International Conference on Trends and Perspectives in Linear Statistical Inference and 21st International Workshop on Matrices and Statistics

Book of Abstracts

July 16 – 20, 2012
Będlewo, Poland

Edited by

Katarzyna Filipiak
Department of Mathematical and Statistical Methods,
Poznań University of Life Sciences, Poland

and

Martin Singull
Department of Mathematics, University of Linköping, Sweden

Printed by

Bogucki Wydawnictwo Naukowe, www.bogucki.com.pl,
from camera ready materials provided by Editors

ISBN: 978-83-63400-12-5

Contents

Part V. Contributed Talks

Part VI. Posters

Part I

Introduction

The International Conference on Trends and Perspectives in Linear Statistical Inference, LinStat'2012, and the 21st International Workshop on Matrices and Statistics, IWMS'2012, will be held on July 16-20, 2012 in the Mathematical Research and Conference Center of the Polish Academy of Sciences at Będlewo near Poznań. This is the follow-up of the 2008 and 2010 editions held in Będlewo, Poland and in Tomar, Portugal.

The purpose of the meeting is to bring together researchers sharing an interest in a variety of aspects of statistics and its applications as well as matrix analysis and its applications to statistics, and offer them a possibility to discuss current developments in these subjects. The format of this meeting will involve plenary talks, special sessions, contributed talks and posters.
The conference will mainly focus on a number of topics: estimation, prediction and testing in linear models, robustness of relevant statistical methods, estimation of variance components appearing in linear models, generalizations to nonlinear models, design and analysis of experiments, including optimality and comparison of linear experiments, and applications of matrix methods in statistics.
The work of young scientists has a special position in the LinStat'2012 to encourage and promote them. The best poster as well as the best talk of Ph.D. students will be awarded. Prize-winning works will be widely publicized and promoted by the conference.

It is expected that many of presented papers will be published, after refereeing, in a Special Issue of each of the journals: *Communications in Statistics - Theory and Methods* and *Communications in Statistics - Simulation and Computation*, associated with this conference. All papers submitted must meet the publication standards of mentioned journals and will be subject to normal refereeing procedure.

Committees and Organizers

The Scientific Committee for LinStat'2012 comprises

- Augustyn Markiewicz (chair, Poznań University of Life Sciences, Poland),
- Anthony C. Atkinson (London School of Economics, UK),
- João T. Mexia (New University of Lisbon, Portugal),
- Simo Puntanen (University of Tampere, Finland),
- Dietrich von Rosen (Swedish University of Agricultural Sciences, Uppsala and Linköping University, Sweden),
- Götz Trenkler (Technical University of Dortmund, Germany),
- Roman Zmyślony (University of Zielona Góra, Poland).

The Scientific Committee for IWMS'2012 comprises

- Simo Puntanen (chair, University of Tampere, Finland),
- George P. H. Styan (honorary chair, McGill University, Montreal, Canada),
- S. Ejaz Ahmed (Brock University, St. Catharines, Canada),
- Jeffrey Hunter (Auckland University of Technology, New Zealand),
- Augustyn Markiewicz (Poznań University of Life Sciences, Poland),
- Dietrich von Rosen (Swedish University of Agricultural Sciences, Uppsala and Linköping University, Sweden),
- Götz Trenkler (Technical University of Dortmund, Germany),
- Júlia Volaufová (Louisiana State University, Health Sciences Center, New Orleans, USA),
- Hans J. Werner (University of Bonn, Germany).

The Organizing Committee comprises

- Katarzyna Filipiak (chair, Poznań University of Life Sciences, Poland),
- Francisco Carvalho (Polytechnic Institute of Tomar, Portugal),
- Małgorzata Graczyk (Poznań University of Life Sciences, Poland),
- Jan Hauke (Adam Mickiewicz University, Poznań, Poland),
- Agnieszka Łacka (Poznań University of Life Sciences, Poland),
- Martin Singull (University of Linköping, Sweden),
- Waldemar Wołyński (Adam Mickiewicz University, Poznań, Poland).

The Organizers are

- Stefan Banach International Mathematical Center, Institute of Mathematics of the Polish Academy of Sciences, Warsaw,
- Faculty of Mathematics and Computer Science, Adam Mickiewicz University, Poznań,
- Institute of Socio-Economic Geography and Spatial Management, Adam Mickiewicz University, Poznań,
- Department of Mathematical and Statistical Methods, Poznań University of Life Sciences.

Call for Papers

We are pleased to announce a special issue of *Communications in Statistics – Theory and Methods* and *Communications in Statistics – Simulation and Computation* (Taylor & Francis) devoted to LinStat-IWMS'2012.

They will include selected papers strongly correlated to the talks of the conference and with emphasis on advances on linear models and inference.

Coordinator-Editor: N. Balakrishnan

Guest Editors of Theory and Methods: Júlia Volaufová and Augustyn Markiewicz

Guest Editors of Simulation and Computation: Simo Puntanen and Katarzyna Filipiak

All papers submitted must meet the publication standards of Communications in Statistics (see: http://www.math.mcmaster.ca/bala/comstat/) and will be subject to normal refereeing procedure. The deadline for submission of papers is **the end of November, 2012**.

Papers should be submitted using the web site

http://mc.manuscriptcentral.com/lsta

If the author does not have account, he should create one. The contributors must choose **"Special Issue – Advances on Linear Models and Inference"** (Theory and Methods) and **"Special Issue – Advances on Linear Models and Inference: Computational Aspects"** (Simulation and Computation) as the manuscript type.

Part II

Program

Program

Sunday, July 15, 2012

14:00 – 19:00 Registration
19:00 – Reception Dinner

Monday, July 16, 2012

7:30 – 8:50 Breakfast

Plenary Session

8:55 – 9:00 *Opening*
9:00 – 9:45 R. A. Bailey: *Optimal design of experiments with very low average replication*
9:45 – 10:30 R. Bhatia: *Geometric mean of matrices*

10:30 – 11:00 Coffee Break

Parallel Session A – Multivariate Analysis part I

11:00 – 11:30 D. von Rosen: *From linear to multilinear models*
11:30 – 11:50 S. Toprak: *A new approach to adaptive spline threshold autoregression by using Tikhonov regularization and continuous optimization*

Parallel Session B – Matrices for Linear Models part I

11:00 – 11:30 H. J. Werner: *On the linear aggregation problem in the general Gauß-Markov model*
11:30 – 11:50 F. S. Kurnaz: *A new Liu-type estimator*

11:50 – 12:00 Break

Parallel Session A – Multivariate Analysis part II

12:00 – 12:20 C. Cuevas-Covarrubias: *Mutual Principal Components, reduction of dimensionality in statistical classification*
12:20 – 12:40 Z. Hanusz: *Simulation study on improved Shapiro-Wilk test of normality*
12:40 – 13:00 D. Klein: *Estimators of serial covariance parameters in multivariate linear models*

Parallel Session B – Mixed Models part I

12:00 – 12:35 J. Volaufová: *On testing linear hypotheses in general mixed models*
12:35 – 13:00 B. Arendacká: *Building stones for inference on variance component*

13:00 – 15:00 Lunch

Parallel Session A – Robust Statistical Methods part I

15:00 – 15:20 A. C. Atkinson: *Very robust regression*
15:20 – 16:00 D. Perrotta: *Considerations on sampling, precision and speed of robust regression estimators*

Parallel Session B – General part I

15:00 – 15:30 A. Lee: *Getting the "correct" answer from survey responses: an application of regression mixture models*
15:30 – 16:00 A. Alin: *Bootstrap confidence regions for multinomial probabilities based on penalized power-divergence test statistics*

16:00 – 16:30 Coffee Break

Parallel Session A – Experimental Designs part I

16:30 – 17:05 S. Gilmour: Q_B*-optimal saturated two-level main effects designs*
17:05 – 17:30 H. Abebe: *On the choice of a prior distribution for Bayesian D-optimal designs for the logistic regression model*

Parallel Session B – Mixed Models part II

16:30 – 17:05 F. Vaida: *Conditional AIC for linear mixed effects models*
17:05 – 17:30 A. Michalski: *On admissibility of decision rules derived from submodels in two variance components model*

17:30 – 17:50 Coffee Break

Parallel Session A – Applications part I

17:50 – 18:20 J.-P. Masson: *About the evolution of the genomic diversity in a population reproducing through partial asexuality*
18:20 – 18:40 K. Bartoszek: *Multivariate linear phylogenetic comparative models and adaptation*

Parallel Session B – Mixed Models part III

17:50 – 18:20 V. Witkovský: *On exact and approximate simultaneous confidence regions for parameters in normal linear model with two variance components*
18:20 – 18:40 N. Demetrashvili *Estimating intraclass correlation and its confidence interval in linear mixed models*
18:40 – 19:00 N. Vahabi: *Multilevel Rasch model and item response theory*

19:00 – Dinner
20:30 – Concert

Tuesday, July 17, 2012

7:30 – 9:00 Breakfast

Plenary Session

9:00 – 9:45 K. M. Prasad: *Partial orders on matrices and the column space decompositions*
9:45 – 10:30 P. C. Rodrigues: *Low-rank approximations and weighted low-rank approximations*

10:30 – 11:00 Coffee Break

Plenary Session

11:00 – 11:45 P. E. Oliveira: *Smoothing discrete distributions*

11:45 – 12:00 Break

Parallel Session A – Mixed Models part IV

12:00 – 12:35 L. R. LaMotte: *On inverse prediction in mixed models*
12:35 – 13:00 Y. Liang: *Variance components estimability in multilevel models with block circular symmetric covariance structure*

Parallel Session B – Experimental Designs part II

12:00 – 12:35 J. P. Morgan: *Weighting, model transformation, and design optimality*
12:35 – 13:00 C. May: *A sequential generalized DKL-optimum design for model selection and parameter estimation in nonlinear nested models*

13:00 – 15:00 Lunch
15:00 – 21:00 Excursion

Wednesday, July 18, 2012

7:30 – 9:00 Breakfast

Plenary Session

9:00 – 9:45 Somnath Datta: *Nonparametric regression for sojourn time distributions in a multistate model*

9:45 – 10:30 O. Bluder: *Investigation of Bayesian Mixtures-of-Experts models to predict semiconductor lifetime*

10:30 – 11:00 Coffee Break

Parallel Session A – Robust Statistical Methods part II

11:00 – 11:50 D. P. Wiens: *Robust model-based sampling designs*

Parallel Session B – High-Dimensional Data part I

11:00 – 11:30 Susmita Datta: *Nonparametric regression using partial least squares dimension reduction in multistate models*

11:30 – 11:50 T. Górecki: *First and second derivative in time series classification using DTW*

11:50 – 12:00 Break

Parallel Session A – Matrices for Linear Models part II

12:00 – 12:30 H. Drygas: *Linear models in the face of Diabetes Mellitus: the influence of physical activity*

12:30 – 13:00 R. Sund: *Muste - editorial environment for matrix computations*

Parallel Session B – High-Dimensional Data part II

12:00 – 12:35 A. Khalili: *Simultaneous fixed and random effect selection in finite mixture of linear mixed-effect models*

12:35 – 13:00 A. Göktaş: *A comparison of logit and probit models for a binary response variable via a new way of data generalization*

13:00 – 15:00 Lunch

Parallel Session A – Model Selection, Penalty Estimation and Applications

15:00 – 15:30 S. E. Ahmed: *Absolute Penalty and Shrinkage Estimation in Weibull censored regression model*

15:30 – 16:00 E. P. Liski: *Model averaging via penalized least squares in linear regression*

16:00 – 16:20 N. Acar: *Model selection in log-linear models by using information criteria*

Parallel Session B – Optimum Design for Mixed Effects Regression Models

15:00 – 15:30 B. Bogacka: *Optimum designs for enzyme kinetic models with co-variates*

15:30 – 16:00 F. Mentré: *Two-stage optimal designs in nonlinear mixed effect models: application to pharmacokinetics in children*

16:00 – 16:20 M. Prus: *Optimal designs for prediction of individual effects in random coefficient regression models*

16:20 – 16:50 Coffee Break

16:50 – **Poster Session**

A. Bachratá: *Using methods of stochastic optimization for constructing optimal experimental designs with cost constraints*

S. Bělašková: *Regression model of AMH*

S. Donevska: *Regression analysis between parts of compositional data*

C. Fernandes: *COBS and stair nesting - segregation and crossing*

A. Göktaş: *Validity of the assumed link functions for some binary choice models based on the bootstrap confidence band with R*

M. Graczyk: *Regular E-optimal spring balance weighing design with correlated errors*

R. Gupta: *Estimation of parameters of structural change under small sigma approximation theory*

D. Kayzer: *Canonical variate analysis of chlorophyll a, b and a+b content in tropospheric ozone-sensitive and resistant tobacco cultivars exposed in ambient air conditions*

P.-H. Liau: *Latin square designs and fractional factorial designs*

D. G. Pereira: *Weighted linear joint regression analysis*

M. Przystalski: *Modeling resistance to oat crown rust in series of oat trials*

P. Ramos: *Estimation of variance components in balanced, staggered and stair nested designs*

Ł. Smaga: *D-optimal chemical balance weighing designs for three objects if n=2 (mod 4)*

T.-S. Tsou: *Is the skew t distribution truly robust?*

M. Tučková: *Design of experiment for regression models with constrains*

R. Zmyślony: *Inference for the interclass correlation in familial data using small sample asymptotics*

19:00 – 20:00 Concert: *Indian Singing by Susmita Datta*

20:00 – Barbecue

Thursday, July 19, 2012

7:30 – 9:00 Breakfast

Plenary Session

9:00 – 9:45 P. Šemrl: *Adjacency preserving maps*

9:45 – 10:30 G. P. H. Styan: *An illustrated introduction to Euler and Fitting factorizations and Anderson graphs for classic magic matrices*

10:30 – 11:00 Coffee Break

Plenary Session – George P. H. Styan's 75th Birthday

11:00 – 11:25 G. Trenkler: *The Luoshu and most perfect pandiagonal magic squares*

11:25 – 11:50 K. Conradsen: *Multivariate analysis of polarimetric SAR images*

11:50 – 12:00 Break

Parallel Session A – Matrices for Linear Models part III

12:00 – 12:20 A. Erkoç: *A graphical evaluation of Robust Ridge Regression in mixture experiments*

12:20 – 12:40 S. Türkan: *Cook's distance for ridge estimator in semiparametric regression*

12:40 – 13:00 K. Nosek: *Change-point detection in two-phase regression with inequality constraints*

Parallel Session B – George P. H. Styan's 75th Birthday

12:00 – 12:30 O. M. Baksalary: *Some comments on joint papers by George P.H. Styan and the Baksalarys*

12:30 – 13:00 C. A. Coelho: *Celebrating George P. H. Styan's 75th birthday and my meetings with him*

13:00 – 15:00 Lunch

Parallel Session A – Multivariate Analysis part III

15:00 – 15:20 M. Onozawa: *Tests for profile analysis based on two-step monotone missing data*

15:20 – 15:40 Ł. Waszak: *Functional discriminant coordinates*

15:40 – 16:00 R. Enomoto: *Normality test based on Song's multivariate kurtosis*

Parallel Session B – George P. H. Styan's 75$^{\text{th}}$ Birthday

15:00 – 15:30 S. J. Haslett: *Equivalence of linear models under changes to data, design matrix, or covariance structure*
15:30 – 15:45 J. Hauke: *Nonnegativity of eigenvalues of sum of diagonalizable matrices*
15:45 – 16:00 C. A. Coelho: *On the distribution of linear combinations of chi-square random variables*

16:00 – 16:30 Coffee Break

Parallel Session A – Multivariate Analysis part IV

16:30 – 17:05 J. Najim: *Eigenvalue estimation of covariance matrices of large dimensional data*
17:05 – 17:30 J. Pielaszkiewicz: *Asymptotic spectral analysis of matrix quadratic forms*

Parallel Session B – George P. H. Styan's 75$^{\text{th}}$ Birthday

16:30 – 17:00 A. J. Scott: *Fitting Generalized Linear Models to sample survey data*
17:00 – 17:30 K. L. Chu: *The magic behind the construction of certain Agrippa-Cardano type magic matrices*

17:30 – 17:50 Coffee Break

Parallel Session A – General part II

17:50 – 18:15 U. Beyaztaş: *Jackknife-after-Bootstrap as logistic regression diagnostic tool*
18:15 – 18:40 D. Haki: *Improved estimation of the mean by using coefficient of variation as a prior information in ranked set sampling*
18:40 – 19:00 S. Takahashi: *Simultaneous confidence intervals among mean components in elliptical distributions*

Parallel Session B – George P. H. Styan's 75$^{\text{th}}$ Birthday

17:50 – 18:15 P. Bertrand: *Study with George Styan*
18:15 – 18:30 P. Loly's video: *Using singular values for comparing and classifying magical squares (natural magic and Latin)*
18:30 – 19:00 S. Puntanen: *Oh, still crazy after all these years?*

19:30 – Conference Dinner
– **After Dinner Speaker:** A. J. Scott
– **After Dessert Speaker:** S. Puntanen
– **YSA Prize Ceremony**

Friday, July 20, 2012

7:30 – 9:00 Breakfast

Plenary Session

9:00 – 9:45 T. Mathew: *Tolerance intervals in general mixed effects models using small sample asymptotics*

9:45 – 10:30 C. Hao: *Influential observations in the extended Growth Curve model with crossover designs*

10:30 – 11:00 Coffee Break

Parallel Session A – Multivariate Analysis part V

11:00 – 11:25 M. Fonseca: *Linear models with doubly exchangeable distributed errors*

11:25 – 11:50 A. Roy: *Classification of higher-order data with separable covariance and structured multiplicative or additive mean models*

Parallel Session B – Applications part II

11:00 – 11:25 T. Kossowski: *The Moran coefficient for non-normal data: revisited with some extensions*

11:25 – 11:50 R. Covas: *Some math on the electricity market by a generalization of the Black-Scholes formula*

11:50 – 12:00 Break

Parallel Session A – General part III

12:00 – 12:30 D. Kosiorowski: *Robust monitoring of multivariate data stream*

12:30 – 13:00 D. İnan: *Modeling multiple time series data using wavelet-based support vector regression*

Parallel Session B – Experimental Designs part III

12:00 – 12:20 L. Filová: *Constructing efficient exact designs of experiments using integer quadratic programming*

12:20 – 12:40 S. D. Georgiou: *Latin hypercube designs and block-circulant matrices*

12:40 – 13:00 S. Stylianou: *Construction and analysis of D-optimal edge designs*

13:00 – 15:00 Lunch

Parallel Session A – Mixed Model part V

15:00 – 15:35 E. Fišerová: *Sensitivity analysis in mixed models*

15:35 – 16:00 F. Carvalho: *Linear and quadratic sufficiency in mixed model*

Parallel Session B – Experimental Designs part IV

15:00 – 15:35 T. Caliński: *On combining information in a generally balanced nested block design*
15:35 – 16:00 A. Łacka: *Analysis of an experiment in a generally balanced nested block design*

16:00 – 16:30 Coffee Break

Parallel Session A – Mixed Model part VI

16:30 – 17:05 B. Schaffrin: *On the Errors-In-Variables Model with singular covariance matrices*
17:05 – 17:30 M. Salehi: *Multilevel linear mixed model for the analysis of longitudinal studies*

Parallel Session B – Experimental Designs part V

16:30 – 17:05 J. Kunert: *Optimal designs for the Michaelis Menten model with correlated observations*
17:05 – 17:30 K. Filipiak: *On universal optimality of circular repeated measurements designs*

17:30 – 17:50 Coffee Break

Parallel Session A – Matrices for Linear Models part IV

17:50 – 18:10 B. Asikgil: *A novel approach for estimation of seemingly unrelated linear regressions with high order autoregressive disturbances*
18:10 – 18:30 B. Eygi Erdogan: *A comparison of different parameter estimation methods in fuzzy linear regression*
18:30 – 18:50 N. Güler: *A study on the equivalence of BLUEs under a general linear model and its transformed models*

Parallel Session B – Experimental Designs part VI

17:50 – 18:25 D. Uciński: *D-optimum hybrid sensor network deployment for parameter estimation of spatiotemporal processes*
18:25 – 18:50 A. Markiewicz: *Optimality of neighbor designs*

18:50 – 19:50 Closing

19:00 – Dinner

Saturday, July 21, 2012

8:00 – 10:00 Breakfast

Part III

Special Sessions

Robust Statistical Methods

Anthony C. Atkinson

London School of Economics, UK

Abstract

Robust statistical methods are intended to behave well in the presence of departures from the model that explains the greater part of the data. A contamination model for the data is that the observations y have density

$$f(y) = (1-\epsilon)f_1(y, \theta_1) + \epsilon f_2(y, \theta_2). \tag{1}$$

The simplest example is when $f_1(y, \theta_1) = \phi(\mu, \sigma^2)$, the normal distribution. When there is no contamination ($\epsilon = 0$) the minimum variance unbiased estimator of μ is the sample mean $\bar{y}$. Now suppose that there is some contamination. In the finite sample case, even with $\epsilon = 1/n$, the sample mean has unbounded bias as the observation from $f_2(.)$ becomes increasingly extreme. The estimate breaks down as the observation goes to $\pm\infty$. Asymptotically (as $n \to \infty$) the sample mean has zero breakdown.

The sample median, on the other hand is not so affected. Asymptotically up to half the observations can be moved arbitrarily far away from μ with the median providing an unbiased estimator. However, the variance of the median is asymptotically $\pi/2$, so that the efficiency of the median as an estimator of location is 0.637, although the breakdown point is 50%. An aim of robust statistics is to find estimators that are unbiased in the presence of contamination whilst achieving the Cramer-Rao lower bound. Of course, such estimators do not exist, but breakdown can be traded against variance inflation.

The trade-off is achieved through the use of M-estimators and their extensions. Given an estimator of μ, say $\tilde{\mu}$, the residuals are defined as

$$r_i(\tilde{\mu}) = y_i - \tilde{\mu}. \tag{2}$$

As is well known, the least squares estimate of μ, which is also the maximum likelihood estimate, minimizes the sum of squares

$$\sum_{i=1}^{n} \{r_i(\mu)/\sigma\}^2. \tag{3}$$

Of course the value of σ is irrelevant.

Traditional robust estimators attempt to limit the influence of outliers by replacing the square of the residuals in (3) by a function ρ of the residuals

which is bounded. The M (Maximum likelihood like) estimate of μ is the value that minimizes the objective function

$$\sum_{i=1}^{n} \rho\{r_i(\mu)/\sigma\}. \tag{4}$$

Of the numerous form that have been suggested for $\rho(.)$ (Adrews et al., 1972, Hampel et al., 1986, Huber and Ronchetti, 2009) perhaps the most popular choice is Tukey's Biweight function

$$\rho(x) = \begin{cases} \frac{x^2}{2} - \frac{x^4}{2c^2} + \frac{x^6}{6c^4} & \text{if } |x| \leq c \\ \frac{c^2}{6} & \text{if } |x| > c, \end{cases} \tag{5}$$

where c is a crucial tuning constant. For small $x, \rho(x)$ behaves like (3). For large $|x|$ the residuals are constant; the effect of extreme observations is mitigated.

In equation (4) it is assumed that σ is known, yielding the estimate $\tilde{\mu}_M(\sigma)$. Otherwise, an M-estimator of scale $\tilde{\sigma}_M$ is defined as the solution to the equation

$$\frac{1}{n}\sum_{i=1}^{n} \rho\{r_i(\mu)/\sigma\} = K_c, \tag{6}$$

where both μ and σ are iteratively jointly estimated. K_c and c are related constants which are linked to the breakdown point of the estimator of μ.

Regression, which will be the subject of two of the talks, is more difficult. If the contamination is only in the y direction, M-estimation is appropriate. However, if the x values may also be outlying, leverage points may be present. Then, not only is ordinary least squares exceptionally susceptible to the presence of outliers, but so are M-estimates. Instead, very robust methods, with an asymptotic breakdown point of 50% of outliers are to be preferred.

Very robust regression was introduced by Rousseeuw (1984) who developed suggestions of Hampel (1975) that led to the Least Median of Squares (LMS) and Least Trimmed Squares (LTS) algorithms.

In the regression model $y_i = x_i^T\beta + \epsilon_i$, the residuals in (2) become $r_i(\tilde{\beta}) = y_i - x_i^T\tilde{\beta}$. The LMS estimator minimizes the hth ordered squared residual $r^2_{[h]}(\beta)$ with respect to β, where $h = \lfloor(n+p+1)/2\rfloor$ and $\lfloor.\rfloor$ denotes integer part.

The convergence rate of $\tilde{\beta}_{\text{LMS}}$ is $n^{-1/3}$. Rousseeuw (1984, p. 876) also suggested Least Trimmed Squares (LTS) which has a convergence rate of $n^{-1/2}$ and so better properties than LMS for large samples. As opposed to minimising the median squared residual, $\tilde{\beta}_{\text{LTS}}$ is found to

$$\text{minimize } SS_{\text{T}}\{\hat{\beta}(h)\} = \sum_{i=1}^{h} e_i^2\{\hat{\beta}(h)\}, \tag{7}$$

where, for any subset $\mathcal{H}$ of size h, the parameter estimates $\hat{\beta}(h)$ are straightforwardly obtained by least squares.

Unlike M-estimation, these procedures do not require an estimate of σ^2. However, the estimate is required for outlier detection. Let the minimum value of (7) be $SS_{\mathrm{T}}(\tilde{\beta}_{\mathrm{LTS}})$. The estimator of σ^2 is based on this residual sum of squares. However, since the sum of squares contains only the central h observations from a normal sample, the estimate needs scaling. The factors come from the general results of Tallis (1963) on elliptical truncation.

The LMS and LTS estimators are least squares estimates from carefully selected subsets of the data, asymptotically one half for LTS. If there are no, or only a few, outliers, such estimates will be inefficient. To increase efficiency, reweighted versions of the LMS and LTS estimators can be computed, using larger subsets of the data. These estimators are found by giving weight 0 to observations which are determined to be outliers when using the parameter estimates from LMS or LTS. Least squares is then applied to the remaining observations.

An alternative to these forms of very robust estimation is deletion of outliers, starting from a fit to all the data (Cook and Weisberg, 1982, Atkinson, 1985). If there are few outliers, the resulting estimators will be based on most of the data and so will be more efficient than those based on smaller subsets. However, in the presence of many outliers these backwards methods can fail. Atkinson and Riani (2000) suggest a Forward Search (FS) in which least squares is used to fit the model to subsets of the data of increasing size. The process stops when all observations not used in fitting are determined to be outliers. See Atkinson et al. (2010) for a recent discussion of the FS.

In LMS and LTS inference is made from models fitted to subsets of the data of one or two sizes, with perhaps subsets of three different sizes for the reweighted versions. Instead, in the FS the model is progressively fitted to subsets of increasing size. The procedure needs both to reject all outliers, in order to provide unbiased estimates of the parameters, and to use as many observations as possible in the fit in order to enhance efficiency. One thread in the session will be the improved properties of the estimates that result from using this flexible, data-dependent subset size for parameter estimation.

A second thread in the session has to do with efficient computation. The LMS and LTS estimates used are approximations found by least squares fitting to many subsets of observations. As a consequence LMS and LTS estimation (and, in general, all algorithms of robust statistics) spend a large part of the computational time in sampling subsets of observations and then computing parameter estimates from the subsets. In addition, each new subset has to be checked as to whether it is in general position (that is, it has a positive determinant). For these reasons, when the total number of possible subsets is much larger than the number of distinct subsets used for estimation, an efficient method is needed to generate a new random subset without checking explicitly if it contains repeated elements. We also need to ensure that the current subset has not been previously extracted. A lexicographic approach can be found that fulfills these requirements.

In addition to data analysis, robust techniques can be employed in the design of experiments. The model is (1) with $f_1(.)$ typically a regression model and $f_2(.)$ a departure, specified to some extent. In Box and Draper (1963) interest is in protecting second-order response surface models from biases from omitted third-order terms. Only the second-order model will be fitted to the data. The methods of optimum experimental design (Fedorov, 1972, Atkinson et al., 2007) require that a model, or models, be specified. In a series of papers Wiens and co-workers (Wiens and Zhou, 1997, Wiens, 1998, Fang and Wiens, 2000, Wiens, 2009) extend optimum design to partially specified situations. For example, Fang and Wiens (2000) bound the departure between the fitted and true models. They also allow for the possibility of heteroscedastic errors, bounding the magnitude of departure from homoscedasticity. With loss function the average mean squared error of prediction, I-optimal (Atkinson et al., 2007, §10.6) designs are obtained when the data are homoscedastic and the polynomial model is correct ($f_2(.) = 0$). When these conditions do not hold, the robust design replaces the support points of the optimum design with clusters of observations at nearby but distinct sites.

References

[1] Andrews, D.F., P.J. Bickel, F.R. Hampel, W.J. Tukey, and P.J. Huber (1972). *Robust Estimates of Location: Survey and Advances.* Princeton, NJ: Princeton University Press.

[2] Atkinson, A.C. (1985). *Plots, Transformations, and Regression.* Oxford: Oxford University Press.

[3] Atkinson, A.C., A.N. Donev, and R.D. Tobias (2007). *Optimum Experimental Designs, with SAS.* Oxford: Oxford University Press.

[4] Atkinson, A.C. and M. Riani (2000). *Robust Diagnostic Regression Analysis.* New York: SpringerŰVerlag.

[5] Atkinson, A.C., M. Riani, and A. Cerioli (2010). The forward search: theory and data analysis (with discussion). *J. Korean Statist. Soc. 39*, 117–134.

[6] Box, G.E.P. and N.R. Draper (1963). The choice of a second order rotatable design. *Biometrika 50*, 335–352.

[7] Cook, R.D. and S. Weisberg (1982). *Residuals and Influence in Regression.* London: Chapman and Hall.

[8] Fang, Z. and D.P. Wiens (2000). Integer-valued, minimax robust designs for estimation and extrapolation in heteroscedastic, approximately linear models. *J. Amer. Statist. Assoc. 95*, 807–818.

[9] Fedorov, V.V. (1972). *Theory of Optimal Experiments.* New York: Academic Press.

[10] Hampel, F., E.M. Ronchetti, P. Rousseeuw, and W.A. Stahel (1986). *Robust Statistics.* New York: Wiley.

[11] Hampel, F.R. (1975). Beyond location parameters: robust concepts and methods. *Bull. Int. Statist. Inst. 46*, 375–382.

[12] Huber, P.J. and E.M. Ronchetti (2009). *Robust Statistics*, 2nd Ed. New York: Wiley.

[13] Rousseeuw, P.J. (1984). Least median of squares regression. *J. Amer. Statist. Assoc. 79*, 871–880.

[14] Tallis, G.M. (1963). Elliptical and radial truncation in normal samples. *Ann. Math. Statist. 34*, 940–944.

[15] Wiens, D.P. (1998). Minimax robust designs and weights for approximately specified regression models with heteroscedastic errors. *J. Amer. Statist. Assoc. 93*, 1440–1450.

[16] Wiens, D.P. (2009). Robust discrimination designs. *J. R. Stat. Soc. Ser. B Stat. Methodol. 71*, 805–829.

[17] Wiens, D.P. and J. Zhou (1997). Robust designs based on the infinitesimal approach. *J. Amer. Statist. Assoc. 92*, 1503–1511.

Experimental Designs

Steven Gilmour

University of Southampton, UK

Abstract

The statistical design of experiments plays a vital role in experimentation in industry, medicine, agriculture, science and engineering. The need to obtain data which will give accurate and precise answers to research questions as economically as possible requires careful planning of experiments before they are run. Statistical methodology for designing experiments has a long history and classical methods continue to be successfully applied. However, as new technologies or business requirements lead to new types of experiment being conducted, research in the design of experiments continues and is currently experiencing an upsurge in activity.

The connection between design of experiments and linear statistical inference is old, with careful randomization providing a robust justification for linear models in many design structures and properties of estimators from linear theory providing the basis for optimal choices of designs. Many problems require computationally intensive optimizations and matrix methods provide the basis for this.

We encourage both invited and contributed papers in the design of experiments for the LinStat 2012 conference. There will be a stream of sessions on this theme, aiming to bring together international leaders in the field as well as early-career researchers to encourage the exchange of ideas and give participants a broad view of the subject. Papers related to any aspect of the design of experiments are encouraged, so that participants can get as broad a view as possible of the subject.

Some particular areas of research which are expected to feature are:

1. Block designs: the idea of blocking experimental units to improve the precision of treatment comparisons is widely used in practice. However, the extension to complex blocking structures continues to be an important area of research. Applications in genomics, proteomics and metabolomics have motivated recent work on optimal designs with very small block sizes, e.g. in experiments using microarrays. Related ideas for controlling variation, such as neighbour-balanced designs in agricultural experiments are increasingly popular and some of the same ideas can be used in experiments on social networks, in which neighbour relationships (or friendships) form less regular networks of experimental units.

2. Nonlinear design: the ideas generated from optimal design for linear models have been extended to cover various forms of nonlinear model. Although the basic theory is worked out, computational limitations mean that the application of nonlinear design is just starting. Experiments in pharmacokinetics and other areas of biological kinetics have motivated research on optimal design for nonlinear mixed models. However, as more is learned, the clearer it becomes that there are difficult problems to overcome and some of the current research will be presented at this conference. The advantages of pseudo-Bayesian design are well-recognised, but considerable research is still going on to find practicable ways of implementing these methods.
3. Factorial and response surface designs: increasing pressure on costs increases the importance of studying many factors in a single experiment and in industrial research the benefits of multifactorial experiments are widely recognised. Much current research focuses on designs which are useful when not all effects of interest can be studied. At one extreme, there has been an explosion of interest in supersaturated designs for screening very large numbers of factors. Research continues on how to analyse the data from such designs, while attention is turning to how to design follow-up experiments, or sequences of supersaturated designs. For more detailed study of processes, response surface methodology is widely used in practice. It has become increasingly recognised that many, perhaps most, industrial experiments have some factors whose levels are harder to set than others. This leads naturally to split-plot and other multi-stratum designs, and this is a topic of ongoing interest.
4. Experiments with discrete responses: most optimal design theory has been developed for linear models, or over-simplified generalized linear models. In most experiments, unit-to-unit variance must be allowed for and this requires the use of generalized linear mixed models and the design of experiments for such models has started to attract interest. Such data are often combined with complex factorial treatment designs and sometimes with multi-stratum structures and this is expected to become an area of active research in the near future.
5. Design for observational systems: Designing spatial sampling schemes and computer experiments are two types of application which have many similarities with design of experiments. They differ in that there is no concept of allocating and randomizing treatments to experimental units, but many of the same concepts of optimal design apply nonetheless. An explosion of research in such areas, and increasing realization that it is very similar to optimal design, will be reflected in the conference programme.

Submissions are encouraged in all of these areas of research, but also in others. The emphasis will be on methodological developments, but applied papers are also of interest.

Multivariate Analysis

Dietrich von Rosen

Swedish University of Agricultural Sciences, Uppsala, Sweden
Linköping University, Sweden

Abstract

Multivariate statistical analysis has a long history, but most of us probably do not have a clear picture of when it really started, what it was in the past and what it is today. In the present introduction we give a few personal reflections about some areas which are connected to the analysis based on the dispersion matrix or the multivariate normal distribution, omitting a discussion of many "multivariate areas" such as factor analysis, structural equations modelling, multivariate scaling, principal components analysis, multivariate calibration, cluster analysis, path analysis, canonical correlation analysis, non-parametric multivariate analysis, graphical models, multivariate distribution theory, and Bayesian multivariate analysis, to mention a few.

To begin with, it is of interest to cite a reply made by T.W. Anderson, concerning a discussion of the 2nd edition of his book on multivariate analysis "For a confident and thorough understanding, the mathematical theory is necessary" (Schervish, 1987). Although these words were written more than 25 years ago, they make even more sense today.

The multivariate normal (Gaussian) distribution was first applied about 200 years ago. Today one possesses substantial knowledge of the distribution: the characteristic function, moments, density, derivatives of the density, characterizations, and marginal distributions, among other topics. Closely connected to the distribution are the Wishart and the inverse Wishart distributions and different types of multivariate beta distributions. When extending the multivariate normal distribution the class of elliptical distributions is sometimes used since it includes the normal distribution. Other types of multivariate normal distributions which share many basic properties with the classical "vector-normal" distribution are the matrix normal, the bilinear normal and the multilinear normal distributions. To some extent they are all special cases of the multivariate normal distribution (classical vector-valued distribution), but in view of the possible applications, there are some advantages to be gained from studying all these different cases.

It is interesting to observe that it is still a relatively open question how to decide if data follows a multivariate normal distribution. The existing tests may be classified either as goodness-of-fit tests or as tests based on characterizations. However, most of the tests are connected with some asymptotic result and the size of the samples needed to make testing interesting is not

obvious. Too large samples will usually lead to the test statistics becoming asymptotically normally distributed, even if the original data is not normal, whereas small samples will mean that there is no power when testing for normality. Here one can envisage computer-intensive methods to becoming beneficial, since they can speed up convergency.
Concerning modelling there has been a tendency to create more and more complicated models: i.e. the parametrization has tended to become more advanced and the distributions have tended to deviate more from the normal distribution. An interesting class to study is skew-symmetric distributions, which include a skew-normal distribution. One natural field of application of skewed distributions is cases when there exist certain detection limits. However, one should not forget that a small change in the parametrization may have drastic inferential consequences, for example, when extending the MANOVA model

$$\mathbf{X} = \mathbf{BC} + \mathbf{E}, \qquad \mathbf{E} \sim N_{p,n}(\mathbf{0}, \mathbf{\Sigma}, \mathbf{I}),$$

where $\mathbf{B}$ and $\mathbf{\Sigma}$ are unknown parameters, to the Growth Curve model

$$\mathbf{X} = \mathbf{ABC} + \mathbf{E}, \qquad \mathbf{E} \sim N_{p,n}(\mathbf{0}, \mathbf{\Sigma}, \mathbf{I}),$$

where $\mathbf{B}$ and $\mathbf{\Sigma}$ are unknown parameters, as in MANOVA, and $\mathbf{A}$ and $\mathbf{C}$ are known design matrices. With the Growth Curve model we actually move from the exponential family to the curved exponential family with significant consequences, e.g. for the Growth Curve model the MLEs of $\mathbf{B}$ are non-linear and the estimators are not independent of the unique MLE of $\mathbf{\Sigma}$. A further generalization is a spatial-temporal setting

$$\mathbf{X} = \mathbf{ABC} + \mathbf{E}, \qquad \mathbf{E} \sim N_{p,nk}(\mathbf{0}, \mathbf{\Sigma}, \mathbf{I} \otimes \mathbf{\Psi}),$$

where $\mathbf{\Sigma}$ models the dependency over time and $\mathbf{\Psi}$ is connected to spatial dependency. In summary, in MANOVA most things work as in the corresponding univariate case, i.e. easily interpretable mean and dispersion estimators are obtained, while in the Growth Curve model explicit estimators are also obtained, but the mean estimators are non-linear and more difficult to interpret. For the spatial-temporal model, no explicit MLEs are available but one has algorithms which deliver unique estimators. Concerning the future we will probably see more articles where for $\mathbf{X} \in N(\mu, \mathbf{\Sigma})$ there are models which state that $\mu \in \mathcal{C}(\mathbf{C}_1) \otimes \mathcal{C}(\mathbf{C}_2) \otimes \cdots \otimes \mathcal{C}(\mathbf{C}_m)$, i.e. a tensor product of $\mathcal{C}(\mathbf{C}_i)$, where $\mathcal{C}(\mathbf{C}_i)$ stands for the space generated by the columns of $\mathbf{C}_i$, and $\mathbf{\Sigma} = \mathbf{\Sigma}_1 \otimes \mathbf{\Sigma}_2 \otimes \cdots \otimes \mathbf{\Sigma}_m$. Another type of generalization which has been taking place for decades is the assumption of different types of dispersion structures, e.g. structures connected to factor analysis, structures connected to spatial relationships, and structures connected to time series, structures connected to random effects models, structures connected to graphical normal models, structures connected to the complex normal and quaternion normal distributions.

