tianshan Rd
Cultural center
上海虹桥文化艺术中心
Loushanguan Road
娄山关路
Hongqiao Tiandu
虹桥天都
Shanghai Tianshan Parkson
上海天山百盛
Bank of Communications
交通银行
Shanghai Bright Eye Hospital
上海普瑞眼科医院
Industrial and Commercial Bank of...
工商银行
Lanyueliang
蓝月亮
Maotai Rd
Brawway Hotel
Zhuang Yuan Lou Restaurant
Gubei Rd
Ziyun W Rd
Jinhongqiao International Center
上海金虹桥国际中心
Will's Fitness Club
威尔士健身俱乐部
上海对外经贸大学古北校区
Shanghai University of International Business...
SUIBE Gubei Campus
Gubei Road 620
Block B, Far East International Plaza...
远东国际广场
Languifang
兰桂坊酒家
Weizang
味酱
Sheng Gao International Building
盛高国际大厦
Xianxia Rd
Shanghai University of Engineering and...
Shanghai University of Engineering Science
上海工程技术大学
HSBC
汇丰银行
Xinghe Chamber
星河会所
Orient International Plaza
东方国际大厦
Wanshan Rd
新虹桥中心大厦
HongQiao New Century Plaza Property...
新世纪广场
Intex Shanghai
上海市国际展览中心有限公司
Renaissance Shanghai Yangtze
上海扬子江万丽大酒店
Shanghai International Trade Center
上海国际贸易中心
Shanghai World Trade Exhibition Hall
上海世贸展馆
Honggu Rd
Hongjing Villa
Hongjing Flower Garden
Yundian Household
Xingyi Rd
Loushanguan Rd
Map data ©2019 United States Terms Send feedback 100 m

Schedule of IWMS-2019

Thursday, 6 June 2019

Opening Session
Thursday, 09:30-10:00
Speech from **Yonglin Xu** (Vice President, SUIBE)
Speech from **Jeffrey Hunter** (Chair, IWMS IOC)
Chair: Yonghui Liu

Followed by Group Photo

Plenary Sessions PS1-PS2
Thursday, 10:00-11:45
Chair: Jeffrey J. Hunter

10:00-10:40 **Rajendra Bhatia** (Ashoka University, New Delhi, India)
"Geometry and means of positive definite matrices"

10:40-11:10 Tea Break

11:10-11:50 **Kai-Tai Fang** (BNU-HKBU United International College, China)
"Representative points of elliptically symmetric distributions"

12:00-13:00 Lunch

Plenary Session PS3
Thursday, 13:00-13:40
Chair: K.M. Prasad

13:00-13:40 **Lixing Zhu** (Hong Kong Baptist University, Hong Kong, China)
"Order determination for large-dimensional matrices"

MS1. Experimental Design
Thursday, 13:40-16:00

Chair and Organizer: Kai-Tai Fang

13:40-14:10 **A.M. Elsawah** (BNU-HKBU United International College, China)
"Building some bridges among various experimental designs"
14:10-14:40 **Yu Tang** (Soochow University, China)
"Uniform design on general domain"

14:40-15:00 Tea Break

15:00-15:30 **Aijun Zhang** (The University of Hong Kong, Hong Kong, China)
"Data-driven Space-filling Design"
15:30-16:00 **Yongdao Zhou** (Nankai University, China)
"Orthogonal uniform composite designs"

Contributory Session CS1
Thursday, 16:00-17:20
Chair: Oskar Maria Baksalary
16:00-16:20 **Mika Mattila** (University of Tampere, Finland)
"Singularity of LCM matrices on GCD closed sets with 9 elements"

16:20-16:40 **Jorge Delgado** (Universidad de Zaragoza, Spain)
"Accurate computations for some subclasses of totally positive matrices"

16:40-17:00 **Dragana Cvetkovic Ilic** (University of Nis, Serbia)
"Completion problems on operator matrices and its various applications"

17:00-17:20 **Kehui Chen** (University of Pittsburgh, Pittsburgh, USA)
"Consistent community detection in multi-layer network data"

Poster by George P.H. Styan (joint work with Ka Lok Chu and Simo Puntanen)
Available during the workshop days.
"Magic squares, prime numbers and postage stamps"

17:30-19:00 Dinner

Friday, 7 June 2019

Plenary Sessions PS4-PS7
Friday, 09:00-12:00
Chair: Jeffrey J. Hunter

09:00-09:40 **Oskar Maria Baksalary** (Adam Mickiewicz University in Poznan, Poland)
"A gaze at recent applications and characterizations of the Moore-Penrose inverse"

09:40-10:20 **K.M. Prasad** (Manipal Academy of Higher Education, Manipal, India)
"Inverse complimentary matrix method and its applications to general linear model"

10:20-10:40 Tea Break

10:40-11:20 **Yongge Tian** (Central University of Finance and Economics, Beijing, China)
"Identifying conditions for multilinear matrix equations to always hold with applications"

11:20-12:00 **Jianxin Pan** (Manchester University, United Kingdom)
"Calibration for non-positive definite covariance matrix"

12:00-13:00 Lunch

MS2. Inference in Parametric Models
Friday, 13:00-14:30
Chair and Organizer: Julia Volaufova

13:00-13:30 **Yuli Liang** (Örebro University, Sweden)
"Two-sample correlation parameter testing in models with a Kronecker product covariance structure"

13:30-14:00 **Lynn Roy LaMotte** (Louisiana State University Health Sciences Center, USA)
"ANOVA SSs and proportional subclass numbers"

14:00-14:30 **Julia Volaufova** (Louisiana State University Health Sciences Center, USA)
"Comment on inference in a simple linear random coefficient model with missing covariates"

14:30-15:00 Tea Break

MS3. Linear Models and Multivariate Analysis
Friday, 15:00-17:00
Chair and Organizer: Simo Puntanen

15:00-15:30 **Kimmo Vehkalahti** (University of Helsinki, Finland)
"Multivariate analysis for data scientists"

15:30-16:00 **Jianwen Xu** (Chongqing University, China)

"Estimating equations in various statistical models and methods"

16:00-16:30 **Simo Puntanen** (University of Tampere, Finland)

"Linear sufficiency: a review and some new results''

16:30-17:00 **Changyu Lu** (Shanghai Lixin University of Accounting and Finance, China)

18:00-21:00 Conference Banquet (Zhuang Yuan Lou Restaurant)

Saturday, 8 June 2019

Plenary Sessions PS8-PS10

Saturday, 09:00-11:20

Chair: Simo Puntanen

09:00-09:40 **Shuangzhe Liu** (University of Canberra, Australia)

"Professor Heinz Neudecker and matrix differential calculus"

09:40-10:20 **Shurong Zheng** (North Eastern Normal University, Changchun, Jilin, China)

"Statistical Inference on High-dimensional Covariance Matrices"

10:20-10:40 Tea Break

10:40-11:20 **Fuzhen Zhang** (Nova Southeastern University, Fort Lauderdale, USA)

"Inequalities for selected eigenvalues of the product of matrices"

Contributory Session CS2

Saturday, 11:20-12:00

Chair: Jianxin Pan

11:20-11:40 **Ji-Eun Choi** (Ewha Womans University of Statistics, Republic of Korea)

"A correlation break test based on self-normalization"

11:40-12:00 **Zhizheng Wang** (Linnaeus University, Växsjö, Sweden)

"A review of Dempster's non-exact test for high-dimensional mean vector"

12:00-13:00 Lunch

MS4-part 1. Predictive Modelling and Diagnostics
Saturday, 13:00-15:00
Chair and Organizer: Shuangzhe Liu

13:00-13:30 **Seng-Huat Ong** (UCSI University, Malaysia)
"A General Method of Computing Mixed Poisson Probabilities by Monte Carlo Sampling"
13:30-14:00 **Lei Shi** (Yunnan University of Finance and Economics, China)
"Sparse local influence analysis"
14:00-14:30 **Tatjana von Rosen** (Stockholm University, Sweden)
"Assessment of influence on the score test statistic in non-linear regression models"
14:30-15:00 **Tiefeng Ma** (Southwestern University of Finance and Economics, Chengdu, China)
"A shape-based multiple segmentation algorithm for change-point detection"

15:00-15:30 Tea Break

MS5. Decompositions of Tensor Spaces with Applications to Multilinear Models
Saturday, 15:30-17:30
Organizer: Dietrich von Rosen, *Chair*: Martin Singull

15:30-16:00 **Jianhua Hu** (Shanghai University of Finance and Economics, China)
"Simultaneous response and predictor selection model and estimation to multivariate linear regression"
16:00-16:30 **Martin Singull** (Linköping University, Sweden)
"Estimation, testing and residual analysis in the GMANOVA-MANOVA model"
16:30-17:00 **Feng Li** (Central University of Finance and Economics, Beijing, China)
"Forecasting with time series imaging"
17:00-17:30 **Chengcheng Hao** (Shanghai University of International Business and Economics, China)
"A bilinear reduced rank model"