High-dimensional statistical analysis is, with today's huge amount of available data, of the utmost interest. Indeed various different high-dimensional approaches are natural extensions of classical multivariate methods. A general characterization of high-dimensional analysis is that in the multivariate setting there are more dependent variables than independent observations. It is driven by theoretical challenges as well as numerous applications such as applications within signal processing, finance, bioinformatics, environmetrics, chemometrics, etc. The area comprises, but is not limited to, random matrices, Gaussian and Wishart matrices with sizes which turn to infinity, free probability, the R-transform, free convolution, analysis of large data sets, various types of p, n-asymptotics including the Kolmogorov asymptotic approach, functional data analysis, smoothing methods (splines); regularization methods (Ridge regression, partial least squares (PLS), principal components regression (PCR), variable selection, blocking); and estimation and testing with more variables than observations.

If one considers the asymptotics with p indicating the number of dependent variables and n the number of independent observations, there are a number of different cases: $p/n \to c$, where c is a known constant, and both p and n go to infinity without any relationship between p and n. The latter case, however, has to be treated very carefully in order to obtain interpretable results. For example, one has to distinguish if first p and then n goes to infinity or vice versa, or $\min(p, n) \to \infty$. When studying proofs of different situations in the literature, it is not obvious which situation is considered and many results can only be viewed as approximations and not as strict asymptotic results, at least on the basis of the presented proofs.

One of the main problems in multivariate statistical analysis as well as high-dimensional analysis occurs when the inverse dispersion matrix, $\mathbf{\Sigma}^{-1}$, has to be estimated. If $\mathbf{\Sigma}$ is known, it often follows from univariate analysis that the statistic of interest is a function of $\mathbf{\Sigma}^{-1}$. Then one tries to replace $\mathbf{\Sigma}^{-1}$ with an estimator. If $\mathbf{S}$ is an estimator of $\mathbf{\Sigma}$, the problem is that $\mathbf{S}^{-1}$ may not exist or may perform poorly due to multicollinearity, for example. If $\mathbf{S}$ is singular, then $\mathbf{S}^{+}$ has been used. Moreover, "ridge type" estimators of the form $(\mathbf{S} + \lambda \mathbf{I})^{-1}$ are in use (Tikhonov regularization). Sometimes a shrinking takes place through a reduction of the eigenspace by removing the part which corresponds to small eigenvalues. A different idea is to use the Cayley-Hamilton theorem and utilize the fact that

$$\mathbf{\Sigma}^{-1} = \sum_{i=1}^{p} c_i \mathbf{\Sigma}^{i-1},$$

where $\mathbf{\Sigma}$ is of the size $p \times p$ and since $\mathbf{\Sigma}$ is unknown the constants c_i are also unknown. Then an approximation of $\mathbf{\Sigma}^{-1}$ is given by

$$\mathbf{\Sigma}^{-1} \approx \sum_{i=1}^{a} c_i \mathbf{\Sigma}^{i-1}, \qquad a \leq p,$$

and an estimator is found via $\widehat{\mathbf{\Sigma}}^{-1} \approx \sum_{i=1}^{a} \widehat{c}_i \mathbf{S}^{i-1}$. When determining c_i a Krylov space method, partial least squares (PLS), is used.
Needless to say, there are many interesting research questions to work on. Computers are nowadays important tools but much more important are ideas which can challenge some fundamental problems. For example in high-dimensional analysis we have parameter spaces which are infinitely large and it is really unclear how to handle and interpret this situation. Hopefully the discussions in this conference will deal with some of the challenging multivariate statistical problems.

References

[1] Schervish, M.J. (1987). A review of multivariate analysis. With discussion and a reply by the author. *Statist. Sci.* **2**, 396–433.

Mixed Models

Júlia Volaufová

LSUHSC School of Public Health, New Orleans, USA

Abstract

Mixed models, simply put, are models of a response that involve fixed and random effects (see, e.g., [1]). Here we give a very superficial and brief coverage of the wide variety of models this term encompasses.
Historically, the most widely-investigated mixed model is the *linear* mixed model. For an n-dimensional response vector $\boldsymbol{Y}$, the model can be expressed as

$$\boldsymbol{Y} = X\beta + Z\boldsymbol{\gamma} + \boldsymbol{\epsilon}, \tag{1}$$

where X and Z are fixed and known matrices of covariates with β a fixed vector parameter and $\boldsymbol{\gamma}$ a vector of random effects. The often-invoked distributional assumptions are $\boldsymbol{\gamma} \sim N_l(0, G)$ and $\boldsymbol{\epsilon} \sim N_n(0, R)$. The matrices G (nnd) and R (pd) are modeled as members of chosen classes (e.g., compound symmetry, AR(1), unstructured), which involve further unknown parameters. In special cases when the matrices G and R depend linearly on a set of unknown scalars, the covariance matrix of the response can be expressed as $\mathrm{Cov}(\boldsymbol{Y}) = \sum_{i=1}^{p} \vartheta_i V_i$ where the parameters ϑ_i are interpreted as *variance-covariance components*. $\boldsymbol{\gamma}$ and $\boldsymbol{\epsilon}$ are assumed to be mutually independent, which implies that $\mathrm{Cov}(\boldsymbol{Y}) = ZGZ' + R$.
This class of models covers a broad range of situations. Here is a partial list.

- In *repeated measures* models (see e.g., [17]), also called *longitudinal* models (see, e.g., [9]), multiple observations are carried out, say over time, on each individual sampling unit.
- In *cluster randomized settings* (see, e.g., [10]), dependencies between observations on sampling units are introduced due to clustering in the randomization process.
- In *hierarchical* or *multilevel* settings, a subset of parameters on a given level is considered to be a random vector whose distribution depends on an additional set of unknown parameters.
- In some situations it is possible to partition the response vector into independent subvectors, as in longitudinal models, but in many cases such partitioning is not straightforward, e.g., in some geodetic or geophysical applications (see e.g., [8] or [2]) when combining experiments with different precisions, each relating to the same mean parameter. In these models it is not obvious and it is not even necessary to identify the latent random effects - the model for the response vector $\boldsymbol{Y}$ is parametrized by

(unknown) fixed vector parameters of the mean and variance-covariance components.

The class of linear mixed models can be viewed within the broader context of *nonlinear* mixed models. There, the response variable Y can be modeled in general as

$$Y = f(\boldsymbol{x}, \boldsymbol{z}, \beta, \boldsymbol{\gamma}) + \epsilon. \tag{2}$$

The function $f(.)$ is a nonlinear function of fixed (β) and random ($\boldsymbol{\gamma}$) (vector) parameters as well as vectors of covariates ($\boldsymbol{x}$ and $\boldsymbol{z}$). Mostly it is assumed that $f(.)$ is differentiable with respect to β and $\boldsymbol{\gamma}$. The distributional assumptions regarding $\boldsymbol{\gamma}$ and ϵ may be the same as in the linear case.
An example of such a model (2) is a *random coefficients* model (see, e.g., [23]) that can be set up in two stages. For stage 1, the model takes the form

$$Y_{ij} = f(t_{ij}, \beta_i) + \varepsilon_{ij},\ i = 1, \ldots, n,\ j = 1, \ldots, T_i, \tag{3}$$

where Y_{ij} is the response for subject i at time j, $f(.)$ is a nonlinear function of the p-vector of subject-specific vector parameters β_i and time (t_{ij}), and ε_{ij} is the error term, which we assume follows a normal distribution with mean zero and variance σ^2. The second stage is at the population level. At this stage the subject-specific parameters are defined by the model:

$$\beta_i = A_i\beta + B_i\boldsymbol{\gamma}_i, \qquad i = 1, \ldots, n. \tag{4}$$

In this model, β is a vector of fixed population parameters and $\boldsymbol{\gamma}_i$ is a vector of random effects for subject i. In most cases the matrix A_i takes the form $A_i = I_p \otimes a_i'$ (see, e.g., [22]), where the vector a_i is the vector of covariates. The matrix B_i is used to determine which elements of β_i have random components and which are fixed. A well-known example of a random coefficient model is a *growth curve model*, a special case of which is when, among other assumptions, the dimension of each subject specific response is the same. It can be expressed in terms of a *multivariate model* for a response vector $\boldsymbol{Y}$, which in general may have expectation $E(\boldsymbol{Y}) = \sum_{i=1}^{p}(C_i \otimes D_i)\boldsymbol{\beta}_i$ and covariance matrix $\mathrm{Cov}(\boldsymbol{Y}) = \Sigma_1 \otimes \Sigma_2$ with the matrices Σ_1 and Σ_2 and the vector β unknown.
Using the *generalized* mixed linear model, we model the transformed mean of Y via a *link* function as a linear function of covariates (see, e.g., [14]). Typically, the conditional distribution of the response belongs to the exponential family, and often there is a functional relationship between the parameters of the mean and the variance-covariance components. The conditional mean of the ith observation, μ_i, is linked via a function, say $g(\mu_i)$, to the covariates and random effects in terms of additive effects as

$$g(\mu_i) = \boldsymbol{x}_i'\beta + \boldsymbol{z}_i'\boldsymbol{\gamma}, \tag{5}$$

where the meaning of β and $\boldsymbol{\gamma}$ is as above. We note that the nonlinear random coefficients model can be perceived as a special case of the generalized linear mixed model.

Although these models have been notoriously studied for many decades, there is a variety of questions still to be addressed. The main aim of inference is estimation and hypotheses testing. Maximum likelihood or quasi-likelihood methods result in point estimates with good large sample properties under generally mild conditions. Point and interval estimation of variance-covariance components has kept statisticians busy for decades (see, e.g., [3,11,12,15], etc.). Ultimately in almost all settings one is interested in testing (linear) hypotheses about the parameter β and/or about variance-covariance components. Usually we see $H_{0_1} : H'\beta = h$ or $H_{0_2} : h'\boldsymbol{\vartheta} = 0$, or simultaneously both. Except for a few special models there is no exact test available for H_{0_1}; a variety of approximate tests has been studied for quite a time (see, e.g., [4], [6], [7], [19], [21], [13], [20], and many others). The hypothesis H_{0_2} also has been extensively studied; however even in models with only two variance-covariance components the question of finding a test with optimal properties (in some sense) in general is still open.

In linear mixed models, the choices of the structures of G and R may be consequential, but often these choices are made arbitrarily and subjectively. Various information criteria have been developed, recommended, and modified for the purpose of informing these choices. Effects of such data-driven choices on inferential procedures for fixed effects are just beginning to be investigated and are related to the broad area of model building (see, e.g., [18]).

Here we invite contributions that address pertinent questions and relate to any aspects of this broad class of mixed models.

References

[1] Demidenko, E. (2004). *Mixed Models. Theory and Applications.* John Wiley & Sons, Inc., New York.

[2] Fišerová, E., L. Kubáček, and P. Kunderová (2007). *Linear Statistical Models. Regularity and Singularities.* Academia, Praha.

[3] Hartley, H.O. and J.N.K. Rao (1967). Maximum-Likelihood Estimation for the Mixed Analysis of Variance Model. *Biometrika 54*, 93–108.

[4] Harville, D.A. and D.R. Jeske (1992). Mean Squared Error of Estimation or Prediction Under a General Linear Model. *J. Amer. Statist. Assoc. 87*, 724–731.

[5] Harville, D.A. (1997). *Matrix Algebra from a Statistician's Perspective.* Springer.

[6] Kenward, M.G. and J.H. Roger (1997). Small Sample Inference for Fixed Effects From Restricted Maximum Likelihood. *Biometrics 53*, 983–997.

[7] Khuri, A.I., T. Mathew, and B.K. Sinha (1998). *Statistical Tests for Mixed Linear Models.* John Wiley & Sons, Inc., New York.

[8] Kubáček, L., L Kubáčková, and J. Volaufová (1995). *Statistical Models with Linear Structures.* Veda, Bratislava.

[9] Laird, N.M. and J.H. Ware (1982). Random-Effects Models for Longitudinal Data. *Biometrics 38*, 963–974.

[10] Laird, N.(2004). Analysis of Longitudinal and Cluster-Correlated Data. *NSF-CBMS Regional Conference Series in Probability and Statistics 8*, Published by IMS, i-ii+1–155.

[11] LaMotte, L.R. (1973). Quadratic estimation of variance components. *Biometrics 29*, 311–330.

[12] LaMotte, L.R. (1973) On nonnegative quadratic unbiased estimation of variance components. *J. Amer. Statist. Assoc. 68*, 728–730.

[13] Livacic-Rojas, P., G. Vallejo, and P. Fernandez (2010). Analysis of Type I Error Rates of Univariate and Multivariate Procedures in Repeated Measures Designs. *Comm. Statist. Simulation Comput. 39*, 624–640.

[14] McCulloch, Ch.E., S.R. Searle, and J.M. Neuhaus (2008). *Generalized, Linear, and Mixed Models.* 2nd Ed. John Wiley & Sons, Inc., New York.

[15] Mathew, T., T. Nahtman, D. von Rosen, and B.K. Sinha (2009). Non-negative Estimation of Variance Components in Heteroscedastic One-way Random-effects ANOVA Models. *Statistics 44*, 557–569.

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

[17] Reinsel, G. (1982). Multivariate Repeated-Measurement or Growth Curve Models with Multivariate Random-Effects Structure. *J. Amer. Stat. Assoc. 77*, 190–195.

[18] Vaida, F. and S. Blanchard (2005). Conditional Akaike information for mixed-effects models. *Biometrika 92*, 351–370.

[19] Volaufová, J. and L.R. LaMotte (2008). Comparison of approximate tests of fixed effects in linear repeated measures design models with covariates. *Tatra Mt. Math. Publ. 39*, 17Ũ-25.

[20] Volaufová, J. and L.R. LaMotte (2012). A Simulation Comparison of Approximate Tests for Fixed Effects in Random Coefficients Growth Curve Models. *Comm. Statist. Simulation Comput.* To appear.

[21] Volaufová, J. (2009). Heteroscedastic ANOVA: old p values, new views. *Statist. Papers 50*, 943–962.

[22] Vonesh, E.F. and R.L. Carter (1987). Efficient Inference for Random-Coefficient Growth Curve Models with Unbalanced Data. *Biometrics 43*, 617–628.

[23] Vonesh, E.F. and V.M. Chinchilli (1997). *Linear and Nonlinear Models for the Analysis of Repeated Measurements.* Marcel Dekker, Inc.

Part IV

Invited Speakers

Optimal design of experiments with very low average replication

Rosemary A. Bailey

Queen Mary, University of London, UK

Abstract

Trials of new crop varieties usually have very low average replication. Thus one possibility is to have a single plot for each new variety and several plots for a control variety, with the latter well spread out over the field. A more recent proposal is to ignore the control, and instead have two plots for each of a small proportion of the new varieties.

Variation in the field may be accounted for by a polynomial trend, by spatial correlation, or by blocking. However, if the experiment has a second phase, such as making bread from flour milled from the grain produced in the first phase, then that second phase usually has blocks.

The optimality criterion used is usually the A criterion: the average variance of the pairwise differences between the new varieties.

I shall compare designs under the A criterion when the average replication is much less than two.

Geometric mean of matrices

Rajendra Bhatia

Indian Statistical Institute, New Delhi, India

Abstract

Positive definite matrices are important in diverse areas like statistics, image processing, quantum information, electrical engineering, elasticity, machine learning etc. An appropriate notion of averaging a family of such matrices has been developed in recent years, and has brought together diverse areas like differential geometry, matrix analysis, numerical analysis and approximation theory. This talk will provide a survey of some of the key ideas.

Nonparametric regression for sojourn time distributions in a multistate model

Somnath Datta and Dogu Lorenz

University of Louisville, USA

Abstract

Multistate models are generalizations of traditional survival data where an individual undergoes different types of events corresponding to transitions to various states of a system. We consider multistate event data that are right censored. Under this setup, inferring on the state waiting (or sojourn) time distribution corresponding to a give transient state j is problematic since neither the entry nor the exit times are fully observed. In this talk, we introduce novel procedures to test the effect of a categorical covariate on the sojourn time distribution. In the later part of the talk, we introduce an Aalen type linear hazard model for the state waiting time distribution that can incorporate both discrete and continuous covariates. The methods are illustrated using a number of real data applications.

Tolerance intervals in general mixed effects models using small sample asymptotics

Thomas Mathew and Gaurav Sharma

University of Maryland Baltimore County, USA

Abstract

The computation of tolerance intervals in mixed and random effects models has not been satisfactorily addressed in a general setting when the data are unbalanced and/or when covariates are present. In the talk, satisfactory one-sided and two-sided tolerance intervals in such a general scenario will be derived, by applying small sample asymptotic procedures. In the case of one-sided tolerance limits, the problem reduces to the interval estimation of a percentile, and accurate confidence limits are derived using small sample asymptotics. In the case of a two-sided tolerance interval, the problem does not reduce to an interval estimation problem; however, it is possible to derive an approximate margin of error statistic that is an upper confidence limit for a linear combination of the variance components. For the latter problem, small sample asymptotic procedures can once again be used in order to arrive at an accurate upper confidence limit. In the talk, balanced and unbalanced data situations will be treated separately, and computational issues will be briefly addressed. Extensive numerical results show that the tolerance intervals derived based on small sample asymptotics exhibit satisfactory performance regardless of the sample size. The results will be illustrated using examples.

Smoothing discrete distributions

Paulo E. Oliveira

University of Coimbra, Portugal

Abstract

We will discuss estimation of probability distributions on discrete, finite or infinite, space using nonparametric methods. This model includes of course, categorical distributions. Although the smoothing implied in nonparametric methods may seem, at first glance, unnatural, smoothing does improve upon the naïve frequency estimator. Discretizations of the kernel estimator and the correspondent characterizations of asymptotic properties are discussed.
When dealing with categorical distributions one is often faced with relatively few observations, when compared to the support size. This leads to considering error criteria better adapted to this sparse estimation problem. Asymptotics with respect to these sparse criteria is discussed. These results do not really fall into the general approach to nonparametric estimation, as they imply that the base space should be updated as the sample size grows. Other error criteria, such as relative errors, are commonly considered in parametric problems. We will adapt relative error criteria to our nonparametric estimation problem. The estimator found can be explicitly written but their asymptotics is harder to describe, in some cases only doable indirectly. However, their finite sample performance is, depending on the properties of the true probability distribution, good. We will also discuss the integration into the estimator of partial known information about the true probability.

Keywords

Discrete distributions, Local polynomial estimator, Relative errors, Asymptotics, Sparse observations.

Partial orders on matrices and the column space decompositions

K. Manjunatha Prasad

Manipal University, India

Abstract

In literature, we have several partial orders on subclasses of rectangular matrices of same size and some which are dominated by known “Minus Partial Order". Star partial order ([3]) on rectangular matrices of size, Sharp order ([6]) on class of square matrices of same size and of index one, and the Core partial order ([2]) are such partial orders dominated by minus partial order to name a few. It is well known that the $m \times n$ matrices B and $A-B$ decomposes the given matrix A under minus partial order (i.e., $B, A-B \leq^- A$) is equivalent to say that the column spaces of B and $A-B$ decomposes the column space of the matrix A (i.e., $\mathcal{C}(B) \oplus \mathcal{C}(A-B) = \mathcal{C}(A)$). The same is true for the row spaces. In fact, there is one to one correspondence between matrix decompositions with reference to minus partial order, column space decompositions and row space decompositions. The characterization of the partial orders such as star partial order and sharp order involve both column space and row space of given matrices. In fact, matrix decomposition $A = B + C$ with reference to star partial order corresponds to decomposition of column space and row space of A orthogonally and similarly other matrix partial orders are characterized by the typical characteristic decompositions of the column space and row spaces. Even while studying the shorted matrices (see [1], [5] and [10]) involves both row space and column spaces of given matrices. Now in the light one to one correspondence between column space decompositions and row space decompositions, we characterize the partial orders with reference to column space decomposition alone. Also, it results in having a new definition of shorted matrix with reference to various partial orders i.e., only with reference to the decomposition of column space decomposition.

Keywords

Matrix partial order, Minus partial order, Star partial order, Sharp partial order, Shorted matrix.

References

[1] Anderson, W.N. (1971). Shorted Operators. *SIAM J. Appl. Math. 20*, 520–525.

[2] Baksalary, O.M. and G. Trenkler (2010). Core inverse of matrices, *Linear Multilinear Algebra 58(6)*, 681–697.

[3] Drazin, M.P. (1978). Natural structures on semigroups with involution. *Bull. Amer. Math. Soc. 84*, 139–141.

[4] Hartwig, R.E. (1980). How to order regular elements?. *SIAM J. Appl.Math. 25*, 1–13.

[5] Mitra, S.K. (1986). The minus partial order and shorted matrix. *Linear Algebra Appl. 81*, 207–236.

[6] Mitra, S.K. (1987). On group inverses and the sharp order. *Linear Algebra Appl. 92*, 17–37.

[7] Mitra, S.K. (1989). Block independance in generalized inverse: a coordinate free look. In: Dodge, Y. (Ed.) *Data Analysis and Inference North-Holland* (pp. 429–443).

[8] Mitra, S.K. (1994). Separation theorems. *Linear Algebra Appl. 208/209*, 238–256.

[9] Mitra, S.K., P. Bhimasankaram, and S.B. Malik (2010). *Matrix Partial Orders, Shorted Operators and Applications*. World Scientific, New Jersey.

[10] Mitra, S.K. and K.M. Prasad (1996). The nonunique shorted matrix. *Linear Algebra Appl. 237/238*, 41–70.

[11] Mitra, S.K. and S. Puntanen (1990-91). The shorted operator statisically interpreted. *Calcutta Statist. Assoc. Bull. 40*, 157–160.

[12] Mitra, S.K., S. Puntanen, and G.P.H. Styan (1995). Shorted matrices and their applications in linear statistical models: a review. In: E.-M. Tiit, T. Kollo, H. Niemi (Eds.), *New Trends in Probability and Statistics, Proceedings of the 5th Tartu Conference on Multivariate Statistics* (pp. 289–311). TEV Ltd., Vilnius, Lithuania.

[13] Mitra, S.K. and M.L. Puri (1979). Shorted operators and generalized inverses of matrices. *Linear Algebra Appl. 25*, 45–56.

[14] Mitra, S.K. and M.L. Puri (1982). Shorted matrix: an extended concept and some applications. *Linear Algebra Appl. 42*, 57–79.

[15] Nambooripad, K.S.S. (1980). The natural partial order on a regular semigroup. *Proc. Edinb. Math. Soc. 23*, 249–260.

[16] Rao, C.R. (1971). Unified theory of linear estimation. *Sankhya, Ser A. 33*, 371–394.

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

[18] Werner, H.J. (1986). Generalized inversion and weak bi-complementarity. *Linear Multilinear Algebra 19*, 357–372.

Adjacency preserving maps

Peter Šemrl

University of Ljubljana, Slovenia

Abstract

Hua's fundamental theorems of geometry of matrices characterize bijective maps on various spaces of matrices preserving adjacency in both directions. We will discuss some recent improvements of these results.

Investigation of Bayesian Mixtures-of-Experts models to predict semiconductor lifetime

Olivia Bluder

Alpen-Adria University Klagenfurt and KAI - Kompetenzzentrum Automobil- und Industrieelektronik GmbH, Villach, Austria

Abstract

Investigating the reliability of a semiconductor device is time and cost consuming, but essential for industry and customers. To save resources, models that predict the lifetime and the valid parameter range dependent on the stress conditions are needed.

The given semiconductor lifetime data show a mixture of two log-normal distributions [1], where the mixture weights of the two components depend on the applied peak temperature. Hence, a Bayesian Mixtures-of-Experts (ME) approach is used [3]. For the component means linear models as well as physical acceleration models [2] are investigated. Under the assumption of informed normal priors for the model parameters and slightly data dependent hierarchical inverse Gamma priors for the variances, the mixture based on two Coffin-Manson models shows the best fit and the best prediction quality. Applying the model to lifetime data from other semiconductor technologies shows that the combined Bayesian ME and Coffin-Manson approach is valid for other designs as well. With the given model parameter ranges for one semiconductor design based on a minimum number of stress tests can be predicted. Hence, resources, especially testing time, can be saved.

Keywords

Bayesian Mixtures-of-Experts models, Semiconductor lifetime prediction, Linear models, Physical acceleration models.

References

[1] Bluder, O., M. Glavanovics, and J. Pilz (2011). Applying Bayesian Mixtures-of-Experts models to statistical description of smart power semiconductor reliability. *Microelectronics Reliab. 51*, 1464–1468.

[2] Escobar, L.A. and W.Q. Meeker (2006). A review of accelerated test models. *Statist. Sci. 21(4)*, 552–577.

[3] Frühwirth-Schnatter, S. (2006). *Finite mixture and Markov switching models* (1st ed.) Springer, Berlin.

Influential observations in the extended Growth Curve model with cross-over designs

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

[1] Stockholm University, Sweden
[2] Swedish University of Agricultural Sciences, Uppsala, Sweden
[3] Linköping University, Sweden

Abstract

Growth Curve model (GCM) and extended GCM are useful tools to model repeated measurements in cross-over designs. [2] and [3] assessed influence of observations on estimating the GCM with unstructured covariance. This work is to propose quantities to detect influential measurements in the extended GCM. It is known that various residuals in the extended GCM can be defined by projecting data matrix onto four orthogonal spaces, see [1]. The relations between the influence quantities and the residuals are surveyed.

Keywords

Extended Growth Curve model, Influence analysis, Repeated measurement design, Statistical diagnostics.

References

[1] Hamid, J.S. and D. von Rosen (2006). Residuals in the Extended Growth Curve model. *Scand. J. Statist. 33*, 121–138.
[2] Pan, J.X. and K.T. Fang (2002). *Growth Curve Models and Statistical Diagnostics.* Springer Verlag.
[3] von Rosen, D. (1995). Influential observations in multivariate linear models. *Scand. J. Statist. 22*, 207–222.

Low-rank approximations and weighted low-rank approximations

Paulo C. Rodrigues

Centro de Matemática e Aplicações, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Portugal

Abstract

Principal component analysis (PCA) is one of the most widely used multivariate techniques. It is usually applied to two-way matrices with individuals in the rows and variables in the columns, and converts the possibly correlated variables into a set of orthogonal variables, the principal components. Several algorithms have been proposed to obtain the least squares estimates for the component scores and for the loadings, being the most used the eigenvalue decomposition of the covariance (or correlation) matrix of the data or the singular value decomposition of the two-way data matrix. In this paper we will be mostly interested in the weighted version of this low-rank approximation. This allows us to give weights to the variables and/or the individuals according to the outcome of a preliminary analysis of the two-way data, e.g., in the case of repeated measurements the weights can be given by the inverse of their error variances. The use of the weighted PCA also increases the robustness when compared with the standard PCA. Applications to genetic and financial data will be presented.

Keywords

Principal component analysis, Additive main effects and multiplicative interaction model, Plant genetics, Public debt.

Part V

Contributed Talks

On the choice of a prior distribution for Bayesian D-optimal designs for the logistic regression model

Haftom Abebe, Frans Tan, Gerard Van Breukelen, Jan Serroyen, and Martijn Berger

Maastricht University, The Netherlands

Abstract

A common way to design a binary response experiment is to design the experiment to be most efficient for a best guess of the parameter values on which the optimal design depends. A design which is optimal for a best guess, however, may not be efficient for other parameter values. The Bayesian optimal design approach is a useful tool to take into account uncertainty of the parameter values. Bayesian D-optimal designs for a logistic regression model with two parameters are investigated. Such designs depend on the choice of a prior distribution. Using numerical search and sampling from normal and uniform priors we show that if we do not have much information about the value of the parameters, a prior distribution with relatively large variance will lead to a Bayesian design which remains highly efficient under other prior distributions. We also compare uniform and normal priors and find that both distributions are approximately equally efficient. Finally, we study the efficiencies of designs with equidistant equally weighted design points against the Bayesian D-optimal designs and find that 4 and 5 equidistant equally weighted design points are highly efficient.

Keywords

Bayesian D-optimal designs, Logistic regression model, Maximin Bayesian D-optimal design, Locally D-optimal designs, Relative efficiency.

References

[1] Atkinson, A.C., A.N. Donev, and R.D. Tobias (2007). *Optimum Experimental Designs, with SAS*. Clarendon, Oxford.

[2] Berger, M.P.F. and W.K. Wong (2009). *An Introduction to Optimal Designs for Social and Biomedical Research*. John Wiley and Sons Ltd.

[3] Braess, D. and H. Dette (2007). On the number of support points of maximin and Bayesian optimal designs. *Ann. Statist. 35*, 772–792.

[4] Chaloner, K. and K. Larntz (1989). Optimal Bayesian designs applied to logistic experiments. *J. Statist. Plann. Inference 18*, 191–208.

[6] Firth, D. and J.P. Hinde (1997). On Bayesian D-optimum design criteria and the equivalence theorem in non-linear models. *J. Roy. Stat. Soc. Ser. B 59*, 793–797.

Model selection in log-linear models by using information criteria

Nihan Acar[1], **Eylem D. Howe**[1], **and Andrew Howe**[2]

[1] Mimar Sinan Fine Arts University, Istanbul, Turkey
[2] Transatlantic Petrolium, Istanbul, Turkey

Abstract

Log-linear models help to reveal association patterns among categorical variables that are widely encountered in sectors such as ecology, medicine and banking. These models are generally used in the analysis of contingency tables. In log-linear models deviance and chi-square statistics are mostly used to select the best model which fits data. Because the chi-square statistic is affected by sample size, information criteria such as AIC-type criteria are used lately in many areas.
In this study we purpose to apply and measure the efficiency of information criteria in log-linear models.

Keywords

Log-linear models, Contingency tables, Information criteria.

References

[1] Dobson, J.A. (2002). *An Introduction to Generalized Linear Models.* Chapman and Hall.
[2] Bedrick, E.J. and K.W. Crandall (2010). Model selection criteria for log-linear models. *Aust. N. Z. J. Stat. 52(4)*, 439–449.
[3] Bozdogan, H. (2000). Akaike's information criterion and recent developments in information complexity. *J. Math. Psych. 44*, 62–91.
[4] Rao, P.R. and H. Toutenburg (1999). *Linear Models: Least Squares and Alternatives.* Springer-Verlag.

Absolute Penalty and Shrinkage Estimation in Weibull censored regression model

S. Ejaz Ahmed

Brock University, St. Catharines, Canada

Abstract

In this talk we address the problem of estimating a vector of regression parameters in the Weibull censored regression model. Our main objective is to provide natural adaptive estimators that significantly improve upon the classical procedures in the situation where some of the predictors may or may not be associated with the response. In the context of two competing Weibull censored regression models (full model and candidate sub-model), we consider an adaptive shrinkage estimation strategy that shrinks the full model maximum likelihood estimate in the direction of the sub-model maximum likelihood estimate. The shrinkage estimators are shown to have higher efficiency than the classical estimators for a wide class of models. Further, we consider a LASSO type estimation strategy and compare the relative performance with the shrinkage estimators. Monte Carlo simulations reveal that when the true model is close to the candidate sub-model, the shrinkage strategy performs better than the LASSO strategy when, and only when, there are many inactive predictors in the model. Shrinkage and LASSO strategies are applied to a real data set from Veteran's administration (VA) lung cancer study to illustrate the usefulness of the procedures in practice.

Bootstrap confidence regions for multinomial probabilities based on penalized power-divergence test statistics

Aylin Alin[1] and Ayanendranath Basu[2]

[1] Dokuz Eylül University, Izmir, Turkey
[2] Indian Statistical Institute, India

Abstract

In general confidence regions for multinomial probabilities are constructed based on the Pearson χ^2 statistic. [1] constructed the bootstrap and asymptotic confidence regions for multinomial parameters based on power-divergence test statistics . In this study, we consider confidence regions for multinomial probabilities based on ordinary and penalized power-divergence test statistics. We built bootstrap and asymptotic confidence regions. We use two types of bootstrap confidence regions. The first type is called percentile interval which is the mostly used version of bootstrap intervals. The second type is Bca interval proposed by [2] as the improved version of percentile interval. We only consider small sample sizes where asymptotic properties fail and the alternative methods are needed mostly. Performances are compared based on average coverage probabilities calculated by designed simulation studies.

Keywords

Bca interval, Bootstrap, Power-divergence test statistics, Penalization.

References

[1] Morales, D., L. Pardo, and L. Santamaría (2004). Bootstrap confidence regions in multinomial sampling. *Appl. Math. Comput. 155*, 295–315.
[2] Efron, B. (1987). Better bootstrap confidence intervals. *J. Amer. Statist. Assoc. 82*, 171–185.

Building stones for inference on variance components

Barbora Arendacká

Physikalisch-Technische Bundesanstalt, Berlin, Germany

Abstract

In his paper, Burch [1] suggested how to make inference on variance components in linear mixed models provided a certain decomposition of the covariance matrix exists and he showed how these ideas apply in some cases of two-way random effects models without interactions. However, he did not show how to derive the requested building stones - independent quadratic forms - in general. We will point out that his approach can be viewed as a generalization of the ANOVA decomposition of the total sum of squares. Then the requirement of independence leads to a decomposition of the $(n-p)$-dimensional space into orthogonal invariant subspaces and in a case most favourable for inference, this immediately suggests an algorithm for derivation of the requested quantities. The presented approach also allows for characterizing designs in which the favourable procedure is applicable as we will illustrate for the case of two-way random effects models.

Keywords

Independent quadratic forms, Variance components, ANOVA, Invariant subspaces.

References

[1] Burch, B.D. (2007). Generalized confidence intervals for proportions of total variance in mixed linear models. *J. Statist. Plann. Inference 137*, 2394–2404.

A novel approach for estimation of seemingly unrelated linear regressions with high order autoregressive disturbances

Baris Asikgil

Mimar Sinan Fine Arts University, Istanbul, Turkey

Abstract

The problem of estimating a system of linear regression equations in which the disturbances are contemporaneously correlated across equations has been investigated in the past years. One of the major problems encountered in the estimation of such system of linear regression equations is the possible existence of serial correlation of the disturbances. [3] modified the original "seemingly unrelated linear regressions" estimation technique known as Zellner's two stage Aitken estimator for the first order autoregressive disturbances in each equation. Also, several alternative estimators given by [2] are compared for small samples.
In this paper, seemingly unrelated linear regressions with high order autoregressive disturbances are considered. A novel approach which includes a polynomial tapering function given by [1] is proposed for high order autoregressive disturbances in order to obtain more efficient parameter estimates. Monte Carlo simulation study is applied to compare this approach with the other estimators for small-sample efficiency.

Keywords

Linear regression, Contemporaneously correlation, Autoregressive disturbances, Tapering procedure.

References

[1] Asikgil, B. (2009). *Nonlinear Regression in the Presence of Autocorrelated Disturbances.* PhD thesis. Mimar Sinan F. A. University, Istanbul, Turkey.
[2] Kmenta, J. and R.F. Gilbert (1970). Estimation of seemingly unrelated regressions with autoregressive disturbances. *J. Amer. Statist. Assoc. 65*, 186–197.
[3] Parks, R.W. (1967). Efficient estimation of a system of regression equations when disturbances are both serially and contemporaneously correlated. *J. Amer. Statist. Assoc. 62*, 500–509.

Very robust regression

Anthony C. Atkinson[1] **and Marco Riani**[2]

[1] London School of Economics, UK
[2] University of Parma, Italy

Abstract

The numerous methods of very robust regression resist up to 50% of outliers. This breakdown point, the maximum that can be achieved, is defined asymptotically as the outlying observations become infinitely far from the regression data. To distinguish between such very robust methods we study their behaviour as a function of the distance between the regression data and the outliers. We introduce a parameter λ that defines a parametric path in the space of models that enables us to study, in a systematic way, the properties of estimators as the groups of data move from being far apart to close together. We examine, as a function of λ, the variance and squared bias of several estimators and we also consider their power when used in the detection of outliers.
The results of our systematic approach are described in Riani et al. ([1]). An algorithm using the forward search (Atkinson and Riani, [2]) has the best properties for both size and power of the outlier tests. The comparisons use new algorithms for Least Trimmed Squares estimators that have increased *computational* efficiency due to improved combinatorial sampling. The efficient sampling method forms part of the subject of the talk by Domenico Perrotta.

Keywords

Distance of outliers, Forward search (FS), Least trimmed squares (LTS), MM estimate, Multiple outliers.

References

[1] Riani, M., A.C. Atkinson, and D. Perrotta (2011). Calibrated very robust regression. Technical Report NI11033-DAE, Isaac Newton Institute, Cambridge, UK.
[2] Atkinson, A.C. and M. Riani (2000). *Robust Diagnostic Regression Analysis.* Springer–Verlag, New York.

Some comments on joint papers by George P.H. Styan and the Baksalarys

Oskar M. Baksalary

Adam Mickiewicz University, Poznań, Poland

Abstract

The paper [2] opens a list of joint publications by George P. H. Styan and Jerzy K. Baksalary. Even though Jerzy passed away prematurely in 2005, the cooperation between George and the Baksalarys has continued up to the present days. It is now Jerzy's son and the author of the present talk who has a privilege and pleasure to work together with George. So far the cooperation between George and the two Baksalarys resulted in 16 papers, including 9 joint papers by George and Jerzy and 8 joint papers by George and Oskar. Thus, there is one joint paper by George, Jerzy, and Oskar, namely [1]. In the talk several comments on the joint publications by George and the Baksalarys will be made.

References

[1] Baksalary, J.K., O.M. Baksalary, and G.P.H. Styan (2002). Idempotency of linear combinations of an idempotent matrix and a tripotent matrix. *Linear Algebra Appl. 354*, 21–34.

[2] Baksalary, J.K., F. Pukelsheim, and G.P.H. Styan (1989). Some properties of matrix partial orderings. *Linear Algebra Appl. 119*, 57–85.

Multivariate linear phylogenetic comparative models and adaptation

Krzysztof Bartoszek

Chalmers University of Technology and the University of Gothenburg, Sweden

Abstract

The need for taking into account evolutionary relationships when analyzing between species data is by now firmly established. However stochastic models allowing for multiple co-evolving traits are extremely limited and essentially do not go beyond a multivariate Brownian motion with a trend. This does not allow one to model adaptation, not even with a definition as weak as convergence in distribution. The linear stochastic differential equation model presented in [3],

$$\mathrm{d}\boldsymbol{Y}(t) = -\mathbf{A}(\boldsymbol{Y}(t) - \boldsymbol{\psi}(t))\mathrm{d}t + \boldsymbol{\Sigma}\mathrm{d}\boldsymbol{W}(t),$$

where $\boldsymbol{W}(t)$ is a standard Brownian motion, allows for modelling adapting traits with such as notion of adaptation but up till now had only partial multivariate implementations [2,4]. In the talk a recently developed R package [1] which nearly completely covers the framework from [3] in multiple dimensions will be presented. The properties of the mean and covariance functions will be discussed in terms of the definition of adaptation as weak convergence. With multiple interacting traits the study of adaption requires one to look at conditional distributions of interest, especially their limiting properties. These will be presented and discussed with an emphasis on their biological interpretation. For example if we consider the multivariate extension of the model from [4],

$$\begin{aligned}\mathrm{d}\boldsymbol{Y}(t) &= -\mathbf{A}(\boldsymbol{Y}(t) - (\boldsymbol{\psi}(t) + \mathbf{B}\boldsymbol{X}(t)))\mathrm{d}t + \boldsymbol{\Sigma}_y\mathrm{d}\boldsymbol{W}_y(t)\\ \mathrm{d}\boldsymbol{X}(t) &= \boldsymbol{\Sigma}_x\mathrm{d}\boldsymbol{W}_x(t),\end{aligned}$$

then if $\mathbf{A}$ has positive real part eigenvalues, the regression coefficient of $\boldsymbol{Y}(t)$ on $\boldsymbol{X}(t)$ will converge to $\mathbf{B}$, but if the $\boldsymbol{X}$ variables are also adapting,

$$\begin{aligned}\mathrm{d}\boldsymbol{Y}(t) &= -\mathbf{A}_y(\boldsymbol{Y}(t) - (\boldsymbol{\psi}_y(t) + \mathbf{B}\boldsymbol{X}(t)))\mathrm{d}t + \boldsymbol{\Sigma}_y\mathrm{d}\boldsymbol{W}_y(t)\\ \mathrm{d}\boldsymbol{X}(t) &= -\mathbf{A}_x(\boldsymbol{X}(t) - \boldsymbol{\psi}_x(t))\mathrm{d}t\boldsymbol{\Sigma}_x\mathrm{d}\boldsymbol{W}_x(t),\end{aligned}$$

then this limit will not in general equal $\mathbf{B}$. This can be interpreted that even if evolution would go on for infinity the $\boldsymbol{Y}$ and $\boldsymbol{X}$ traits would never evolve to the optimal relationship between them. These concepts will be illustrated by an example re–analysis of the Cervidae dataset [5].

Keywords

Ornstein–Uhlenbeck process, Phylogenetic comparative methods, Multivariate models, Evolution, Adaptation.

References

[1] Bartoszek, K., J. Pienaar, P. Mostad, S. Andersson, and T.F. Hansen. A comparative method for studying multivariate adaptation. *J. Theor. Biol.* Submitted.

[2] Butler, M.A. and A.A. King (2004). Phylogenetic comparative analysis: a modelling approach for adaptive evolution. *Am. Nat. 164*, 683–695.

[3] Hansen, T.F. (1997). Stabilizing selection and the comparative analysis of adaptation. *Evolution 51*, 1341–1351.

[4] Hansen, T.F., J. Pienaar, and S.H. Orzack (2008). A comparative method for studying adaptation to a randomly evolving environment. *Evolution 62*, 1965–1977.

[5] Plard, F., C. Bonenfant, and J.M. Gaillard (2011). Revisiting the allometry of antlers among deer species: male–male sexual competition as a driver. *Oikos 120*, 601–606.

Study with George Styan

Philip Bertrand

Solihull, UK

Abstract

George and I met 56 years ago in 1956. He was studying Pure Mathematics, I Mathematical Physics. The first lecture of a new Professor of Mathematical Statistics at the University, Henry Daniels, was more interesting than any other lecture I had received there in my previous two years. Henry Daniels, changed the direction of both George and I to Mathematical Statistics. Henry showed that the subject of Mathematical Statistics is 'the application of the scientific method to the study of any subject'. He demonstrated the logic of this assertion. To understand and develop the subject a student needs to study the topics in pure mathematics including complex variable theory, group theory, statistical distribution theory, geometry, algebra, calculus, stochastic processes, determinants and matrices. These subjects I shared in classes with George Styan between 1957 to 1959. In 1959 George moved to Oxford University to do a masters degree. I continued studying a postgraduate diploma in mathematical statistics under Henry Daniels. In 1960 George and I both worked in London where we met frequently for social discussions with other friends. We also frequently discussed our different statistical problems. George moved on to North America whilst I continued to work in Britain. We met again around 1990 when George came to give a talk in our department.

Our friendship continues from then on.

Keywords

Algebra, Complex variables, Group theory, Determinants and matrices.

Jackknife-after-Bootstrap as logistic regression diagnostic tool

Ufuk Beyaztaş and **Aylin Alin**

Dokuz Eylül University, Izmir, Turkey

Abstract

Jackknife-after-Bootstrap (JaB) has first been proposed by [1] then used by [2] and [3] to detect influential observations in linear regression models. In this study, we propose using JaB to detect influential observations in logistic regression model. Performance of the proposed method will be compared with the traditional method for standardized Pearson residuals, Cook's distance, change in the Pearson chi-square statistic and change in the deviance by both real world examples and simulation study. The results reveal that under considered scenarios proposed method performs better than traditional method and is more robust to masking and swamping effects.

Keywords

Logistic regression, Bootstrap, Jackknife, Logistic regression diagnostics.

References

[1] Efron, B. (1992). Jackknife-after-Bootstrap standard errors and influence functions. *J. R. Stat. Soc. Ser. B Stat. Methodol. 54* 83–127.

[2] Martin, M.A. and S. Roberts (2010). Jackknife-after-Bootstrap regression influence diagnostics. *J. Nonparametr. Stat. 22* 257–269.

[3] Beyaztaş, U. and A. Alin (2012). Jackknife-after-Bootstrap method for detection of influential observations in linear regression models. *Comm. Statist. Simulation Comput.* In press.

Optimum designs for enzyme kinetic models with co-variates

Barbara Bogacka[1], **Mahbub Latif**[2], **and Steven Gilmour**[3]

[1] Queen Mary, University of London, UK
[2] University of Dhaka, Bangladesh
[3] University of Southampton, UK

Abstract

In this talk we consider a population optimum design of experiments for non-linear models and in the specific application to enzyme kinetic studies. In the early stage of drug development pharmaceutical companies are interested in whether the new candidate medicinal product interacts with other drugs. Since most of the drugs are metabolized in human liver, these early stage pharmacokinetic experiments are conducted at different levels of concentration of the new compound applied to liver tissues representing "subjects" in the study. Also, the liver tissues differ in some systematic way what can be incorporated in the model as a function of co-variates. In our studies, which are based on a set of real data, we find that some of the parameters of this function differ across the population and so are treated as random. The question is about the choice of the liver tissues as well as the levels of concentration of the new medicinal product so that all the model parameters are estimated with high precision. Although it is set in a specific application it prompts several methodology questions in the optimum design theory to be answered.

Keywords

Mixed-effects model, D-optimality, Transform-both-sides model.

On combining information in a generally balanced nested block design

Tadeusz Caliński

Poznań University of Life Sciences, Poland

Abstract

Nested block designs are quite often used in practice, particularly in agricultural experimentation. Their statistical properties have been considered in many papers, as reviewed by Bailey (1999). Of special interest are those nested block designs which satisfy the general balance property introduced by Nelder (1965) and discussed by several authors, by Bailey (1994) and by Bogacka and Mejza (1994) in particular.
The purpose of the present paper is to give explicit formulae for analyzing an experiment carried out in a nested block design having the general balance property of some desirable pattern. The results follow from a randomization-derived mixed model, decomposed into stratum submodels. Attention is confined here to the combined analysis allowing the information from different strata to be joined together, following Nelder (1968). The paper is essentially an extension of some results presented in Chapter 5 of Caliński and Kageyama (2000).

Keywords

Combined analysis, General balance property, Nested block design, Randomization-derived model, Stratum submodels.

References

[1] Bailey, R.A. (1994). General balance: Artificial theory or practical relevance? In: T. Caliński, R. Kala (Eds.), *Proc. Int. Conf. on Linear Statist. Inference LINSTAT'93* (pp. 171–184). Kluwer Acad. Publ., Dordrecht.

[2] Bailey, R.A. (1999). Choosing designs for nested blocks. *Biom. Lett. 36*, 85–126.

[3] Bogacka, B. and S. Mejza (1994). Optimality of generally balanced experimental block designs. In: T. Caliński, R. Kala (Eds.), *Proc. Int. Conf. on Linear Statist. Inference LINSTAT'93* (pp. 185–194). Kluwer Acad. Publ., Dordrecht.

[4] Caliński, T. and S. Kageyama (2000). *Block Designs: A Randomization Approach, Volume I: Analysis.* Lecture Notes in Statistics, Volume 150. Springer, New York.

[5] Nelder, J.A. (1965). The analysis of randomized experiments with orthogonal block structure. *Proc. Roy. Soc. Lond. Ser. A 283*, 147–178.

[6] Nelder, J.A. (1968). The combination of information in generally balanced designs. *J. Roy. Statist. Soc. Ser. B 30*, 303–311.

Linear and quadratic sufficiency in mixed model

Francisco Carvalho[1,3], Augustyn Markiewicz[2], and João T. Mexia[1,4]

[1] Centro de Matemática e Aplicações, Universidade Nova de Lisboa, Portugal
[2] Poznań University of Life Sciences, Poland
[3] Instituto Politécnico de Tomar, Portugal
[4] Departamento de Matemática, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Portugal

Abstract

Up to now, see e.g. [3] and [1], linear and quadratic sufficiency have been mainly used in obtaining BLUE and BQUE for models with one variance component.
We use the orthogonal structure of variance-covariance matrix of models with orthogonal block structure to extend the use of linear and quadratic sufficiency, as defined in [2], in obtaining best linear unbiased estimators and best quadratic unbiased estimators.
We will consider the model

$$M_\sigma : \mathbf{Y} = \mathbf{X}\beta + \mathbf{X}_1\beta_1 + \varepsilon$$

where β is fixed and β_1 and ε are independent with null mean vectors and variance-covariance matrices $\sigma_1^2\mathbf{I}_{c_1}$ and $\sigma^2\mathbf{I}_n$.

Keywords

Linear sufficiency, Quadratic sufficiency, Variance components, Mixed model.

References

[1] Groß, J. (1998). A note on the concepts of linear and quadratic sufficiency. *J. Statist. Plann. Inference 70(1)*, 69–76.
[2] Mueller, J. (1987). Sufficiency and completeness in the linear model. *J. Multivariate Anal. 21(2)*, 312–323.
[3] Neudecker, H. (1980). Best quadratic unbiased estimation of the variance matrix in normal regression. *Statist. Papers 21(3)*, 239–243.

The magic behind the construction of certain Agrippa–Cardano type magic matrices

Ka Lok Chu[1], George P. H. Styan[2], and Götz Trenkler[3]

[1] Dawson College, Westmount, Canada
[2] McGill University, Montreal, Canada
[3] Dortmund University of Technology, Germany

Abstract

We build on results [5] presented at the International Workshop on Combinatorial Matrix Theory and Generalized Inverses of Matrices (Manipal University, January 2012). In this talk we will present procedures for the construction of certain Agrippa–Cardano [1,2] type magic matrices using *magic-basis matrices*. We generate classic rank-3 $n \times n$ Agrippa–Matlab [1,4] magic matrices with n doubly-even, and classic nonsingular Agrippa–Fermat magic matrices [1,3] with n singly-even. We investigate some matrix-theoretic properties and present some interesting findings.

References

[1] Agrippa von Nettesheim, H.C. (1533). *De occulta philosophia libri tres.* Soter, Köln [Cologne].

[2] Cardano, G. (1539). *Practica arithmetice et mensurandi singularis: in qua que preter alias coninentur, versa pagina demonstrabit*, medico mediolanonsis. Mediolani: Io. Antonins Castellioneus medidani imprimebat, impensis Bernardini Calusci, 312 pp.

[3] Fermat, P. de (1894). *Œuvres de Fermat, tome deuxième: correspondance.* Gauthier-Villars, Paris.

[4] Moler, C. (1993). Cleve's Corner: MATLAB's magical mystery tour, *The MathWorks Newsletter 7(1)*, 8–9.

[5] Styan, G.P.H., G. Trenkler, and K.L. Chu (2012). An introduction to Yantra magic squares and Agrippa–Cardano type magic matrices: Lecture notes. In: R.B. Bapat, S.J. Kirkland, K.M. Prasad, S. Puntanen (Eds.), *Lectures on Matrix and Graph Methods* (pp. 159–220). Manipal University Press. [Expanded version: Report 2011-07, Dept. of Mathematics and Statistics, McGill University, 129 pp., 13 June 2012.]

Celebrating George P. H. Styan's 75th birthday and my meetings with him

Carlos A. Coelho

Departamento de Matemática and Centro de Matemática e Aplicações, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Portugal

Abstract

The first time I met George Styan was in July 2004 in Lisbon when he was on his way to the 11th ILAS Conference in Coimbra.
But George had already been in Portugal before and I learned how much he was fond of Conventual, a very fine and nice old style restaurant in Lisbon. Then I also learned that George really is an appreciator of good food and a very well-educated wine drinker. With this detail in common it was really easy to become a good friend with George.
Since then we met a number of times, the most significant of which was at the time of the 17th IWMS held in Tomar, Portugal, in 2008.
Before this event, during a short stay of George and Evelyn in Lisbon, we had the opportunity to go to some nice spots like Sintra and to hang around a few nice places near Lisbon and even to attend a Leonard Cohen concert, together with some friends.
Actually, even more than good food and a good wine, and more than a good mathematical challenge, George enjoys the company of his family and his friends. We may even say that more than Mathematics, it is his family and his friends that play and have always played a central role in his life. Everybody knows well how much he cares about Evelyn, the great woman behind the great man, and also everybody knows the looks in George's face when he meets the ones he cares about.
Inevitably, besides addressing some of George's honors and also his scientific work and his interest in mathematics related stamps, it is based on a number of pictures, either taken by the author or by other friends and a couple of them even taken by George himself, that this little contribution to the celebration of George Styan's 75th birthday will be indeed more a celebration of the way George enjoys and nurtures the company of the ones he loves.

On the distribution of linear combinations of chi-square random variables

Carlos A. Coelho

Departamento de Matemática and Centro de Matemática e Aplicações, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Portugal

Abstract

The distribution of linear combinations of independent chi-square random variables is intimately related with the distribution of quadratic forms in normal random variables [1,6–11,13,14] and thus it also appears as the limit distribution of quadratic forms in non-normal random variables. As such, this distribution has been studied by many authors [2,4–15]. However, there is still much room left for improvement, since while some simpler approximations do not yield sufficiently good results, other approximations which show a better performance are sometimes too complicated to be implemented in practical terms.

In this paper the exact distribution of linear combinations of independent chi-square random variables is obtained, for some particular cases, in a closed finite highly manageable form, while for the general case a near-exact approximation [3] is obtained, which is able to yield very manageable and well-performing approximations.

Keywords

Characteristic function, Gamma distribution, Generalized integer gamma distribution, Generalized near-integer gamma distribution, Mixtures.

References

[1] Baksalary, J.K., J. Hauke, and G.P.H. Styan (1994). On some distributional properties of quadratic forms in normal variables and on some associated matrix partial orderings. In: T.W. Anderson, K.T. Fang, I. Olkin (Eds.), *Multivariate Analysis and its Applications, IMS Lecture Notes-Monograph Series* (pp. 111-121). Institute of Mathematical Statistics, Hayward, California, vol. 24.

[2] Castaño-Martínez, A. and F. López-Blázquez (2005). Distribution of a sum of weighted noncentral chi-square variables. *Test 14*, 397–415.

[3] Coelho, C.A. (2004). The generalized near-integer Gamma distribution: a basis for 'near-exact' approximations to the distribution of statistics which are the product of an odd number of independent Beta random variables. *J. Multivariate Anal. 89*, 191–218.

[4] Davis, A.W. (1977). A differential equation approach to linear combinations of independent chi-squares. *J. Amer. Statist. Assoc. 72*, 212–214.

[5] Gabler, S. and C. Wolff (1987). A quick and easy approximation to the distribution of a sum of weighted chi-square variables. *Statist. Papers 28*, 317–325.

[6] Imhof, J.P. (1961). Computing the distribution of quadratic forms in normal variables. *Biometrika 48*, 419–426.

[7] Jensen, D.R. and H. Solomon (1972). A Gaussian approximation to the distribution of a definite quadratic form. *J. Amer. Statist. Assoc. 67*, 898–902.

[8] Johnson, N.L. and S. Kotz (1968). Tables of distributions of positive definite quadratic forms in central normal variables. *Sankhya B 30*, 303–314.

[9] Kotz, S. and N.L. Johnson (1967). Series representations of distributions of quadratic forms in normal variables. I. Central case. *Ann. Math. Statist. 38*, 823–837.

[10] Kotz, S. and N.L. Johnson (1967). Series representations of distributions of quadratic forms in normal variables. II. Non-central case. *Ann. Math. Statist. 38*, 838–848.

[11] Lu, Z.-H. (2006). The numerical evaluation of the probability density function of a quadratic form in normal variables. *Comput. Statist. Data Anal. 51*, 1986–1996.

[12] Moschopoulos, P.G. and W.B. Canada (1984). The distribution function of a linear combination of chi-squares. *Comput. Math. Appl. 10*, 383–386.

[13] Robbins, H. (1948). The distribution of a definite quadratic form. *Ann. Math. Statist. 19*, 266–270.

[14] Robbins, H. and E.J.G. Pitman (1949). Application of the method of mixtures to quadratic forms in normal variates. *Ann. Math. Statist. 20*, 552–560.

[15] Solomon, H. and M.A. Stephens (1977). Distribution of a sum of weighted chi-square variables. *J. Amer. Statist. Assoc. 72*, 881–885.

Multivariate analysis of polarimetric SAR images

Knut Conradsen

Technical University of Denmark, Lyngby, Denmark

Abstract

The author first met G.P.H. Styan at a meeting in Greece 40 years ago. During the years, they have shared the interest in matrices and multivariate statistics, GPHS from a mathematical perspective, KC an applied do. In the presentation those perspectives are combined in some applications of the multivariate complex Wishart distribution in the analysis of radar images.
Due to its all-weather mapping capability independently of e.g. cloud cover, synthetic aperture radar (SAR) data holds a strong potential for change detection studies in remote sensing applications. The radar backscattering is sensitive to the dielectric properties of the vegetation and the soil, to the plant structure (i.e., the size, shape, and orientation distributions of the scatterers), to the surface roughness, and to the canopy structure (e.g., row direction and spacing, and cover fraction). The polarimetric SAR measures the amplitude and phase of backscattered signals in four combinations of the linear receive and transmit polarizations: HH, HV, VH, and VV. These signals form the complex scattering matrix. The inherent speckle in the SAR data is reduced by spatial averaging (at the expense of loss of spatial resolution). In this so-called multi-look case a more appropriate representation of the backscattered signal is the covariance matrix in which the average properties of a group of resolution cells can be expressed in a single matrix. This averaged covariance matrix follows a complex Wishart distribution.
In [2,4] change detection was analyzed on bi-temporal data. In [3] these results are extended to multitemporal data. A good survey on the relevant theory on multivariate analysis in the complex normal setting is given in [1].

References

[1] Coelho, C.A., F.J. Marques, and B.C. Arnold (2011). A general approach to the exact and near-exact distributions of the main likelihood ratio test statistics used in the complex multivariate Normal setting. Preprint. [http://www.dm.fct.unl.pt/sites/www.dm.fct.unl.pt/files/preprints/2011/19_11.pdf]

[2] Conradsen, K., A.A. Nielsen, J. Schou, and H. Skriver (2003). A test statistic in the complex Wishart distribution and its application to change detection in polarimetric SAR data. *IEEE Trans. Geoscience Remote Sensing 41(1)*, 4–19.

[3] Conradsen, K., A.A. Nielsen, and H. Skriver (2012). A test statistic for equality of several complex covariance matrices applied to change detection in truly multi-temporal, full and dual polarization SAR data. Work in progress.

[4] Schou, J., H. Skriver, A.A. Nielsen, and K. Conradsen (2003). CFAR edge detector for polarimetric SAR images. *IEEE Trans. Geoscience Remote Sensing 41(1)*, 20–32.

Some math on the electricity market by a generalization of the Black-Scholes formula

Ricardo Covas

Centro de Matemática e Aplicações, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Portugal

Abstract

Electricity is an interesting commodity for mathematicians to work on. In fact, the variety of financial and real options traded are far from being plain vanilla and, nevertheless, being most quite exotic, they have been priced with standard tools. No doubt, the literature is sparse and it's a growing subject. One of the interesting options which, nowadays, exist in Europe is the different electricity prices between countries. If one has the ability to trade energy across countries, these electricity spreads are spread options. Since this possibility to trade is limited in time and capacity, the existing spread option is, somehow, unique.

Selling energy across countries implies offer and demand bids for electricity in each country, thus having positive probabilities of negative cashflows, by which the option to use transfer capacities cannot be priced with the Black-Scholes Formulas. In this work it's proposed a new way of pricing the daily electricity transfer capacity, where we take into account the traders daily operation (which has changed since [2]) and, therefore, all the inherent risks factors not included in the Black & Scholes world.

Keywords

Electricity market, Spread option, Conditional expectation.

References

[1] Black, F. and M. Scholes (1973). The pricing of options and corporate liabilities. *J. Political Econ. 81(3)*, 637–654.

[2] Covas, R. and L. Pascual (2008). Pricing daily electricity transfer capacities in the Spain-France Interconnection. *IEEE.*

[3] Eydeland, A. and K. Wolyniec (2003). *Energy and Power Risk Management.* Wiley and Sons.

[4] Margrabe, W. (1978). The value of an option to exchange one asset for another. *J. Finance 33(1)*, 177–186.

Mutual Principal Components, reduction of dimensionality in statistical classification

Carlos Cuevas-Covarrubias

Anahuac University, Mexico

Abstract

Linear discriminant analysis (LDA) and principal components analysis (PCA) are two fundamental tools of multivariate statistics. Given a p-dimensional random variable $\mathbf{X}$, PCA finds its optimal representation in a lower dimensional space. LDA assumes that the sample space of $\mathbf{X}$ is partitioned into two different categories. Given $\mathbf{x}$, a particular realization of $\mathbf{X}$, LDA lets us infer whether $\mathbf{x}$ comes from one category or the other. We present an original combination of PCA and LDA where the area under the ROC curve appears as the link between both methods; we call this *Mutual Principal Components.* Our objective is to represent $\mathbf{X}$ in terms of a small number of non correlated factors and maximum separability. Assuming that $\mathbf{X}$ is distributed according to a Gaussian mixture, a parametric approach selects those components with maximum contribution to the area under the ROC curve of an optimal linar discriminant function. A distribution free alternative shows that this principle is equivalent to maximize the square cosine between this discriminant function and the vector space generated by the colums of the resulting principal components transformation matrix.

Keywords

Classification, Linear score, ROC curve, PCA, Reduction of dimensionality.

References

[1] Anderson, T.W. and R.R. Bahadur (1962). Classification into two multivariate normal distributions with different covariance matrices. *A*nn. Math. Statist. 33, 420–431.

[2] Anderson, T.W. (1984). *An Introduction to Multivariate Statistical Analysis* (2nd ed). John Wiley and Sons.

[3] Chang, W.C. (1983). Using Principal Components before separating a mixture of two multivariate normal distributions. *Appl. Statist. 32(3)*, 267–275.

[4] Krzanowski, W.J. and D.J. Hand (2009). *ROC Curves for Continuous Data.* CRC Press.

Nonparametric regression using partial least squares dimension reduction in multistate models

Susmita Datta

University of Louisville, USA

Abstract

We introduce a method of constructing non-parametric regression estimators of state occupation probabilities in a multistate model. In order to tackle potentially large number of predictors in modern genomic and proteomic data sets we use partial least squares to compute estimated latent factors from the transition times along with the covariates which are then used in an additive model in order to avoid curse of dimensionality. We illustrate the methodology using simulated and real data sets.

Estimating intraclass correlation and its confidence interval in linear mixed models

Nino Demetrashvili and Edwin van den Heuvel

University of Groningen, University Medical Center Groningen, The Netherlands

Abstract

The methodology proposed in this study is motivated by an example from the medical field. Oncologists delineate organs for radiotherapy and it is essential that the measurements agree in these procedures. To assess the consistency of measurements among oncologists, on a random sample of subjects, the intraclass correlation (ICC) would yield a suitable estimate for studying the agreement.

In technical terms, the ICC is a ratio of sum of variances that are related to differences among measured subjects and the total variance. What variance is considered relevant depends on the design of agreement study; respectively, the number of variance components changes in the numerator and the denominator of the ICC. For statistical inference, it is important but challenging to determine the distribution of estimators of such ratios and to construct the confidence intervals. In most literature, the ICC has been studied for one-way and two-way analysis of variance only. Most proposed approximate methods are based on functions of the mean squares which are model-specific (e.g. two factorial) and lack generalization to higher order (e.g. three factorial) models. The objective of this study is to extend the construction of confidence intervals for the linear mixed models, but in particular to our three-way mixed models for delineation of organs. The generalization will coincide with existing methods for two-way and one-way mixed effects models. To obtain an approximate upper and lower confidence limits, we approximate the ICC with a function of F-distributed variable and a Beta distribution. Our proposed methodology is supported by simulation studies.

Keywords

Linear mixed model, Confidence interval, Intraclass correlation, Small sample.

References

[1] Burdick, R.K. and F.A. Graybill (1992). *Confidence Intervals on Variance Components.* Marcel Dekker.

[2] Donner, A. and G. Wells (1986). A comparison of confidence interval methods for the intraclass correlation coefficient. *Biometrics 57(2)*, 401–412.

[3] Knight, K. (2000). *Mathematical Statistics*. Chapman & Hall.

[4] Müller, R. and P. Büttner (1994). A critical discussion of intraclass correlation coefficients. *Stat. Med. 13*, 2465–2476.

[5] Searle, S.R., G. Casella, and C.E. McCulloch (2006). *Variance Components*. Wiley.

[6] Srivastava, M.S. (1993). Estimation of the intraclass correlation coefficient. *Ann. Human Genetics 57*, 159–165.

[7] Swallow, W.H. and J.F. Monahan (1984). Monte Carlo comparison of ANOVA, MIVQUE, REML, and ML estimators of variance components. *Technometrics 26*, 47–57.

Linear models in the face of Diabetes Mellitus: the influence of physical activity

Hilmar Drygas

University of Kassel, Germany

Abstract

A linear model for Diabetes Mellitus is described. The influence factors are nutrition, time and physicals activity. Two models are compared, one with moderate physical activity and another one with strong physical activity. The question is whether strong physical activity leads to a significant reduction of the blood-sugar. It is shown that there are substantial reductions of blood-sugar due to physical activity, but due to a high variance significance can only be achieved in very rare cases.

Normality test based on Song's multivariate kurtosis

Rie Enomoto[1], **Naoya Okamoto**[2], **and Takashi Seo**[1]

[1] Tokyo University of Science, Japan
[2] Tokyo Seiei College, Japan

Abstract

In statistical analysis, the test for normality is an important problem. The most widely applied tests of multivariate normality are based on Mardia's multivariate generalization of skewness and kurtosis. Mardia [1], Srivastava [3] and Song [2] gave definitions of the multivariate sample kurtosis. We consider the multivariate normality test based on the sample measure of multivariate kurtosis defined by Song [2]. We derive expectation and variance of Song's kurtosis and a new test statistic for assessing multivariate normality. Moments of Song's kurtosis are calculated easily using independency of random vectors. We investigate the accuracies of upper percentiles, type I error and of power for the test statistic via a Monte Carlo simulation for selected values of parameters.

Keywords

Multivariate normality test, Multivariate kurtosis, Asymptotic expansion.

References

[1] Mardia, K.V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika 57*, 519–530.

[2] Song, K.-S. (2001). Rényi information, loglikelihood and an intrinsic distribution measure. *J. Statist. Plann. Inference 93*, 51–69.

[3] Srivastava, M.S. (1984). A measure of skewness and kurtosis and a graphical method for assessing multivariate normality. *Statist. Probab. Lett. 2*, 263–267.

A graphical evaluation of Robust Ridge Regression in mixture experiments

Ali Erkoç[1] **and Kadri U. Akay**[2]

[1] Mimar Sinan Fine Arts University, Istanbul, Turkey
[2] University of Istanbul, Turkey

Abstract

In mixture experiments, estimation of the parameters is generally based on Ordinary Least Squares (OLS). However, in the presence of multicollinearity and outlier, OLS can result in very poor estimates. In this case, effects due to the combined outlier-multicollinearity problem can be reduced to certain extent by using alternative approaches. One of these approaches is to use biased-robust regression techniques for the estimation of the parameters. In this paper, we suggest the use of robust ridge regression based on M-estimator in the cases where there is multicollinearity and outliers during the analysis of mixture experiments. Also, for selection of biasing parameter, we use a new graphical approach for evaluating the effect of the robust ridge regression estimator with respect to the scaled prediction variance and fraction of design space plots. The suggested graphical approaches are illustrated on hot-melt adhesive data set.

Keywords

Experiments with mixture, Robust regression, Robust Ridge Regression, Multicollinearity, Scaled prediction variance, Fraction of design space plot.

References

[1] Pfaffenberger, R.C. and T.E. Dielman (1990). A comparison of regression estimators when both multicollinearity and outliers are present. In: K.D. Lawrence, J.L. Arthur (Eds.), *Robust Regression: Analysis and Applications* (pp. 243–270). Marcel Dekker, Inc.

[2] Silvapulle, M.J. (1991). Robust ridge regression based on M-estimator. *Austral. J. Statist. 33(3)*, 319–333.

[3] Smith, W.F. (2005). *Experimental Design for Formulation.* ASA-SIAM Series on Statistics and Applied Probability, SIAM, Philadelphia, ASA, Alexandria, VA.

A comparison of different parameter estimation methods in fuzzy linear regression

Birsen Eygi Erdogan and **Fatih Erduvan**

Marmara University, Istanbul, Turkey

Abstract

Fuzzy logic is the concept that concerns people's thinking with imprecise statements. It is easy to work with accurate data through the classic linear regression analysis. However, it is inevitable to use fuzzy linear regression if the dependent or independent variables or the relation between them are fuzzy. The estimation of the fuzzy linear regression parameters generally are gained by two approaches. The first one includes the methods that are based on linear programming. The second one is based on the methods of the fuzzy least squares. The main object of this paper is to apply and compare the performance of the different fuzzy logic approximation methods using a real world data set (the Ataþehir district housing prices).

Keywords

Fuzzy logic regression, House pricing forecast, Linear programming.

References

[1] Diamond, P. (1988). Fuzzy least squares. *Inform. Sci. 46*, 141–157.
[2] Peters, G. (1994) Fuzzy linear regression with fuzzy intervals. *Fuzzy Sets and Systems 63*, 45–55.
[3] Hojati, M., C.R. Bector, and K. Simimou (2005). A simple method for computation of fuzzy linear regression. *European J. Oper. Res. 166*, 172–184.
[4] Kim, B. and R.R. Bishu (1998). Evaluation of fuzzy linear regression models by comparing membership functions. *Fuzzy Set and Systems 100*, 342–352.
[5] Zadeh, L.A. (1965). Fuzzy Sets. *Information and Control 8*, 338–353.

On universal optimality of circular repeated measurements designs

Katarzyna Filipiak

Poznań University of Life Sciences, Poland

Abstract

Our aim is to characterize the universally optimal design among the class of circular repeated measurements designs. We show, that some circular weakly neighbor balanced designs defined by Filipiak and Markiewicz [2] for an interference model, which are uniform on periods, are universally optimal under the model of repeated measurements design. Our results correspond to the work of Magda [3] and Kunert [2].

Keywords

Repeated measurements designs, Uniform design, Circular balanced design, Universal optimality.

References

[1] Filipiak, K. and A. Markiewicz, A. (2012). On universal optimality of circular weakly neighbor balanced designs under an interference model. *Comm. Statist. Theory Methods 41*, 2356–2366.
[2] Kunert, J. (1984). Designs balanced for circular residual effects. *Comm. Statist. Theory Methods 13(21)*, 2665–2671.
[3] Magda, C.G. (1980). Circular balanced repeated measurements designs. *Comm. Statist. Theory Methods 9*, 1901–1918.

Constructing efficient exact designs of experiments using integer quadratic programming

Lenka Filová and Radoslav Harman

Comenius University in Bratislava, Slovakia

Abstract

We propose a method of computing exact experimental designs by integer quadratic programming. The key idea is a suitable quadratic approximation of the criterion of D-optimality in the neighbourhood of the approximate D-optimal information matrix, which we call the criterion of Q-optimality. We demonstrate on several examples that the D-efficiency of the exact Q-optimal designs is usually very high. An important advantage of the method is that it can be applied to situations with marginal and cost constraints on the design.

Keywords

D-optimal design, Q-optimal design, Exact design, Marginal restrictions, Cost restrictions, Integer quadratic programming.

Sensitivity analysis in mixed models

Eva Fišerová

Palacký University Olomouc, Czech Republic

Abstract

Statistical models for experiments in geodesy, biology, environmental research, etc. usually involve unknown parameters not only in a regression function but also in a covariance matrix (variance components). For measurement it is used two or more different measurement devices. Since it is not known whether precision of measurement specified in certificates is true, the variance components must be estimated, e.g. by minimum norm quadratic unbiased estimator (MINQUE) [1], [3], and plug-in estimators for the regression parameters can be used. To find statistical properties of plug-in estimators is rather difficult. In some cases the sensitivity approach can be used. If we know that the true value of the variance components is with sufficiently high probability in so-called insensitivity region, then the plug-in estimator is almost the best linear unbiased estimator [2]. Consequently, approximations of variance components can destroy the optimum quality of statistical inference, e.g. confidence and significance levels, what can also be analyzed by sensitivity approach. In the contribution the sensitivity analysis will be applied on geodetical example.

Keywords

Plug-in estimator, Insensitivity region, Variance components, MINQUE.

References

[1] Fišerová, E., L. Kubáček, and P. Kunderová (2007). *Linear Statistical Models: Regularity and Singularities.* Academia, Praha.

[2] Fišerová, E. and L. Kubáček (2009). Insensitivity region for deformation measurement on a dam. *Environmetrics 20*, 776–789.

[3] Rao, C.R. and J. Kleffe (1988). *Estimation of Variance Components and Applications.* North Holland, Amsterdam.

Inference in linear models with doubly exchangeable distributed errors

Miguel Fonseca[1] **and Anuradha Roy**[2]

[1] Centro de Matemática e Aplicações, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Portugal
[2] University of Texas at San Antonio, USA

Abstract

We study the general linear model (GLM) with doubly exchangeable distributed error for m observed random variables. The doubly exchangeable general linear model (DEGLM) arises when the $m-$dimensional error vectors are "doubly exchangeable" (defined later), jointly normally distributed, which is a much weaker assumption than the independent and identically distributed error vectors as in the case of GLM or classical GLM (CGLM). We estimate the parameters in the model and also find their distributions. We show that the testings of intercept and slope are possible in DEGLM as a particular case using parametric bootstrap as well as multivariate Satterthwaite approximation.

Keywords

Doubly exchangeable covariance structure, Linear model, Parametric bootstrap, Multivariate Satterthwaite approximation.

References

[1] Anderson, T.W. (2003). *An Introduction to Multivariate Statistical Analysis*, 3rd ed. New Jersey: Wiley.
[2] Arnold, S.F. (1979). Linear models with exchangeably distributed errors. *J. Amer. Statist. Assoc. 74*, 194–199.
[3] Gupta, A.K. and D.K. Nagar (2000). *Matrix Variate Distributions*. Chapman & Hall/CRC, Florida.
[4] Khuri, A.I. (2010). *Linear Model Methodology*. Chapman & Hall/CRC, Florida.
[5] Khuri, A.I., T. Mathew, and B.K. Sinha (1998). *Statistical Tests for Mixed Linear Models*. Wiley Series in Probability and Mathematical Statistics, Applied Section, New York: Wiley.
[6] Kollo, T. and D. von Rosen (2005). *Advanced multivariate statistics with matrices*. Dordrecht, The Netherlands: Springer.
[7] Harville, D.A. (1997). *Matrix Algebra from a Statistician's Perspective*. New York: Springer-Verlag.
[8] Leiva, R. (2007). Linear discrimination with equicorrelated training vectors. *J. Multivariate Anal. 98*, 384–409.