17:30-19:00 Dinner

Sunday, 9 June 2019

MS6: Statistical Modelling for Complex Data
Sunday, 9:00 – 10:40
Chair and Organizer: Rui LI

9:00-9:20 **Guanfu Liu** (Shanghai University of International Business and Economics, China)
"Semi-parametric homogeneity test and sample size calculation for a two-sample problem under an inequality constraint"
9:20-9:40 **Caiyun Fan** (Shanghai University of International Business and Economics, China)
"Concordance-assisted learning for estimating optimal individualized treatment regimes"
9:40-10:00 **Yan Fan** (Shanghai Univ. of International Business and Economics, China)
"Single-Index-Based CoVaR With Very High-Dimensional Covariates"
10:00-10:20 **Hongmei Lin** (Shanghai University of International Business and Economics, China)
"Direct Local Linear Estimation for Sharpe Ratio Function in Heteroscedastic Regression Models"
10:20-10:40 **Huiling Yuan** (Shanghai University of International Business and Economics, China)
"Forecasting security's volatility using low-frequency historical data, high-frequency historical data and option-implied volatility"

10:40-11.00 Tea Break

MS4-part 2. Predictive Modelling and Diagnostics
Sunday, 11:00 -12:00
Chair and Organizer: Shuangzhe Liu

11:00-11:30 **Fukang Zhu** (Jilin University, China)
"Robust quasi-likelihood estimation for the negative binomial integer-valued GARCH(1,1) model with an application to transaction counts"
11:30-12:00 **Shimizu Kunio** (The Institute of Statistical Mathematics, Tokyo, Japan)
"A Wicksell-Kibble Type Distribution on a Hyper-Cylinder with an Application to Wind Direction and Speed Data"

Closing Session
Sunday, 12:00-12:15
Chair: Yonghui Liu

12:15-13:30 Lunch

Abstracts

Geometry and means of positive definite matrices

Rajendra Bhatia

Ashoka University, India

Abstract

We will describe two Riemannian distances on the space of positive definite matrices. One of these, the Riemann-Cartan distance, is a matrix version of the Fisher-Rao metric, and the other, the Bures-Wasserstein distance is a matrix version of the Hellinger (Bhattacharyya) distance. Connections with diverse areas like Riemannian geometry, statistics, optimal transport, quantum information and matrix analysis will be indicated.

Representative Points of Elliptically Symmetric Distributions

Kai-Tai Fang

BNU-HKBU United International College, Hong Kong, China

Abstract

The problem of selecting a given number of representative points (RPs) to retain as much information of the population as possible arises in many situations. One approach is proposed by Cox (1957) who proposed the mean square error (MSE) criterion and gave a table of RPs of the univariate normal distribution for $k \leq 6$. In general, this approach is definedas follows: a set of k RPs for the distribution of a random vector $\in \mathbf{R^p}$ is set of k points minimizing the expected squared distance between and the nearest point in the set.

There are different motivation for defining representative points. Max (1960) seeks to quantize the univariate normal distribution, Bofinger (1970) studied the question of grouping a continuous bivariate distribution by intervals on the marginals thereby obtaining a discrete bivariate distribution. In order to the standardize clothes, suppose taking p measurements of the body of each n individuals (in general, n is sufficiently large), and project these p dimensional data onto a $q(q < p)$ dimensional space by principal components analysis or by some other method. They wish to select k points that best represent the data in the q-dimensional space (see Fang (1976)).Motivated by this problem Fang and He (1982) proposed the question based on the standardize clothes how to choose k points under MSE. A similar background, Flury (1990) studied a project of the Swiss Army which wanted to design new protection masks. To put the construction of the new protection masks on a good empirical grounds, a group of anthropologists was hired to measure the heads of 900 Swiss soldiers. He and his coauthors found that when $k = 2, p > 2$ the representative points of an elliptical symmetric distribution on the direction of the eigenvector associated the largest eigenvalue of the covariance of . Therefore, they propose the name "principal points". In this talk I first review historical development of the representative points for univariate and multivariate cases and applications in resampling and density estimation.

Some new results for the elliptical symmetric distributions will be presented. Some comparisons among the representative points for different dimension p, the number of representative points k and different subclasses of elliptically symmetric distributions: normal, Kotz type, Pearson Type II and Pearson Type VII are given.

Keywords

Representative Points, Elliptically Symmetric Distributions, Principal points

Order determination for large-dimensional matrices

Lixing Zhu[1]

[1]*Hong Kong Baptist University, Hong Kong, China*

Abstract

This research aims to attack two longstanding problems in determining the model order dimensionality for those eigen-decomposition-based criteria. First, due to the existence of some dominating eigenvalues compared to other nonzero eigenvalues, the true order is often underestimated. Second, the estimation accuracy of any existing method often relies on the uniqueness of minimum/maximum of the criterion. To handle these problems, we propose a thresholding double ridge ratio criterion. Unlike all the existing eigen decomposition-based criteria, this criterion can dfine a consistent estimate even when there are several local minima. This generic strategy is readily applied to many fields. As the examples, we give the details about sufficient dimension reduction in regressions with fixed and divergent dimensions; about when the number of projected covariates can be consistently estimated, when cannot if a sequence of regression models converges to a limiting model with fewer projected covariates; about ultra-high dimensional approximate factor models and about spiked population models. Numerical studies are conducted to examine the finite sample performance of the method and real data are analysed for illustration.

Keywords

Representative Points, Elliptically Symmetric Distributions, Principal points

A gaze at recent applications and characterizations of the Moore–Penrose inverse

Oskar Maria Baksalary[1] **and Götz Trenkler**[2]

[1] *Adam Mickiewicz University in Poznań, Poland*
[2] *Dortmund University of Technology, Germany*

Abstract

The Moore–Penrose inverse is to celebrate its 100th birthday in 2020, as the notion standing behind the term was first defined by Moore in 1920 [1]. Its rediscovery by Penrose in 1955 [2] can be considered as a caesura after which the inverse attracted the attention it deserves and has henceforth been exploited in various research areas of applied origin. During the talk we will discuss several examples of recent applications of the Moore–Penrose inverse demonstrating that the notion continues to play a role of a valuable tool to cope with the current research problems.

A part of the talk will be devoted to the results concerned with the representations of the Moore–Penrose inverse of matrices. The topic has attracted a considerable attention over the years and several different approaches were exploited so far. In the talk we will recall some of the available results (concerned e.g., with matrices modified by matrices of rank-one, partitioned matrices, functions of other generalized inverses, or functions of a square matrix represented by the Hartwig–Spindelböck decomposition) and shed light on selected problems considered by the authors.

Keywords

Generalized inverses of matrices, Partitioned matrices.

References

[1] Moore, E.H. (1920). On the reciprocal of the general algebraic matrix. *Bull. Amer. Math. Soc. 26*, 394-395.

[2] Penrose, R. (1955). A generalized inverse for matrices. *Math. Proc. Cambridge Philos. Soc. 51*, 406-413.

Inverse Complimentary Matrix Method and its Applications to General Linear Model

Manjunatha Prasad Karantha[1], **Nayan Bhat K**[1] and **Eagambram Narayanan**[2]

[1]*Department of Statistics, Manipal Academy of Higher Education, Manipal, India 576104*

[2]*Former Deputy Director General, Indian Statistical Services, Government of India*

Abstract

In this presentation, we revisit the concept of 'Inverse Complimented Matrix Method' introduced by Eagambaram (2018) and obtain new applications of Inverse Complimented Matrix Method. Some of well known generalized inverses and outer inverses of given matrix are characterized by identifying appropriate compliment. Also, we exhibit that the method helps to decompose the matrices V and X in the general linear model $\left(Y, X\beta, \sigma^2 V\right)$ and provide a representation of the model. An explicit expression for Admissible Linear Estimator of an estimable $A\beta$ is also obtained by this method.

Keywords

generalized inverse, shorted matrix, general linear model

References

[1] Baksalary, J.K. and Markiewicz, A (1988). *Admissible linear estimators in the general Gauss-Markov model. Journal of Statistical Planning and Inference*, 19(3): 349-359.

[2] Ben-Israel, A. and Greville, T.N.E. (1974). *Generalized Inverses: Theory and Applications*. Wiley-Interscience, New York.

[3] Dengupta, D. and Jammalamadaka, S.R. (2003). *Linear models: an integrated approach*. World Scientific.

[4] Eagambaram, N. (2018). *Disjoint sections and generalized inverses of matrices. Bulletin of Kerala Mathematical Society*, 16(1): 153-161.

[5] Mitra, S.K. (1986). *The minus partial order and shorted matrix. Linear Algebra Appl.*, 83: 1-27.

[6] Rao, C.R., and Mitra, S.K. (1971). *Generalized Inverse of Matrices and Its Applications*. Wiley, New York, 1971.