[9] Mitra, P.K. and B.K. Sinha (2007). A generalized p-value approach to inference on common mean. *J. Stat. Plann. Inference 137*, 3634–3642.

[10] Pan, J.X. and K.T. Fang (2002). *Growth Curve Models and Statistical Diagnostics*. New York: Springer-Verlag.

[11] Roy, A. and R. Leiva (2007). Discrimination with jointly equicorrelated multilevel multivariate data. *Adv. Data Anal. Classif. 1*, 175–199.

[12] Tan, W.Y. and R.P. Gupta (1983). On approximating a linear combination of central Wishart matrices with positive coefficients. *Comm. Statist. Theory Methods 12*, 2589–2600.

[13] Weerahandi, S. (1996). *Exact Statistical Methods for Data Analysis*. Berlin: Springer-Verlag.

Latin hypercube designs and block-circulant matrices

Stelios D. Georgiou

University of the Aegean, Samos, Greece

Abstract

Computer simulations are usually needed to study a complex physical process. In this paper, some procedures for constructing orthogonal block-circulant Latin hypercube designs are proposed. The basic concept of these methods is to use vectors with a constant periodic autocorrelation function to obtain suitable block-circulant Latin hypercube designs. Using this method one is able to construct orthogonal and near-orthogonal Latin hypercube designs with favorable properties. Orthogonal Latin hypercube designs (OLHDs) with fixed number of factors and flexible run sizes can be constructed using a slightly modified technique. Some new multiplication structures and constructions are also provided. For example, it is shown how one may obtain orthogonal Latin hypercube designs with $(runs, factors) = (2n\ell + s, m\ell)$, for $\ell = 12, 16, 20, 24$ and $s = 0, 1$ by using an $OLHD(n, m)$. The properties of the generated designs are further investigated and a brief comparison with known designs is given.

Keywords

Computer experiments, Fold-over designs, Circulant matrices, Autocorrelation function, Orthogonal designs, Construction.

References

[1] Bingham, D., R.R. Sitter, and B. Tang (2009). Orthogonal and nearly orthogonal designs for computer experiments. *Biometrika 96*, 51–65.
[2] Butler, N.A. (2001). Optimal and orthogonal Latin hypercube designs for computer experiments. *Biometrika 88*, 847–857.
[3] Cioppa, T.M. and T.W. Lucas (2007). Efficient nearly orthogonal and spaced-filling Latin hypercubes. *Technometrics 49*, 45–55.
[4] Fang, K.T., R. Li, and A. Sudjianto (2006). *Design and Modeling for Computer Experiments*. New York: CRC press.
[5] Georgiou, S.D. (2009). Orthogonal Latin hypercube designs from generalized orthogonal designs. *J. Statist. Plann. Inference 139*, 1530–1540.
[6] Lin, C.D., D. Bingham, R.R. Sitter, and B. Tang (2010). A new and flexible method for constructing designs for computer experiments. *Ann. Statist. 38*, 1460–1477.

[7] Lin, C.D., R. Mukerjee, and B. Tang (2009). Construction of orthogonal and nearly orthogonal Latin hypercubes. *Biometrika 96*, 243–247.

[8] Owen, A.B. (1994). Controlling correlations in Latin hypercube samples. *J. American Statist. Assoc. 89*, 1517–1522.

[9] Pang, F., M.-Q. Liu, and D.K.J. Lin (2009). A construction method for orthogonal Latin hypercube designs with prime power levels. *Statist. Sinica 19*, 1721–1728

[10] Steinberg M. and D.K.J. Lin (2006). A construction method for orthogonal Latin hypercube designs. *Biometrika 93*, 279–288.

[11] Sun, F., M.-Q. Liu, and D.K.J. Lin (2009). Construction of orthogonal Latin hypercube designs. *Biometrika 96*, 971–974.

[12] Tang, B. (1993). Orthogonal array-based Latin hypercubes. *J. American Statist. Assoc. 88*, 1392–1397.

[13] Tang, B. (1998). Selecting hypercubes using correlation criteria. *Statist. Sinica 8*, 965–977.

[14] Ye, K.Q. (1998). Orthogonal column Latin hypercubes and their application in computer experiments. *J. American Statist. Assoc. 93*, 1430–1439.

Q_B-optimal saturated two-level main effects designs

Steven Gilmour[1] **and Pi-Wen Tsai**[2]

[1] University of Southampton, UK
[2] National Taiwan Normal University, Taiwan

Abstract

We provide a general framework that incorporates experimenters' prior beliefs into the design selection process for the study of saturated two-level main effects designs, which are commonly used for screening experiments. We show that under the sets of priors with more weights on models of small size, p-efficient designs should be recommended; when models with more parameters are of interest, D-optimal designs would be better. Also, we present new classes designs which can be found between these two designs under different sets of priors. The way in which the choice of designs depends on experimenters' prior beliefs will be demonstrated for the cases when $N \equiv 2 \bmod 4$. Some constructions using conference matrices will also be discussed.

A comparison of logit and probit models for a binary response variable via a new way of data generalization

Özge Akkuş[1], Atilla Göktaş[1], and Selen Çakmakyapan[2]

[1] Muğla University, Turkey
[2] Istanbul Medeniyet University, Turkey

Abstract

Logit and probit models are two members of generalized linear models family that are widely used especially when the dependent variable is observed to be binary. The properties that make a difference for these two models for the same data set are resulted from the assumptions they use and their mathematical functions. There is no study specifying a certain judgment on the preference of these models to make a decision which model is better in what condition. In this study, a new data generalization technique has been proposed for the simulation study conducted to make a comparison of the model fits to binary logit and probit models for the generated data set under certain conditions to reach an end to which condition is better.
In the process of the simulation study, a dependent and explanatory variables are generated from multivariate normal distribution which is very much different from the ordinary generating procedure. As is already known, this procedure uses the information of the interested model itself. Hence the generation of this type would always be in favor of the interested model not the alternative and there would be no sense to make a comparison from such data generalization. In the proposed generating process since the generated dependent variable is always continuous, it should be classified as binary to make the dataset usable for logit and probit models. After fitting logit and probit models to the generated data sets, goodness-of-fit-test results related to both models, residuals, deviances and some pseudo R^2's used for binary dependent variables have been obtained to make significant comparisons. These procedures have been performed for two different cut points used to classify response variables, three different relationship levels among variables (high, medium, none) and five different sample sizes. For each cut point, relationship level and sample size the simulation has been replicated for a thousand time. Since the obtained estimated probabilities from both models are considerably close, it is found that there has been no statisticaly significant difference among most pseudo R^2's. However, when the residuals are taken into account, probit model has a priority to be used for a sample size that is less than 200, whereas the Logit model is superior for a sample

size that is greater than 200. Another remarkable finding is that the different cut-off levels and relationship have not any effect on the choice of the model.

Keywords

Binary logit, Binary probit, Pseudo R-square, Deviance.

References

[1] Agresti, A. (2002). *Catagorical Data Analysis* (Second Edition). Wiley, New Jersey.

[2] Aldrich, J.H. and F.D. Nelson (1984). *Linear Probability, Logit, and Probit Models.* (pp. 397–402). Sage Publications, London.

[3] Anderson-Sprecher, R. (1994). Model comparisons and R squares. *Amer. Statist. 48*, 113–117.

[4] Cameron, A.C. and A.G. Windmeijer (1997). An R-squared measure of goodness of fit for some common nonlinear regression models. *J. Econometrics 77*, 329–342.

[5] Cox, D.R. and N. Wermuth (1992). A comment on the coefficient of determination for binary responses. *Amer. Statist. 46*, 1–4.

[6] Hosmer, D.W., T. Hosmer, S. Le Cessie, and S. Lemeshow (1997). A comparison of goodness-of-fit tests for the logistic regression model. *Stat. Med. 16*, 965–980.

[7] Uçar, Ö. (2004). Nitel Verilerin Analizinde Lojit ve Probit Modeller, *Yüksek-lisans Tezi, Hacettepe Üniversitesi*, Ankara-2004.

[8] Veall, M.R. and K.F. Zimmermann (1994). Evaluating pseudo-R^2's for binary probit models. *Quality & Quantity 28*, 151–164.

[9] Veall, M.R. and KF. Zimmermann (1996). Pseudo-R^2's measures for some common limited dependent variable models. *Sunderforschungsbereich 386*, 1–34.

[10] Windmeijer, F.A.G. (1995). Goodness of fit measures in binary choice models. *Econometric Rev. 14*, 101–116.

[11] Winkelmann, R. and S. Boes (2006). *Analysis of Microdata.* Springer, Berlin.

[12] Zelner, B.A. (2008). Using simulation to interpret and present logit and probit results. *Working paper.*

First and second derivative in time series classification using DTW

Tomasz Górecki[1] and Maciej Łuczak[2]

[1] Adam Mickiewicz University, Poznań, Poland
[2] Koszalin University of Technology, Poland

Abstract

In our previous work [2] we developed some new distance function based on a derivative and showed that our algorithm is efficient. In contrast to well-known measures from the literature, our approach considers the general shape of a time series rather than point-to-point function comparison. The new distance was used in classification with the nearest neighbor rule. Now, we improve on our previous technique adding second derivative. In order to provide a comprehensive comparison, we conducted a set of experiments, testing effectiveness on 20 time series data sets from a wide variety of application domains. Our experiments show that our method provides a significant higher quality of classification.

Keywords

Dynamic time warping, Derivative dynamic time warping, Data mining, Time series.

References

[1] Ding, H., G. Trajcevski, P. Scheuermann, X. Wang, and E. Keogh (2008). Querying and Mining of Time Series Data: Experimental Comparison of Representations and Distance Measures. In: *Proc. 34th Int. Conf. on Very Large Data Bases*, 1542–1552.

[2] Górecki, T. and M. Łuczak (2012). Using derivatives in time series classification. *Data Min. Knowl. Discov.* DOI: 10.1007/s10618-012-0251-4.

[3] Keogh, E. and M. Pazzani (2001). Dynamic Time Warping with Higher Order Features. In: *First SIAM International Conference on Data Mining (SDM'2001), Chicago, USA*.

A study on the equivalence of BLUEs under a general linear model and its transformed models

Nesrin Güler

Sakarya University, Turkey

Abstract

The general linear model $\mathcal{A} = \{y, X\beta, \sigma^2 V\}$ known as full model and its transformed model $\mathcal{T} = \{Fy, FX\beta, \sigma^2 FVF'\}$ are considered. The expression for the difference between the best linear unbiased estimator (BLUE) of $FX\beta$ under the full model and its BLUE under the transformed model is given. The necessary and sufficient conditions between the equality of BLUEs of $FX\beta$ are obtained under the full and transformed models. Furthermore, some results are given for the special choices of the transformation matrix F. The results obtained in this study are based on a generalized inverse of a symmetric matrix which is obtained from the Pandora's Box equation called by [10].

Keywords

BLUE, General linear model, Transformed models, Sub-sample models, Reduced models.

References

[1] Alalouf, I.S. and G.P.H. Styan (1979). Characterizations of estimability in the general linear model. *Ann. Statist. 7*, 194–200.

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

[3] Baksalary, J.K. and T. Mathew (1990). Rank invariance criterion and its application to the unified theory of least squares. *Linear Algebra Appl. 127*, 393–401.

[4] Farebrother, R.W. (1979). Estimation with aggregated data. *J. Econometrics 10*, 43–55.

[5] Gross, J. and S. Puntanen (2000). Estimation under a general partitioned linear model. *Linear Algebra Appl. 321*, 131–144.

[6] Gross, J., G. Trenkler, and H.J. Werner (2001). The equality of linear transformations of the ordinary least squares estimator and the best linear unbiased estimator. *Sankhya A 63*, 118–127.

[7] Hall, F.J. and C.D.Jr. Meyer (1975). Generalized inverses of the fundamental bordered matrix used in linear estimation. *Sankhya A 37*, 428–438. [Corrigendum (1978), 40, 399.]

[8] Isotalo, J., S. Puntanen, and G.P.H Styan (2008). A useful matrix decomposition and its statistical applications in linear regression. *Comm. Statist. Theory Methods 37*, 1436–1457.

[9] Lucke, B. (1991). On BLU-estimation with data of different periodicity. *Econom. Lett. 35*, 173–177.

[10] Rao, C.R. (1971). Unified theory of linear estimation. *Sankhya A 33*, 371–394. [Corrigendum (1972), 34, 194, 477.]

[11] Rao, C.R. (1972). A note on the IPM method in the unified theory of linear estimation. *Sankhya A 34*, 285–288.

[12] Rao, C.R. (1973). Representations of best linear unbiased estimators in the Gauss-Markov model with a singular dispersion matrix. *J. Multivariate Anal. 3*, 276–292.

[13] Tian, Y., M. Beisiegel, E. Dagenais, and C. Haines (2008). On the natural restrictions in the singular Gauss-Markov model. *Statist. Papers 49*, 553–564.

[14] Tian, Y. and S. Puntanen (2009). On the equivalence of estimations under a general linear model and its transformed models. *Linear Algebra Appl. 430*, 2622–2641.

[15] Tian, Y. and J. Zhang. (2011). Some equalities for estimations of partial coefficients under a general linear regression model. *Statist. Papers 52*, 911–920.

[16] Zhang, B. (2007). The BLUE and MINQUE in Gauss-Markoff model with linear transformation of the observable variables. *Acta Math. Sci. Ser. B 27*, 203–210.

[17] Zhang, B., B. Liu, and C. Lu (2004). A study of the equivalence of the BLUEs between a partitioned singular linear model and its reduced singular linear models. *Acta Math. Sin. (Engl. Ser.) 20(3)*, 557–568.

Improved estimation of the mean by using coefficient of variation as a prior information in ranked set sampling

Duygu Haki[1], **Özlem Ege Oruç**[2], **and Müjgan Tez**[1]

[1] Marmara University, Istanbul, Turkey
[2] Dokuz Eylül University, Izmir, Turkey

Abstract

Estimation of population parameters is considered by several statisticians when additional information such as coefficient of variation, kurtosis or skewness is known. This estimation technique is called improved estimation. Searls (1964), Khan (1968) and Arnholt and Hebert (1995) utilized the known coefficient of variation on improved estimating the population mean.
RSS, which has been developed by McIntyre (1952), is a sampling procedure that can be viewed as a generalization of the simple random sample (SRS). This method is applied for situations in which measuring a variable is costly or difficult, but where ranking in small subsets is easy. As it was proved by McIntyre, mean of this sample is an unbiased estimator of the population mean. Additionally, it is well-known that population parameters can be estimated more efficiently using a RSS as opposed to a SRS. This paper is concerned with the improved estimation of the population mean by using coefficient of variation as a prior information in ranked set sampling (RSS). Compare it with the estimator of the mean in RSS, the estimator of the mean in Simple Random Sampling (SRS) and improved estimator of the mean in Simple Random Sampling (SRS) in the sense of Mean Square Errors (MSE). It is observed that the proposed RSS estimator is more efficient than others.

Keywords

Ranked set sample, Improved estimation, Efficiency, Mean squared error, Coefficient of variation.

References

[1] Arnholt, A.T. and J.L. Hebert (1995). Estimating the mean with known coefficient of variation. *Amer. Statist. 49*, 367–369.
[2] Khan, R.A. (1968). A note on estimating the mean of a normal distribution with known coefficient of variation. *J. Amer. Statist. Assoc. 63*, 1039–1041.
[3] McIntyre, G.A. (1952). A method for unbiased selective sampling using ranked sets. *Aust. J. Agric. Res. 3*, 385–390.

[4] Patil, G.P. (1995). Editorial: Ranked set sampling. *Environ. Ecol. Stat. 2*, 271–285.

[5] Searls, D.T. (1964). The utilization of a known coefficient of variation in the estimation procedure. *J. Amer. Statist. Assoc. 59*, 1225–1226.

[6] Takahasi, K. and K. Wakimoto (1968). On unbiased estimates of the population mean based on the sample stratified by means of ordering. *Annals of the Institute of Statistical Mathematics, 20*, 1–31.

[7] Wolfe, D.A. *Ranked Set Sampling: An approach to more efficient data collection.* Department of Statistics Ohio State University. http://www.stat.osu.edu./ amd/papers/wolfe-ranked-set-sampling.pdf

Simulation study on improved Shapiro-Wilk test of normality

Zofia Hanusz and **Joanna Tarasińska**

University of Life Sciences in Lublin, Poland

Abstract

The W statistic proposed by Shapiro and Wilk ([2]) is frequently used for testing of the univariate and multivariate normality. However, the table of coefficients in the W statistic, its critical values and also constants in Johnson's S_B transformation to normal distribution ([3]), are not correct. Royston ([1]) gave an approximation for coefficients in the W statistic and use them to evaluate proper critical values of the Shapiro-Wilk test. In the paper, we determine new constants for the W statistic and Johnson's S_B transformation. Empirical significant levels of the improved Shapiro-Wilk test and the power against chosen alternatives are evaluated via simulation study.

Keywords

Multivariate normality, Empirical significant level, Power of the test.

References

[1] Royston, P. (1992). Approximating the Shapiro-Wilk W-test for non-normality. *Stat. Comput. 2*, 117–119.

[2] Shapiro, S.S. and M.B. Wilk (1965). An analysis of variance test for normality (complete samples). *Biometrika 52*, 591–611.

[3] Shapiro, S.S. and M.B. Wilk (1968). Approximations for the null distribution of the W statistic. *Technometrics 10*, 861–866.

Equivalence of linear models under changes to data, design matrix, or covariance structure

Stephen J. Haslett

Massey University, Palmerston North, New Zealand

Abstract

For the mixed linear model, there is a collection of results giving conditions under which fixed parameter estimates, and/or random parameter predictors remain unchanged. Some of these results were initially developed for models with only fixed parameters, others include situations where at least some parameters are random. These equivalence results cover a range of situations - the covariance structure of error processes, design matrices, and even data may be altered. Covariance structure changes have a broad range, from conditions under which ordinary least squares estimates (OLSE) are best linear unbiased estimates (BLUE) ([9], [2]), to conditions for two sets of BLUEs and/or two sets of BLUPs to be equivalent ([10], [11], [1], [6], [7]). Changes in design structure link to adding or deleting regressors or parameters ([5]). Data changes are related to data cloning techniques ([3]), and to adding new observations ([8], [4]). These types of model modification will be discussed and various possible applications will be outlined.

References

[1] Baksalary, J.K. and T. Mathew T. (1990). Rank invariance criterion and its application to the unified theory of least squares. *Linear Algebra Appl. 127*, 393–401.

[2] Baksalary, J.K., S. Puntanen, and G.P.H. Styan (1990). On T. W. Anderson's contributions to solving the problem of when the ordinary least-squares estimator is best linear unbiased and to characterizing rank additivity of matrices. In: G.P.H. Styan (Ed.), *The Collected Papers of T. W. Anderson: 1943-1985* (pp. 1579–1591). Wiley, New York.

[3] Haslett, S. and K. Govindaraju (2012). Data cloning: data visualisation, smoothing, confidentiality, and encryption. *J. Statist. Plann. Inference 142(2)*, 410–422.

[4] Haslett, S. and S. Puntanen (2010a). A note on the equality of BLUPs for new observations under two linear models. *Acta Comment. Univ. Tartu. Math. 14*, 27–33.

[5] Haslett, S. and S. Puntanen (2010b). Effect of adding regressors on the equality of the BLUEs under two linear models. *J. Statist. Plann. Inference 140*, 104–110.

[6] . Haslett, S. and S. Puntanen (2010c). Equality of the BLUEs and/or BLUPs under two linear models using stochastic restrictions. *Statist. Papers 51*, 465–475.

[7] Haslett, S. and S. Puntanen (2011). On the equality of BLUPs under two linear mixed models. *Metrika 74(3)*, 381–395.

[8] Isotalo, J. and S. Puntanen (2006). Linear prediction sufficiency for new observations in the general GaussŰMarkov model. *Comm. Statist. Theory Methods 35*, 1011–1023.

[9] Puntanen, S. and G.P.H. Styan (1989). The equality of the ordinary least squares estimator and the best linear unbiased estimator (with discussion). *Amer. Statist. 43*, 151–161. [Commented by Oscar Kempthorne on pp. 161Ű162 and by Shayle R. Searle on pp. 162Ű163, Reply by the authors on p. 164].

[10] Rao, C.R. (1968). A note on a previous lemma in the theory of least squares and some further results *Sankhya A 30*, 245–252.

[11] Rao, C.R. (1973). Representations of best linear unbiased estimators in the Gauss-Markov model with singular dispersion matrix. *J. Multivariate Anal. 3*, 276–292.

Nonnegativity of eigenvalues of sum of diagonalizable matrices

Charles R. Johnson[1], Jan Hauke[2], and Tomasz Kossowski[2]

[1] College of William and Mary, Williamsburg, USA
[2] Adam Mickiewicz University, Poznań, Poland

Abstract

Properties of eigenvalues of matrices used in statistical analysis provide an important base in the description of statistical properties within analyzed problem, see e.g., [1] and [2]. The paper extends some characterizations of diagonalizable matrices whose sum has nonnegative eigenvalues. In the paper there are presented some general comments as well as examples of matrices from specific subsets.

Keywords

Nonnegative eigenvalues, Diagonalizable matrices.

References

[1] Baksalary, J.K., J. Hauke, and G.P.H. Styan (1994). On some distributional properties of quadratic forms in normal variables and some associated matrix partial orderings. In: *Proceedings of the International Symposium on Multivariate Analysis and Its Applications, Hong Kong, March 1992.* IMS Lecture Notes - Monograph Series 24, 111–121.

[2] Puntanen, S., G.P.H. Styan, and J. Isotalo (2012). *Matrix Tricks for Linear Statistical Models: Our Personal Top Twenty.* Springer.

Modeling multiple time series data using wavelet-based support vector regression

Deniz İnan and **Birsen Eygi Erdogan**

Marmara University, Istanbul Turkey

Abstract

In recent years, support Vector Regression (SVR) has been applied in various fields such as financial time series prediction and engineering applications. Different from the classical regression approach, SVR attempts to minimize the generalization error bound instead of minimizing the observed training error. This paper deals with the application of wavelet-based support vector regression (WSVR) on multiple time series data. WSVR is the straightforward extension from linear regression to nonlinear regression using the wavelet kernel. The main objective of this paper is to examine the feasibility of WSVR in time series forecasting by comparing it with generalized least squares (GLS) approach.

Keywords

Support vector regression, Wavelet kernel, Generalized least squares, Time series.

References

[1] Hamel, L.(2009). *Knowledge Discovery with Support Vector Machines.* Wiley.

[2] Makridakis, S., S.C. Wheelwright, and R.J. Hyndman (1998). *Forecasting Methods and Applications (Third Edition).* Wiley.

[3] Palancz, B., L. Völgyesi, and G. Popper (2005). Support Vector Regression Via Mathematica. *Periodica PolyTechnica Civ. Eng. 49*, 59–84.

[4] Myers, R.H., D.C. Montgomery, and G.G. Vinning (2002). *Generalized Linear Models.* Wiley.

[5] Tay, F.E.H. and L. Cao (2001). Application of support vector machines in financial time series forecasting. *Omega: Int. J. Management Sci. 29*, 309–317.

[6] Wang, L. and X. Jin (2011). Stock exchange index prediction based on wavelet-based adaptive support vector regression algorithm. *J. Inf. Comput. Sci. 8*, 4053–4059.

Simultaneous fixed and random effect selection in finite mixture of linear mixed-effect models

Abbas Khalili, Yeting Du, Russell Steele, and Johanna Neslehova

McGill University, Montreal, Canada

Abstract

Linear mixed-effects (LME) models are frequently employed for modeling longitudinal data. One complicating factor in the analysis of such data is that samples are sometimes obtained from a population with significant underlying heterogeneity, which would be hard to capture by a single LME model. Such problems may be addressed by a finite mixture of linear mixed-effects (FMLME) models, which segments the population into subpopulations and models each subpopulation by a distinct LME model. Often in the initial stage of a study, a large number of covariates are introduced. However, their associations to the response variable vary from one component to another of the FMLME model. To enhance predictability and to obtain a parsimonious model, it is of great practical interest to identify the important effects, both fixed and random, in the model. Traditional variable selection techniques such as stepwise deletion and subset selection are computationally expensive as the number of covariates and components in the mixture model increases. In this article, we introduce a new penalized likelihood approach for simultaneous selection of fixed and random effects in FMLME models. We also propose a nested EM algorithm for efficient numerical computations. The estimators are shown to possess consistency and sparsity properties and asymptotic normality. We illustrate the performance of our method through simulations and a real data example.

Keywords

Linear mixed-effect models, Mixture models, Regularization methods, EM algorithm.

Estimators of serial covariance parameters in multivariate linear models

Daniel Klein and **Ivan Žežula**

Šafárik University, Košice, Slovakia

Abstract

The basic model we consider is the multivariate linear model with serial correlation structure:

$$Y = XB + \mathbf{e}, \quad \operatorname{vec}(\mathbf{e}) \sim N\left(0, \Sigma \otimes I_n\right), \quad \Sigma = \sigma^2 \rho^{|i-j|}.$$

Here $Y_{n\times p}$ is a matrix of independent p-variate observations, $X_{n\times m}$ is a design matrix and $\mathbf{e}_{n\times p}$ is a matrix of random errors. As for the unknown parameters, $B_{m\times r}$ is an location parameters matrix, and σ^2, ρ are (scalar) covariance parameters.
Our aim is to estimate the unknown parameters of matrix Σ. We propose a method for obtaining explicit estimators of both σ^2 and ρ and we discuss some properties of the derived estimators.

Keywords

Multivariate linear model, Serial structure, Explicit estimators.

References

[1] Kollo, T. and D. von Rosen (2005). *Advanced Multivariate Statistics with Matrices.* Springer, Dordrecht, The Netherlands.
[2] Lee, J.C. (1988). Prediction and estimation of growth curves with special covariance structures. *J. Amer. Statist. Assoc. 83*, 432–440.
[3] Ye, R.D. and S.G. Wang (2009). Estimating parameters in extended growth curve models with special covariance structures. *J. Statist. Plann. Inference 139*, 2746–2756.
[4] Žežula, I. (2006). Special variance structures in the growth curve model. *J. Multivariate Anal. 97*, 606–618.

Robust monitoring of multivariate data stream

Daniel Kosiorowski, Małgorzata Snarska, and Oskar Knapik

Cracow University of Economics, Poland

Abstract

A data stream could be defined as continuous sequence of ordered observations of indeterminate length. Because the data are arriving continuously and there is no known end to it, the classical approach of reading in all data and then processing them is not feasible. Data stream carry signals that appear randomly, are irregularly spaced and the time duration between successive signals is not deterministic but random. Additionally data streams generally are generated by multivariate non-stationary models of unknown form.
In this paper we present three approaches to robust analysis of multivariate economic data streams. We use the information-theoretic approach proposed by [1] based on the relative Kullback-Leibler entropy and bootstrapping to extract possible changes and Kulldorff's spatial scan statistics to identify regions where large changes have occurred. Second approach is a modification of a concept of a statistical coherence (frequency domain approach) for regression analysis of time series. It is shown how spectral coherence can be used to examine the relation between two signals and to detect the change in these relationship. Our proposal is to estimate spectral coherence using robust version of Welch approach. In our third proposition appealing to a data depth concept we use multivariate Wilcoxon statistic and robust semi-parametric regression to monitor linear relationship between multivariate data stream components.

Keywords

Data stream, Robust spectral analysis, Depth function, Relative entropy.

References

[1] Dasu, T., S. Krishnan, S. Venkatasubramanian, and K. Yi (2006). An information-theoretic approach to detecting changes in multi-dimensional data streams. In: *Proceedings of the 38th Symposium on the Interface of Statistics, Computing Science, and Applications (Interface '06), Pasadena, CA.*

[2] Wellman, R. and Ch.H. Müller (2010). Depth notions for orthogonal regression. *J. Multivariate Anal. 101*, 2358–2371.

The Moran coefficient for non-normal data: revisited with some extensions

Daniel A. Griffith[1], Jan Hauke[2], and Tomasz Kossowski[2]

[1] University of Texas, Dallas, USA
[2] Adam Mickiewicz University, Poznań, Poland

Abstract

The distributional properties of the Moran coefficient index (MC) measuring spatial autocorrelation were investigated by many authors, see e.g. [1]. The properties of MC for non-normal random variables were analysed by Griffith in [2]. The general idea of that paper was to extend Pitman-Koopmans theorem for the mean and the variance of this index. The principal conclusion was that under independence assumption and big enough sample size the Pitman-Koopmans theorem results can be extended to some non-normal random variables. The independence and identically distributed property reduced the necessary sample size for this extension, as did the properties of symmetry and normal approximation.
In the paper we continue the analysis performing simulations for randomly generated variables for the following distributions: beta, gamma, hypergeometric, inverse hypergeometric, log-normal, exponential, negative binomial, and t Student, as well as their mixtures and using Box-Cox power transformation.

Keywords

Moran coefficient, Normality.

References

[1] Griffith, D.A. (2003). *Spatial Autocorrelation and Spatial Filtering.* Springer.
[2] Griffith, D.A. (2010). The Moran coefficient for non-normal data. *J. Statist. Plann. Inference 140*, 2980–2990.

Optimal designs for the Michaelis Menten model with correlated observations

Holger Dette[1] and Joachim Kunert[2]

[1] Ruhr University Bochum, Germany
[2] Technical University of Dortmund, Germany

Abstract

In this paper we investigate the problem of designing experiments for weighted least squares analysis in the Michaelis Menten model. We study the structure of exact D-optimal designs in a model with an autoregressive error structure. Explicit results for locally D-optimal are derived for the case where 2 observations can be taken per subject. Additionally standardized maximin D-optimal designs are obtained in this case. The results illustrate the enormous difficulties to find exact optimal designs explicitly for nonlinear regression models with correlated observations.

Keywords

Autoregressive errors, Michaelis Menten model, Exact designs, Locally D-optimal designs, Standardized maximin optimal design.

References

[1] Dette, H. and S. Biedermann (2003). Robust and efficient designs for the Michaelis-Menten model. *J. Amer. Statist. Assoc. 98*, 679–686.

[2] Dette, H., J. Kunert, and A. Pepelyshev (2008). Exact optimal designs for weighted least squares analysis with correlated errors. *Statist. Sinica 18*, 135–154.

[3] Pepelyshev, A. (2007). Optimal Designs for the Exponential Model with Correlated Observations. In: J. Lopez-Fidalgo, J.M. Rodriguez-Diaz, B. Torsney (Eds.), *MODA 8, Advances in Model-Oriented Design and Analysis* (pp. 165–172), Physica Verlag.

A new Liu-Type Estimator

Fatma S. Kurnaz[1] **and Kadri U. Akay**[2]

[1] Karadeniz Technical University, Trabzon, Turkey
[2] University of Istanbul, Turkey

Abstract

Ridge regression (RR) and Liu estimators, which include single biasing parameter, specially depend on ordinary least squares (OLS) estimator. Due to the effects of multicollinearity on the OLS estimator, it have recently been proposed biased estimators include the two biasing parameter. Estimating biasing parameters of estimators include two biasing parameters are usually based on the methods proposed to Ridge and Liu estimators. But, very complicated equations may occur, when these methods are applied to estimators proposed. In this paper, we introduce a general new Liu-type estimator includes estimators with two biasing parameters as special cases. Also, necessary and sufficient conditions according to the mean squared error matrix criterion are derived, to show the superiority of the new estimator over the OLS, RR, Liu estimator, and the other estimators which include two biasing parameters. Lastly, the superiority to other estimators of the new Liu-type estimator is illustrated both theoretically and graphically on dataset Portland cement is widely used in the literature.

Keywords

Biased regression, Mean squared error, Multicollinearity, Ridge Regression, Liu estimator

References

[1] Liu, K. (2003). Using Liu-Type Estimator to Combat Collinearity. *Comm. Statist. Theory Methods 32*, 1009–1020.

[2] Özkale, M. R. and S. Kaçıranlar (2007). The Restricted and Unrestricted Two-Parameter Estimators. *Comm. Statist. Theory Methods 36*, 2707–2725.

[3] Sakallıoğlu S. and S. Kaçıranlar (2008). A New Biased Estimator based on ridge estimation. *Statist. Papers 49*, 669–689.

[4] Yang, H. and X. Chang (2010). A New Two-Parameter Estimator in Linear Regression. *Comm. Statist. Theory Methods 39*, 923–934.

Analysis of an experiment in a generally balanced nested block design

Agnieszka Łacka

Poznań University of Life Sciences, Poland

Abstract

The aim of the study is to present practical aspects related to the analysis of an experiment carried out in a nested block design. The considered experimental design is characterized by the orthogonal block structure and has the property of general balance. Analysis of the experiment was based on the theorem presented by prof. Tadeusz Caliński in his paper: "On combining information in a generally balanced nested block design" and includes both stratum analysis based on basic contrasts and combined analysis allowing for combining data from a number of strata.
The experiment data used in the presented example originate from the experiment concerning evaluation of efficiency of some chemical substances in various concentrations in reduction of plant damages caused by slugs, *A. Lusitanicus*, and studying their influence on behavioral and physiological reactions of these slugs. The calculations were made with the use of the R platform.

Keywords

Nested block design, Stratum analysis, Combining information.

References

[1] Caliński, T. and S. Kageyama (2000). *Block Designs: A Randomization Approach, Volume I: Analysis.* Lecture Notes in Statistics, Vol. 150. Springer, New York.

[2] Łacka, A., M. Kozłowska, and J. Kozłowski (2009). Some optimal block designs with nested rows and columns for research on alternative methods of limiting slug damage. *Statist. Papers 50(4)*, 837–846.

[3] Nelder, J.A. (1965). The analysis of randomized experiments with orthogonal block structure. *Proc. Roy. Soc. London Ser. A 283*, 147–178.

[4] Nelder, J.A. (1968). The combination of information in generally balanced designs. *J. Roy. Statist. Soc. Ser. B 30*, 303–311.

On inverse prediction in mixed models

Lynn R. LaMotte

LSUHSC School of Public Health, New Orleans, USA

Abstract

This talk will present a general approach to inverse prediction in the context of mixed models. Given training data on Y at x and covariates z, and a mystery specimen with $Y = y_*$ and $Z = z_*$, the objective is to construct a confidence set on the subject's unknown x_*. Simulation results will be presented for three different settings: (1) heteroscedastic linear regression, (2) classification, in which x is categorical, and (3) categorical response Y.

Keywords

Multivariate calibration, Categorical response.

Getting the "correct" answer from survey responses: an application of regression mixture models

Nicholas Fisher[1] **and** **Alan Lee**[2]

[1] University of Sydney, Australia
[2] University of Auckland, New Zealand

Abstract

This talk addresses a problem that can arise in surveys, in which some respondents misinterpret the rating method and so assign high ratings when they intended to assign low ratings, and vice versa. We present a method, based on fitting regression mixture models, that allows these misinterpretations to be corrected with high probability, and more meaningful conclusions drawn. The method is illustrated with data from a community value survey.

Keywords

Community Value Survey, Missing data, EM algorithm, Regression mixture.

Variance components estimability in multilevel models with block circular symmetric covariance structure

Yuli Liang[1], **Tatjana von Rosen**[1], **and Dietrich von Rosen**[2,3]

[1] Stockholm University, Sweden
[2] Swedish University of Agricultural Sciences, Uppsala, Sweden
[3] Linköping University, Sweden

Abstract

The multilevel model with the block circular symmetric covariance structure is considered. We established the spectral properties of this patterned covariance matrix. It has been shown that the explicit maximum likelihood estimators (MLEs) of variance-covariance components do not exist in this model, unless we put restrictions on the parameter space.
It is shown that by putting restrictions on the spectrum of the block circular covariance matrices, some natural reparameterization conditions (e.g. sum-to-zero) are derived. Sufficient conditions of obtaining explicit estimators for variance-covariance components are presented. Different restricted models are discussed in order to obtain explicit estimators, get interpretable model reparameterizations and keep invariant properties of the block circular symmetric covariance structure.
In the class of restricted models, it gives us the flexibility to choose the reasonable constraints among them according to different data, which is quite advantageous.

Keywords

Circular block symmetry, restricted model, explicit maximum likelihood estimator, variance components.

References

[1] Liang, YL., D. von Rosen, and T. von Rosen (2011). On estimation in multilevel models with block circular symmetric covariance structure. *Acta Comment. Univ. Tartu. Math.* Provisionally accepted.
[2] Nahtman, T. (2006). Marginal permutation invariant covariance matrices with applications to linear models. *Linear Algebra Appl. 417*, 183–210.
[3] Nahtman, T. and D. von Rosen (2008). Shift permutation invariance in linear random factor models. *Math. Meth. Statist. 17*, 173–185.

[4] Olkin, I. and S.J. Press (1969). Testing and estimation for a circular stationary model. *Ann. Math. Statist. 40*, 1358–1373.
[5] Olkin, I. (1972). Testing and estimation for structures which are circularly symmetric in blocks. In: D.G. Kabe, R.P. Gupta (Eds.), *Multivariate Statistical Inference* (pp. 183–195). North-Holland, Amsterdam.
[6] Perlman, M.D. (1987). Group symmetry covariance models. *Statist. Sci. 2*, 421–425.
[7] Szatrowski, T.H. and J.J. Miller (1980). Explicit maximum likelihood estimates from balanced data in the mixed model of the analysis of variance. *Ann. Statist. 8*, 811–819.
[8] Szatrowski, T.H. (1980). Necessary and sufficient conditions for explicit solutions in the multivariate normal estimation problem for patterned means and covariances. *Ann. Statist. 8*, 802–810.

Model averaging via penalized least squares in linear regression

Antti Liski[1] and Erkki P. Liski[2]

[1] Tampere University of Technology, Finland
[2] University of Tampere, Finland

Abstract

We consider parameter estimation under model uncertainty by averaging across least squares estimates obtained from a set of models. Existing model averaging methods usually require estimation of a single weight for each candidate model. However, in applications the number of candidate models may be huge. Then the approach based on estimation of single weights becomes computationally infeasible. Utilizing a connection between shrinkage estimation and model weighting we present an accurate and computationally efficient model averaging estimation method. The performance of our estimators is displayed in simulation experiments which utilize a realistic set up based on real data.

Keywords

Shrinkage estimation, Model selection, Mean square error, Efficiency bound, Simulation experiment.

References

[1] Hansen, B.E. (2007). Least squares model averaging. *Econometrika 75*, 1175–1189.

[2] Magnus, J.R., O. Powell, and P. Prüfer (2010). A comparison of two model averaging techniques with an application to growth empirics. *J. Econometrics 154*, 139–153.

[3] Sund, R., M. Juntunen, P. Lüthje, T. Huusko, M. Mäkelä, M. Linna, A. Liski, and U. Häkkinen (2006). *PERFECT - Hip Fracture, Performance, Effectiveness and Cost of Hip Fracture Treatment Episodes* (In Finnish), National Research and Development Centre for Welfare and Health, Helsinki.

Optimality of neighbor designs

Augustyn Markiewicz

Poznań University of Life Sciences, Poland

Abstract

The concept of neighbor designs was introduced and defined in [5] along with some methods of their construction. Henceforth many methods of construction of neighbor designs as well as of their generalizations are available in the literature; cf. [3] and [4]. However there are only few results on their statistical properties. Therefore the aim of the talk is to give an overview of study on their optimality. It will include recent results on optimality of some neighbor designs under various linear models; cf. [1] and [2].

Keywords

Neighbor design, Circular block design, Universal optimality, Interference model.

References

[1] Filipiak, K. (2012). Universally optimal designs under an interference model with equal left- and right-neighbor effects. *Statist. Probab. Lett. 82*, 592–598.

[2] Filipiak, K. and A. Markiewicz (2012). On universal optimality of circular weakly neighbor balanced designs under an interference model. *Comm. Statist. Theory Methods 41*, 2356–2366.

[3] Hwang, F.K. (1973). Constructions for some classes of neighbor designs, *Ann. Statist. 1*, 786–790.

[4] Hwang, F.K. and S. Lin (1977). Neighbor designs. *J. Combin. Theory 23*, 302–313.

[5] Rees, D.H. (1967). Some designs of use in serology. *Biometrics 23*, 779–791.

About the evolution of the genomic diversity in a population reproducing through partial asexuality

Solenn Stoeckel and Jean-Pierre Masson

INRA BIO3P Rennes, France

Abstract

Reproductive systems define how the genetic diversity is transmitted through generations and thus highly constraint the genetic evolution of species. Many species of relevant interests for human activities and ecosystems can reproduce both through sexual or asexual events during their life. Despite their widespread interests, we have few tools to predict the evolution of the genetic diversity within those partially asexual species. Moreover, the scarce previous models propose contradictory or unclear results. We thus formalized the exact probabilities of evolving genotypic states through generations using transition probabilities as function of the rate of asexuality and embedded them within a Markov chain. The model takes into account for mutation and drift forces, giving the opportunity to assess the distributions of any expected genetic index at a locus that did not experiment selection. Such model computation relies on fat matrices because of the number of genotypic states. We used massive parallelized algorithm to compute them. It provided unseen results, enabled precise predictions and clarified some controversial biological points.

Keywords

Markov chain, Rate of clonality, Mutation, Genetic drift, Matrix calculus, Maximum likelihood.

References

[1] Ewens, W.J. (2004). *Mathematical Population Genetics. I. Theoretical Introduction* (2nd ed). Springer.

[2] Feller, W. (1951). Diffusion processes in Genetics. In: J. Neyman (Ed.), *2nd Berkeley Symp. on Math. Stat. and Prob.* (pp. 227–246). Berkeley: University of California Press.

[3] Oksendal, B. (1998). *Stochastic Differential Equations: An Introduction with Applications* (5th ed). Springer.

A sequential generalized DKL-optimum design for model selection and parameter estimation in non-linear nested models

Caterina May[1] **and Chiara Tommasi**[2]

[1] University of Eastern Piedmont, Novara, Italy
[2] University of Milan, Italy

Abstract

A sequential procedure is proposed to select the best model among several nested non-linear models and to estimate efficiently the parameters of the chosen model. The procedure is based on an adaptive generalized DKL-optimum design, which is optimal for the double goal of model selection and parameter estimation. The proposed sequential scheme selects the best non-linear model with probability converging to one; moreover it estimates efficiently its parameters, since the adaptive sequential DKL-optimum designs converge to the D-optimum design for the "true" model. These results are proved by means of asymptotic theory arguments for argmin of convex random functions.

Keywords

DKL-optimality, Sequential design of experiments, Stochastic convergence, Semi-continuity, Argmin processes, Convexity.

Two-stage optimal designs in nonlinear mixed effect models: application to pharmacokinetics in children

Cyrielle Dumont[1,2], Marylore Chenel[2], and France Mentré[1]

[1] University Paris Diderot, Paris, France
[2] Institut de Recherches Internationales Servier, Suresnes, France

Abstract

Nonlinear mixed effect models (NLMEM) are used in pharmacometrics to analyse longitudinal data through models. Approaches based on the Fisher information matrix (M_F) can be used to optimise the design of these studies. A first-order linearization of the model was proposed to evaluate MF for these models [7] and is implemented in the R function PFIM [1]. Local optimal design needs some *a priori* values of the parameters which might be difficult to guess. Adaptive designs are useful to provide some exibility and were applied in pharmacometrics [6,9]. However, two articles in other contexts [2,5] discussed that two-stage designs could be more efficient than fully adaptive designs. Moreover, two-stage designs are easier to implement in clinical practice. We implemented in a working version of PFIM the optimisation of the determinant of M_F for two-stage designs in NLMEM. We evaluated the approach by simulation. The example concerns a drug in development for which a pharmacokinetic study in children is needed and will be analysed through NLMEM as recommended [4,8]. For the first stage, parameters were estimated using predictions from pharmaco-chemical properties of the drug [3]. We evaluated one and two-stage designs assuming that some parameter(s) is (are) different than the initial one(s). We evaluated the impact of the size of each cohort on the precision of population parameters estimation.

Keywords

Adaptive design, Design optimisation, Fisher information matrix, Nonlinear mixed effect models, PFIM, Population pharmacokinetics.

References

[1] Bazzoli, C., S. Retout, and F. Mentré (2010). Design evaluation and optimisation in multiple response nonlinear mixed effect models: PFIM 3.0. *Comput. Methods and Programs Biomed. 98(1)*, 55–65.

[2] Chen, T.T. (1997). Optimal three-stage designs for phase II cancer clinical trials. *Stat. Med. 16(23)*, 2701–2711.

[3] Dumont, C., M. Chenel, and F. Mentré (2011). Design optimisation of a pharmacokinetic study in the paediatric development of a drug. *Population Approach Group in Europe*, Abstr 2160 [www.pagemeeting.org/?abstract=2160].

[4] EMEA (2006). Guideline on the role of pharmacokinetics in the development of medicinal products in the paediatric population. *Scientific Guideline*. 2006.

[5] Federov, V., Y. Wu, and R. Zhang (2012). Optimal dose-finding designs with correlated continuous and discrete responses. *Stat. Med. 31(3)*, 217–234.

[6] Foo, L.K. and S. Duffull (2012). Adaptive optimal design for bridging studies with an application to population pharmacokinetic studies. *Pharmac. Res. 29(6)*, 1530–1543.

[7] Mentré, F., A. Mallet, and D. Baccar (1997). Optimal design in random-effects regression models. *Biometrika 84(2)*, 429–442.

[8] Tod, M., V. Jullien, and G. Pons (2008). Facilitation of drug evaluation in children by population methods and modelling. *Clinical Pharmacokinetics 47(4)*, 231–243.

[9] Zamuner, S., V.L. Di Iorio, J. Nyberg, R.N. Gunn, V.J. Cunningham, R. Gomeni, and A.C. Hooker (2010). Adaptive-optimal design in PET occupancy studies. *Clinical Pharmacology & Therapeutics 87(5)*, 563–571.

On admissibility of decision rules derived from submodels in two variance components model

Andrzej Michalski

Wrocław University of Environmental and Life Sciences, Poland

Abstract

The statistical inference on model parameters (e.g. for models ANOVA) is often conducted through the combined analysis using the information from independent submodels obtained by orthogonal decomposition of the observed vector ([7], [6], [1]). The statistical decision rules obtained in this way are uniquelly given and have under corresponding submodels the desirable statistical properties (e.g. inter- and intra- block estimators of variance components for a mixed linear model corresponding to a randomized block design). Only, in special cases (see [2]), estimation and testing under the overall randomization model are relevant. Generally, the estimators of variance components derived from submodels are inadmissible in the class of all invariant quadratic unbiased estimators (e.g. estimator of variance of block effects, see [4]). In reference to tests concerning variance components the ratio tests allowing the information from different submodels (strata) have a structure of Wald's test and generally are admissible, although the tests have weak statistical properties (cf. [6], where author shows how to recover the intra-block information to improve tests of hypotheses concerning inter-block parameters, see also [5]). In this article author presents a subclass of admissible bayesian invariant quadratic unbiased estimators (cf. [3]) which uniformly dominate the unbiased inter-block estimator of the variance of block effects proposed by Caliński and Kageyama in 1991. It will be illustrated by numerical examples for some connected and disconnected orthogonal block designs. Besides, author gives some results concerning admissibility of biased bayesian quadratic estimators of inter-block variance component in mixed linear model with two variance components corresponding to block designs.

Keywords

Admissibillity, Block designs, Variance components, Inter- and intra-block estimators, Invariant quadratic unbiased bayesian estimators, Testing of hypotheses.

References

[1] Caliński, T. and S. Kageyama (1991). On the randomization theory of intra-block and inter-block analysis. *Biom. Lett. 28*, 97–122.

[2] Caliński, T., S. Gnot, and A. Michalski (1998). On admissibility of the intra-block and inter-block variance component estimators. *Biom. Lett. 35(1)*, 11–26.

[3] Gnot, S. and J. Kleffe (1983). Quadratic estimation in mixed linear models with two variance components. *J. Statist. Plann. Inference 8*, 267–279.

[4] Gnot, S. and A. Michalski (1991). Linear and quadratic estimation from inter- and intra-block sources of information. *Statistics 22(1)*, 17–32.

[5] Gnot, S. and A. Michalski (1994). Tests based on admissible estimators in two variance components models. *Statistics 25*, 213–223.

[6] Portnoy, S. (1973). On recovery of intra-block information. *J. Amer. Statist. Assoc. 68*, 384–391.

[7] Rao, C.R. (1956). On the recovery of inter-block information in variety trials. *Sankhya 17*, 105–114.

Weighting, model transformation, and design optimality

John P. Morgan and **J. W. Stallings**

Virginia Polytechnic Institute and State University, Blacksburg, Virginia, United States

Abstract

Traditional design optimality criteria place equal emphasis on estimable functions of model parameters. Use of weighted criteria allows experiments to be designed so to place increased emphasis on estimation of those functions of the parameters that are of greater interest. Here design weighting is investigated for the linear model $y = A_d\tau + L\beta + e$ in which A_d (whose column space contains the all-one vector) is the design matrix to be selected, the parameters of interest are τ, the matrix L is fixed by the experimental setup, and β is comprised of nuisance parameters including an intercept. If C_d is the information matrix for estimation of τ, then $C_{dW} = W^{-1/2}C_dW^{-1/2}$ is a weighted information matrix that for any conventional criterion Φ induces a weighted criterion Φ_W via $\Phi_W(C_d) = \Phi(C_{dW})$. The weight matrix W can be any symmetric, positive definite matrix. Among the results established are: (i) for any desired assignment of (positive) weights to any full rank set of linearly independent, estimable functions of τ there is a corresponding weight matrix W; (ii) every admissible design is weighted E-optimal with respect to some weighting; (iii) optimal design for a reparameterized model is equivalent to weighted optimality for the original model. Result (iii) demonstrates, for instance, why orthogonal arrays need not be optimal fractions under a baseline parametrization (see [2]). Families of weight matrices W are explored according to features they encompass. Among these families are the diagonal weight matrices employed in [1].

Keywords

Design admissibility, Design optimality, Weighted optimality criterion.

References

[1] Morgan, J.P. and X. Wang (2010). Weighted optimality in designed experimentation. *J. Amer. Statist. Assoc. 105*, 1566–1580.

[2] Mukerjee, R. and B. Tang (2012). Optimal fractions of two-level factorials under a baseline parameterization. *Biometrika 99*, 71–84.

Eigenvalue estimation of covariance matrices of large dimensional data

Jamal Najim[1], **Jianfeng Yao**[1], **Abla Kammoun**[1], **Romain Couillet**[2], and **Mérouane Debbah**[2]

[1] Télécom Paristech and CNRS, Paris, France
[2] Supélec, Gif-sur-Yvette, France

Abstract

In many recent applications, one has to face high-dimensional datasets, where the number of available samples is of the same order as the dimension of each observation (although usually larger).
In this presentation, we shall address the problem of estimating the covariance matrix associated to such a dataset. Of course, in such a case, the traditional empirical estimator of the covariance matrix fails to be consistent and we shall rely on techniques based on large random matrix theory. We will present results associated to parametrized covariance matrices, where the number of distinct eigenvalues is known. We will also present estimation results of specific linear statistics. The main motivations come from wireless communication issues and will be briefly presented if time permits.

Keywords

Estimation, Large random matrix, Wireless communication applications.

Change-point detection in two-phase regression with inequality constraints

Konrad Nosek

AGH University of Science and Technology, Cracow, Poland

Abstract

Two-phase regression models with inequality constraints on the regression parameters and with a small number of measurements are considered. Tests for the presence of a change-point are constructed. The tests procedure are based on the likelihood ratio in a linear model with inequality constraints. Numerical approximations to the powers against various alternatives are given and compared with the powers of the likelihood ratio tests in the two-phase regression models without inequality constraints and with the powers of some other tests.

Keywords

Change-point, Two-phase regression, Linear regression model with inequality constraints, Likelihood ratio test.

References

[1] Chen, J. and A.K. Gupta (2000). *Parametric Statistical Change Point Analysis.* Boston: Birkhaeuser.

[2] Julious, S.A. (2001). Inference and estimation in changepoint regression problem. *Statistician 50*, 51–61.

[3] Kim, H.-J. and D. Siegmund (1989). The likelihood ratio tests for a change-point in simple linear regression. *Biometrika 76*, 409–423.

[4] Silvapulle, M.J. and P.K. Sen (2005). *Constrained Statistical Inference: Inequality, Order, and Shape Restrictions.* New York, Wiley.

[5] Smith, A.F.M. and D.G. Cook (1980). Straight lines with a change-point: a Bayesian analysis of some renal transplant data. *Appl. Statist. 29*, 180–189.

Tests for profile analysis based on two-step monotone missing data

Mizuki Onozawa, Sho Takahashi, and Takashi Seo

Tokyo University of Science, Japan

Abstract

We consider profile analysis when the data has two-step monotone missing observations. For two-sample profile analysis, there are three hypotheses of interest in comparing the profiles of two samples: two profiles are parallel, two profiles are same level, and two profiles are flat. The T^2-type statistics and their asymptotic null distributions for the three hypotheses are given. We propose the approximate upper percentiles of these test statistics. When the data dose not have the missing observations, the test statistics reduce to the usual test statistics given, for example, in Morrison ([1]). Further, we consider a parallel profile model for several groups when the data has two-step monotone missing observations. Under the assumption of non-missing data, the likelihood ratio test procedure are derived by Srivastava ([2]). We derive the test statistic based on the likelihood ratio. Finally the accuracy of the approximate values are investigated by Monte Carlo simulation for some selected values of parameters.

Keywords

Profile analysis, Two-step monotone missing data.

References

[1] Morrison, D.F. (2005). *Multivariate Statistical Methods* (4th Ed.) Duxbury.
[2] Srivastava, M.S. (1987). Profile analysis of several groups. *Comm. Statist. Theory Methods 16*, 909–926.

Considerations on sampling, precision and speed of robust regression estimators

Domenico Perrotta[1], Marco Riani[2], and Francesca Torti[3]

[1] European Commission, Joint Research Centre, Ispra, Italy
[2] University of Parma, Italy
[3] University of Milan Bicocca, Milan, Italy

Abstract

Methods of very robust regression, which resist up to 50% of outliers, spend a large part of the computational time in sampling subsets of observations and then computing parameter estimates from the subsets. The precision of the estimates depends on the amount of sampling, as we have to find solutions of non-smooth functions with lot of local minima. For example, Least Trimmed Squares (LTS) estimators try to minimize the sum of the h smallest squared residuals, where h is typically $(n-p+1)/2$ and the amount of sampling may vary from one to three thousands of subsets depending on the problem size (see e.g. [2]). To address large datasets, say with $1000 < n < 100.000$ units and $p = 10$ variables, [1] proposed a fast algorithm that can use fewer subsets, but applies c-steps to get approximations with lower objective function value. Moreover, to reduce for large datasets the applications of c-steps, which are $O(n)$, a divide and conquer strategy that partitions the dataset in smaller blocks of 300 observations is used.
We will show how Least Trimmed Squares (LTS) estimators can be made faster with an improved combinatorial sampling approach [3]. Then, we will illustrate the effect of increasing the amount of sampling on the precision of the estimates obtained with the traditional and fast LTS strategies.

Keywords

Least Trimmed Squares, Efficient random samples generation.

References

[1] Rousseeuw, P.J. and K. van Driessen (2006). Computing LTS Regression for Large Data Sets. *Data Min. Knowl. Discov. 12*, 29–45.
[2] Rousseeuw, P.J. and M. Hubert (1997). Recent developments in PROGRESS. *IMS Lecture Notes Monogr. Ser. 31*, 201–214.
[3] Torti, F., D. Perrotta, A.C. Atkinson, and M. Riani (2012). Benchmark testing of algorithms for very robust regression: FS, LMS and LTS. *Comput. Statist. Data Anal.* In press (doi:10.1016/j.csda.2012.02.003).

Asymptotic spectral analysis of matrix quadratic forms

Jolanta Pielaszkiewicz

Linköping University, Sweden

Abstract

The asymptotic spectral distribution of a sum of matrix quadratic forms

$$Q = AA' + \sum_{i=1}^{k} \frac{1}{n} X_i X_i',$$

where A is non-random and $X \sim N_{p,n}(0, \Sigma, \Psi)$, p and n are, respectively, the number of variables and observations, and $\frac{p}{n} \to c > 0$ will be discussed. Early results of Marchenko and Pastur will be related to theorems of Girko and von Rosen ([2]), and Silverstein and Bai ([3]). Then, after a short introduction to free-probability theory and justification of free-independence of the quadratic forms, results regarding the use of the R-transform for asymptotic spectral analysis of Q will be presented.

Keywords

Asymptotic distribution, Distribution function of eigenvalues, Random matrix, Matrix quadratic form, R-transform, Stieltjes transform, Free-probability.

References

[1] Girko, V. (1990). *Theory of Random Determinants.* Kluwer Academic Publishers.

[2] Girko, V. and D. von Rosen(1994). Asymptotics for the normalized spectral function of matrix quadratic form. *Random Oper. Stochastic Equations 2(2)*, 153–161.

[3] Silverstein J.W. and Z.D. Bai (1995). On the empirical distribution of eigenvalues of a class of large dimensional random matrices. *J. Multivariate Anal. 54*, 175–192.

[4] Voiculescu, D.V., K.J. Dykema, and A. Nica (1992). *Free random variables.* CRM Monograph Series 1, Amer. Math. Soc., Providence, RI.

Optimal designs for prediction of individual effects in random coefficient regression models

Maryna Prus and **Rainer Schwabe**

Otto-von-Guericke University, Magdeburg, Germany

Abstract

In the last years random coefficient regression models have become popular in many application fields, especially in biosciences. Besides the estimation of population parameters describing the mean behavior across all individuals a prediction of the individual response or the individual deviations for the specific individuals under investigation may be of interest, the latter for example in selection studies.

For the determination of optimal designs for estimating the population parameters some analytical and practical results may be found in the literature. Concerning prediction of the individual responses the theory developed by Gladitz and Pilz [1] for optimal designs requires the prior knowledge of the population parameters.

We develop the theory and solutions for prediction of individual response and individual deviations for the practical relevant situation of unknown population parameters. While the optimal designs for individual response will differ from the Bayesian designs proposed by Gladitz and Pilz [1], the Bayesian designs turn out to remain their optimality, if only the individual deviations are of interest, as long as all individuals are treated under the same regime. The obtained theoretical results will be illustrated by a simple example.

Keywords

Individual designs, Prediction, Individual parameters, Random coefficient regression models, Linear mixed models.

References

[1] Gladitz, J. and J. Pilz (1982). Construction of optimal designs in random coefficient regression models. *Statistics 13*, 371–385.

Oh, still crazy after all these years?

Simo Puntanen

University of Tampere, Finland

Abstract

Yeah, I think so.

Keywords

Best linear unbiased estimation, Cauchy–Schwarz inequality, Column space, Eigenvalue decomposition, Estimability, Gauss–Markov model, Generalized inverse, Idempotent matrix, Linear model, Linear regression, Löwner ordering, Matrix inequalities, Oblique projector, Ordinary least squares, Orthogonal projector, Partitioned linear model, Partitioned matrix, Rank cancellation rule, Reduced linear model, Schur complement, Singular value decomposition.

From linear to multilinear models

Dietrich von Rosen

Swedish University of Agricultural Sciences and Linköping University, Sweden

Abstract

The presentation is based on a number of figures illustrating appropriate linear spaces. The start is the classical Gauss-Markov model from where we jump into the multivariate world, i.e. MANOVA. The next stop will be the Growth Curve model and then a quick exposure of extended growth curves will take place. The tour is ended with some comments on multilinear models

Keywords

Multilinear models, Growth Curve models, Extended Growth Curve models.

Classification of higher-order data with separable covariance and structured multiplicative or additive mean models

Anuradha Roy[1] and Ricardo Leiva[2]

[1] The University of Texas at San Antonio, USA
[2] National University of Cuyo, Mendoza, Argentina

Abstract

Although devised in 1936 by Fisher [1], discriminant analysis is still rapidly evolving, as the complexity of contemporary data sets grows exponentially. Our classification rules explore these complexities by modeling various correlations in higher order data. Moreover, our classification rules are suitable to data sets where the number of response variables is comparable or larger than the number of observations. We assume that the higher-order observations have a separable covariance matrix and two different Kronecker product structures on the mean vector ([2], [3]). In this article we consider quadratic discrimination among g different populations where each individual has κth order ($\kappa \geq 2$) measurements.

Keywords

Higher-order data, Separable covariance structure, Structures on mean vector.

References

[1] Fisher, R.A. (1936). The use of multiple measurements in taxonomic problems. *Ann. Eugenics 7*, 179–188.
[2] Leiva, R. and A. Roy (2011). Linear discrimination for multi-level multivariate data with separable means and jointly equicorrelated covariance structure. *J. Statist. Plann. Inference 141*, 1910–1924.
[3] Leiva, R. and A. Roy (2012). Linear discrimination for three-level multivariate data with separable additive mean vector and doubly exchangeable covariance structure. *Comput. Statist. Data Anal. 56*, 1644-1661.

Multilevel linear mixed model for the analysis of longitudinal studies

Masoud Salehi[1] **and Farid Zayeri**[2]

[1] Tehran University of Medical Science, Iran
[2] Shahid Beheshti University of Medical Science, Tehran, Iran

Abstract

The use of longitudinal studies (studies in which the response of each individual is observed on two or more occasions) has been considered a lot over the last decades. Longitudinal studies beyond the cross-sectional studies in several ways: longitudinal study gives the opportunity for controlled and more reliable measurement of exposure history. Also longitudinal study gives information about individual change over time and factors that affected this change. Finally, this study provides more efficient estimates of parameters than cross-sectional study with the same number of individuals. A number of methods and statistical models on the analysis of hierarchical and longitudinal data have used in most researches, including traditional approaches such as repeated measurements analysis and multivariate analysis of variance. But new approaches, including multilevel linear mixed models, also known as hierarchical linear models, random coefficient models, and mixed-effect models, have become an increasingly important strategy for analyzing longitudinal data.

The observations within an individual are assumed to be correlated in such data and multilevel linear mixed models include the subject-specific profile in the model structure, therefore, these models should be well suited to describe longitudinal data. Recently, multilevel linear mixed models have applied in a few medical literatures, while this field has the potential and possibilities of these models. In this paper we introduce multilevel linear mixed model for the analysis of longitudinal data and interpretation of the parameters of the model at each level. As an example, the data from a sample of dental composites will analyze using SAS PROC MIXED.

Keywords

Multilevel linear mixed model, longitudinal study, dental composites.

References

[1] Diggle, P.J., K.Y. Liang, and S.L. Zeger (2000). *Analysis of Longitudinal Data.* Oxford University Press.

[2] Fitzmaurice, G.M., M. Davidian, G.Verbeke, and G. Molenberghs (2009). *Longitudinal Data Analysis.* Chapman & Hall/CRC.
[3] Fitzmaurice, G.M., N.M. Laird, and J.H. Ware (2009). *Applied Longitudinal Analysis.* Wiley.
[4] Goldstein, H. (1986). Multilevel mixed linear model analysis using iterative generalized least squares. *Biometrika 78*, 43–56.
[5] Goldstein, H. (1979). *The Design and Analysis of Longitudinal Studies.* Academic Press.
[6] Hedeker, D. and R.D. Gibbons (2006). *Longitudinal Data Analysis.* Wiley.
[7] Liang, K.Y. and S.L. Zeger (1986). Longitudinal data analysis using generalized linear models. *Biometrika 73*, 13–22.
[8] Molenberghs, G. and G. Verbeke (2005). *Models for Discrete Longitudinal Data* (2nd ed.) Springer.
[9] Verbeke, G. and G. Molenberghs (2000). *Linear Mixed Models for Longitudinal Data.* Springer.
[10] Ware, J.H. (1985). Linear models for the analysis of longitudinal studies. *Amer. Statist. 39*, 95–101.

On the Errors-In-Variables Model with singular covariance matrices

Burkhard Schaffrin[1], **Kyle Snow**[1,2], **and Frank Neitzel**[1,3]

[1] The Ohio State University, Columbus, USA
[2] Topcon Positioning Systems, Inc., Columbus, USA
[3] Berlin Institute of Technology, Germany

Abstract

Over the last few years, Total Least-Squares (TLS) estimation within Errors-In-Variables (EIV) Models has been extended not only to the case of element-wise weighted observations (corresponding to diagonal weight matrices), but also – and more importantly – to the case of arbitrary positive-definite weight matrices (defined as inverse covariance matrices), in which case "Mahboub's algorithm" provides the Weighted TLS Solution after a few iterations. Yet, the case of an EIV-Model with singular covariance matrices has not been considered in much detail, although unique TLS solutions may exist that take the (uninvertible) covariance matrices into proper account. Here, a generalization of "Mahboub's algorithm" will be developed for this purpose, followed by its application to a typical geodetic example (such as the 2-D Helmert transformation).

Keywords

Errors-In-Variables (EIV) Models, Total Least-Squares (TLS), Singular covariance matrices, 2-D coordinate transformations.

References

[1] Mahboub, V. (2012). On Structured Weighted Total Least-Squares for geodetic transformations. *J. Geodesy*. Submitted.
[2] Markovsky, I. and S. van Huffel (2007). Overview of Total Least-Squares Methods. *Signal Processing 87(10)*, 2283–3202.
[3] Neitzel, F. (2010). Generalization of Total Least-Squares on example of unweighted and weighted similarity transformation. *J. Geodesy 84*, 751–762.
[4] Schaffrin, B. and A. Wieser (2009). Empirical affine reference frame transformations by weighted multivariate TLS adjustment. In: H. Drewes (Ed.), *Geodetic Reference Frames, IAG-Symp. 134*, (pp. 213–218). Springer.

Fitting Generalized Linear Models to sample survey data

Alastair Scott and Thomas Lumley

University of Auckland, New Zealand

Abstract

Data from large complex surveys like NHANES are being used increasingly to build regression models. To give some idea of the extent of this, a call to Google Scholar comes up with more than 30,000 papers containing both "NHANES" and "regression model". Unfortunately complexities such as variable selection probabilities and multi-stage sampling mean that the assumptions underlying standard statistical methods for model-building are not even approximately valid for survey data. The problem of parameter estimation has been largely solved through the use of weighted estimating equations, and software for fitting GLMs to survey data is now available in most major statistical packages. The big gap in the output from these packages is an analogue of the deviance and related quantities like AIC. It turns out to be straightforward to extend the results in Rao & Scott (1984) for loglinear models in contingency tables to arbitrary GLMs. We show that the asymptotic distribution of the log-likelihood ratio is a linear combination of chi-squared random variables whose coefficients are eigenvalues of a matrix product that does not involve the inverse of the estimated covariance matrix. We then use results from Scott & Styan (1985) to obtain usable approximations to this asymptotic distribution using only information that is routinely available in large public-release surveys.

An illustrated introduction to Euler and Fitting factorizations and Anderson graphs for classic magic matrices

Miguel A. Amela[1], Ka Lok Chu[2], Amir Memartoluie[3], George P. H. Styan[4], and Götz Trenkler[5]

[1] General Pico, Argentina
[2] Dawson College, Montreal, Canada
[3] University of Waterloo, Canada
[4] McGill University, Montreal, Canada
[5] Dortmund University of Technology, Germany

Abstract

We build on results [8] about Euler factorizations of magic matrices presented at the LINSTAT'2008 Conference in celebration of Tadeusz Caliński's 80th Birthday. Our classic magic matrices are $n \times n$ with entries $0, 1, \ldots, n^2-1$ in some order. These matrices are fully-magic in that the numbers in all rows, columns, and the two main diagonals all add up to the same *magic sum*. The 4×4 classic magic matrix $\mathbf{M}$ has *Euler factorization* $\mathbf{M} = 4\mathbf{L}_1 + \mathbf{L}_2$, where the first *Euler component matrix* $\mathbf{L}_1 = \left[\frac{1}{4}\mathbf{M}\right]$ is the 4×4 matrix with entries which are the integer parts of entries in $\frac{1}{4}\mathbf{M}$. We also build on seminal results [5] by Friedrich Fitting (1862–1945) and his son Hans Fitting (1906–1938) about the factorization $\mathbf{M} = 8\mathbf{B}_1 + 4\mathbf{B}_2 + 2\mathbf{B}_3 + \mathbf{B}_4$, where the binary *Fitting component matrices* $\mathbf{B}_1 = \left[\frac{1}{8}\mathbf{M}\right]$, $\mathbf{B}_2 = \left[\frac{1}{4}\mathbf{M} - 2\mathbf{B}_1\right]$ and $\mathbf{B}_3 = \left[\frac{1}{2}\mathbf{M} - 4\mathbf{B}_1 - 2\mathbf{B}_2\right]$.

We believe that the proof that there are precisely 880 essentially distinct classic 4×4 fully-magic squares was first given by Fitting [5], though in [6] Bernard Frénicle de Bessy (c. 1605–1675) enumerated and classified these 880 matrices over 200 years earlier. Fitting [5] also showed that precisely 528 of these 880 have all four binary component matrices $\mathbf{B}_1$, $\mathbf{B}_2$, $\mathbf{B}_3$, $\mathbf{B}_4$ fully-magic, while very recently Amela [1] and Setsuda [7] have shown that 128 more, and so precisely 656 of these 880 have both Euler component matrices $\mathbf{L}_1$, $\mathbf{L}_2$ fully-magic. Brigadier-General F.J. Anderson (1860–1920) observed in [2] that certain 4×4 classic magic matrices have symmetric graphs.

In this talk we present a new and interesting classification of these 880 matrices using the 5 Frénicle–Amela patterns [1], [6], the 12 Dudeney types [3], and the Anderson symmetric graph property [2]. We have tried to illustrate our findings as much as possible, and whenever feasible with images of postage stamps or other philatelic items, with special emphasis on those associated with Leonhard Euler (1707–1783), Friedrich Fitting (1862–1945),

Hans Fitting (1906–1938), and Brigadier-General Sir Francis James Anderson CB KBE (1860–1920).

References

[1] Amela, M.A. (2009). Modeles structurels de carrés magiques 4×4: a la mémoire de Bernard Frénicle de Bessy. 17 pp., June 2009. (See also 4×4 *fully magic squares with Euler and binary factorization*, Personal communication from M.A. Amela to G.P.H. Styan, 20 April 2012).

[2] Anderson, Brigadier-General F.J. (1918). The 34 supermagic square. *Science Progress in the Twentieth Century: A Quarterly Review of Scientific Work, Thought & Affairs 13*, 86–96.

[3] Dudeney, H.E. (1910). The magic square of sixteen. *The Queen: The LadyŠs Newspaper and Court Chronicle*, 125–126.

[4] Euler, L. (1849). De quadratis magicis (in Latin). Talk delivered to the St. Petersburg Academy of Sciences on October 17, 1776. *Leonhardi Euleri Commentationes arithmeticae collectae, auspiciis Academiae imperialis scientiarum petropolitanae 2*, 593–602.

[5] Fitting, F. (1931). Rein mathematische Behandlung des Problems der magischen Quadrate von 16 und von 64 Feldern. *Jahresbericht der Deutschen Mathematiker-Vereinigung 40*, 177–199. [Translated into English by Götz Trenkler, in Research Report 2012-02, Dept. of Mathematics & Statistics, McGill University, pp. 3–45.]

[6] Frenicle [Bernard Frénicle de Bessy (c.1605–1675)] (1693). Des quarrez ou tables magiques; avec table générale des quarrez magiques de quatre de costé. In: P. de la Hire (Ed.), *Divers ouvrages de mathématique et de physique, par Messieurs de l'Académie Royale des Sciences* (pp. 423–507). De l'Imprimerie Royale par Jean Anisson, Paris.

[7] Setsuda, K. (2012). *Various Arts and Tools for Studying Magic Squares*, Section 7-4: Compose all the solutions of 3 types of magic squares of order 4 only by the 4-th increment number system. http://homepage2.nifty.com/KanjiSetsuda/pages/epages/S74C4P4.pdf.

[8] Styan, G.P.H., C. Boyer, and Ka Lok Chu (2009). Some comments on Latin squares and on Graeco-Latin squares, illustrated with postage stamps and old playing cards. *Statist. Papers 50*, 917–941.

Construction and analysis of D-optimal edge designs

Stella Stylianou

University of the Aegean, Samos, Greece

Abstract

Edge designs are screening experimental designs that allow a model-independent estimate of the set of relevant variables, thus providing more robustness than traditional designs. In this paper, new classes of D-optimal edge designs are constructed. This construction uses weighing matrices of order n and weight k together with permutation matrices of order n to obtain D-optimal edge designs. Linear and quadratic simulated screening scenarios are studied and compared using linear regression and edge designs analysis. An alternative method for constructing and analyzing expanded edge designs is introduced. This method provides a model-independent estimate of the set of active factors and also gives a linearity test for the underlying model.

Keywords

Screening, Linear models, Regression analysis, Conference matrices, Weighing matrices, Simulation experiments.

References

[1] Aboukalam, M.A.F. and A.A. Al-Shiha (2001). A robust analysis for unreplicated factorial experiments. *Comput. Statist. Data Anal. 36*, 31–46.

[2] Box, G.E.P. and R.D. Meyer (1986). An analysis for unreplicated fractional factorials. *Technometrics 28*, 11–18.

[3] Elster, C. and A. Neumaier (1995). Screening by conference designs. *Biometrika 82*, 589–602.

[4] Lenth, R.S. (1989). Quick and easy analysis of unreplicated factorials. *Technometrics 31*, 469–73.

Muste – editorial environment for matrix computations

Reijo Sund[1] and **Kimmo Vehkalahti**[2]

[1] National Institute for Health and Welfare, Helsinki, Finland
[2] University of Helsinki, Finland

Abstract

Practical application of multivariate statistical methods requires appropriate tools for the analyses. Such tools should provide flexible and powerful instruments to perform matrix computations. We present an editorial environment for matrix computations that allows to freely mix natural language and computation schemes. The presented matrix interpreter is one part of whole integrated system intended for statistical computing and related tasks. The history of this system dates back to early 1960s when *Seppo Mustonen* developed a library of matrix subroutines for the Elliott 803 computer [2]. This library was expanded to a statistical programming language `SURVO 66` [1]. Innovative editorial environment was introduced in `SURVO 76` [3] and the current version of the matrix interpreter (created by Mustonen in 1985) is based on the first C language version `SURVO 84C` [4,5]. Following `SURVO 98` and `SURVO MM`, the newest generation of the system is called `Muste`. Muste is an open source implementation of Survo and developed as a multiplatform `R` package [6]. It is freely available from the R-forge development platform.
We demonstrate the use of Muste implementation of the Survo matrix interpreter in the case of so called direct factor analysis in which exploratory factor analysis is considered as a specific data matrix decomposition with fixed unknown matrix parameters. In this recent approach all model unknowns including common and unique factor scores are estimated simultaneously by minimizing a specific object function with an alternating least squares (ALS) algorithm utilizing singular value decomposition (SVD) of data matrices. Such technique also allows to generalize factor analysis into cases with more variables than observations [7].

Keywords

Survo, Muste, R-project, Factor analysis, Singular value decomposition, Alternating least squares.

References

[1] Alanko, T., S. Mustonen, and M. Tienari (1968). A statistical programming language SURVO 66. *BIT 8*, 69–85.

[2] Mustonen, S. (1963). *SMS, a System of Matrix Subroutines for use with the 803 autocode*. National Elliott computer application program LM21. Elliott Computing Division, Borehamwood.
http://www.survo.fi/publications/Mustonen1963.pdf.

[3] Mustonen, S. (1981). On Interactive Statistical Data Processing. *Scand. J. Statist. 8*, 129–136.

[4] Mustonen, S. (1987). Editorial approach in statistical computing. In: T. Pukkila, S. Puntanen (Eds.) *Proceedings of the Second International Tampere Conference in Statistics* (pp. 205–224). University of Tampere, Finland.

[5] Mustonen, S. (1999). Matrix computations in Survo. In: *Proceedings of the 8th International Workshop on Matrices and Statistics (IWMS)*. University of Tampere, Finland.
http://www.helsinki.fi/survo/matrix99.html.

[6] Sund, R. (2011). *Muste – the R implementation of Survo*. Yearbook of the Finnish Statistical Society 2010, pp. 133–146.
http://www.survo.fi/muste/publications/sund2011.pdf.