Identifying conditions for multilinear matrix equations to always hold with applications

Yongge Tian

Shanghai Business School, Shanghai, China & Central University of Finance and Economics, Beijing, China

Abstract

Any algebraic expression that involves variables may vary with respect to the choice of the variables. Thus one of the fundamental problems in algebra is to determine conditions under which a given algebraic expression does not change with respect to the choice of variables in it. In my talk, I introduce a block matrix representation method to display necessary and sufficient conditions for the following two general multilinear matrix equations

$$(A_1 + B_1X_1C_1)(A_2 + B_2X_2C_2)\cdots(A_k + B_kX_kC_k) = M,$$

$$(A_1 + B_1X_1C_1 + D_1Y_1E_1)\cdots(A_k + B_kX_kC_k + D_kY_kE_k) = M$$

to always hold respectively with respect to all variable matrices $X_1, \ldots, X_k$ and $Y_1, \ldots, Y_k$. I then present some concrete examples on establishing such kinds of multilinear matrix identities in matrix theory with emphasis on characterizing numerous matrix identities and matrix set inclusions composed by generalized inverses.

Keywords

multilinear matrix equations, block matrix, matrix set inclusions, generalized inverses.

Calibration for non-positive definite covariance matrix

Chao Huang[1], **Daniel Farewell**[2] **and Jianxin Pan**[3]

[1] *South East Wales Trials Unit, Cardiff University, Cardiff, CF14 4YS, UK*
[2] *School of Medicine, Cardiff University, Cardiff, CF14 4YS, UK*
[3] *School of Mathematics, University of Manchester, M13 9PL, UK*

Abstract

Covariance matrices that fail to be positive definite arise often in covariance estimation. Approaches addressing this issue exist, but are not well supported theoretically. In this paper, we propose a unified statistical and numerical matrix calibration method, finding the optimal positive definite surrogate in the sense of Frobenius norm. The proposed method is well supported theoretically and the proposed algorithm can be directly applied to any estimated covariance matrix. Numerical simulation results show that the calibrated matrix is typically closer to the true covariance, while making only limited changes to the original covariance structure. The proposed method is also applied to a real data analysis for illustration.

Keywords

Covariance matrix calibration, Nearness problem, Non-positive definiteness, Spectral decomposition

References

[1] Diggle, P. J. (1988). An approach to the analysis of repeated measures. *Biometrics*, **44**, 959-971.

[2] Diggle, P.J. and Verbyla, A. P. (1998). Nonparametric estimation of covariance structure in longitudinal data. *Biometrics*, **54**, 401-415.

[3] Higham, N.J. (1988). Computing a nearest symmetric positive semidefinite matrix. *Linear Algebra and Appl.*, **103**, 103-118.

[4] Huang C., Farewell D. and Pan J. (2017). A calibration method for non-positive definite covariance matrix in multivariate data analysis. *Journal of Multivariate Analysis*, **157**, Pages 45-52.

[5] Zarowski., Christopher J. (2004). An Introduction to Numerical Analysis for Electrical and Computer Engineers. Wiley.

Professor Heinz Neudecker and matrix differential calculus

Shuangzhe Liu[1]

[1] *University of Canberra, Australia*

Abstract

The late Professor Heinz Neudecker is regarded as the founding father of matrix differential calculus. He laid the foundation for the theory and practice of matrix differential calculus and his contributions were compiled in the standard work Magnus and Neudecker [1]. The methods developed in his work are still used by contemporary econometricians and statisticians today in analysing multivariate models.

In this talk, we present the fundamental idea and notation in matrix differential calculus based on differentials (rather than derivatives). We discuss some results, with a focus on its applications to topics in deep learning, predictive modelling, sensitivity analysis and statistical diagnostics.

Keywords

Matrix inequalities, Matrix products, Jacobian, Hessian, Optimisation

References

[1] Magnus, J., H. Neudecker (1988, 1999, 2019). *Matrix Differential Calculus with Applications in Statistics and Econometrics.* Chichester: Wiley.

Inequalities for selected eigenvalues of the product of matrices

Bo-Yan Xi[1] **and Fuzhen Zhang**[2]

[1] *Inner Mongolia University for Nationalities, China*
[2] *Nova Southeastern University, United States*

Abstract

The product of a Hermitian matrix and a positive semidefinite matrix has only real eigenvalues. We present bounds for sums of eigenvalues of such a product.

Keywords

Eigenvalue, Hermitian matrix, inequality, positive semidefinite matrix

Statistical Inference on High-dimensional Covariance Matrices

Shurong Zheng[1]

[1]*Northeast Normal University, China*

Abstract

With the rapid development of computer science, it is possible to collect, store and analyze high-dimensional data. But some classical statistical methods become invalid. For example, the log-likelihood ratio test for testing the identity of covariance matrix has the Type I errors tending to one as the data dimensiona and sample size tend to infinity proportionally. This talk will introduce some estimation methods and testing methods to deal with high-dimensional covariance matrices.

Building some bridges among various experimental designs

A. M. Elsawah

BNU-HKBU United International College, China

Abstract

Designing their experiments is the significant problem that experimenters face. Maximin distance designs, supersaturated designs, minimum aberration designs, uniform designs, minimum moment designs and orthogonal arrays are arguably the most exceedingly used designs for many real-life experiments. From different perspectives, several criteria have been proposed for constructing these designs for investigating either quantitative or qualitative factors. Each of those criteria has its pros and cons and thus an optimal criterion does not exist, which may confuse investigators searching for a suitable criterion for their experiment. Some logic questions are now arising such as, are these designs consistent?, can an optimal design via a specific criterion perform well based on another criterion? and can an optimal design for screening quantitative factors be optimal for screening qualitative factors?. Through theoretical justification, this paper tries to answer these interesting questions by building some bridges among various criteria and their corresponding designs. Some conditions under which these designs agree with each other are given. Since some of those criteria have conceptual simplicity and tremendous computational advantages over others, recommended criteria in specific circumstances are given via these bridges that are used to effectively study some hard problems, such as detection of (combinatorial/geometrical) non-isomorphism among designs and selection of good designs. Benchmarks for reducing the computational complexity are given.

Uniform design on general domain

Yu Tang[1]

[1] *Soochow University, China*

Abstract

Uniform design aims to scatter points as evenly as possible on certain domain. Although in real applications, the experimental domain is often quite arbitrary, the discrepancies frequently used to measure the uniformity of experimental designs are normally defined on unit cube. In this paper, we will introduce a unified framework to measure the uniformity of an experimental design on general domain. We will also give some examples to illustrate the construction of uniform design on some specific domains.

Data-driven Space-filling Design

Aijun Zhang[1], **Mei Zhang**[2] **and Yong-Dao Zhou**[3]

[1]*Department of Statistics and Actuarial Science, The University of Hong Kong, Hong Kong SAR, China*
[2] *College of Mathematics, Sichuan University, China*
[3] *School of Statistics and Data Science, Nankai University, China*

Abstract

The quest for a small data to represent a big data is important for data compression, exploration and subsampled modeling. We consider the data-driven space-filling design with the criterion of empirical F-discrepancy. Asymptotic optimality is established for an inversion construction method based on existing uniform experimental designs. When the small data is required to be a subset of the big data, we develop an effective subdata selection algorithm based on the proposed data-driven space-filling design. Such algorithm has potential applications in large-scale machine learning in both supervised and unsupervised settings.

Keywords

Space-filling design, Empirical F-discrepancy, Big data subsampling, Large-scale machine learning.

Orthogonal Uniform Composite Designs

Xue-Ru Zhang, Min-Qian Liu and Yong-Dao Zhou[1]

School of Statistics and Data Science, LPMC & KLMDASR, Nankai University, Tianjin 300071, China

Abstract

Composite designs are frequently utilized for fitting response surfaces in practice. This paper proposes a new type of composite designs, orthogonal uniform composite designs (OUCDs), which combine orthogonal arrays and uniform designs. Such designs not only inherit the advantages of orthogonal-array composite designs such as high estimation efficiencies and ability for multiple analysis for cross validation, but also have more flexible run sizes than central composite designs and orthogonal-array composite designs. Moreover, OUCDs are more robust than other types of composite designs under certain conditions. Some construction methods for OUCDs under the maximin distance criterion are provided and their properties are also studied.

Keywords

Central composite design, Maximin distance criterion, Orthogonal-array composite design, Robustness, Uniform design.

[1]Corresponding author, e-mail: ydzhou@nankai.edu.cn

Two-sample correlation parameter testing in models with a Kronecker product covariance structure

Yuli Liang[1], **Chengcheng Hao**[2] **and Zhengtao Li**[2]

[1] *Orebro University School of Business, Sweden*
[2]*Shanghai University of International Business and Economics, China*

Abstract

Under a model having a Kronecker product covariance structure with compound symmetry, two-sample hypothesis testing for a correlation is investigated. Several tests are suggested and practical recommendations are made based on their type I error probabilities and powers.

Keywords

Covariance matrix, Transformations, Longitudinal data.