[7] Unkel, S. and N.T. Trendafilov (2012). Zig-zag exploratory factor analysis with more variables than observations. *Comput. Statist.* In press.

Simultaneous confidence intervals among mean components in elliptical distributions

Sho Takahashi, Takahiro Nishiyama, and Takashi Seo

Tokyo University of Science, Japan

Abstract

We consider simultaneous confidence intervals for pairwise comparisons among components of mean vector. Such a situation arises, for example, in multiple comparisons of the components of repeated measurements of the same quantity in different conditions. Actually, in order to construct the simultaneous confidence intervals, it is required to give the upper percentiles of $F^2_{\max\cdot p}$ statistic. However, in general, it is difficult to find the exact values even under normality. So the approximate upper percentiles of $F^2_{\max\cdot p}$ statistic have been discussed by many authors (see, e.g., [1]). In this study, we consider approximation to the upper percentiles of $F^2_{\max\cdot p}$ statistics based on Bonferroni's inequality in elliptical distributions. Further, in order to evaluate the accuracy of the approximations, some numerical results by Monte Carlo simulations are given.

Keywords

Asymptotic expansion, Bonferroni's inequality, Elliptical distributions, Monte Carlo simulation, Pairwise comparisons.

References

[1] Seo, T. (1995). Simultaneous confidence procedures for multiple comparisons of mean vectors in multivariate normal populations. *Hiroshima Math. J. 25*, 387–422.

A new approach to adaptive spline threshold autoregression by using Tikhonov regularization and continuous optimization

Secil Toprak and Pakize Taylan

Dicle University, Diyarbakır, Turkey

Abstract

This paper investigates the use of conic adaptive spline treshold autoregression (C-ASTAR) which was developed using adaptive spline treshold autoregression (ASTAR) and conic quadratic programming (CQP).
MARS, a modern technology in statistical learning, has importance in regression and classification [1]. MARS is very useful for high dimensional problems and shows a great promise for fitting nonlinear multivariate functions. MARS technique does not impose any particular class of relationship between the predictor variables and outcome variable of interest. In other words, a special advantage of MARS lies in its ability to estimate the contribution of the basis functions so that both the additive and interaction effects of the predictors are allowed to determine the response variable.
By letting the predictor variables in the MARS algorithm be lagged in values of a time series system, one obtains a univariate ASTAR model for nonlinear autoregressive threshold modeling and analysis of time series, thereby extending the threshold autoregression (TAR) time series methodology [2]. ASTAR consists of two complementary algorithms as MARS. To estimate the model function, as MARS algorithm, ASTAR has two stepwise algorithms, which provide to determinate basis functions stand in the model and to get the best appropriate model. Because the model obtained with the forward stepwise algorithm used in the first step has a very complex structure in the second step using backward stepwise algorithm basis functions remove in turn to reach optimum model.
In this study, a new approach was applied for the second stepwise algorithm of ASTAR. With this approach, ASTAR model turned to the Tikhonov regularization problem was transformed to CQP problem. When bounds of this optimization problem are determined using multiobjective optimization approach, too many solutions can be obtained. Thus, it is aimed to attain an optimum solution.
In conclusion, linear regression, ASTAR algorithm and C-ASTAR algorithm were applied to two different time series data sets, and these approaches performances were compared by using different measures.

Keywords

Time series, Multivariate adaptive regression splines (MARS), Adaptive splines treshold autoregression (ASTAR), Tikhonov regularization, Multiobjective optimization, Conic quadratic programming (CQP).

References

[1] Friedman, J.H. (1991). Multivariate adaptive regression splines. *Ann. Statist. 19(1)*, 1–67.

[2] Stevens, J.G. (1991). *An Investigation of Multivariate Adaptive Regression Splines for Modeling and Analysis of Univariate and Semi-Multivariate Time series Systems.* Ph.D. Thesis, Naval Postgraduate School, California, USA.

The Luoshu and most perfect pandiagonal magic squares

Götz Trenkler[1] **and Dietrich Trenkler**[2]

[1] Dortmund University of Technology, Germany
[2] University of Osnabrück, Germany

Abstract

First the structure of 3×3 magic squares is investigated. It is shown that these squares can be represented by dyadic products of three mutually orthogonal vectors. Their Moore-Penrose inverse, numerical range and polar decomposition are derived. In the second part 4×4 pandiagonal magic squares are studied. Based on a simple representation with four mutually orthogonal vectors, many features of these magic squares like EP-ness, normality, symmetry and associatedness are considered. The talk is highlighted by a 4×4 pandiagonal magic square with numerous patterns, consisting of prime numbers only.

Cook's distance for ridge estimator in semiparametric regression

Semra Türkan and **Oniz Toktamis**

Hacettepe University, Ankara, Turkey

Abstract

The detection of influential observations has attracted a great deal of attention in last few decades. Most of the ideas of determining influential observations are based on single-case diagnostics with ith case deleted. The Cook's distance are most commonly used among the other single-case diagnostics and successfully applied to various statistical models. In this article, we propose Cook's distance for the ridge regression estimator of the parametric component in the semiparametric regression model to detect influential observations. We investigate the performance of proposed diagnostic to detect influential observations by using real data and simulation data.

Keywords

Semiparametric regression model, Ridge regression estimator, Cook's distance, Influential observations.

References

[1] Belsley, D.A. (1991). *Conditioning Diagnostics: Collinearity and Weak Data in Regression.* Wiley.

[2] Hu, H. (2005). Ridge estimation of semiparametric regression model. *J. Comput. Appl. Math. 176*, 215–222.

[3] Ruppert, D., M.P. Wand, R.C. Carroll, and R. Gill (2003). *Semiparametric Regression.* Cambridge.

[4] Roozbeh, M., M. Arashý, and H.A. Nýroumand (2010). Semiparametric ridge regression approach in partially linear models. *Comm. Statist. Simulation Comput. 39*, 449–460.

[5] Walker, E. and J.B. Birch (1988). Influence measures in ridge regression. *Technometrics 30*, 221–227.

D-optimum hybrid sensor network deployment for parameter estimation of spatiotemporal processes

Dariusz Uciński

University of Zielona Góra, Poland

Abstract

Process control often requires models in which non-negligible spatial dynamics has to be included in addition to the temporal one. Modelling then involves partial differential equations and a major difficulty in model calibration is the impossibility to measure process variables over the entire spatial domain. This leads to the question of how to optimally place sensors. Many sensor placement strategies have been developed [2]. They usually exploit the *Fisher information matrix* associated with the parameters to be identified. A revived interest in optimal sensor location is correlated with advances in *Sensor Networks* (SNs) which highly increase the flexibility of observation systems [1].
In this talk, a SN is considered which includes a number of mobile nodes which can move in a given spatial domain and, therefore, we would like their trajectories to be optimal in a sense. In addition to that, the data from mobile sensors are to be complemented by the ones gathered by a given number of nodes selected from among a greater number of nodes whose locations in space are fixed. Therefore, a decision must be made about which subset of non-mobile sensors is to be activated. Mathematically, the problem is a mixed discrete optimal control one and, due to its potential high dimensionality, naive solutions are deemed to failure. We apply the branch-and-bound method to drastically reduce the search space. The key idea behind it is alternation between two relaxed problems, namely a discrete optimization one related to stationary sensors and an optimal control one associated with moving sensors.

Keywords

D-optimality, Spatiotemporal process, Sensor network, Branch and bound.

References

[1] Li, X. (2008). *Wireless Ad Hoc and Sensor Networks Theory and Applications.* Cambridge University Press.
[2] Uciński, (2005). *Optimal Measurement Methods for Distributed Parameter System Identification.* CRC Press.

Multilevel Rasch model and item response theory

Nassim Vahabi[1], Mahmoud R. Gohari[1], and Ali Azarbar[2]

[1] Tehran University of Medical Science, Iran
[2] Alborz University, Iran

Abstract

The analysis of response data to test items requires psychometric methods to investigate characteristics of items and individuals that answer those items. Item Response Models (IRMs) consider that a latent variable explains these responses. The applications of IRT modeling have increased considerably in recent years because of its utility in developing of measuring instruments. Often the relations between the items and latent variable are of interest. Some procedures (factor analysis, discriminant analysis) allow the links between the items and the latent variables to be defined, but none of them make direct estimation of latent variable.
In 1960 Georg Rasch suggested a statistical Rasch Model (RM) that makes it possible to define these links and obtain scales with a good fit of an IRM. It transforms the cumulative raw scores (achieved by a subject across items or by an item across subjects) into linear continuous measures of ability of person and difficulty of item. Unidimensionality is a primary assumption of the Rasch model, that is, responses to the items should measure a single construct so the Rasch model is a unidimensional IRM. In Rasch model, raw data from a rating scale is converted to an equal interval scale measured in logits (log odd units) that allows one to use more variant parametric statistics instead of nonparametric statistics.
RM actually is a member of Hierarchical Generalized Linear Model (HGLM). In the simpler formulation of this model it is possible to consider a dichotomous RM as a two-level multilevel logistic model with random intercept, where the items and subjects are the level-1 and level-2 units, respectively and also item parameter is fixed and the person parameter is random. So with RM it is possible to incorporate a nested structure of the data and to include covariates at different hierarchical levels.
In this paper, we will present Item Response Theory and Multilevel Rasch Model, and will show the results on the basis of a data set of quality of life (SF36) by running WINSTEPS software.

Keywords

Multilevel Rasch model, Item response theory, Quality of life.

References

[1] Bond, T.G. and C.M. Fox (2007). *Applying the Rasch Model: Fundamental Measurement in the Human Sciences.* Mahwah NJ: Eribaum.

[2] Christensen, K.B., J.B. Bjorner, S. Kreiner, and J.H. Petersen (2004). Latent regression in loglinear Rasch models. *Comm. Statist. Theory Methods 33*, 1295–313.

[3] Davier, M. and C.H. Carstensen (2007). *Multivariate and Mixture Distribution Rasch Models.* Springer.

[4] Fischer, G. and I. Molenar (1995). *Rasch Models.* Springer.

[5] Hagquist, C. and D. Andrich (2009). Using the Rasch model in nursing research: an introduction and illustrative example. *Int. J. Nursing Studies 46*, 380–393.

[6] Henson, S., J. Bland, and J. Cranfield (2010). Difficulty of healthy eating: a Rasch model approach. *Soc. Sci. Med. 70*, 1574–80.

[7] Huber, C., M. Mesbah, and M. Nikulin (2008). *Mathematical Methods in Survival Analysis and Quality of Life.* Wiley.

[8] Mesbah, M. (2004). Measurement and analysis of health related quality of life and environmental data. *Environmetrics 15*, 473–81.

[9] Rasch, G. (1960). *Probabilistic Models for Some Intelligence and Attainment Tests.* Danish Institute for Educational Research.

[10] Wu, M. and R. Adams (2007). *Applying the Rasch Model to Psycho-social Measurement: A Practical Approach.* Educational Measurement Solutions, Melbourne.

Conditional AIC for linear mixed effects models

Florin Vaida

University of California, San Diego, USA

Abstract

We show that for a linear mixed effects model where the question of interest concerns cluster-specific inference the commonly-used definition for AIC is not appropriate. We propose a new definition for this context, which we call the conditional Akaike information criterion (cAIC). The cAIC is obtained from first principles, and we show that the penalty for the random effects is related to the effective number of parameters, rho, proposed by Hodges and Sargent; rho reflects a level of complexity between a fixed-effects model with no cluster effects, and a corresponding model with fixed cluster-specific effects. We provide finite-sample results for the linear mixed-effects model with known random effects variances, and an asymptotic approximation for a special case with unknown random effects variances. We compare the conditional AIC with the marginal AIC (in current standard use), and we argue that the latter is only appropriate when the inference is focused on the marginal, population-level parameters. A pharmacokinetics data application is used to illuminate the distinction between the two inference settings, and the usefulness of the conditional AIC. Extensions to generalized linear mixed model and proportional hazards mixed effects models, based on asymptotic arguments, are also considered.

On testing linear hypotheses in general mixed models

Júlia Volaufová and Jeffrey Burton

LSUHSC School of Public Health, New Orleans, USA

Abstract

Testing linear hypotheses about parameters of the mean (fixed effects) in linear mixed models has been studied extensively for decades. The methodology developed for linear mixed models can be adapted to nonlinear mixed models. Here we look into existing tests and discuss adjustment of a test based on a correction (approximation) of the estimated covariance matrix of fixed effects estimators. The correction takes into account the estimates of variance-covariance components, and the development is similar to the one done by Kackar-Harville and Kenward-Roger. The Satterthwaite approximation is used for calculation of degrees of freedom. The approach via first order approximation and via two-stage estimation for a nonlinear random-coefficient regression model is investigated.

Keywords

Nonlinear mixed model, ML and REML estimation, Adjusted F-test.

References

[1] Kackar, R.N. and D.A. Harville (1984). Approximations for standard errors of estimators of fixed and random effect in mixed linear models. *J. Amer. Statist. Assoc. 79*, 853–862.

[2] Kenward, M.G. and J.H. Roger (1997). Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics 53*, 983–997.

[3] Retout, S., F. Mentré, and R. Bruno (2002). Fisher information matrix for non-linear mixed-effects models: evaluation and application for optimal design of enoxaparin population pharmacokinetics. *Stat. Med. 21*, 2623–2639.

[4] Retout, S. and F. Mentré (2003). Further developments of the Fisher information matrix in nonlinear mixed effects models with evaluation in population pharmacokinetics. *J. Biopharm. Statist. 13*, 209–227.

[5] Sheiner, L.B. and S.L. Beal (1980). Evaluation of methods for estimating population pharmacokinetic parameters. I. Michaelis-Menten model: Routine pharmacokinetic data. *J. Pharmacokin. Biopharm. 8*, 553–71.

[6] Vonesh, E.F. and V.M. Chinchilli (1997). *Linear and Nonlinear Models for the Analysis of Repeated Measurements.* Marcel Dekker, Inc.

Functional discriminant coordinates

Tomasz Górecki, Mirosław Krzyśko and Łukasz Waszak

Adam Mickiewicz University, Poznań, Poland

Abstract

Let be y_{lij} the observed value of the tested statistical feature on the the i-th individual belogning to the l-th class in the j-th time point, where $i = 1, 2, \ldots, N_l$, $j = 1, 2, \ldots, J_i$, $l = 1, 2, \ldots, L$, $N_1 + N_2 + \ldots + N_L = N$.
The moments of observation t_{lij} of the statistical feature can vary from individual to individual and intervals between observation moments need not be identical. Then our data consist of pairs $\{t_{lij}, y_{lij}\}$, where $t_{lij} \in I$, $i = 1, 2, \ldots, N_l$, $j = 1, 2, \ldots, J_i$, $l = 1, 2, \ldots, L$.
We convert discrete data $\{t_{lij}, y_{lij}\}$ to functional data:

$$\{x_{li}(t),\ i = 1, 2, \ldots, N_l,\ l = 1, 2, \ldots, L,\ t \in I\},\ \text{where}$$

$$x_{li}(t) = \sum_{k=0}^{N-1} c_k \varphi_k(t),\ t \in I,$$

$\{\varphi_k(t)\}$ is the chosen orthonormal base system. The coefficients c_k are estimated from the data by least squares method.
The method of construction of discriminant coordinates in $L_2(I)$-space for functional data is described in the mongraph [1]. In this paper we propose a new method of construction of discriminant coordinates and its kernel variant.

Keywords

Functional data, Orthonormal basis, Discriminant coordinate, Reproducing kernel Hilbert space, Kernel.

References

[1] Ramsay, J.O. and B.W. Silverman (2005). *Functional Data Analysis* (2nd Edition). Springer.

On the linear aggregation problem in the general Gauß-Markov model

Fikri Akdeniz[1] and Hans J. Werner[2]

[1] Çukurova University, Adana, Turkey
[2] University of Bonn, Germany

Abstract

We consider the linear aggregation problem in the general possibly singular Gauß-Markov model. For the true underlying micro relations, which explain the micro behavior of the individuals, no restrictive rank conditions are assumed. We investigate several estimators for certain linear transformations of the systematic part of the corresponding macro relations and discuss their properties.

Keywords

Aggregation bias, Best linear unbiased estimator, Linear aggregation, Micro relation, Macro relation.

Robust model-based sampling designs

Douglas P. Wiens

University of Alberta, Edmonton, Canada

Abstract

I will describe some work currently being carried out with Alan Welsh at Australian National University. The problem addressed is to draw a sample, from which to estimate a population total. The data are completely known covariates, to which the unknown response variable is related. Difficulties to be overcome are that the relationship between these variables is only approximately, and perhaps erroneously, specified; similarly the variance/covariance structure of the data must be anticipated at the design stage. We derive minimax designs, and a genetic algorithm for computing the designs.

Keywords

Design, Genetic algorithm, Minimax, Robust.

On exact and approximate simultaneous confidence regions for parameters in normal linear model with two variance components

Viktor Witkovský[1] **and Júlia Volaufová**[2]

[1] Slovak Academy of Sciences, Bratislava, Slovakia
[2] LSUHSC School of Public Health, New Orleans, USA

Abstract

We consider normal linear regression model with two variance-covariance components

$$Y \sim N_n(X\beta, \sigma^2 V(\lambda)),$$

where X is known $(n \times p)$ matrix, $\beta \in R^p$ is unknown vector of parameters and $\sigma^2 V(\lambda) = \sigma^2(I_n + \lambda V)$ is the variance-covariance matrix, with known n.n.d. matrix V, which depends on unknown parameters $\sigma^2 > 0$ and $\lambda \geq 0$. We will present a brief overview the standard LRT/RLRT test statistics and will present the form and properties of their exact and/or approximate distributions under null hypothesis, see e.g. [1,2], which could be used for construction of the simultaneous confidence regions for some combinations of the parameters $\theta, \lambda, \sigma^2$, where $\theta = H'\beta$, H being a known matrix such that $R(H) \subseteq R(X')$, based on inverting the exact (restricted) likelihood ratio tests of the following null hypotheses:

$$H_0 : \theta = \theta_0 \text{ and } \lambda = \lambda_0 \tag{1}$$

$$H_0 : \theta = \theta_0 \text{ and } \lambda = \lambda_0 \text{ and } \sigma^2 = \sigma_0^2 \tag{2}$$

$$H_0 : \lambda = \lambda_0 \tag{3}$$

$$H_0 : \lambda = \lambda_0 \text{ and } \sigma^2 = \sigma_0^2. \tag{4}$$

Keywords

Linear regression model with two variance components, Exact likelihood ratio test, Simultaneous confidence regions.

References

[1] Crainiceanu, C.M. and D. Ruppert (2004). Likelihood ratio tests in linear mixed models with one variance component. *J. R. Stat. Soc. Ser. B Stat. Methodol. 66*, 165–185.

[2] Volaufová, J. and V. Witkovský (2012). On exact inference in linear models with two variance-covariance components. Submitted to *Tatra Mt. Math. Publ.* Proceedings from Probastat 2011.

Part VI

Posters

Using methods of stochastic optimization for constructing optimal experimental designs with cost constraints

Alena Bachratá and Radoslav Harman

Comenius University in Bratislava, Slovakia

Abstract

We propose a stochastic optimization method related to simulated annealing for constructing efficient designs of experiments under a broad class of linear constraints on the design weights.
The linear constraints can represent restrictions on various types of "costs" associated with the experiment.
To illustrate the method we computed D_A-optimal designs for estimating a set of treatment contrasts in the case of block experiments with blocks of size two.

Regression model of AMH

T. Rumpikova[1], Silvie Bělašková[2], D. Rumpik[1], and J. Loucky[3]

[1] Clinic for Reproductive Medicine and Gynaecology Zlin, Czech Republic
[2] Tomas Bata University in Zlin, Czech Republic
[3] Imalab s.r.o. Zlin, Czech Republic

Abstract

Anti-Mullerian hormone (AMH), which is also known as Mullerian inhibitory substance (MIS), is produced in the ovary by granulosa cells in pre-antral and small-antral follicles. AMH is a marker for ovarian reserve and it has been shown to be a good predictor of the number of oocytes retrieved from patiens undergoing IVF. There is a relationship between AMH levels and ovarian response during IVF. Many studies found a high level of correlation between the AMH level and the number of oocytes retrieved. Women with lower levels of AMH have lower count of the antral follicules and produce a lower number of oocytes. Unlike other levels of hormonal biomarkers - FSH, estradiol, inhibin B - AMH has a relatively stable expression during the menstrual cycle therefore the AMH test can be done on any day of womanťs cycle. Along with the evaluation of the age, basal FSH, inhibin B, antral follicle counts by ultrasound AMH allows much more precise estimate of ovarian reserve - fertility potential, ovarian response and estimates the chances of pregnancy success with IVF treatment. The objective of the study was to determine how AMH levels affect probability of the fertility.

Keywords

Anti-Mullerian hormone, Follicle-stimulating hormone, Logistic regression, Probability of the fertility.

Calibration between log-ratios of parts of compositional data

Sandra Donevska, Eva Fišerová, and Karel Hron

Palacký University Olomouc, Czech Republic

Abstract

Compositional data are multivariate observations carrying only relative information, popularly represented as proportions or percentages. Consequently, only ratios between parts of compositional data are informative [1,4]. They are characterized by the simplex sample space with the Aitchison geometry that has Euclidean vector space structure. Thus, since compositional data have different nature from the standard multivariate observations that rely on the Euclidean geometry in real space, they need to be expressed in real space using proper log-ratio transformation before standard statistical analysis is applied.
In the contribution we will perform calibration between parts of compositional data. One possible way to solve this problem is to apply orthogonal regression to all log-ratios of pairs of compositional parts. We will focus on some properties and interpretation on matrices of predicted averages and residual variances as results for all the mentioned combinations of log-ratios. The corresponding statistical inference will be performed using a linear regression model with type-II constraints [2,3].

Keywords

Compositional data, Log-ratio transformation, Orthogonal regression, Linear model with type-II constraints.

References

[1] Aitchison, J. (1986). *The Statistical Analysis of Compositional Data.* Chapman and Hall, London.
[2] Donevska, S., E. Fišerová, and K. Hron (2011). On the equivalence between orthogonal regression and linear model with type-II constraints. *Acta Univ. Palacki. Olomuc., Fac. rer. nat. Math. 50*, 19–27.
[3] Fišerová, E. and K. Hron (2010). Total least squares solution for compositional data using linear models. *J. Appl. Stat. 37*, 1137–1152.
[4] Pawlowsky-Glahn, V. and A. Buccianti (2011). *Compositional Data Analysis: Theory and Applications.* Wiley, Chichester.

COBS and stair nesting - segregation and crossing*

Célia Fernandes[1], **Paulo Ramos**[1], **and João T. Mexia**[2]

[1] Lisbon Superior Engineering Institute, Portugal
[2] Centro de Matemática e Aplicações, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Portugal

Abstract

Stair nesting leads to very light models since the number of their treatments is additive on the numbers of observations in which only the level of one factor various. These groups of observations will be the steps of the design. In stair nested designs we work with fewer observations when compared with balanced nested designs and the amount of information for the different factors is more evenly distributed. We now integrate these models into a special class of models with orthogonal block structure for which there are interesting properties.

Keywords

COBS, Stair nesting, Segregation, Cross.

References

[1] Carvalho, F., J.T. Mexia, and M. Oliveira (2008). Canonic inference and commutative orthogonal block structure. *Discuss. Math. Probab. Stat. 28*, 171–181.
[2] Fernandes, C., P. Ramos, and J.T. Mexia (2007). Variance components estimation in generalized orthogonal models. *Discuss. Math. Probab. Stat. 27*, 99–115 (2007).
[3] Fernandes, C., P. Ramos, and J.T. Mexia (2010a). Balanced and step nesting designs - Application for cladophylls of asparagus. *J. Biostat. 4*, 279–287.
[4] Fernandes, C., P. Ramos, and J.T. Mexia (2010b). Algebraic structure of step nesting designs. *Discuss. Math. Probab. Stat. 30*, 221–235.
[5] Fernandes, C., P. Ramos, S. Saraiva, and J.T. Mexia (2005). Optimization of nested step designs. *Biom. Lett. 42*, 143–151.
[6] Fonseca, M., J.T. Mexia, and R. Zmyślony (2006). Binary operations on Jordan algebras and orthogonal normal models. *Linear Algebra Appl. 417*, 75–86.
[7] Vanleeuwen, D., J. Seely, and D. Birkes (1998). Sufficient conditions for orthogonal designs in mixed linear models. *J. Statist. Plann. Inference 73*, 373–389.

* This work was partially supported by the Portuguese Foundation for Science and Technology through PEst-OE/MAT/UI0297/2011 (CMA).

[8] Vanleeuwen, D., D. Birkes, and J. Seely (1999). Balance and orthogonality in designs for mixed classification models. *Ann. Statist. 27*, 1927–1947.

[9] Zmyślony, R. (1978). *A Characterization of Best Linear Unbiased Estimators in the General Linear Model*. Lectures Notes in Statistics 2, Springer-Verlag, Berlin.

Validity of the assumed link functions for some binary choice models based on the bootstrap confidence band with R

Özge Akkuş, Serdar Demır, and Atilla Göktaş

Muğla University, Turkey

Abstract

In this study, we have introduced the commands testing the validity of the assumed link functions of the binomial logit, probit and complementary log-log models based on the bootstrap confidence bands in R. Some parts of the commands are designed to be optional and provide users to have the results with respect to the different parametric models such as logit, probit and complementary log-log model and different semiparametric estimators such as the semiparametric maximum likelihood estimator and the weighted semiparametric least square estimator. Researchers studying in this area could easily test the accuracy of the assumed parametric link functions for binary outcomes using the commands in R, which is free, widely used and a very popular statistical package as well. The applicability of the codes was supported over hypothetical and real data sets.

Keywords

Bootstrap confidence band, Validity test, Binomial choice, R package.

References

[1] Davison, A.C. and D.V. Hinkley (1997). *Bootstrap Methods and Their Application.* Cambridge University Press, New York.

[2] Hardle, W. (1990). *Applied Nonparametric Regression.* Cambridge University Press, New York.

[3] Härdle, W., M. Müller, S. Sperlich, and A. Werwatz (2004). *Nonparametric and Semiparametric Models.* Springer-Verlag, New York.

[4] Horowitz, J.L. (1998). *Semiparametric Methods in Econometrics.* Springer-Verlag, New York.

[5] Horowitz, J.L. and W. Härdle (1994). Testing a parametric model against a semiparametric alternative. *Econometric Theory 10*, 821–848.

[6] Hayfield, T. and J.S. Racine (2008). Nonparametric Econometrics: The np Package. *J. Stat. Software 27(5)*, 1–32.

[7] Ichimura, H. (1993). Semiparametric least squares (SLS) and weighted SLS estimation of single-index models. *J. Econometrics 58*, 71–120.

[8] Klein, W. and R.H. Spady (1993). An efficient semiparametric estimator for binary response models. *Econometrica 61*, 387–421.
[9] Nadaraya, E.A. (1964). On estimating regression. *Theory Probab. Appl. 10*, 186–190.
[10] Watson, G.S. (1964). Smooth regression analysis. *Sankhyā A 26*, 359–372.

Regular E-optimal spring balance weighing design with correlated errors

Bronisław Ceranka and Małgorzata Graczyk

Poznań University of Life Sciences, Poland

Abstract

The problems linked with an E-optimal spring balance weighing design with correlated errors are discussed. The concept of the paper is the generalization of ideas of optimal designs presented in [1] and [2]. The topic is focus on the determining the maximal eigenvalue of the information matrix for the design. There is given the lowest of the eigenvalue and the conditions under which the lowest bound is fulfill. The constructing method of the E-optimal design, based on the incidence matrices of balanced incomplete block designs, is presented.

References

[1] Jacroux M. and W. Notz (1983). On the optimality of spring balance weighing designs. *Ann. Statist. 11*, 970–978.

[2] Neubauer M.G. and W. Watkins (2002). E-optimal of spring balance weighing designs for $n \equiv -1 (\mathrm{mod}\, 4)$ objects. *Siam J. Matrix Anal. Appl. 24*, 91–105.

Estimation of parameters of structural change under small sigma approximation theory

Romesh Gupta

University of Jammu, India

Abstract

In this paper, the structural change in a linear regression model over two different periods of time is estimated. The ordinary least squares and Stein-rule estimators are employed to estimate the structural change. Their efficiency properties are derived using the small sigma theory and dominance conditions are derived.

Keywords

Structural change, Ordinary least squares, Stein-rule estimators.

Canonical variate analysis of chlorophyll a, b and $a+b$ content in tropospheric ozone-sensitive and resistant tobacco cultivars exposed in ambient air conditions

Dariusz Kayzer, **Klaudia Borowiak**, **Anna Budka**, **and Janina Zbierska**

Poznań University of Life Sciences, Poland

Abstract

Tropospheric ozone effects negatively crop plants causing the biomass and yield losses, which might be connected with plant photosynthesis activity decrease. Chlorophyll content has been discovered as one of the parameters, which responses for higher ozone concentrations. However, these results were usually obtained during fully controlled conditions. Hence, it is necessary to conduct investigations in ambient air conditions to confirm these findings.
Ozone-sensitive and -resistant tobacco cultivars were employed in presented investigations. Plants were exposed in 6 sites for 7 two-week series in growing season of 2006. Simultaneously, one site was located in control conditions with no ozone. Chlorophyll a, b and $a+b$ in fresh and dry weight content were measured after every exposure series with using the extraction by DMSO method.
The aim of presented study was to examined if ozone affects chlorophyll content in these two cultivars exposed in various sites in several series. As well as, the determination differences in leaf response for further choice the best leaf to physiological plant investigations. For these purposes canonical variate analyses was employed. Graphical presentation of obtain results is presented here. Experimental objects were placed in space of canonical variates, while points described the chlorophyll content were located in dual space of canonical variates.
The results revealed differences between chlorophyll content measured in different exposed series, although there was no differences between sites, except control and site located in the city centre. Probably, sites of exposure did not differ the ozone effect due to small differences in tropospheric ozone concentrations. While higher differences were noted between certain series, which might be connected with favorable meteorological conditions for ozone creation as well as for plant photosynthesis activity and chlorophyll creation. Moreover, both tobacco cultivars responded similarly for ozone occurrence in the ambient air, which might be a very good indicator of ozone effect without

visible symptoms. Additionally, the obtained results pointed out the best leaf for further investigations.

Latin square designs and fractional factorial designs

Pen-Hwang Liau and Pi-Hsiang Huang

National Kaohsiung Normal University, Taiwan

Abstract

A Latin square of order s is an arrangement of the s letters in an $s \times s$ square so that every letter appears exactly once in every row and exactly once in every column. The fractional factorial designs, a subset of the full factorial, are widely used in industrial research or other fields to reduce the cost of the experiment. In fact, Latin squares may also be used for fractional factorial designs, and there are some relationships between these two kinds of design. [5] used two examples to show that a Latin square can be chosen such that it corresponds to a fractional factorial design. In this presentation, we are going to study this topic more precisely. Furthermore, we will explore the relationship between fractional factorial design and Latin square design in general, where s is a prime or a power of a prime.

References

[1] Bose, R. and K. Bush (1952). orthogonal arrays of strength two and three. *Ann. Math. Statist. 23*, 508–524.
[2] Box, G.E.P. and J.S. Hunter (1961). The 2^{k-p} fractional factorial designs. *Technometrics 3*, 311–351 and 449–458.
[3] Box, G.E.P., W.G. Hunter, and J.S. Hunter (1978). *Statistics for Experimenters.* Wiley.
[4] Brownlee, K.A., B.K. Kelly, and P.K. Loraine (1948). Fractional replication arrangements for factorial experiments with factors at two levels. *Biometrika 35*, 268–276.
[5] Copeland, K.A.F. and P.R. Nelson (2000). Latin squares and two-level fractional factorial designs. *J. Qual. Technol. 32*, 432–439.
[6] Finney, D.J. (1945). The fractional replication of factorial arrangements. *Annals of Eugenics 12*, 291–301.
[7] Mann, H. (1942). The construction of orthogonal Latin squares. *Ann. Math. Statist. 13*, 418–423.
[8] Shao, J. and W. Wei (1987). A formula for the number of Latin squares. *Discrete Math. 110*, 293–296.

Weighted linear joint regression analysis

Dulce G. Pereira[1], **Paulo C. Rodrigues**[2], **and João T. Mexia**[2]

[1] University of Évora, Portugal
[2] Centro de Matemática e Aplicações, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Portugal

Abstract

The introduction of weights in the phenotypic trait (e.g. yield) for different genotypes in different environments enables us to generalize the joint regression analysis (JRA) [1,2], for the case when the error variance is not homogeneous across environments. Moreover it is possible to use incomplete blocks, while the environments are "better" represented accordingly to the accuracy in the measurements.
To fit the regressions for the weighted linear JRA, an algorithm is derived to minimize the sum of sums of weighted residuals [3,4]. An application with data sets from spring barley (*Hordeum vulgare* L.) breeding programme carried out in Czech Republic is presented and the results are compared with the standard JRA.

References

[1] Finlay, K.W. and G.N. Wilkinson (1968). The analysis of adaptation in a plant breeding programme. *Aust. J. Agr. Res. 14*, 742–754.
[2] Gusmão, L. (1985). An adequate design for regression analysis of yield trials. *Theor. Appl. Genet. 71*, 314–319.
[3] Pereira, D.G. and J.T. Mexia (2008). Selection proposal of spring barley in the years from 2001 to 2004, using Joint Regression Analysis. *Plant Breeding 127*, 452–458.
[4] Pereira, D.G., P.C. Rodrigues, S. Mejza, and J.T. Mexia (2011). A comparison between joint regression analysis and AMMI: a case study with barley. *J. Stat. Comput. Simul. 82*, 193–207.

Modeling resistance to oat crown rust in series of oat trials

Marcin Przystalski[1], Piotr Tokarski[1], and Wiesław Pilarczyk[1,2]

[1] Research Center for Cultivar Testing, Słupia Wielka, Poland
[2] Poznań University of Life Sciences, Poland

Abstract

Based on the results of post registration variety trials a recommendation for farmers is produced which varieties should be sown. In trials on spring oat one of the observed characteristics is resistance to oat crown rust. This is main disease which affects all regions of crop growth ([2]). Crown rust reduces oat yield and causes thin kernels with low weight. Moderate to severe epidemics can reduce grain yield by 10 to 40%.
To answer the question which oat varieties in the Polish post registration trial system are the best in terms of resistance to crown rust, we analyzed series of 40 oat field trials from two consecutive years 2009 and 2010. For this purpose the generalized linear mixed model ([3]) with single variance component representing variety×site interaction was applied. The most resistant varieties were identified and significant differences were detected. One of the varieties was also more resistant to crown rust then standard (the combination of three varieties pointed by specialist as standard varieties). Maximum likelihood estimates were obtained using Laplace transformation to compute likelihood function. All computations were performed using R package ordinal ([1]).

Keywords

Generalized linear mixed model, Multinomial distribution, Ordinal data, Oat crown rust.

References

[1] Christensen, R.H.B. (2012). ordinal – Regression Models for Ordinal Data R package version 2012.01-19. http://www.cran.r-project.org/package=ordinal/.
[2] Simmons, M.D. (1985). Crown rust. In: A. P. Roelfs, W. R. Bushnell (Eds.), *The Cereal Rusts 2* (pp. 131–172). Academic Press.
[3] Tutz G. and W. Hennevogl (1996). Random effects in ordinal regression models. *Comput. Statist. Data Anal. 22*, 537–557.

Estimation of variance components in balanced, staggered and stair nested designs*

Paulo Ramos[1], Célia Fernandes[1], and João T. Mexia[2]

[1] Lisbon Superior Engineering Institute, Portugal
[2] Centro de Matemática e Aplicações, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Portugal

Abstract

Traditional balanced nested designs are the most popular form of nesting but we are forced to divide repeatedly the plots and we have few degrees of freedom for the first levels. Meanwhile the number of treatments increases rapidly with the number of factors and the number of levels in each factor. These designs are orthogonal and the estimators of the variance components are independent. As an alternative we have the unbalanced nesting. The most popular unbalanced nested design is the staggered nested design. This design requires less observations than the balanced case and the degrees of freedom are almost the same for the different factors. However this design is not orthogonal. Another alternative is the stair nested design. In this design we can work with fewer observations than the balanced case, the amount of information for the different factors is more evenly distributed and the number of degrees of freedom is not very different among the factors. However this design have an orthogonal structure unlike the staggered nested designs so they retain the simplicity associated with orthogonality in balanced nested designs. In this work we compare the results obtained for the estimators of the variance components using these three designs.

Keywords

Balanced nested designs, Staggered nested designs, Stair nested designs, Variance components.

References

[1] Bainbridge, T. (1965). Staggered nested designs for estimating variance components. *Industrial Quality Control 22*, 12–20.
[2] Cox, D. and P. Solomon (2003). *Components of Variance*. Chapman and Hall, New York.

* This work was partially supported by the Portuguese Foundation for Science and Technology through PEst-OE/MAT/UI0297/2011 (CMA).

[3] Fernandes, C., P. Ramos, and J.T. Mexia (2010). Algebraic structure of step nesting designs. *Discuss. Math. Probab. Stat. 30*, 221–235.

[4] Fernandes, C., P. Ramos, S. Saraiva, and J.T. Mexia (2007). Variance components estimation in generalized orthogonal models. *Discuss. Math. Probab. Stat. 27*, 99–115.

[5] Fonseca, M., J.T. Mexia, and R. Zmyślony (2006). Binary operations on Jordan algebras and orthogonal normal models. *Linear Algebra Appl. 417*, 75–86.

[6] Khuri, A., T. Mathew, and B. Sinha (1998). *Statistical Tests for Mixed Linear Models*. John Wiley and Sons, New York.

[7] Ojima, Y. (1998). General formulae for expectations, variances and covariances of the mean squares for staggered nested designs. *J. Appl. Stat. 25*, 785–799.

[8] Ojima, Y. (2000). Generalized staggered nested designs for variance components estimation. *J. Appl. Stat. 27*, 541–553.

D-optimal chemical balance weighing designs for three objects if $n \equiv 2 \pmod 4$

Krystyna Katulska and Łukasz Smaga

Adam Mickiewicz University, Poznań, Poland

Abstract

In this paper, chemical balance weighing design problem for three objects and the errors between the observations follow a first-order autoregressive process is considered. From such assumptions, the covariance matrix of error components depends on the known parameter ρ. We prove the D-optimality of some designs in the class of designs for three objects, when the number of observations $n \equiv 2 \pmod 4$ and some $\rho \geq 0$. Some necessary and sufficient conditions under which the design is D-optimal in considered class of designs are also proved.

References

[1] Jacroux, M., C.S. Wong, and J.C. Masaro (1983). On the optimality of chemical balance weighing designs. *J. Statist. Plann. Inference 8*, 231–240.

[2] Katulska, K. and Ł. Smaga (2012a). D-optimal chemical balance weighing designs with $n \equiv 0 (\mathrm{mod}\, 4)$ and 3 objects. *Comm. Statist. Theory Methods.* DOI:10.1080/03610926.2011.608587.

[3] Katulska, K. and Ł. Smaga (2012b). D-optimal chemical balance weighing designs with autoregressive errors. *Metrika.* DOI:10.1007/s00184-012-0394-8

[4] Li, C.H. and S.Y. Yang (2005). On a conjecture in D-optimal designs with $n \equiv 0 \pmod 4$. *Linear Algebra Appl. 400*, 279–290.

[5] Yeh, H.G. and M.N. Lo Huang (2005). On exact D-optimal designs with 2 two-level factors and n autocorrelated observations. *Metrika 61*, 261–275.

Is the skew *t* distribution truly robust?

Tsung-Shan Tsou[1,2] and Wei-Cheng Hsiao[1]

[1] National Central University, Jhongli, Taiwan
[2] Cathay Medical Research Institute, Taipei, Taiwan

Abstract

The skew t distribution is considered by many a flexible model for modeling general asymmetric data. The model parameters are believed to be able to properly capture the skewness and kurtosis possessed in data.
The alleged robustness property of the skew t distribution is inspected in details in the independent and identically distributed and regression situations. It is found that the skew t distribution is robust only when the extent of asymmetry is mild and the magnitude of kurtosis is small.
We recommend using an existing parametric robust likelihood approach to analyze data when one is uncertain about the distribution underlying the data.

Design of experiment for regression models with constraints

Michaela Tučková and **Lubomír Kubáček**

Palacký University in Olomouc, Czech Republic

Abstract

We consider the linear regression model with parameter constraints, i.e. regression model with constraints of the type I. For this model we present the exact form of the criterial function and the iterative computation of optimum designs.
We focused on the criterion of local A-optimality, the criterion of local D-optimality and the criterion of local C-optimality. The aim of the presented contribution is to show the exact form of the gradient of the considered local criterion functions.

Keywords

Regression model with parameter constrains, A-optimal design, D-optimal design, C-optimal design.

References

[1] Pázman, A. (2002). Optimal design of nonlinear experiments with parameter constrains. *Metrika 56*, 113–130.

Inference for the interclass correlation in familial data using small sample asymptotics

Miguel Fonseca[1], João T. Mexia[1], Thomas Mathew[2], and Roman Zmyślony[3,4]

[1] Centro de Matemática e Aplicações, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Portugal
[2] University of Maryland, Baltimore County, USA
[3] University of Zielona Góra, Poland
[4] University of Opole, Poland

Abstract

Inference on the parent-offspring correlation coefficient is an important problem in the analysis of familial data, and point estimates and likelihood based inference are available in the literature. In this work, corrections for the signed log-likelihood ratio test statistics are proposed, based on small sample asymptotics, in order to achieve accurate small sample performance. The corrected statistic can be used for hypothesis testing as well as for interval estimation. Numerical results are reported to show that the resulting tests and confidence intervals exhibit satisfactory performance regardless of the sample sizes. The results are illustrated using an example.

Keywords

Correlation coefficient, High order asymptotics, Likelihood ratio.

Part VII

George P. H. Styan

George P. H. Styan's Editorial Positions and Publications

Carlos A. Coelho

Departamento de Matemática and Centro de Matemática e Aplicações, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Portugal

George P. H. Styan has a long and honorable career in Mathematics. He has served in the Editorial Boards of many scientific journals and published ten books and over one hundred fifty papers. In the following pages we attempt to compile an almost complete listing of his editorial positions and publications (books and papers) up to now.

Many thanks to Simo Puntanen and George Styan himself for the help in gathering the information.

I. Editorial Positions

Editor–in–Chief:

- *Chance* Magazine: vol. 9, no. 1 (1996)–vol. 11, no. 4 (1998);
- *Forthcoming Events/Activités Prévues*: supplement to *The Canadian Journal of Statistics/La Revue Canadienne de Statistique*: 1979–1984;
- *Image–The Bulletin of the International Linear Algebra Society*: no. 13 (July 1994)–no. 30 (April 2003) jointly with Steven J. Leon: no. 13 (July 1994)–no. 18 (Winter/Spring 1997, with Hans Joachim Werrner: no. 25 (October 2000)–no. 30 (April 2003);
- *The IMS Bulletin*: vol. 16 (1987)–vol. 21 (1992);
- *International Calendar of Statistical Events*: In *The IMS Bulletin*, vol. 16 (1987)–vol. 25 (1996); *Supplément à La Gazette des Sciences Mathématiques du Québec*: 1977–1978;
- *Statistical Science Association of Canada/Association Canadienne de Science Statistique Newsletter*: 1972.

Abstracting Editor:

- *Current Index to Statistics:* 2000–to date; pub. American Statistical Association (Alexandria, Virginia) & Institute of Mathematical Statistics (Beachwood, Ohio), ISBN 1–931586–15–2, ISSN 1094–7469. [Bibliography of over 250,000 items referencing the literature in statistics and probability.]

Managing Editor:

– *The Canadian Journal of Statistics/La Revue Canadienne de Statistique:* 1977–1984, 1999–2009, vol. 5, no. 2 (June 1977)–vol. 12, no. 2 (June 1984), vol. 27, no. 1 (March 1999)–vol. 36, no. 4 (December 2008); Editor: McGill University Reports from the Department of Mathematics and Statistics (ISSN 0824–4944): 1982–1984 (jointly with W. O. J. Moser), 2002–to date.

Associate Editor:

– *Far East Journal of Mathematics:* 2007–to date;
– *Journal of Inequalities in Pure and Applied Mathematics:* 1999–2005;
– *Journal of Statistics & Management Systems:* 2005–to date;
– *Mathematical Inequalities & Applications:* 1997–to date;
– *Communications in Statistics*: 1977–2000;
– *Journal of Statistical Planning and Inference*: 1992–2000;
– *Istatistik–Journal of the Turkish Statistical Association*: 1995–1999;
– *Linear Algebra and its Applications*: 1986–2004;
– *SIAM Journal on Matrix Analysis and Applications*: 1995–1997;
– *SSC Liaison–The Newsletter of the Statistical Society of Canada*: 1987–1999 & Production Editor: 2005–2007.

Book Reviews Editor:

– *The Canadian Journal of Statistics*: 1985–1988.

Corresponding Editor:

– *The IMS Bulletin*: 1993–2000.

Joint Editor:

– (with A. J. Wright), *European Aviation News*: February 1958–1963;
– (with L. F. Sarjeant), *Overseas Civil Register News*: 1956–1957;
– *European Aviation News*: January 1958.

Advisory Editor:

– *Chance* Magazine: 1999–2007.

Member of the International Editorial Board:

– Research Group on Mathematical Inequalities and Applications (RGMIA), Melbourne, Australia: 1999– .

II. Special Issues Guest Edited

1. I. Olkin, C. R. Rao & G. P. H. Styan, eds. (1985–1986). First Special Issue on Linear Algebra and Statistics: *Linear Algebra Appl. 67* (June 1985), 279 pp.; vol. 70 (October 1985), 369 pp. & vol. 82 (October 1986), pp. 143–279.

2. M. D. Perlman, F. Pukelsheim & G. P. H. Styan, eds. (1990). Second Special Issue on Linear Algebra and Statistics: *Linear Algebra Appl. 127* (January 1990), viii + 656 pp.

3. J. K. Baksalary & G. P. H. Styan, eds. (1992). Third Special Issue on Linear Algebra and Statistics: *Linear Algebra Appl. 176* (November 1992), viii + 289 pp. [Includes papers presented at the International Workshop on Linear Models, Experimental Designs & Related Matrix Theory: Tampere, Finland, August 1990.]

4. J. K. Baksalary & G. P. H. Styan, eds. (1993). Special Issue, Papers Presented at the International Workshop on Linear Models, Experimental Designs, and Related Matrix Theory. *J. Statist. Plann. Inference 36(2,3)* (August/September 1993), ii + pp. 127–432. [Zbl 783.00013; 24 research papers presented at the Workshop held in Tampere, Finland, 6–8 August 1990.]

5. F. Pukelsheim, G. P. H. Styan, H. Wolkowicz & I. Zaballa, eds. (1994). Special Issue Honoring Ingram Olkin. *Linear Algebra Appl. 199* (March 1994), viii + 445 pp. [Special Issue in Honor of Ingram Olkin's 70th Birthday.]

6. J. J. Hunter, S. Puntanen & G. P. H. Styan, eds. (1994). Fourth Special Issue on Linear Algebra and Statistics: *Linear Algebra Appl. 210* (October 1994), ii + 273 pp.

7. R. B. Bapat, G. P. H. Styan & H. J. Werner, eds. (1996). Fifth Special Issue on Linear Algebra and Statistics: In Celebration of the 75th Birthday of C. R. Rao. *Linear Algebra Appl. vol. 237/238* (April 1996), vii + 592 pp. [MR1382661.]

8. R. W. Farebrother, S. Puntanen, G. P. H. Styan & H. J. Werner, eds. (1997). Sixth Special Issue on Linear Algebra and Statistics: *Linear Algebra Appl. 264* (October 1997), ix + 506 pp. [Zbl 881.00016.]

9. R. W. Farebrother, S. Puntanen, G. P. H. Styan & H. J. Werner, eds. (1999). Seventh Special Issue on Linear Algebra and Statistics: *Linear Algebra Appl. 289* (March 1999), iv + 344 pp. [Includes papers from the Sixth International Workshop on Matrices and Statistics, Istanbul, Turkey, August 16–17, 1997. MR 99i:00019, Zbl 928.00027.]

10. S. Puntanen, G. P. H. Styan & H. J. Werner, eds. (2000). Eighth Special Issue on Linear Algebra and Statistics. *Linear Algebra Appl. 321* (15 December 2000), xi + 412 pp. [MR1799634, Zbl 0972.00016.]

11. S. Puntanen, G. P. H. Styan & H. J. Werner, eds. (2002). Ninth Special Issue on Linear Algebra and Statistics. *Linear Algebra Appl. 354* (15 October 2002), xii + 291 pp. [MR1927644.]

12. S. Puntanen, G. P. H. Styan & H. J. Werner, eds. (2004). Tenth Special Issue (Part 1) on Linear Algebra and Statistics. *Linear Algebra Appl. 388* (1 September 2004), 400 pp.

13. S. Puntanen, G. P. H. Styan & H. J. Werner, eds. (2005). Tenth Special Issue (Part 2) on Linear Algebra and Statistics. *Linear Algebra Appl. 410* (15 November 2005), 290 pp.

14. S. E. Ahmed, J. J. Hunter, G. P. H. Styan & G. Trenkler, eds. (2009). Special Issue devoted to selected papers presented at the 16th International Workshop on Matrices and Statistics (IWMS–2007): Windsor, Ontario, Canada, June 1-3, 2007. *Linear Algebra Appl. 430(10)*, pp. 2563–2834 (1 May 2009).

III. Publications

Books

1. T. W. Anderson, S. Das Gupta & G. P. H. Styan (1972, 1977). *A Bibliography of Multivariate Statistical Analysis.* Oliver & Boyd, Edinburgh, Scotland, x+642 pp., ISBN 0–05–002548–1. Reprinted by Halsted Press, New York, 1972, ISBN 0–470–02650–2 & by R. E. Krieger, Huntington, New York, 1997, ISBN 0–88275–477–7. [MR56: 1585; Zbl 263.62001, 421.62033.]

2. S. Puntanen & G. P. H. Styan (1988). *A Personal Guide to the Literature in Matrix Theory for Statistics and Some Related Topics.* Report A 205, Dept. of Mathematical Sciences, University of Tampere, iii+157 pp., December 1988, ISBN 951–44–2385–2, ISSN 0356–3134.

3. G. P. H. Styan, ed. (1990). *Abstracts of Papers Presented in Uppsala, Sweden, 13–18 August 1990 (2nd World Congress of the Bernoulli Society for Mathematical Statistics and Probability, 53rd Annual Meeting of the Institute of Mathematical Statistics).* Bernoulli Society for Mathematical Statistics and Probability & Institute of Mathematical Statistics, [iii +] 217 pp.

4. G. P. H. Styan, ed. (1990). *The Collected Papers of T. W. Anderson: 1943–1985. With commentaries.* John Wiley & Sons, New York, vol. 1: xlvi + 825 pp., vol. 2: pp. i–viii & 827–1681, ISBN 0–471–62442–5. [MR 91j: 01064.]

5. G. P. H. Styan, ed. (1998). *Three Bibliographies and a Guide.* Prepared for the Seventh International Workshop on Matrices and Statistics, in celebration of T. W. Anderson's 80th birthday (Fort Lauderdale, Florida, December 1998). 100 pp. [Includes: A bibliography on the distribution of quadratic forms in normal variables, with special emphasis on the Craig–Sakamoto theorem and on Cochran's theorem (with M. Dumais), pp. 1–9; A bibliography on the Laguerre–Samuelson inequality and on some related inequalities (with S. T. Jensen), pp. 10–16; A third bibliography

on the Frucht–Kantorovich inequality and on some related inequalities (with G. Alpargu), pp. 17–26.

6. E. P. Liski, J. Niemelä, J. Isotalo, S. Puntanen & G. P. H. Styan, eds. (2006). *Festschrift for Tarmo Pukkila on his 60th Birthday*, Dept. of Mathematics, Statistics and Philosophy, University of Tampere, 383 pp., ISBN–13:978–951–44–6620–5, ISBN–10:951–44–6620–9.

7. S. Puntanen, G. P. H. Styan & J. Isotalo (2011). *Matrix Tricks for Linear Statistical Models: Our Personal Top Twenty.* Springer, xvii+486 pages. ISBN: 978–3–642–10472–5. Expanded version of *Matrix Tricks for Linear Statistical Models: Our Personal Top Sixteen*, 3rd Edition. Report A 363, Dept. of Mathematics, Statistics & Philosophy, University of Tampere, Tampere, Finland, 207 pp., December 2005. [Original version by S. Puntanen: Report A 302, May 1996; 2nd edition by S Puntanen & G. P. H. Styan, Report A 330, Dept. of Mathematics, Statistics & Philosophy, University of Tampere.]

8. Y. Tian & G. P. H. Styan. *Rank Equalities and Inequalities Related to Generalized Inverses and Their Applications.* Research monograph in progress.

9. S. Wang, G. P. H. Styan, Z. Jia & Y. Tian. *Inequalities in Matrix Theory.* Translation from the Chinese (Anhui Educational Press, viii + 340 pp., 1994, ISBN 7–5336–1386–4), in progress.

10. S. Puntanen, G. P. H. Styan & J. Isotalo (2012). *Formulas Useful for Linear Regression Analysis and Related Matrix Theory: It's Only Formulas But We Like Them*, v+125 pages. Springer. To appear.

11. S. Puntanen, G. P. H. Styan & J. Isotalo (2012). *Estimation, Prediction and Testing in Linear Models*, c. 125 pages. Springer. To appear.

Papers in peer–reviewed Journals and Collections/Edited Books

1. G. P. H. Styan & H. Smith, Jr. (1964). Markov chains applied to marketing. *J. Marketing Research 1(1)*, 50–55. [Translated into Spanish: "Cadenas de Markov aplicadas a marketing", mimeo, 13 pp.]

2. G. E. Sharpe & G. P. H. Styan (1965). Circuit duality and the general network inverse. *IEEE Transactions on Circuit Theory CT–12*, 22–27. [Abstract: *IEEE Spectrum*, 2 (1965), 135.]

3. G. E. Sharpe & G. P. H. Styan (1965). A note on the general network inverse. *IEEE Transactions on Circuit Theory CT–12*, 632–633.

4. R. L. Wolf, M. Mendlowitz, J. Roboz, G. P. H. Styan, P. Kornfeld & A. Weigl (1966). Treatment of hypertension with spironolactone: double–blind study. *J. Amer. Medical Assoc. 198*, 1143–1149. [Abstract: *Biometrics*, 23 (1967), 607.]

5. G. E. Sharpe & G. P. H. Styan (1967). A note on equicofactor matrices. *Proceedings of the IEEE*, 55, 1226–1227.

6. G. P. H. Styan (1970). Notes on the distribution of quadratic forms in singular normal variables. *Biometrika 57*, 567–572. [RZMat 1971#5 V134, STMA 14:128, Zbl 264:62006.]

7. F. Harary, B. Lipstein & G. P. H. Styan (1970). A matrix approach to nonstationary chains. *Oper. Res. 18*, 1168–1181. [MR 43:8127, Zbl 229:60045.]

8. G. Marsaglia & G. P. H. Styan (1972). When does rank(A+B) = rank(A) + rank(B)? *Canad. Math. Bull. 15*, 451–452. [MR 47:236, Zbl 252:15002.]

9. S. I. Grossman & G. P. H. Styan (1972). Optimality properties of Theil's BLUS residuals. *J. Amer. Statist. Assoc. 67*, 672–673. [MR 52:7030, Zbl 265:62020.]

10. T. Cacoullos & G. P. H. Styan (1973). A bibliography of discriminant analysis. In: T. Cacoullos (ed.), *Discriminant Analysis and Applications: Proceedings of the NATO Advanced Study Institute on Discriminant Analysis and Applications*, Athens, Greece, June 8–20, 1972. Academic Press, New York, pp. 375–434. [MR 57:17959, Zbl 297:62037.]

11. J. Keilson & G. P. H. Styan (1973). Markov chains and M–matrices: inequalities and equalities. *J. Math. Anal. Appl. 41*, 439–459. [MR 47:3422, Zbl 255:15016.]

12. G. H. Golub & G. P. H. Styan (1973). Numerical computations for univariate linear models. *J. Stat. Comput. Simul. 2*, 253–274. [MR 51:11840, Zbl 283:62062.]

13. G. P. H. Styan (1973). Hadamard products and multivariate statistical analysis. *Linear Algebra Appl. 6*, 217–240. [MR 47:6724, Zbl 255:15002.]

14. G. P. H. Styan (1973). When does least squares give the best linear unbiased estimate? In: D. G. Kabe & R. P. Gupta (eds.), *Multivariate Statistical Inference: Proceedings of the Research Seminar at Dalhousie University*, Halifax, Nova Scotia, March 23–25, 1972. North–Holland, Amsterdam & American Elsevier, New York, pp. 241–246. [MR 47:7861, 52:9518; Zbl 264:62028.]

15. G. H. Golub & G. P. H. Styan (1973). Some aspects of numerical computations for linear models. In: W. J. Kennedy (ed.), *Proceedings of the Computer Science and Statistics Seventh Annual Symposium on the Interface: Iowa State University*, Ames, October 18–19, 1973. Statistical Numerical Analysis and Data Processing Section, Statistical Laboratory, Iowa State University, Ames, Iowa, pp. 189–192.

16. N. J. Pullman & G. P. H. Styan (1973). The convergence of Markov chains with nonstationary transition probabilities and constant causative matrix. *Stochastic Process. Appl. 1*, 279–285. [MR 53:14665, Zbl 263:60026.]

17. D. A. S. Fraser, I. Guttman & G. P. H. Styan (1974). Serial correlation and distributions on the sphere. *Comm. Statist. Theory Methods 5*, 97–118. [MR 56:13460, Zbl 341:62023.]

18. G. Marsaglia & G. P. H. Styan (1974). Equalities and inequalities for ranks of matrices. *Linear Multilinear Algebra 2*, 269–292. [MR 52:5711, Zbl 297:15003.]

19. G. Marsaglia & G. P. H. Styan (1974). Rank conditions for generalized inverses of partitioned matrices. *Sankhyā A 36*, 437–442. [MR 52:5699, Zbl 309:15002.]

20. C. C. Paige, G. P. H. Styan & P. G. Wachter (1975). Computation of the stationary distribution of a Markov chain. *J. Stat. Comput. Simul. 4*, 173–186. [Zbl 331:60040.]

21. M. A. Giguère & G. P. H. Styan (1976). Comparisons between maximum likelihood and naïve estimators in a multivariate normal population with data missing on one variate. *Bulletin de l'Institut International de Statistique: Proceedings of the 40th Session (Warsaw, 1975) 46 (3)*, 303–308. [MR 58:8024, Zbl 398:62038.]

22. T. Mäkeläinen & G. P. H. Styan (1976). A decomposition of an idempotent matrix where nonnegativity implies idempotence and none of the matrices need be symmetric. *Sankhyā A 38*, 400–403. [MR 57:12554, Zbl 412:15018.]

23. T. Papaioannou, G. P. H. Styan & L. L. Ward (1976). A comparison of BMD, SAS and SPSS. [With discussion by J. P. Miller, J. W. Frane, J. B. Fry, N. Van Eck & S. Robinovitz, and by J. Goodnight, and with reply by the authors.] In: J. Horwich & E. R. Horwich (eds.), *SAS.ONE: Proceedings of First International S.A.S. Users Conference: Kissimee, Florida, January 26–28, 1976*. S.A.S. Users Group, Raleigh, North Carolina, pp. 361–397.

24. W. T. Dent & G. P. H. Styan (1978). Uncorrelated residuals from linear models. *J. Econometrics 7*, 211–225. [MR 80m:62064, Zbl 337:62081.]

25. M. A. Giguère & G. P. H. Styan (1978). Multivariate normal estimation with missing data on several variates. In: *Transactions of the Seventh Prague Conference on Information Theory, Statistical Decision Functions, Random Processes and of the Eighth European Meeting of Statisticians (Technical University, Prague, August 18–23, 1974)*, Academia [Publishing House of the Czechoslovak Academy of Sciences], Prague & D. Reidel, Dordrecht, vol. B, pp. 129–139. [MR 80a:62005, 80e:62037, Zbl 406:62033.]

26. F. Pukelsheim & G. P. H. Styan (1979). Nonnegative definiteness of the estimated dispersion matrix in a multivariate linear model. *Bulletin de l'Académie Polonaise des Sciences, Série des Sciences Mathématiques 27*, 327–330. [MR 80j:62050, Zbl 413:62035.]

27. I. S. Alalouf & G. P. H. Styan (1979). Estimability and testability in restricted linear models. *Mathematische Operationsforschung und Statistik Series Statistics 10*, 189–201. [MR 82g:62088, Zbl 417: 62055.]

28. H. Wolkowicz & G. P. H. Styan (1979). Extensions of Samuelson's inequality. *Amer. Statist. 33*, 143–144. [MR 80h:62038.]

29. I. S. Alalouf & G. P. H. Styan (1979). Characterizations of estimability in the general linear model. *Ann. Statist. 7*, 194–200. [MR 80g:62044, Zbl 398:62053.]

30. V. Seshadri & G. P. H. Styan (1980). Canonical correlations, rank additivity and characterizations of multivariate normality. In: B. Gyires (ed.), *Analytic Function Methods in Probability Theory: Proceedings of the Colloquium on the Methods of Complex Analysis in the Theory of Probabil-*

ity and Statistics held at the Kossuth L. University, Debrecen, Hungary, August 29–September 2, 1977. Colloquia Mathematica Societatis János Bolyai, vol. 21, János Bolyai, Budapest & North–Holland, Amsterdam, pp. 331–344. [MR 80m:60002, 81h:62089, Zbl 419:62049.]

31. H. Wolkowicz & G. P. H. Styan (1980). Bounds for eigenvalues using traces. *Linear Algebra Appl. 29*, 471–506. [MR 81k:15015, Zbl 435:15015 sic.]

32. H. Wolkowicz & G. P. H. Styan (1980). More bounds for eigenvalues using traces. *Linear Algebra Appl. 31*, 1–17. [MR 81k:15016, Zbl 434:15003.]

33. H. Wolkowicz & G. P. H. Styan (1980). Reply: Letter to the Editor "Extensions of Samuelson's inequality". *Amer. Statist. 34*, 250–251.

34. G. P. H. Styan (1981). On Lavoie's determinantal inequality. *Linear Algebra Appl. 37*, 77–80. [MR 82k:15009, Zbl 433:15006.]

35. M. C. Lewis & G. P. H. Styan (1981). Equalities and inequalities for conditional and partial correlation coefficients. In: M. Csörgő , D. A. Dawson, J. N. K. Rao & A. K. Md. E. Saleh (eds.), *Statistics and Related Topics: Proceedings of the International Symposium on Statistics and Related Topics: Ottawa, May 1980.* North–Holland, Amsterdam, pp. 57–65. [MR 83j:62005, 83k:62073; Zbl 506:62040.]

36. T. Mäkeläinen, K. Schmidt & G. P. H. Styan (1981). On the existence and uniqueness of the maximum likelihood estimate of a vector–valued parameter in fixed–size samples. *Ann. Statist. 9*, 758–767. [MR 83b:62053, Zbl 473:62004.]

37. G. P. H. Styan (1982). The Canadian Journal of Statistics/La Revue Canadienne de Statistique. In: S. Kotz, N. L. Johnson & C. B. Read (eds.), *Encyclopedia of Statistical Sciences, Volume 1: A to Circular Probable Error.* Wiley, New York, pp. 352–354.

38. J. M. Borwein, G. P. H. Styan & H. Wolkowicz (1982). Some inequalities involving statistical expressions: Solution to Problem 81–10 [posed] by L. V. Foster. *SIAM Review 24*, 340–342. [Reprinted in: M. S. Klamkin (ed.), *Problems in Applied Mathematics: Selections from SIAM Review.* SIAM, Philadelphia, 1990, pp. 373–375.]

39. T. W. Anderson & G. P. H. Styan (1982). Cochran's theorem, rank additivity and tripotent matrices. In: G. Kallianpur, P. R. Krishnaiah & J. K. Ghosh (eds.), *Statistics and Probability: Essays in Honor of C. R. Rao.* North–Holland, Amsterdam, 1–23. [MR 83h:15002, Zbl:62030.] Reprinted in The Collected Papers of T. W. Anderson: 1943–1985 (G. P. H. Styan, ed.), Wiley, New York, [vol. 2,] pp. 1307–1329 (1990). [MR 91j: 01064.]

40. G. P. H. Styan (1983). Generalized inverses. In: S. Kotz, N. L. Johnson & C. B. Read (eds.), *Encyclopedia of Statistical Sciences, Volume 3: Faà di Bruno's Formula to Hypothesis Testing.* Wiley, New York, pp. 334–337.

41. G. P. H. Styan (1983). On some inequalities associated with ordinary least squares and the Kantorovich inequality. In: *Festschrift for Eino*

Haikala on his Seventieth Birthday, Acta Universitatis Tamperensis, Series A, vol. 153, pp. 158–166. [MR 85g:62120.]

42. J. K. Merikoski, G. P. H. Styan & H. Wolkowicz (1983). Bounds for ratios of eigenvalues using traces. *Linear Algebra Appl. 55*, 105–124. [MR 85a:15019, Zbl 522:15008.]

43. F. Pukelsheim & G. P. H. Styan (1983). Convexity and monotonicity properties of dispersion matrices of estimators in linear models. *Scand. J. Statist. 10*, 145–149. [MR 85h:62092, Zbl 539:62078.]

44. G. P. H. Styan & A. Takemura (1983). Rank additivity and matrix polynomials. In: S. Karlin, T. Amemiya & L. A. Goodman (eds.), *Studies in Econometrics, Time Series, and Multivariate Statistics in Honor of Theodore W. Anderson*. Academic Press, New York, 545–558. [MR 85f:62004, 86e:15003; Zbl 586:15002.]

45. J. O. Ramsay, J. ten Berge & G. P. H. Styan (1984). Matrix correlation. *Psychometrika 49*, 403–423. [MR 86c: 62069, Zbl 581:62048.]

46. I. S. Alalouf & G. P. H. Styan (1984). Characterizations of the conditions for the ordinary least squares estimator to be best linear unbiased. In: Y. P. Chaubey & T. D. Dwivedi (eds.), *Topics in Applied Statistics: Proceedings of the Statistics '81 Canada Conference: Concordia University, Montréal, April–May 1981*. Concordia University, Montréal (Québec), pp. 331–344.

47. A. J. Scott & G. P. H. Styan (1985). On a separation theorem for generalized eigenvalues and a problem in the analysis of sample surveys. *Linear Algebra Appl. 70*, 209–224. [MR 87i:62100, Zbl 587: 62023.]

48. G. P. H. Styan (1985). Schur complements and linear statistical models. In: T. Pukkila & S. Puntanen (eds.), *Proceedings of the First International Tampere Seminar on Linear Statistical Models and their Applications: Tampere, Finland, August–September 1983*. Dept. of Mathematical Sciences, University of Tampere, pp. 37–75.

49. D. Latour & G. P. H. Styan (1985). Canonical correlations in the two–way layout. In: T. Pukkila & S. Puntanen (eds.), *Proceedings of the First International Tampere Seminar on Linear Statistical Models and their Applications: Tampere, Finland, August–September 1983*. Dept. of Mathematical Sciences, University of Tampere, pp. 225–243.

50. Y. Thibaudeau & G. P. H. Styan (1985). Bounds for Chakrabarti's measure of imbalance in experimental design. In: T. Pukkila & S. Puntanen (eds.), *Proceedings of the First International Tampere Seminar on Linear Statistical Models and their Applications: Tampere, Finland, August–September 1983*. Dept. of Mathematical Sciences, University of Tampere, pp. 323–347.

51. R. E. Hartwig & G. P. H. Styan (1986). On some characterizations of the "star" partial ordering for matrices and rank subtractivity. *Linear Algebra Appl. 82*, 145–161. [MR 88b:15014a, Zbl 603:15001.]

52. G. P. H. Styan (1986). Canonical correlations in the three–way layout. In: I. S. Francis, B. F. J. Manly & F. C. Lam (eds.), *Pacific Statistical*

Congress: Auckland, New Zealand, May 1985. Elsevier Science Publishers B. V., Amsterdam, pp. 433–438.

53. R. E. Hartwig & G. P. H. Styan (1987). Partially ordered idempotent matrices. In: T. Pukkila & S. Puntanen (eds.), *Proceedings of the Second International Tampere Conference in Statistics: Tampere, Finland, June 1987.* Dept. of Mathematical Sciences, University of Tampere, pp. 361–383.

54. D. Latour, S. Puntanen & G. P. H. Styan (1987). Equalities and inequalities for the canonical correlations associated with some partitioned generalized inverses of a covariance matrix. In: T. Pukkila & S. Puntanen (eds.), *Proceedings of the Second International Tampere Conference in Statistics: Tampere, Finland, June 1987.* Dept. of Mathematical Sciences, University of Tampere, pp. 541–553.

55. H. Wolkowicz & G. P. H. Styan (1988). Samuelson–Nair inequality. In: S. Kotz, N. L. Johnson & C. B. Read (eds.), *Encyclopedia of Statistical Sciences, Volume 8: Regressograms – St. Petersburg Paradox.* Wiley, New York, 258–259.

56. J. K. Baksalary, F. Pukelsheim & G. P. H. Styan (1989). Some properties of matrix partial orderings. *Linear Algebra Appl. 119*, 57–85; erratum: vol. 220, page 3 (1995). [MR 90h:15022, 96b:15032.]

57. G. P. H. Styan (1989). Three useful expressions for expectations involving a Wishart matrix and its inverse. In: Y. Dodge (ed.), *Statistical Data Analysis and Inference: Papers from the International Conference on Recent Developments in Statistical Data Analysis and Inference in Honor of C. Radhakrishna Rao held in Neuchâtel, August 21–24, 1989.* North–Holland, Amsterdam, 283–296. [MR 91i:62004, 92f:62072.]

58. S. Puntanen & G. P. H. Styan (1989). The equality of the ordinary least squares estimator and the best linear unbiased estimator [with comments by O. Kempthorne & by S. R. Searle and with "Reply" by the authors; further discussion in #61 below]. *Amer. Statist. 43*, 153–164. [MR 92e:62125.]

59. J. K. Baksalary, K. Nordström & G. P. H. Styan (1990). Löwner–ordering antitonicity of generalized inverses of Hermitian matrices. *Linear Algebra Appl. 127*, 171–182. Reprinted in *Contributions to the Comparison of Linear Models and to the Löwner–Ordering Antitonicity of Generalized Inverses* by K. Nordström, Tilastotieteellisiä Tutkimuksia [Statistical Studies] vol. 12, Finnish Statistical Society, Helsinki, x + 89 pp. (1990). [MR 91f:15014, 94g:62143; Zbl 697:15007.]

60. J. K. Baksalary, S. Puntanen & G. P. H. Styan (1990). A property of the dispersion matrix of the best linear unbiased estimator in the general Gauss–Markov model. *Sankhyā A 52*, 279–296. [MR 93f:62089, Zbl 727:62072.]

61. S. Puntanen & G. P. H. Styan (1990). "Reply" [to Letters to the Editor by R. W. Farebrother, R. Christensen & D. A. Harville on #58 above]. *Amer. Statist. 44*, 192–193.

62. J. K. Baksalary, S. Puntanen & G. P. H. Styan (1990). On T. W. Anderson's contributions to solving the problem of when the ordinary least–squares estimator is best linear unbiased and to characterizing rank additivity of matrices. In: G. P. H. Styan (ed.), *The Collected Papers of T. W. Anderson: 1943–1985*. Wiley, New York, vol. 2, pp. 1579–1591. [MR 91j:01064.]

63. K. J. Worsley, G. P. H. Styan & J. Bérubé (1991). Genstat ANOVA efficiency factors and canonical efficiency factors for non–orthogonal designs. *Genstat Newsletter 26*, 11–21.

64. J. Bérubé & G. P. H. Styan (1992). On certain inequalities for average efficiency factors associated with the three–way layout of experimental design. In: S. Schach & G. Trenkler (eds.), *Data Analysis and Statistical Inference: Festschrift in Honour of Prof. Dr. Friedhelm Eicker*. Josef Eul Verlag GmbH, Bergisch Gladbach, 421–434. [MR 94j:62003, 94k:62114; Zbl 789.62062.]

65. J. K. Baksalary & G. P. H. Styan (1993). Around a formula for the rank of a matrix product with some statistical applications. In: R. S. Rees (ed.), *Graphs, Matrices, and Designs: Festschrift in Honor of Norman J. Pullman on his Sixtieth Birthday. Lecture Notes in Pure and Applied Mathematics*, vol. 139, Marcel Dekker, New York, 1–18. [MR 93i:05002, 93m:15001; Zbl 850:62628.]

66. J. Bérubé & G. P. H. Styan (1993). Decomposable three–way layouts. *J. Statist. Planning Inference 36*, 311–322. [MR 94j:62160, Zbl 785.62080.]

67. J. Bérubé, R. E. Hartwig & G. P. H. Styan (1993). On canonical correlations and the degrees of non–orthogonality in the three–way layout. In: K. Matusita, M. L. Puri & T. Hayakawa, eds.), *Statistical Sciences and Data Analysis: Proceedings of the Third Pacific Area Statistical Conference: Makuhari (Chiba, Tokyo), Japan, December 11–13, 1991*. VSP International Science Publishers, Utrecht, The Netherlands, pp. 247–252. [MR 96a:62003, 96m:62117; Zbl 858:62045.]

68. J. K. Baksalary, J. Hauke & G. P. H. Styan (1994). On some distributional properties of quadratic forms in normal variables and on some associated matrix partial orderings. In: T. W. Anderson, K. T. Fang & I. Olkin (eds.), *Multivariate Analysis and its Applications. IMS Lecture Notes–Monograph Series*, Institute of Mathematical Statistics, Hayward, California, vol. 24, pp. 111–121. [MR 98e:62010.]

69. S. W. Drury, G. P. H. Styan & G. E. Subak–Sharpe (1994). On a fundamental upper limit for the open–circuit resistance measurable between any two terminals of a positive resistance network. In: *Proceedings of the 1994 IEEE International Symposium on Circuits and Systems: London, England, May 30–June 2, 1994*, vol. 5, pp. 17–20.

70. S. W. Drury & G. P. H. Styan (1995). The singular value decomposition of the square roots of the identity matrix: Solution to Problem 93.3.7 (proposed by R. W. Farebrother). *Econometric Theory 11*, 650–653.

71. S. K. Mitra, S. Puntanen & G. P. H. Styan (1995). Shorted matrices and their applications in linear statistical models: a review. In: E.-M. Tiit, T. Kollo & H. Niemi (eds.), *Multivariate Statistics and Matrices in Statistics: Proceedings of the Fifth Tartu Conference, Tartu–Pühajärve, Estonia, 23–28 May 1994. New Trends in Probability and Statistics*, vol. 3, VSP International Science Publishers, Zeist (Utrecht), The Netherlands & TEV Ltd., Vilnius, Lithuania, pp. 289–311. [MR 99h:62090.]

72. S. Puntanen & G. P. H. Styan (1996). An equivalence relation for two symmetric idempotent matrices: First Solution to Problem 95.3.3 (proposed by S. Liu & W. Polasek). *Econometric Theory 12*, 590–591.

73. F. Zhang & G. P. H. Styan (1996). An equivalence relation for two symmetric idempotent matrices: Second Solution to Problem 95.3.3 (proposed by S. Liu & W. Polasek). *Econometric Theory 12*, 591–592.

74. S. Puntanen & G. P. H. Styan (1996). Matrix results associated with Aitken's generalization of the Gauss–Markov theorem: Solution to Problem 95.3.5 (proposed by R. W. Farebrother). *Econometric Theory 12*, 593–595.

75. S. Puntanen & G. P. H. Styan (1996). The Moore–Penrose generalized inverse of a symmetric matrix: Solution to Problem 95.4.3 (proposed by R. W. Farebrother). *Econometric Theory 12*, 748–749.

76. S. Puntanen & G. P. H. Styan (1996). A brief biography and appreciation of Calyampudi Radhakrishna Rao, with a bibliography of his books and papers. *Linear Algebra Appl. 237/238*, 1–40. [MR 1382662, Zbl 846:01017.]