ANOVA SSs and Proportional Subclass Numbers

Lynn Roy LaMotte[1]

[1]*Louisiana State University Health Sciences Center – New Orleans USA*

Abstract

Soon after Fisher introduced analysis of variance for effects of two factors, it was clear that "the addition law" didn't work in unbalanced models unless the cell sample sizes had the "proportional subclass numbers" property (psn), that $n_{ij} = n_{i\cdot}n_{\cdot j}/n_{\cdot\cdot}$. If not, then SS_{AB}, computed as Fisher described, was not a true SS in the usual sense, and it could take negative values. This led to the still-continuing ambivalence about the appropriate SSs for testing factor main effects in unbalanced models without psn. Consistently, though, textbooks have taught that there is no problem in models having psn: psn is the same as balanced.

In this talk I'll note that this is not true. In unbalanced models that don't have psn, the classical ANOVA SSs test hypotheses that are unrelated to the ANOVA definition of main effects. SS_{AB} tests the right hypothesis iff the model has psn. SS_A tests the right hypothesis iff $n_{ij} = n_{i\cdot}/b$, an additional requirement beyond psn.

The process of examining these properties reveals relations among Types I, II, III, and marginal-means (MM) SSs. For example, the Type II noncentrality parameter for A main effects can be 0 even though there are differences (arbitrarily great) among the A marginal means. Type III SSs, on the other hand, always test at least the estimable part of the corresponding effect contrasts, and MM SSs test exactly the estimable part. These are illustrated by examples.

Keywords

Unbalanced two-factor models, Orthogonal sums of squares, Type I-III SSs.

Comment on inference in a simple linear random coefficient model with missing covariates

Julia Volaufova[1] **and Paige Fisher**[1]

[1]*Biostatistics Program, School of Public Health, LSU Health-NO New Orleans, Louisiana, USA*

Abstract

Missing observations may have a large impact on statistical inference. Since approximately mid 70s of the last century, estimation and prediction have been extensively studied in a variety of rather complex statistical models under the assumption that some observations are not available, they are missing. However, much less is studied on statistical inference when, say, covariates are missing. Here, we consider a simple linear random coefficient model with possibly missing covariates. We briefly review the available methods for estimation and testing of hypotheses about fixed effects parameters. Our focus here is on the approximation to the estimated covariance matrix of the estimator of the hypothesized parameter. Fisher information matrix and its observed version are used as the basis for investigations.

Keywords

Random coefficient model, covariates missing at random, approximate variance, testing hypotheses.

References

[1] Chen, Q., Ibrahim, J.G., Chen M-H., and Senchaudhuri, P. (2008). Theory and inference for regression models with missing responses and covariates. *Journal of Multivariate Analysis 99*, 1302–1331.

[2] Dempster, A.P., Laird, N., and Rubin, D.B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B, 39(1)*, 1–38.

[3] Louis, T.A. (1982). Finding the observed information matrix when using EM algorithm. *Journal of the Royal Statistical Society, Series B, 44(2)*, 226–233.

Multivariate Analysis for Data Scientists

Kimmo Vehkalahti[1]

[1] *University of Helsinki, Finland*

Abstract

The ongoing "Data revolution" [1] sets more requirements for the researchers on all fields of science. One could say – without exaggerating too much – that *we should all be data scientists.*

A special pressure is put on the fields of social and behavioral sciences, where the phenomena, the measurements, and the data are affected by endless sources of uncertainty and may often be much more complex than in many applications of, say, natural sciences.

Hence, the question is: *What should be included in a data scientist's "toolbox" in social and behavioral sciences?*

Our suggestion would be a good combination of classical and modern skills that are covered, for example, by a recent textbook on multivariate analysis [3]: A wide range of methods for visualizing data, linear and generalized linear mixed (and fixed) models, various methods of multivariate analysis (both exploratory and confirmatory), a bit (or a byte) of matrices behind the methods (even without a maths background), programming and using statistical software (preferably R [2]), algorithmic thinking in general, as well as documenting and sharing the code and data on open platforms such as GitHub. In this talk, we discuss some of these topics in more detail.

Keywords

Multivariate analysis, Linear models, Data science, Matrices, Statistics.

References

[1] Kitchin, R. (2014). *The Data Revolution: Big Data, Open Data, Data Infrastructures & Their Consequences*. London: SAGE.

[2] R Core Team (2018). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/

[3] Vehkalahti, K., and Everitt, B. S. (2019). *Multivariate Analysis for the Behavioral Sciences*, 2nd edition. Boca Raton, Florida: Chapman and Hall/CRC. https://github.com/KimmoVehkalahti/MABS/

Estimating equations in various statistical models and methods

Jianwen Xu

Department of Statistics and Actuarial Science, Chongqing University, China

Abstract

In this talk, I will introduce the estimating equations for various models and methods, including the linear models, non-linear models, generalized linear models, quasi-likelihood method and marginal models for longitudinal data analysis. In particular, all these models and methods will be demonstrated to have the same expressions of estimating equations for unknown regression coefficients. Thus, the estimating equations could be regarded as bridges among these statistical models and methods.

Keywords

Estimating equations, Linear and non-linear models, Generalized linear models, Longitudinal data analysis.

References

[1] Raymond H. Myeres, Douglas C. Montgomery, G. Geoffrey Vining, Timothy J. Robinson. (2010). Generalized Linear Models with Applications in Engineering and the Sciences. John Wiley and Sons.

[2] Mccullagh,P. Nelder, J.A. (1989). Generalized Linear Models (Second edition), Chapman and Hall.

[3] Rao,C.R., Toutenburg, H., Shalabh and Heumann, C. (2008). Linear Models and Generalizations: Least Squares and Alternatives, Springer.

Linear sufficiency: a review and some new results

Simo Puntanen

University of Tampere, Finland

Abstract

We consider the general linear model $y = Xb + e$, denoted as $M = \{y, Xb, V\}$, supplemented with the new unobservable random vector y_*, coming from $y_* = X_*b + e_*$. A linear statistic Fy is called linearly sufficient for estimable X_*b if there exists a matrix A such that AFy is the best linear unbiased estimator, BLUE, for X_*b. The concept of linear sufficiency with respect to a predictable random vector is defined in the corresponding way but considering the best linear unbiased predictor, BLUP, instead of BLUE. In this talk, we consider the linear sufficiency of Fy with respect to y_*, X_*b, and e_*. Particular attention is being paid on the connection between the linear sufficiency concept and the equality of the multipliers of y providing BLUEs and BLUPs in the original and in the transformed model $T = \{Fy, FXb, FVF'\}$.

Keywords

BLUE, BLUP, Linear sufficiency, Linear model with new observations, Transformed linear model.

References

[1] Baksalary, J.K. & R. Kala (1981). Linear transformations preserving best linear unbiased estimators in a general Gauss–Markoff model. *Ann. Stat.*, 9, 913–916.

[2] Baksalary, J.K. & R. Kala(1986). Linear sufficiency with respect to a given vector of parametric functions. *J. Stat. Plan. Inf.*, 14, 331–338.

[3] Haslett, S.J., X.-Q. Liu, A. Markiewicz & S. Puntanen (2017). Some properties of linear sufficiency and the BLUPs in the linear mixed model. *Stat. Pap.*, available online.

[4] Kala, R., S. Puntanen & Y. Tian (2017). Some notes on linear sufficiency. *Stat. Papers*, 58, 1–17.

[5] Markiewicz, A. & S. Puntanen (2018). Further properties of linear prediction sufficiency and the BLUPs in the linear model with new observations. *Afrika Statistika*, 13, 1511–1530.

A General Method of Computing Mixed Poisson Probabilities by Monte Carlo Sampling

Seng-Huat Ong[1], **Wen–Jau Lee**[2] **and Yeh-Ching Low**[3]

[1]*Department of Actuarial Science and Applied Statistics, UCSI University, Malaysia*
[2]*Invantest DSG, Suntech@Penang Cybercity, Lintang Mayang Pasir 3,Malaysia*
[3]*Sunway University, Malaysia*

Abstract

Mixed Poisson distributions form an important class of distributions in applications. However, the application of many of these mixed Poisson distributions are hampered by the complicated probability distributions. The paper examines Monte Carlo sampling as a general technique for computation of mixed Poisson probabilities which is applicable to any mixed Poisson distribution with arbitrary mixing distribution. The accuracy and computational speed of this method is illustrated with the Poisson-inverse Gaussian distribution. The proposed method is then applied to compute probabilities of the Poisson-lognormal distribution, a popular species abundance model, It is also shown that in the maximum likelihood estimation of Poisson-lognormal parameters by E-M algorithm, the application of the proposed Monte Carlo computation in the algorithm avoids numerical problems.

Keywords

Gamma distribution, Poisson-lognormal, species abundance, multivariate mixed Poisson, variance reduction, quasi-Monte Carlo, antithetic variate, maximum likelihood; EM algorithm

Sparse local influence analysis

Jun Lu[1] **and Lei Shi**[1]

[1] *School of Statistics and Mathematics, Yunnan University of Finance and Economics, Kunming, 650221 China*

Abstract

Cook's (1986) local influence method is useful for identifying influential observations in statistical diagnostics and sensitivity analysis. However, it is often criticized for lack of a rigorous criterion to judge the influence magnitude from the elements of the main diagnostic. In this paper, we propose a new method, namely sparse local influence analysis, to detect the influential observations. We use the connection of local influence analysis with sparse principal component analysis and produce the modified local diagnostic with sparse elements, i.e. diagnostic with very few nonzero elements. With this method, influential observations can be efficiently detected by the remaining nonzero elements of the modified diagnostic. Two real data sets are used for illustration and a simulation example is conducted to confirm the efficiency of the proposed methodology.

Keywords

Local influence analysis, Influential observations, Sparseness

References

[1] Cook, R. D. (1986). Assessment of local influence. *Journal of the Royal Statistical Society, B* 48,133-169.

[2] Qi,X., Luo, R., Zhao, H. (2013). Sparse principal component analysis by choice of norm. *Journal of Multivariate Analysis*, 114,127-160.

[3] She,Y., Owen, A. (2011). Unmasking multivariate outliers and leverage points. *Journal of the American Statistical Association*, 85,633-639.

[4] Shi, L., and Huang, J. (2008). Outlier detection using nonconvex penalized regression. *Jounal of Multivariate Analysis*,106,626-639.

[5] Shi, L. (1997). Local influence in principal component analysis. *Biometrika*, 84,175-186.

[6] Tibshirani, R. (1996). Regression shrinkage and selection via Lasso. *Journal of the Royal Statistical Society, B*, 58,267-288.

[7] Zou, H., Hastie, T., Tibshirani, R. (2006). Sparse principal component analysis. *Journal of Computational and Graphical Statistics*, 15,265-286.

Assessment of influence on the score test statistic in non-linear regression models

Tatjana von Rosen[1] **and Karin Stål**[2]

[1]*Department of Statistics, Stockholm University, Sweden*
[2]*The National Agency for Education, Sweden*

Abstract

Regression analysis is a statistical technique for exploring the relationships between variables. Frequently, regression models are used to describe the dependence between a response variable and one or several explanatory variables. The parameters in the regression model are estimated based on observed data. However, some observations have a greater impact on the estimated model than others. The regression model considered in this work is the nonlinear model with an additive error term

$$\mathbf{y} = \mathbf{f}(\mathbf{X}, \boldsymbol{\theta}) + \boldsymbol{\epsilon},$$

where $\mathbf{f}(\mathbf{X}, \boldsymbol{\theta}) = (\mathbf{f}(\mathbf{X}_1, \boldsymbol{\theta}), \ldots, \mathbf{f}(\mathbf{X}_n, \boldsymbol{\theta}))^T = (\mathbf{f}_1(\boldsymbol{\theta}), \ldots, \mathbf{f}_n(\boldsymbol{\theta}))^T$, $\mathbf{X}$ is a $n \times p-$matrix of known explanatory variables, $\mathbf{y}$ is the n-vector of responses, $\boldsymbol{\theta}$ is a q-vector of unknown parameters, $\boldsymbol{\epsilon} \sim N(\mathbf{0}, \sigma^2\mathbf{I}_n)$, and $\mathbf{I}_n$ denote the identity matrix of size n.

A well known example of a nonlinear model is the Michaelis-Menten model

$$y = \frac{\theta_1 x}{\theta_2 + x} + \varepsilon,$$

which is used in enzyme kinetics. It relates the initial velocity of an enzymatic reaction, y, to the substrate concentration, x. The parameter θ_1 is the maximum velocity of the enzymatic reaction, representing the asymptotic value of f as $x \to \infty$; θ_2 is the half-velocity parameter, representing the value of x when the velocity of the reaction reaches one-half of its ultimate value.

The existing influence measures in regression analysis are constructed to measure the impact of observations on the parameter estimates or the fitted values. However, it is of interest to assess the influence of observations on hypothesis testing. We will derive a diagnostic measure for assessing the influence of single and multiple observations on the score test statistic [**?**], both in linear and nonlinear regression. The proposed diagnostic measure is derived using the differentiation approach.

Keywords

non-linear regression, influential observations, Score test.

References

[1] Chen, C.F. (1983). Score tests for regression models. *JASA 78*, 158–161.

[2] Chen, C.F. (1985). Robustness aspects of score tests for generalized linear and partially linear regression models. *Technometrics 27*, 277–283.

[3] Li, B. (2001). Sensitivity of Rao's score test, the Wald test and the likelihood ratio test to nuisance parameters. *J Stat Plan Inference 97*, 57–66.

[4] Lustbader, E.D. and Moolgavkar, S.H. (1985). A diagnostic statistic for the score test. *JASA 80*, 375–379.

[5] Rao, C.R. (1948). Large sample tests of statistical hypotheses concerning several parameters with applications to problems of estimation. *Proceedings of the Cambridge Philosophical Society 44*, 50–57.

[6] Vanegas, L.H., Rondón, L.M. and F.J.A. Cysneiros (2013). Assessing robustness of inference in symmetrical nonlinear regression models. *Commun. Stat. - Theory Methods 42*, 1692–1711.

A shape-based multiple segmentation algorithm for change-point detection

Tiefeng Ma

Southwestern University of Finance and Economics, China

Abstract

We will describe two Riemannian distances on the space of positive definite matrices. One of these, the Riemann-Cartan distance, is a matrix version of the Fisher-Rao metric, and the other, the Bures-Wasserstein distance is a matrix version of the Hellinger (Bhattacharyya) distance. Connections with diverse areas like Riemannian geometry, statistics, optimal transport, quantum information and matrix analysis will be indicated.

Robust quasi-likelihood estimation for the negative binomial integer-valued GARCH(1,1) model with an application to transaction counts

Fukang Zhu

Jilin University, China

Abstract

For count time series analysis, the Poisson integer-valued generalized autoregressive conditional heteroscedastic model is very popular but is not usually suitable in the existence of potential extreme observations. Maximum likelihood estimator is commonly used to estimate parameters, but it is highly affected by the outliers. This paper has three main aims. First, we apply the negative binomial model in our study for count time series analysis and consider the maximum likelihood estimation of this model. Second, we extend the Mallows' quasi-likelihood method proposed in the generalized linear models to our situation. Besides, we establish the consistency and asymptotic normality for the resulting robust estimators under some regularity conditions. Third, the performances of these robust estimators in the presence of transient shifts and additive outliers are investigated via simulations. We apply the robust estimator to two stock-market data sets and their prediction performances are assessed by in-sample and out-of-sample predictions.

A Wicksell-Kibble Type Distribution on a Hyper-Cylinder with an Application to Wind Direction and Speed Data

Kunio Shimizu

School of Statistical Thinking, The Institute of Statistical Mathematics, Japan

Abstract

Observation values of wind speed (linear) at 6:00 a.m. and 12:00 noon may be modeled by bivariate gamma, lognormal, and inverse Gaussian distributions. One can predict wind speed values at noon from data at 6:00 using a regression and furthermore may expect a better statistical modeling if wind direction (circular) data at 6:00 are available. In this talk we use the Wicksell-Kibble distribution as a bivariate gamma distribution to construct a hyper-cylindrical distribution with two linear and one circular variables.

The Wicksell-Kibble bivariate gamma distribution considered here has four-parameters with the role of shape (one), scale (two), and correlation coefficient (one). The proposed hyper-cylindrical distribution involves one more parameter as a circular location. The parameter of correlation coefficient for the Wicksell-Kibble distribution controls not only dependence of the three variables but also circular concentration for the hyper-cylindrical distribution. Several properties such as marginal and conditional distributions and their moments are studied, and the regression curve and surface are obtained. Random number generation for the proposed distribution is possible as the joint density is expressed by a multiplication of conditional and marginal densities, and random numbers for each of conditional and marginal distributions can be generated. In particular, the conditional density of one linear variable given other linear and circular variables has a mixture expression of a Poisson probability and a gamma density.

An illustrative example is given for wind direction and speed at 6:00 a.m. and wind speed at 12:00 noon data observed in Tokyo. A comparison between the proposed and existing models is made. It is clear that the model has potential applications to a combination of any one circular and two linear measurements.

Keywords

Directional statistics, Random number generation, Regression surface

Simultaneous Response and Predictor Selection Model and Estimation to Multivariate Linear Regression

Jianhua Hu[1], **Jian Huang**[2] **and Feng Qiu**[3]

[1]*Shanghai University of Finance and Economics, China*
[2] *University of Iowa, Iowa City, Iowa USA*
[3] *Zhejiang Agriculture and Forestry University, China*

Abstract

The response best subset selection model [4] addresses the problems in which predictors are known and only responses need to be selected in multivariate linear models. In practice, we couldn't make sure not only responses but also predictors and need variable selection simultaneously for responses and predictors, for which no research has been found. In this paper, we propose a novelty simultaneous response and predictor selection (SRPS) model, which is motivated by applications where some responses or predictors are unimportant in multivariate linear regression analysis. We simultaneously investigate variable selection both for responses and predictors and estimation to regression coefficients in the standard multivariate linear regression, group adaptive lasso and the response best subset selection contexts. We also establish model consistency, consisting of response selection, predictor selection and coefficient estimation, and the oracle property of coefficient estimators. Our simulation studies suggest that the proposed method is pronouncedly efficient. We also apply our methodology to study a real data set.

Keywords: Group adaptive lasso, model consistency, multivariate linear regression, response best subset selection model, response selection, simultaneous response and predictor selection model

References

[1] Anderson, T.W. (2003). An Introduction to Multivariate Statistical Analysis (3rd Edition). John Wiley & Sons, Inc., Publication.

[2] Breheny, P. and Huang, J. (2009). Penalized methods for bi-level variable selection. *Statistics and Its Interface*, 2(3):369-380.

[3] Fan, J. and Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American Statistical Association*, 96(456):1348-1360.

[4] Hu, J., Huang, J. and Qiu, F. (2018). Response best subset selection model and efficient estimation in multivariate linear regression. *Technical Report*, Shanghai University of Finance and Economics.

[5] Huang, J., Ma, S., Xie, H., and Zhang, C.-H. (2009). A group bridge approach for variable selection. *Biometrika*, 96(2):339-355.

[6] Knight, K. and Fu, W. (2000). Asymptotics for lasso-type estimators. *The Annals of Statistics*, 28(5):1356-1378.

[7] Li, R., Zhong, W., and Zhu, L. (2012). Feature screening via distance correlation learning. *Journal of the American Statistical Association*, 107(499):1129-1139.

[8] Muirhead, R.J. (1982). *Aspects of Multivariate Statistical Theory*. Hohn Wiley & Sons, Inc.

[9] Su, Z., Zhu, G., Chen, X. and Yang, Y. (2016). Sparse envelope model: efficient estimation and response variable selection in multivariate linear regression. *Biometrika*, 103(3), 579-593.

[10] Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 58(1):267-288.

[11] Wang, H. and Leng, C. (2008). A note on adaptive group lasso. *Computational Statistics and Data Analysis*, 52: 5277-5286.

[12] Yuan, M. and Lin, Y. (2006). Model selection and estimation in regression with grouped variables. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 68(1):49-67.

[13] Zhang, C.-H. (2010). Nearly unbiased variable selection under minimax concave penalty. *The Annals of Statistics*, 38(2):894-942.

[14] Zhang, H.H. and Lu, W. (2007). Adaptive lasso for Coxs proportional hazard model. *Biometrika*, 94:691-703.

[15] Zou, H. (2006). The adaptive lasso and its oracle properties. *Journal of the American Statistical Association*, 101(476):1418-1429.

[16] Zou, H. and Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(2):301- 320.

Estimation, testing and residual analysis in the GMANOVA-MANOVA model

Béatrice Byukusenge[1,2], Dietrich von Rosen[2,3] and Martin Singull[2]

[1] *University of Rwanda, Rwanda*
[2] *Linköping University, Sweden*
[3] *Swedish University of Agricultural Sciences, Sweden*

Abstract

In this talk we will consider the GMANOVA-MANOVA model, which is a special case of an extended growth curve model, with no assumption of the nested subspace condition. We derive two types of residual, establish their properties and give interpretations. We also discuss their use in bilinear hypothesis testing for the MANOVA model. Finally, a small simulation study is performed to validate the theoretical results and a numerical example on a real data set from a study that was conducted to investigate two treatments for patients suffering from multiple sclerosis is given.

Keywords

GMANOVA-MANOVA, growth curve model, residual analysis

Forecasting with time series imaging

Feng Li[1]

[1] *Central University of Finance and Economics, China*

Abstract

Feature-based time series representation has attracted substantial attention in a wide range of time series analysis methods. Recently, the use of time series features for forecast model selection and model averaging has been an emerging research focus in the forecasting community. Nonetheless, most of the existing approaches depend on the manual choice of an appropriate set of features. Exploiting machine learning methods to automatically extract features from time series becomes crucially important in the state-of-the-art time series analysis. In this paper, we introduce an automated approach to extract time series features based on images. Time series are first transformed into recurrence images, from which local features can be extracted using computer vision algorithms. The extracted features are used for forecast model selection and model averaging. Our experiments show that forecasting based on automatically extracted features, with less human intervention and a more comprehensive view of the raw time series data, yields comparable performances with the top best methods proposed in the largest forecasting competition M4.

Forecasting with time series imaging

Dietrich von Rosen[1], **Feng Li**[2] **and Chengcheng Hao**[3]

[1] *Swedish University of Agricultural Sciences, Sweden*
[2] *Central University of Finance and Economics, China*
[3] *Shanghai University of International Business and Economics, China*

Abstract

This article considers a bilinear model which includes two different latent effects. The first effect has a direct effect on the response variable whereas the second latent effect is assumed to first effect other latent variables which in turn effect the response variable. In this article latent variables are modelled via rank restrictions on unknown parameters. To have one latent effect which directly effect a response variable has a long history and is often referred to as reduced rank regression. This article presents a likelihood based approach which ends up in explicit estimators. In our model the latent variables act as covariate variables which we know exists but their impact is very vague and will therefore not be considered in detail. One example is if we observe hundreds of weather variables but we cannot say which or how these variables effect, for example, plant growth.

Semi-parametric homogeneity test and sample size calculation for a two-sample problem under an inequality constraint

Guanfu Liu[1], Yan Fan[1], Yang Liu[2] and Yukun Liu[2]

[1] *School of Statistics and Information, Shanghai University of International Business and Economics*
[2] *School of Statistics, East China Normal University*

Abstract

In medical researches such as case-control studies with contaminated controls, frequently encountered is a particular two-sample testing problem in which one sample has a mixture structure. It is very common the case that the exposure in a case-control study may have a positive (or negative) effect on the response variable if the effect exists. This is often ignored by existing tests, which would lead to potentially power loss. Meanwhile, it is of much practical importance to determine a minimal sample size to reach a target power. Based on empirical likelihood and density ratio model, we develop a new EM-test by incorporating the inequality information in the alternative. We show that the proposed EM-test has a mixture of zero and χ^2_1 limiting distribution under the null hypothesis. Its local power analysis and sample size calculations are also investigated. A simulation study and two real data analyses are provided to illustrate the proposed EM-test and sample size formula.

Keywords

Empirical likelihood, EM-test, Sample size calculation, Two-sample problem

Concordance-assisted learning for estimating optimal individualized treatment regimes

Caiyun Fan[1]

[1] *Shanghai University of International Business and Economics, China*

Abstract

We propose new concordance-assisted learning for estimating optimal individulalized treatment regimes. We first introduce a type of concordance function for prescribing treatment and propose a robust rank regression method for estimating the concordance function. We then find treatment regimes, up to a threshold, to maximize the concordance function, named the prescriptive index. Finally, within the class of treatment regimes that maximize the concordance function, we find the optimal threshold to maximize the value function. We establish the rate of convergence and asymptotic normality of the proposed estimator for parameters in the prescriptive index. An induced smoothing method is developed to estimate the asymptotic variance of the estimator. We also establish the $n^{1/3}$-consistency of the estimated optimal threshold and its limiting distribution. In addition, a doubly robust estimator of parameters in the prescriptive index is developed under a class of monotonic index models. The practical use and effectiveness of the methodology proposed are demonstrated by simulation studies and an application to an acquired immune deficiency syndrome data set.

Single-Index-Based CoVaR With Very High-Dimensional Covariates

Yan Fan[1], **Wolfgang Karl Härdle**[2], **Weining Wang**[3] and **Lixing Zhu**[4]

[1]*Shanghai University of International Business and Economics, China*
[2] *Humboldt-Universität zu Berlin, 10178 Berlin, Germany*
[3]*Kings College London, London, United Kingdom*
[4]*Hong Kong Baptist University, Hong Kong, China*

Abstract

Systemic risk analysis reveals the interdependencies of risk factors especially in tail event situations. In applications the focus of interest is on capturing joint tail behavior rather than a variation around the mean. Quantile and expectile regression are used here as tools of data analysis. When it comes to characterizing tail event curves one faces a dimensionality problem, which is important for CoVaR (Conditional Value at Risk) determination. A projection-based single-index model specification may come to the rescue but for ultrahigh-dimensional regressors one faces yet another dimensionality problem and needs to balance precision versus dimension. Such a balance is achieved by combining semiparametric ideas with variable selection techniques. In particular, we propose a projection-based single-index model specification for very high-dimensional regressors. This model is used for practical CoVaR estimates with a systemically chosen indicator. In simulations we demonstrate the practical side of the semiparametric CoVaR method. The application to the U.S. financial sector shows good backtesting results and indicate market coagulation before the crisis period. Supplementary materials for this article are available online.

Keywords

Composite quasi-maximum likelihood estimation, CoVaR, Lasso, Minimum average contrast estimation, Model selection, Quantile single-index regression.

Direct Local Linear Estimation for Sharpe Ratio Function in Heteroscedastic Regression Models

Hongmei Lin[1], **Yuedong Wang**[2]

[1] *Shanghai University of International Business and Economics, China*
[2] *University of California Santa Barbara, United States*

Abstract

The heteroscedastic regression model has been widely used in financial econometrics which allows us to deal with nonlinearity and heteroscedasticity in financial time series. As the ratio of the mean and volatility functions, the Sharpe ratio is one of the most widely used risk/return measures in finance. In this paper we propose a new nonparametric method to estimate the Sharpe ratio function directly using local linear regression. We establish the asymptotic normality for the proposed estimator. Monte Carlo simulation studies show the proposed estimator has excellent finite-sample performance and outperform existing indirect method. We illustrate our method with a real data example.

Keywords

Heteroscedastic regression model; Sharpe ratio function; local linear regression; local likelihood estimation

Forecasting security's volatility using low-frequency historical data, high-frequency historical data and option-implied volatility

Huiling Yuan[1], **Yong Zhou**[2,3], **Zhiyuan Zhang**[4] and **Xiangyu Cui**[4]

[1] *Shanghai University of International Business and Economics, China*
[2] *East China Normal University, China*
[3] *Academy of Mathematics and Systems Sciences, Chinese Academy of Sciences, China*
[4] *Shanghai University of Finance and Economics.*

Abstract

Low-frequency historical data, high-frequency historical data and option data are three major sources, which can be used to forecast the underlying security's volatility. In this paper, we propose two econometric models, which integrate three information sources. In GARCH-Itô-OI model, we assume that the option-implied volatility can influence the security's future volatility, and the option-implied volatility is treated as an observable exogenous variable. In GARCH-Itô-IV model, we assume that the option-implied volatility can not influence the security's volatility directly, and the relationship between the option-implied volatility and the security's volatility is constructed to extract useful information of the underlying security. After providing the quasi-maximum likelihood estimators for the parameters and establishing their asymptotic properties, we also conduct a series of simulation analysis and empirical analysis to compare the proposed models with other popular models in the literature. We find that when the sampling interval of the high-frequency data is 5 minutes, the GARCH-Itô-OI model and GARCH-Itô-IV model has better forecasting performance than other models.

Keywords

High-frequency historical data, Low-frequency historical data, Option-implied volatility, Quasi-maximum likelihood estimators, Forecasting power

Singularity of LCM matrices on GCD closed sets with 9 elements

Mika Mattila

Tampere University, Finland

Abstract

In 1976 H. J. S. Smith defined an LCM matrix as follows: let $S = \{x_1, x_2, \ldots, x_n\}$ be a set of positive integers with $x_1 < x_2 < \cdots < x_n$. The LCM matrix $[S]$ on the set S is the $n \times n$ matrix with $\operatorname{lcm}(x_i, x_j)$ as its ij entry.

During the last 30 years singularity of LCM matrices has interested many authors. In 1992 Bourque and Ligh ended up conjecturing that if the GCD closedness of the set S (which means that $\gcd(x_i, x_j) \in S$ for all $i, j \in \{1, 2, \ldots, n\}$), suffices to guarantee the invertibility of the matrix $[S]$. However, a few years later this conjecture was proven false first by Haukkanen et al. [3] and then by Hong [4]. It turned out that the conjecture holds only on GCD closed sets with at most 7 elements but not in general for larger sets. However, the given counterexamples did not give much insight on why does the conjecture fails exactly in the case when $n = 8$. This situation was improved in articles [5] and [6], where a new lattice theoretic approach is introduced (the method is based on the fact that because the set S is assumed to be GCD closed, the structure $(S, |)$ actually forms a meet semilattice). In article [7] this lattice-theoretic method is then developed even further.

Since the cases $n \leq 8$ have been thoroughly studied in the above mentioned articles, the next natural step is to apply the methods to the case $n = 9$. This was done by Altinisik and Altintas in [1] as they consider the different lattice structures of $(S, |)$ that can result as a singular LCM matrix $[S]$. However, their investigation leaves two open questions, and the main purpose of this presentation is to provide solutions to them.

Keywords

LCM matrix, Bourque-Ligh conjecture, GCD-closed set, meet semilattice, Möbius function.

References

[1] Altinisik, E. and T. Altintas (2017). A note on the singularity of LCM matrices on GCD-closed sets with 9 elements. *Journal of Science and Arts 3(40)*, 413–422.

[2] Bourque, K. and S. Ligh (1992). On GCD and LCM matrices, *Linear Algebra Appl. 174* 65–74.

[3] Haukkanen, P., J. Wang and J. Sillanpää (1997). On Smith's determinant, *Linear Algebra Appl. 258*, 251–269.

[4] Hong, S. (1999) On the Bourque-Ligh conjecture of least common multiple matrices, *J. Algebra 218* 216–228.

[5] Korkee, I., M. Mattila and P. Haukkanen (2018). A lattice-theoretic approach to the Bourque-Ligh conjecture. *Linear Multilinear Algebra*, DOI: 10.1080/03081087.2018.1494695.

[6] Mattila, M., P. Haukkanen and J. Mäntysalo (2015). Studying the singularity of LCM-type matrices via semilattice structures and their Möbius functions. *J. Comb. Theory Series A 135*, 181–200.

[7] Mattila, M., P. Haukkanen and J. Mäntysalo. Studying the inertias of LCM matrices and revisiting the Bourque-Ligh conjecture. Submitted to *J. Comb. Theory Series A*, available at https://arxiv.org/abs/1809.02423.

[8] Smith, H. J. S. (1876). On the value of a certain arithmetical determinant, *Proc. London Math. Soc. 7* 208–212.

Accurate computations for some subclasses of totally positive matrices

Jorge Delgado[1], **Héctor Orera**[1] **and Juan Manuel Peña**[1]

[1] *Universidad de Zaragoza, Spain*

Abstract

In this talk, methods for the accurate computation of the bidiagonal decompositions of some totally positive matrices will be recalled. Then, we will show how these accurate bidiagonal decompositions allow us to calculate with high relative accuracy the eigenvalues, singular values, and inverses of the corresponding matrices.

Keywords

Algebraic computations, High relative accuracy, Total positivity.

Completion problems on operator matrices and its various applications

Dragana Cvetković Ilić[1]

[1] *Faculty of Sciences and Mathematics, University of Nis, Serbia*
[2]

Abstract

We will discuss certain results addressing some completion problems on operator matrices as well as their application to the questions such as that of existence of a linear bounded operator X for which A+CX is of a certain type, when A and C are given bounded linear operators, properties of the sum of operators and reverse order laws for the operators.

Keywords

completion problems, operator matrices, Fredholm operator

References

[1] D.S. Cvetković-Ilić, Invertible and regular completions of operator matrices, *Electronic Journal of Linear Algebra*, 30 (2015) 530-549.

[2] V. Pavlović, D.S. Cvetković-Ilić, Applications of completions of operator matrices to reverse order law for 1-inverses of operators on Hilbert spaces, *Linear Algebra Appl.*, 484 (2015) 219-236.

[3] V. Pavlović, D.S. Cvetković-Ilić, Completions of upper-triangular operator matrices to Kato nonsingularity, *Journal of Math. Anal. Appl.*, 433(2) (2016) 1230-1242.

[4] D.S. Cvetković-Ilić, Completions of upper-triangular matrices to left-Fredholm operators with non-positive index, *Journal of Operator Theory*, 76:1(2016), 101-114.

Consistent community detection in multi-layer network data

Jing Lei[1], **Kehui Chen**[2] **and Brian Lynch**[3]

[1]*Department of Statistics and Data Science, Carnegie Mellon University, USA*

[2,3] *Department of Statistics, University of Pittsburgh, USA*

Abstract

We consider multi-layer network data where the relationships between pairs of elements are reflected in multiple modalities and may be described by multivariate or even high-dimensional vectors. Under the multi-layer stochastic block model framework, we derive consistency results for a least squares estimation of memberships. Our theorems show that, as compared to single-layer community detection, a multi-layer network provides much richer information that allows for consistent community detection from a sparser network, with required edge density reduced by a factor of the square root of the number of layers. In this talk, we will present the theoretical analysis, including the development of a new spectral bound for tensor network data, as well as simulations and a data example.

Keywords

Community detection; Consistency; Sparse network; Tensor concentration bound.

A correlation break test based on self-normalization

Ji-Eun Choi[1] **and Dong Wan Shin**[2]

[1]*Ewha Womans University of Statistics, Republic of Korea*
[2] *Ewha Womans University of Statistics, Republic of Korea*

Abstract

We propose a new CUSUM-type test for a correlation break based on the self-normalization method. The self-normalization test is implemented much simpler than the existing tests based on the long-run variance which need to specify bandwidths and to evaluate complicated consistent estimators for the long-run variances. The limiting null distribution and consistency of the proposed test under an alternative are established. A Monte Carlo simulation demonstrates that the self-normalization test has reasonable size and comparable power, but the existing tests have severe size distortions for serially correlated and/or conditionally heteroscedastic samples. An analysis of returns and realized volatilities of some US, Europe and Japan stock prices by the proposed test indicates absence of correlation break during the period of global financial crisis while those by the existing tests indicate presence of it.

Keywords

Conditional heteroscedasticity; Correlation break; CUSUM test; Self-normalization; Serial dependence.

A review of Dempster's non-exact test for high-dimensional mean vector

Zhizheng Wang

Linnaeus university, Sweden

Abstract

In this paper, we study the problem of testing for mean vectors with a particular focus on the so-called Dempster's non-exact test (Dempster 1958, 1960[1, 2]). The recent development around the topic is reviewed. The asymptotic null and alternative distributions of the test statistic are summarized. Dempster's test is applied to an empirical study - gene-set testing. Size and power simulation results are presented for the one-sample problem.

Keywords

Dempster's test, High-dimensional data, Asymptotic distribution, Kolmogorov condition, MANOVA, Hypothesis testing.

References

[1] Dempster, A.P. (1958). A high dimensional two sample significance test. *The Annals of Mathematical Statistics*, 995–1010.

[2] Dempster, A.P. (1960). A significance test for the separation of two highly multivariate small samples. *Biometrics, 16*(1), 41–50.

List of Participants

	Name	University	Country / Area
1	Oskar Maria Baksalary	Adam Mickiewicz University	Poland
2	Rajendra Bhatia	Ashoka University	India
3	Kehui Chen	University of Pittsburgh	United States
4	Ji-Eun Choi	Ewha Womans University of Statistics	Republic of Korea
5	Jorge Delgado Gracia	Universidad de Zaragoza	Spain
6	A.M. Elsawah	BNU-HKBU United International College	Hong Kong, China
7	Caiyun Fan	Shanghai University of International Business and Economics	China
8	Yan Fan	Shanghai University of International Business and Economics	China
9	Kai-Tai Fang	BNU-HKBU United International College	Hong Kong, China
10	Jing Fu	Chongqing University of Science and Technology	China
11	Ya Gao	Shanghai University of International Business and Economics	China
12	Chengcheng Hao	Shanghai University of International Business and Economics	China
13	Guikai Hu	East China University of Technology	China
14	Jianhua Hu	Shanghai University of Finance and Economics	China
15	Jeffrey Hunter	Auckland University of Technology	New Zealand
16	Dragana Cvetkovic-Ilic	University of Nis	Serbia
17	Bo Jiang	Shandong Institute of Business and Technology	China
18	Hongjie Jiang	Guangxi University for Nationalities	China
19	Dandan Jiang	Xi'an Jiaotong University	China
20	Shimizu Kunio	The Institute of Statistical Mathematics	Japan
21	Hanyu Li	Chongqing University	China
22	Feng Li	Central University of Finance and Economics	China
23	Zhengtao Li	Shanghai University of International Business and Economics	China
24	Yuli Liang	Örebro University	Sweden
25	Hongmei Lin	Shanghai University of International Business and Economics	China
26	Yonghui Liu	Shanghai University of International Business and Economics	China
27	Hui Liu	Shanghai University of International Business and Economics	China
28	Shuangzhe Liu	University of Canberra	Australia

29	Guanfu Liu	Shanghai University of International Business and Economics	China
30	Rui Li	Shanghai University of International Business and Economics	China
31	Changyu Lu	Shanghai Lixin University of Accounting and Finance	China
32	Tiefeng Ma	Southwestern University of Finance and Economics	China
33	Mika Mattila	University of Tampere	Finland
34	Lynn Roy LaMotte	Louisiana State University Health Sciences Center	United States
35	Seng-Huat Ong	UCSI University	Malaysia
36	Jianxin Pan	Manchester University	United Kingdom
37	Wenjie Pan	Shanghai University of International Business and Economics	China
38	Yiluo Peng	Shanghai University of International Business and Economics	China
39	Karantha Manjunatha Prasad	Manipal Academy of Higher Education	India
40	Simo Puntanen	University of Tampere	Finland
41	Tatjana von Rosen	Stockholm University	Sweden
42	Jian Shi	Hebei University	China
43	Lei Shi	Yunnan University of Finance and Economics	China
44	Jiaqi Shi	Shanghai University of International Business and Economics	China
45	Yu Shi	Shanghai University of International Business and Economics	China
46	Martin Singull	Linköping University	Sweden
47	Cihai Sun	Shanghai University of International Business and Economics	China
48	Ling Tang	Chongqing University	China
49	Yu Tang	Soochow University	China
50	Yongge Tian	Central University of Finance and Economics	China
51	Kimmo Vehkalahti	University of Helsinki	Finland
52	Julia Volaufova	Louisiana State University Health Sciences Center	United States
53	Hongxing Wang	Guangxi University for Nationalities	China
54	Lijuan Wang	Shanghai University of International Business and Economics	China
55	Zhizheng Wang	Linnaeus University	Sweden
56	Ya Wang	Shanghai University of International Business and Economics	China
57	Bohui Wang	Shanghai University of International Business	China

		and Economics	
58	Qiuting Xian	Shanghai University of International Business and Economics	China
59	Chuanfu Xiao	Peking University	China
60	Jianwen Xu	Chongqing University	China
61	Shan Xue	Shanghai University of International Business and Economics	China
62	Yajie Yang	Shanghai University of International Business and Economics	China
63	Huiling Yuan	Shanghai University of International Business and Economics	China
64	Aijun Zhang	The University of Hong Kong	Hong Kong, China
65	Fuzhen Zhang	Nova Southeastern University	United States
66	Xiaohui Zhao	Shanghai University of International Business and Economics	China
67	Shurong Zheng	North Eastern Normal University	China
68	Xiaodong Zhou	Shanghai University of International Business and Economics	China
69	Yongdao Zhou	Nankai University	China
70	Yuanyang Zhu	Chongqing University	China
71	Lixing Zhu	Hong Kong Baptist University	Hong Kong, China
72	Fukang Zhu	Jilin University	China

Introduction to Shanghai University of International Business and Economics

Founded in 1960, the Shanghai University of International Business and Economics (former Shanghai Institute of Foreign Trade) was affiliated to the Ministry of Foreign Trade and Economic Cooperation (now the Ministry of Commerce) until September 1994, since when it has been under the administration of the Shanghai Municipal People's Government. In April 2013, it was granted by the Ministry of Education (MOE) a university status and renamed the Shanghai University of International Business and Economics (SUIBE).

SUIBE consists of the School of Business, School of Foreign Languages, School of Finance, School of Law, School of Management, School of Accounting, School of Tourism and Event Management, School of Statistics and Information, School of International Studies, School of Continuing Education, School of Trade Negotiation. SUIBE offers 32 undergraduate programs. Master Degrees are conferred in 7 of its first-level disciplines and 33 of its second-level disciplines. There are over 1000 faculty and staff members working at SUIBE and above 400 of them are professors and associate professors. More than 10,000 students are studying in the undergraduate and graduate programs and some 2000 international students in various programs.

SUIBE gives top priority to the development of students and attaches equal importance to their knowledge acquisition, skills building, and qualities upbringing. It is guided by such education philosophy that SUIBE is committed to improving students' cross-cultural communication abilities and practical skills in their professional fields. Jobs created by SUIBE alumni who started up their own businesses have outnumbered SUIBE's graduates put together in the years since its establishment.

Dedicated to serving social and economic needs through scientific research, SUIBE faculty members have undertaken hundreds of research projects funded at the state and ministerial level. SUIBE made positive contributions to China's resumption of its contracting party status to the GATT and China's accession to the WTO. SUIBE is among the first institutions across the world selected for hosting a WTO Chair.

SUIBE has been an active player in international exchanges and cooperation in education and established extensive partnerships with its overseas counterparts from more than 100 countries. Three Confucius Institutes have been established in Croatia and Slovenia respectively.

Taking talent development, scientific studies, social services, and cultural continuity and innovation as its missions, SUIBE is making every effort to build itself into an internationalized cutting-edge university with a diversity of disciplines while building on its traditional strengths in international business and economics so as to make contributions to the development of Shanghai, China and the world at large.

Introduction to the School of Statistics and Information
Shanghai University of International Business and Economics

School of Statistics and Information (former School of Business Information) at SUIBE was established in 2008, which united the research and teaching of statistics, applied mathematics and computer science into one large school at SUIBE.

The School offers four undergraduate degrees (BSc or BA) in Statistics, Economic Statistics, Information Management & Information Systems and Data Science & Big Data Technology. At the postgraduate level, there are five Master's programs: Mathematical Economics, Statistics, Business Information Management, Master of Applied Statistics and Master of International Business. It consists of the Department of Statistics, the Department of Information Management, the Department of Applied Mathematics, the Department of Data Science and the Center for Big Data in Business.

The School has a highly qualified, highly experienced and passionate faculty. Among 69 faculty members, over 60% of the faculties are professors or associate professors, over 80% hold a doctoral degree, and 72% are aged under 45. Included in the faculty are two Youth Eastern Scholars in Shanghai, three Shanghai Pujiang Scholars, four Baogang Excellent Teachers, an Oversea Famous Teacher, a Siyuan Scholar and two distinguished professors.

During the past three years, the faculty have been awarded over twenty research grants at the state and ministerial level, and have published more than 50 high quality papers in peer refereed journals, including Journal of Royal Statistical Society Series B (Statistical Methodology), Journal of Business & Economic Statistics, Insurance Mathematics and Economics, etc.