77. J. E. Pečarić, S. Puntanen & G. P. H. Styan (1996). Some further matrix extensions of the Cauchy–Schwarz and Kantorovich inequalities, with some statistical applications. *Linear Algebra Appl. 237/238*, 455–476. [MR 97c:15035, Zbl 860:15021.]

78. J. K. Baksalary, P. Šemrl & G. P. H. Styan (1996). A note on rank additivity and range additivity. *Linear Algebra Appl. 237/238*, 489–498. [MR 97b:15026, Zbl 856.47001.]

79. R. E. Hartwig, M. Omladič, P. Šemrl & G. P. H. Styan (1996). On some characterizations of pairwise star orthogonality using rank and dagger additivity and subtractivity. *Linear Algebra Appl. 237/238*, 499–507. [MR 97c:15004, Zbl 848.15013.]

80. G Alpargu & G. P. H. Styan (1996). Some remarks and a bibliography on the Kantorovich inequality. In: A. K. Gupta & V. L. Girko (eds.), *Multidimensional Statistical Analysis and Theory of Random Matrices: Proceedings of the Sixth Eugene Lukacs Symposium, Bowling Green, OH, USA, March 29–30 1996.* VSP International Science Publishers, Zeist (Utrecht), The Netherlands, pp. 1–13. [MR 98b:62002, 98h:15033; Zbl 879:60015.]

81. S. Puntanen, P. Šemrl & G. P. H. Styan (1996). Some remarks on the parallel sum of two matrices. In: L. Kavalieris, F. C. Lam, L. A. Roberts & J. A. Shanks (eds.), *Proceedings of the A. C. Aitken Centenary Conference (incorporating the 3rd Pacific Statistical Congress, the annual meeting of the New Zealand Statistical Association and the 1995 New Zealand Mathematics Colloquium, 28 August–1 September 1995): Otago Conference Series No. 5*. University of Otago Press, Dunedin, New Zealand, pp. 243–256.

82. S. Puntanen & G. P. H. Styan (1997). Orthogonal projectors: Solution to Problem 96.4.3 (proposed by J. Groß & G. Trenkler). *Econometric Theory 13*, 764–765.

83. G. P. H. Styan & G. E. Subak–Sharpe (1997). Inequalities and equalities associated with the Campbell–Youla generalized inverse of the indefinite admittance matrix of resistive networks. *Linear Algebra Appl. 250*, 349–370. [MR 97k:94095, Zbl 867:15003.]

84. G. S. Watson, G. Alpargu & G. P. H. Styan (1997). Some comments on six inequalities associated with the inefficiency of ordinary least squares with one regressor. *Linear Algebra Appl. 264*, 13–53. [MR 98i:15023, Zbl 948.62046.]

85. S. Puntanen & G. P. H. Styan (1998). A fundamental matrix result on scaling in multivariate analysis: Solution to Problem 97.5.3 (proposed by H. Neudecker, A. Satorra & M. van de Velden). *Econometric Theory 14*, 693–695.

86. S. Puntanen, G. P. H. Styan & G. E. Subak–Sharpe (1998). Mahalanobis distance for multinomial data: Solution to Problem 97.5.4 (proposed by H. Neudecker). *Econometric Theory 14*, 695–698.

87. M. Nurhonen, S. Puntanen, G. P. H. Styan & H. Yanai (1998). Simplified matrix proofs related to the deletion of an observation in [the] general linear model. In: S. P. Mukherjee, S. K. Basu & B. K. Sinha (eds.), *Frontiers in Probability and Statistics*. Narosa Publishing House, New Delhi, pp. 267–275. [One of 38 selected papers (out of 101) presented at the Second International Triennial Calcutta Symposium on Probability and Statistics (Calcutta, India, December 30, 1994–January 2, 1995). MR 2000a:62156, Zbl 926.62056.]

88. G. Alpargu, S. W. Drury & G. P. H. Styan (1998). Some remarks on the Bloomfield–Watson–Knott inequality and on some other inequalities related to the Kantorovich inequality. In: *Proceedings of the Conference in Honor of Shayle R. Searle, August 9–10, 1996*. Biometrics Unit, Cornell University, Ithaca, New York, pp. 125–143.

89. S. T. Jensen & G. P. H. Styan (1999). Some comments and a bibliography on the Laguerre–Samuelson inequality with extensions and applications to statistics and matrix theory. In: T. M. Rassias & H. M. Srivastava (eds.), *Analytic and Geometric Inequalities and Applications. Mathematics and Its Applications*, Volume 478, Kluwer Academic Publishers, Dordrecht, pp. 151–181. [MR 2001h:15013; Zbl 0980.15016.]

90. G. P. H. Styan & H. J. Werner (1999). Upper bounds for eigenvalues of nonnegative definite matrices: Solution to Problem 98.2.2 (proposed by E. I. Im). *Econometric Theory 15*, 261–262.

91. S. Puntanen, G. P. H. Styan & H. J. Werner (1999). A determinantal identity: Solution to Problem 98.4.1 (proposed by H. Neudecker & M. van de Velden). *Econometric Theory 15*, 632–633.

92. S. Kaçiranlar, S. Sakallioğlu, F. Akdeniz, G. P. H. Styan & H. J. Werner (1999). A new biased estimator in linear regression and a detailed analysis of the widely–analysed dataset on Portland cement. *Sankhyā B 61*, 443–459.

93. G. P. H. Styan (1999). Comment on "A history of the Statistical Society of Canada: the formative years" [by D. R. Bellhouse & C. Genest, *Statistical Science 14*, 80–125, 1999]. *Statist. Sci. 14*, page 125.

94. S. Puntanen, G. P. H. Styan & H J. Werner (2000). The eigenvalue decomposition of a symmetric matrix: Solution 1 to Problem 99.3.1 (proposed by R. W. Farebrother). *Econometric Theory 16*, 289–294. [Solution 2 by G. Dhaene is on pp. 292–294; combined references are on page 294.]

95. G. Alpargu & G. P. H. Styan (2000). Some comments and a bibliography on the Frucht–Kantorovich and Wielandt inequalities. In: R. D. H. Heijmans, D. S. G. Pollock & A. Satorra (eds.), *Innovations in Multivariate Statistical Analysis: A Festschrift for Heinz Neudecker.* Kluwer Academic Publishers, Dordrecht, pp. 1–38.

96. S. Puntanen, G. P. H. Styan & H. J. Werner (2000). Two matrix–based proofs that the linear estimator Gy is the best linear unbiased estimator. *J. Statist. Plann. Inference 88*, 173–179. [MR 2001h:62120; Zbl 0964.62054.]

97. G. E. Subak–Sharpe & G. P. H. Styan (2000). A necessary condition for the realization of a resistive n–port based on network size and on the concept of weighted terminal valency. *Proceedings of the ISCAS 2000–IEEE International Symposium on Circuits and Systems* (Geneva, Switzerland, May 28–31, 2000), vol. I, pp. I–487–I–490.

98. S. Puntanen, G. P. H. Styan & H. J. Werner (2000). Letter to the Editor about "Simple forms of the best linear unbiased predictor in the general linear regression model" by S. N. Elian [*Amer. Statist. 54*, 25–28, 2000]. *Amer. Statist. 54*, 326–327.

99. S. Puntanen, G. P. H. Styan & H. J. Werner (2001). Determinant of a skew–symmetric matrix: Solution 1 to Problem 00.1.1 (proposed by S. Lawford). *Econometric Theory 17*, 277.

100. Y. Tian & G. P. H. Styan (2001). How to establish universal block–matrix factorizations. *Electron. J. Linear Algebra 8*, 115–127. http://www.math.technion.ac.il/iic/ela [MR 2002f:15017; Zbl 0979.15012.]

101. Y. Tian & G. P. H. Styan (2001). Rank equalities for idempotent and involutory matrices. *Linear Algebra Appl. 335*, 101–117. [MR 2002f:15001; Zbl 0988.15002.]

102. Y. Tian & G. P. H. Styan (2002). A new rank formula for idempotent matrices with applications. *Comment. Math. Univ. Carolin. 43*, 379–384. [MR 2003f:15005.]

103. Y. Tian & G. P. H. Styan (2002). When does rank(ABC)= rank(AB) + rank(BC) − rank(B) hold? *Internat. J. Math. Ed. Sci. Tech. 33*, 127–137. [MR 1880569, Zbl 1015.15001.]

104. G. P. H. Styan (2002). Harold Ruben: 1923–2001. *J. Roy. Statist. Soc. Ser. D Statist. 106*, 568–570.

105. J. K. Baksalary, O. M. Baksalary & G. P. H. Styan (2002). Idempotency of linear combinations of an idempotent matrix and a tripotent matrix. *Linear Algebra Appl. 354*, 21–34. [MR 2003h:15006, Zbl 1016.15027.]

106. J. K. Baksalary & G. P. H. Styan (2002). Generalized inverses of partitioned matrices in Banachiewicz–Schur form. *Linear Algebra Appl. 354*, 41–47. [MR1927646 (2003h:15006), Zbl 1022.15006.]

107. S. W. Drury, S. Liu, C.-Y. Lu, S. Puntanen & G. P. H. Styan (2002). Some comments on several matrix inequalities with applications to canonical correlations: historical background and recent developments. *Sankhyā A 64*, 453–507. [MR1981768 (2004e:62111).]

108. G. P. H. Styan & H. J. Werner (2003). A particular symmetric idempotent matrix: solution to Problem 02.1.2 (proposed by H. Neudecker). *Econometric Theory 19*, 227–228.

109. K. L. Chu, J. Isotalo, S. Puntanen & G. P. H. Styan (2004). On decomposing the Watson efficiency of ordinary least squares in a partitioned weakly singular linear model. *Sankhyā 66*, 634–651. [MR 2205814.]

110. S. Puntanen & G. P. H. Styan (2004). Historical introduction: Issai Schur and the early development of the Schur complement. Chapter 0 and Bibliography in: F. Zhang (ed.), *The Schur Complement and Its Applications*. Springer Science+Business Media, pp. 1–16, 259–288.

111. S. Puntanen & G. P. H. Styan (2004). Schur complements in statistics and probability. Chapter 6 and Bibliography in: F. Zhang (ed.), *The Schur complement and Its Applications*. Springer Science+Business Media, pp. 163–226, 259–288.

112. S. Puntanen, G. P. H. Styan & Y. Tian (2005). Three rank formulas associated with the covariance matrices of the BLUE and the OLSE in the general linear model. *Econometric Theory 21*, 659–663. [MR 2162764 (2006g:62066), Zbl 1072.62049.]

113. O. M. Baksalary & G. P.H. Styan, eds. (2005). Some comments on the life and publications of J. K. Baksalary (1944–2005). *Linear Algebra Appl. 410*, 3–53.

114. Y. Tian & G. P. H. Styan (2005). Cochran's statistical theorem for outer inverses of matrices and matrix quadratic forms. *Linear Multilinear Algebra 53*, 387–392. [MR 2156647 (2006d:62058), Zbl 1083.15007.]

115. K. L. Chu, J. Isotalo, S. Puntanen & G. P. H. Styan (2005). Some further results concerning the decomposition of the Watson efficiency in partitioned linear models. *Sankhyā 67*, 74–89. [MR 2204850.]

116. F. Akdeniz, G. P. H. Styan & H. J. Werner (2006). The general expressions for the moments of the stochastic shrinkage parameters of the Liu–type estimator. *Comm. Statist. Theory Methods 35*, 423–437. [Zbl 1084.62046.]

117. S. Puntanen & G. P. H. Styan (2006). Some comments about Issai Schur (1875–1941) and the early history of Schur complements. In: P. Brown, S. Liu & D. Sharma (eds.), *Contributions to Probability and Statistics: Applications and Challenges – Proceedings of the International Statistics Workshop, University of Canberra, 4–5 April 2005.* World Scientific, Singapore, pp. 28–66.

118. J. Isotalo, S. Puntanen & G. P. H. Styan (2006). Matrix tricks for linear statistical models: a short review of our personal top fourteen. In: P. Brown, S. Liu & D. Sharma (eds.), *Contributions to Probability and Statistics, Applications and Challenges: Proceedings of the International Statistics Workshop, University of Canberra, 4–5 April 2005.* World Scientific, Singapore, pp. 113–128.

119. S. Puntanen & G. P. H. Styan (2006). A conversation with Tarmo Mikko Pukkila. In: E. P. Liski, J. Niemelä, J. Isotalo, S. Puntanen & G. P. H. Styan (eds.), *Festschrift for Tarmo Pukkila on his 60th Birthday.* Dept. of Mathematics, Statistics and Philosophy, University of Tampere, pp. 13–44. [Zbl 1138.01339.]

120. S. Puntanen & G. P. H. Styan (2006). Some comments on the research publications of Tarmo Mikko Pukkila. In: E. P. Liski, J. Niemelä, J. Isotalo, S. Puntanen & G. P. H. Styan (eds.), *Festschrift for Tarmo Pukkila on his 60th Birthday.* Dept. of Mathematics, Statistics and Philosophy, University of Tampere, pp. 45–62. [Zbl 1138.01340.]

121. J. Isotalo, S. Puntanen & G. P. H. Styan (2006). On the role of the constant term in linear regression. In: E. P. Liski, J. Niemelä, J. Isotalo, S. Puntanen & G. P. H. Styan (eds.), *Festschrift for Tarmo Pukkila on his 60th Birthday.* Dept. of Mathematics, Statistics and Philosophy, University of Tampere, pp. 243–259. [Zbl 1145.62350.]

122. Y. Tian & G. P. H. Styan (2006). Rank equalities for idempotent matrices with applications. *J. Comput. Appl. Math. 191*, 77–97. [MR 2217786, Zbl pre05024169.]

123. Y. Tian & G. P. H. Styan (2006). Cochran's statistical theorem revisited. *J. Statist. Plann. Inference 136*, 2659–2667. [Zbl pre05037741.]

124. S. Puntanen & G. P. H. Styan (2007). Chapter 52: Random vectors and linear statistical models. In: L Hogben (ed.), *Handbook of Linear Algebra.* Chapman & Hall/CRC, Boca Raton, pp. 52.1–52.17.

125. S. Puntanen, G. A. F. Seber & G. P. H. Styan (2007). Chapter 53: Multivariate statistical analysis. In: L Hogben (ed.), *Handbook of Linear Algebra.* Chapman & Hall/CRC, Boca Raton, pp. 53.1–53.15.

126. J. Isotalo, S. Puntanen & G. P. H. Styan (2007). Effect of adding regressors on the equality of the OLSE and BLUE. *Int. J. Statist. Sci.*

6, 193–201. [Invited paper in the Second Special Issue in Felicitation of Professor Mir Masoom Ali on the Occasion of his 70th Birthday.]

127. G. P. H. Styan & G. Trenkler (2007). A philatelic excursion with Jeff Hunter in probability and matrix theory. *J. Appl. Math. Decis. Sci.*, 2007, article ID 13749, 10 pp., doi:10.1155/2007/13749. (Invited paper in the Special Issue on Statistics and Applied Probability: A Tribute to Jeffrey J. Hunter; G. C. Wake & P. Cowpertwait, eds.) [Zbl pre05304408.]

128. K. L. Chu, J. Isotalo, S. Puntanen & G. P. H. Styan (2007). The efficiency factorization multiplier for the Watson efficiency in partitioned linear models: some examples and a literature review. *J. Statist. Plann. Inference 137*, 3336–3351. (Invited paper in the Special Issue in Celebration of the Centennial of the Birth of Samarendra Nath Roy (1906–1964), G. S. Mudholkar, A. D. Hutson & M. P. McDermott, eds.) [Zbl 1119.62065.]

129. G. P. H. Styan (2007). A philatelic introduction to magic squares and Latin squares for Euler's 300th birthyear. In: A. Cupillari (ed.), *Proceedings of the Canadian Society for History and Philosophy of Mathematics/Société Canadienne d'Histoire et de Philosophie des Mathématiques*, vol. 20, pp. 306–319. [ISSN 0825–5924. Contributed paper at the 32nd Annual Meeting, Montreal, July 27–29, 2007.]

130. S. Puntanen & G. P. H. Styan (2008). Stochastic stamps: a philatelic introduction to chance. *Chance 21(3)*, 36–41. [MR 2507100.]

131. J. Isotalo, S. Puntanen & G. P. H. Styan (2008). A useful matrix decomposition and its statistical applications in linear regression. *Comm. Statist. Theory Methods 37(8–10)*, 1436–1457. [MR 2440447 (2009m:62211).]

132. O. M. Baksalary & G. P. H. Styan (2008). Some comments on the life and publications of J. K. Baksalary (1944–2005). *Discuss. Math. Probab. Stat. 28(1)*, 5–64. (Invited paper in the Special Issue in Honour of J. K. Baksalary.) [MR 2475197.]

133. O. M. Baksalary & G. P. H. Styan (2008). Some comments on the diversity of Vermeer paintings depicted on postage stamps. *Discuss. Math. Probab. Stat. 28(1)*, 65–83. (Invited paper in the Special Issue in Honour of J. K. Baksalary.) [MR 2475198.]

134. S. Puntanen & G. P. H. Styan (2008). Foreword [to the Special Issue in Honour of J. K. Baksalary]. *Discuss. Math. Probab. Stat. 28(1)*, 85–90. [MR 2475199.]

135. J. Isotalo, S. Puntanen & G. P. H. Styan (2008). Decomposing matrices with J. K. Baksalary. *Discuss. Math. Probab. Stat. 28(1)*, 91–111. [MR 2475200 (2010b:62254).]

136. K. L. Chu, J. Isotalo, S. Puntanen & G. P. H. Styan (2008). Inequalities and equalities for the generalized efficiency function in orthogonally partitioned linear models. In: T. M. Rassias & D. Andrica (eds.), *Inequalities and Applications*. Cluj University Press, Cluj, Romania, pp. 13–69.

137. C. C. Paige, G. P. H. Styan, B. Y. Wang & F. Zhang (2008). Hua's matrix equality and Schur complements. *Int. J. Inf. Syst. Sci. 4(1)*, 124–135. [MR 2401768 (2010f:15037), Zbl pre05347976.]

138. J. Isotalo, S. Puntanen & G. P. H. Styan (2008). The BLUE's covariance matrix revisited: a review. *J. Statist. Plann. Inference 138(9)*, 2722–2737. [Invited paper in the Special Issue in Honor of Theodore Wilbur Anderson, Jr., on the Occasion of his 90th Birthday; T. L. Lai, I. Olkin & R. Velu, eds. MR 2422395, Zbl 1141.62325.]

139. J. Isotalo, S. Puntanen & G. P. H. Styan (2009). Some comments on the Watson efficiency of the ordinary least squares estimator under the Gauss–Markov model. *Calcutta Statist. Assoc. Bull. 61*, 1–15. [MR 2554134 (2010k:62278).]

140. O. M. Baksalary, K. L. Chu, S. Puntanen & G. P. H. Styan (2009). Some comments on Fisher's α index of diversity and on the *Kazwini Cosmography.* In: B. Schipp & W. Krämer (eds.), *Statistical Inference, Econometric Analysis and Matrix Algebra: Festschrift in Honour of Götz Trenkler*. Physica-Verlag, Heidelberg, pp. 369–394.

141. S. E. Ahmed, J. J. Hunter, G. P. H. Styan & G. Trenkler (2009). Preface to the Proceedings of the 16th International Workshop on Matrices and Statistics, Windsor 2007. Held at the University of Windsor, Windsor, ON, June 1–3, 2007. *Linear Algebra Appl. 430(10)*, 2563–2565. [MR 2509840.]

142. K. E. Gustafson & G. P. H. Styan (2009). Superstochastic matrices and magic Markov chains. *Linear Algebra Appl. 430(10)*, 2705–2715. [MR 2509852 (2010c:15031).]

143. Y. Tian & G. P. H. Styan (2009). On some matrix equalities for generalized inverses with applications. *Linear Algebra Appl. 430(10)*, 2716–2733. [MR 2509853 (2010a:15055).]

144. O. M. Baksalary, G. P. H. Styan & G. Trenkler (2009). On a matrix decomposition of Hartwig and Spindelböck. *Linear Algebra Appl. 430(10)*, 2798–2812. [MR 2509859 (2010d:15022).]

145. K. L. Chu, S. Puntanen & G. P. H. Styan (2009). Some comments on philatelic Latin squares from Pakistan. *Pakistan J. Statist. 25*, 427–471. [MR 2750609.]

146. G. P. H. Styan, C. Boyer & K. L. Chu (2009). Some comments on Latin squares and on Graeco–Latin squares, illustrated with postage stamps and old playing cards. *Statist. Papers 50*, 917–941. [MR 2551361 (2010i:05059).]

147. A. Markiewicz, S. Puntanen & G. P. H. Styan (2010). A note on the interpretation of the equality of OLSE and BLUE. *Pakistan J. Statist. 26*, 127–134. [MR 2756730.]

148. S. Puntanen & G. P. H. Styan (2010). Best linear unbiased estimation in linear models. *StatProb: The Encyclopedia Sponsored by Statistics and Probability Societies*. Available at http://statprob.com /encyclopedia/BestLinearUnbiasedEstimatinInLinearModels.html

149. P. D. Loly & G. P. H. Styan (2010). Comments on 4×4 philatelic Latin squares. *Chance: A Magazine for People Interested in the Analysis of Data 23(1)*, 57–62.

150. P. D. Loly & G. P. H. Styan (2010). Comments on 5×5 philatelic Latin squares. *Chance: A Magazine for People Interested in the Analysis of Data 23(2)*, 58–62.

151. P. D. Loly & G. P. H. Styan (2010/2011). Philatelic Latin squares. In: *Proceedings of the Canadian Society for History and Philosophy of Mathematics/Société Canadienne d'Histoire et de Philosophie des Mathématiques 23*, 273–297.

152. K. L. Chu, S. W. Drury, G. P. H. Styan & G. Trenkler (2011). Magic Moore–Penrose inverses and philatelic magic squares with special emphasis on the Daniels–Zlobec magic square. *Croatian Operational Research Review 2*, 4–13.

153. K. L. Chu, S. Puntanen & G. P. H. Styan (2011). Solution to Problem 1/SP09 "Inverse and determinant of a special symmetric matrix" (Problem proposed by H. Neudecker, G. Trenkler & S. Liu). *Statist. Papers 52*, 258–260.

154. S. Puntanen & G. P. H. Styan (2011). Best linear unbiased estimation in a linear model. In: M. Lovric (ed.), *International Encyclopedia of Statistical Science*. Springer, Part 2, pp. 141–144. ISBN: 978–3–642–04897–5.

155. S. Puntanen, G. P. H. Styan & J. Isotalo (2012). Matrix tricks for linear statistical models: a quick look at our personal top ones. In: R. B. Bapat, S. J. Kirkland, K. M. Prasad & S. Puntanen (eds.), *Lectures on Matrix and Graph Methods*. Manipal University Press, pp. 91–112.

156. G. P. H. Styan, G. Trenkler & K. L. Chu (2012). An introduction to Yantra magic squares and Agrippa–Cardano type magic matrices: Lecture notes. In: (R. B. Bapat, S. J. Kirkland, K. M. Prasad & S. Puntanen (eds.), *Lectures on Matrix and Graph Methods*. Manipal University Press, pp. 159–220. [Based on invited talk given at the International Workshop and Conference on Combinatorial Matrix Theory and Generalized Inverses of Matrices, Manipal University, Manipal (Karnataka), India, 2–7 & 10–11 January 2012: video (updated 17 February 2011) online http://vimeo.com/37284121 at Vimeo. Expanded version: Report 2011-07, Dept. of Mathematics and Statistics, McGill University, 129 pp., 13 June 2012.]

157. S. Puntanen & G. P. H. Styan (2012). A conversation with Sujit Kumar Mitra in 1993. In: (R. B. Bapat, S. J. Kirkland, K. M. Prasad & S. Puntanen (eds.), *Lectures on Matrix and Graph Methods*. Manipal University Press, pp. 221–244.

158. G. P. H. Styan (2012). Caïssan squares: the magic of chess. Special talk presented at *The 9th Tartu Conference on Multivariate Statistics & The 20th International Workshop on Matrices and Statistics*, Tartu, Estonia, 26 June–1 July 2011. Accepted for publication in *Acta Comment. Univ. Tartu. Math.*, 36 pp.

159. S. Puntanen & G. P. H. Styan (2012). Chapter 52: Random Vectors and Linear Statistical Models. In: L. Hogben (ed.), *Handbook of Linear Algebra*, 2nd Edition. Chapman & Hall (in press).

160. S. Puntanen, G. A. F. Seber & G. P. H. Styan (2012). Chapter 53: Multivariate Statistical Analysis. In: L. Hogben (ed.), *Handbook of Linear Algebra*, 2nd Edition. Chapman & Hall (in press).

161. G. P. H. Styan (2012). An illustrated introduction to some magic squares from India. Invited paper for: R. B. Bapat, S. J. Kirkland, K. M. Prasad & S. Puntanen (eds.), *Combinatorial Matrix Theory and Generalized Inverses of Matrices*. Springer, in progress. [Based on talks given at the Annual Meeting of the Canadian Society for the History and Philosophy of Mathematics (CSHPM), University of Waterloo, 27–29 May 2012, and at the International Workshop and Conference on Combinatorial Matrix Theory and Generalized Inverses of Matrices, Manipal University, Manipal (Karnataka), India, 2–7 & 10–11 January 2012: video (updated 17 February 2011) online http://vimeo.com/37291712 at Vimeo.]

162. G. E. Subak-Sharpe, S. W. Drury & G. P. H. Styan (2012). Some comments on the properties of the impedance matrices of resistive electrical networks and on the n–port problem. Invited paper for publication in: *Measurement Science Review*. Preprint: 15 pp., 3 June 2012. [Based on talk given by G. E. Subak-Sharpe at the 18th International Workshop on Matrices and Statistics (IWMS-18), Smolenice Castle, Slovakia, 23–27 June 2009.]

163. G. Trenkler & G. P. H. Styan (2012). A purely mathematical treatment of the problem of magic squares with 16 and 64 cells, by F. Fitting [Friedrich Fitting (1862–1945)]. Translated from the German "Rein mathematische Behandlung des Problems der magischen Quadrate von 16 und von 64 Feldern" [*Jahresbericht der Deutschen Mathematiker-Vereinigung,* vol. 40, pp. 177–199 (1931)] by G. Trenkler; edited and commentary by G. P. H. Styan. Report 2012-02: Department of Mathematics and Statistics, McGill University, Montréal, 109 pp., 4 June 2012.

164. M. A. Amela, K. L. Chu, A. Memartoluie, G. P. H. Styan & G. Trenkler (2012). An illustrated introduction to Euler and Fitting factorizations and Anderson graphs for classic magic matrices. [In preparation for presentation as an invited talk in the Special Session to celebrate G. P. H. Styan's 75th Birthday at the International Conference on Trends and Perspectives in Linear Statistical Inference (LINSTAT-2012) & 21st International Workshop on Matrices and Statistics (IWMS-2012), Mathematical Research and Conference Center of the Institute of Mathematics of the Polish Academy of Sciences, Będlewo (near Poznań), Poland, 16–20 July 2012.]

Biographical Publications, including Obituaries

1. G. P. H. Styan, ed. (1987). In Memoriam: George L[ewis] Edgett (1900–1986). *The IMS Bulletin 16*, 95.

2. G. P. H. Styan, ed. (1987). In Memoriam: John Van Ryzin (1935–1987). *The IMS Bulletin 16*, 158.

3. G. P. H. Styan, ed. (1987). In Memoriam: Paruchuri R[ama] Krishnaiah (1932–1987). *The IMS Bulletin 16*, 264.

4. G. P. H. Styan, ed. (1987). The Emperor [Joseph II: 1741–1790]. *The IMS Bulletin 16*, 305–306. [Reprinted in *The IMS Bulletin 21* (1992), 642–643.]

5. G. P. H. Styan, ed. (1988). Louis Guttman: 1916–1987 [Obituary]. *The IMS Bulletin 17*, 284.

6. G. P. H. Styan, ed. (1988). The Reverend Thomas Bayes, F.R.S. 1701?–1761. *The IMS Bulletin 17*, 276–278, 482–483, *20* (1991), 226. [Reprinted in *The IMS Bulletin 21* (1992), 644–649.]

7. G. P. H. Styan (1989). Linnaeus and Celsius. *The IMS Bulletin 18*, 326. [Reprinted in *The IMS Bulletin 21* (1992), 217.]

8. G. P. H. Styan, ed. (1990). Lord [Nathaniel Mayer Victor] Rothschild, GBE, GM, FRS: 1910–1990 [Obituary]. *The IMS Bulletin 19*, 501.

9. G. P. H. Styan, ed. (1991). Thornton Carl Fry: 1892–1991 [Obituary]. *The IMS Bulletin 20*, 2.

10. G. P. H. Styan, ed. (1991). Wassily Hoeffding: 1914–1991 [Obituary]. *The IMS Bulletin 20*, 98.

11. G. P. H. Styan, ed. (1991). Allen Thornton Craig: 1904–1978 [Obituary]. *The IMS Bulletin 20*, 389.

12. G. P. H. Styan, ed. (1991). The Allen T. Craig Lectures at the University of Iowa. *The IMS Bulletin 20*, 389–390.

13. G. P. H. Styan, ed. (1992). José Tiago da Fonseca Oliveira: 1928–1992 [Obituary]. *The IMS Bulletin 21*, 467.

14. G. P. H. Styan (1992). Six–Year Index to Obituaries, PhDs in the Statistical Sciences [and] Photographs [in *The IMS Bulletin 16–21* (1987–1992)]. *The IMS Bulletin 21*, 650–653.

15. G. P. H. Styan, ed. (1999). Feliks Ruvimovich Gantmakher and The Theory of Matrices. *Image 22* (April 1999), 12–13.

16. G. P. H. Styan, ed. (1999). More about W. Spottiswoode. *Image 23* (October 1999), 4–5.

17. G. P. H. Styan, ed. (1999). IMAGE Philatelic Corner [Takakazu Seki Kôwa]. *Image 23* (October 1999), 8.

18. R. W. Farebrother & G. P. H. Styan (2000). A genealogy of the Spottiswoode family: 1510–1900. *Image 25* (October 2000), 19–21.

19. R. W. Farebrother, S. T. Jensen & G. P. H. Styan (2000). Charles Lutwidge Dodgson: A biographical and philatelic note. *Image 25* (October 2000), pp. 22–23.

20. J. Grala, A. Markiewicz & G. P. H. Styan (2000). Tadeusz Banachiewicz: 1882–1954. *Image 25*, 24.

21. R. W. Farebrother & G. P. H. Styan (2001). Some observations on "A genealogy of the Spottiswoode family". *Image 27* (October 2001), 2.

22. R. W. Farebrother, S. T. Jensen & G. P. H. Styan (2002). Sir Thomas Muir and nineteenth–century books on determinants. *Image 28* (April 2002), 6–15.

23. G. P. H. Styan (2002). Harold Ruben: 1923–2001. [An updated version (3 June 2012) of the full obituary is online at http://www.math.mcgill.ca/styan/Ruben-3june12.pdf (with photographs and a complete bibliography). Slightly different versions of the original obituary published in *Amstat News #296*, 26–27, February 2002, *The IMS Bulletin 31(2)*, 17, March/April 2002, *The ISI Newsletter 26(1)*, issue #76, 8–9, February 2002, and in *SSC Liaison 16(1)*, 29–30, February 2002.]

24. G. P. H. Styan (2002). Biographies, portraits and stamps on the Web. *Image 28* (April 2002), 17.

25. R. W. Farebrother, G. P. H. Styan & G. J. Tee (2003). Gottfried Wilhelm von Leibniz: 1646–1716. *Image 30* (April 2003), 13–16.

26. S. Puntanen & G. P. H. Styan (2006). A conversation with Sujit Kumar Mitra in 1993 and some comments on his research publications. Report A 372, Dept. of Mathematics, Statistics and Philosophy, University of Tampere, 61 pp.

27. S. Puntanen & G. P. H. Styan (2006). A photo album for Tarmo Mikko Pukkila. In: E. P. Liski, J. Niemelä, J. Isotalo, S. Puntanen & G. P. H. Styan (eds.), *Festschrift for Tarmo Pukkila on his 60th Birthday.* Dept. of Mathematics, Statistics and Philosophy, University of Tampere, pp. 367–383. [Zbl 1138.01341.]

28. P. Macdonald, A. Tamhane & G. P. H. Styan (2007). Charles W. Dunnett (1921–2007). *SSC Liaison 21(2–3)*, p. E–23 (English), p. F–24 (French).

29. C. Genest, G. P. H. Styan & D. B. Wolfson (2009). Keith John Worsley (1951–2009). *SSC Liaison 23(2)*, 36–37 (in English and in French).

Celebrating George P. H. Styan's 75th birthday and my meetings with him

Carlos A. Coelho

Departamento de Matemática and Centro de Matemática e Aplicações, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Portugal

Cheers Geo(rge)!

The first time I met George Styan was in July 2004 in Lisbon when he was on his way to the 11th ILAS Conference in Coimbra.
But George had already been in Portugal before and I learned how much he was fond of Conventual, a very fine and nice old style restaurant in Lisbon. Then I also learned about George's taste for good food and good wine. With this detail in common it was really easy to become a good friend with George. Since then we met a number of times, the most significant of which was at the time of the 17th IWMS held in Tomar, Portugal, in 2008.
Before this event, during a short stay of George and Evelyn in Lisbon, we had the opportunity to go to some nice spots like Sintra, to hang around a few nice places near Lisbon and even to attend a Leonard Cohen concert, together with some friends.
It was at that time that when going for some beer, which we decided to 'convert' into a nice white wine that I took the picture in Figure 1 at Hennessy's in downtown Lisbon, not far from the Tagus river.

Figure 1 – George Styan and Evelyn at Hennessy's in Lisbon (2008)

By that time I had no idea that this picture came out so appropriately. It not only seems that indeed George is having one of his bright ideas (look at the

lamp that seems to sit on top of his head) but also the saying that 'behind a great man there is always a great woman' seems to be most adequate.
That wherever George is there is cheer and joy is well documented in the pictures in Figure 2, which I had a chance to take during a boat trip at Castelo-do-Bode dam, near Tomar, at the time of IWMS'08, with his friends Mike Perlman and T.W. Anderson.

Figure 2 – Triptic: Michael Perlman, T. W. Anderson, George Styan
right before lunch on a boat at Castelo-do-Bode dam
(Portugal – IWMS'08)

George's almost mythic appreciation of good table makes it easy to picture good moments around a table, and as documented in Figure 3, it almost seems that he carefully chooses his friends as people with the same interests.

Figure 3 – From left to right in a clockwise manner:
The author - George Styan - T. W. Anderson - Michael Perlman
lunch time on the boat at Castelo-do-Bode dam
(Portugal – IWMS'08)

The capacity George has to change things, for the better, with his presence is well documented in the group picture for IWMS'07 in Figure 4. Nothing remains the same after his arrival and actually when we look at the first picture we feel that there is something, more precisely, someone missing.

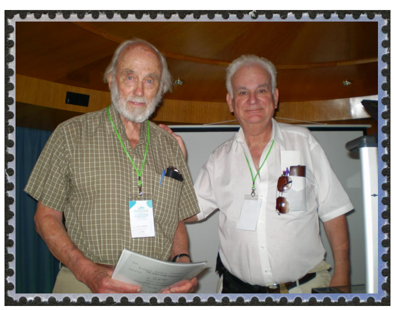
Figure 4 – IWMS'07 - Windsor, Canada
The group picture, before, during and after the arrival of George Styan

Everybody who knows George also knows about his interest in recent years for stamps related with mathematical aspects. This was one of the reasons why in IWMS'08 the organizers presented him with a stamp from the Portuguese post depicting him in his nice outfit from the promotion for his honorary degree from the University of Tampere in 2000, with some of the most celebrated buildings from Tomar in the background. As such I though most adequate to try to build a gallery of some of the existing stamps depicting great mathematicians of all times, from several countries around the world.
This gallery, for sure incomplete, is in Figures 5-7. Many of the stamps were taken from the extraordinary web-site
http://www.mlahanas.de/Stamps /Data/Mathematician/.
Actually, besides the stamp where George is depicted, other pictures would give very good stamps as the ones in Figure 8.

Figure 5 – Stamps of great Mathematicians: A-D

Figure 6 – Stamps of great Mathematicians: D-L

Figure 7 – Stamps of great Mathematicians: L-W



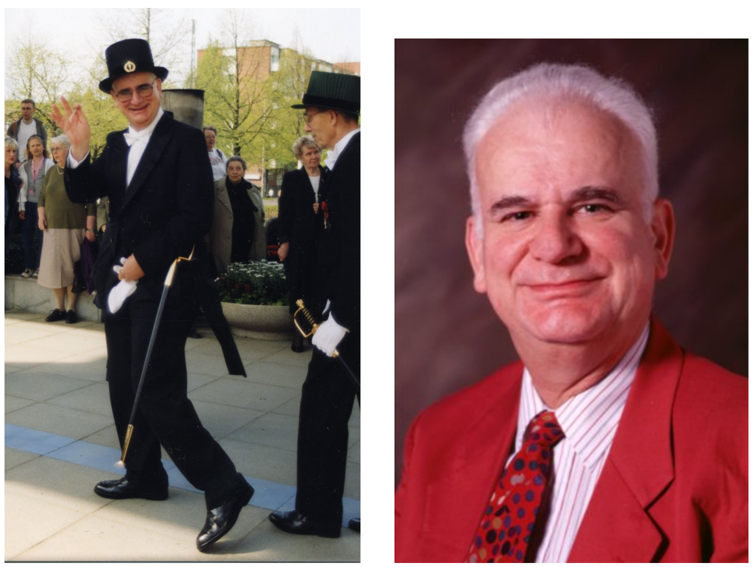
T.W. Anderson & George Styan
on a boat at C.-do-Bode dam

George Styan, Roman Zmyślony
and T.W. Anderson

the author, Ravindra Bapat,
Stephen Haslett, Ejaz Ahmed,
Bertie, George Styan
before dinner in Tomar

Figure 8 – Three pictures from IWMS'08 which would give good stamps

George also got more important honors than the stamp awarded to him at IWMS'08, as his Doctor *Honoris Causa* degree in Tampere, Finland (2000) and his nomination as Honorary Member of the Statistical Society of Canada, in June 2009, documented in Figure 8.
As the statement of his own University, The McGill University, concerning this latter award says: "Honorary Membership of the Statistical Society of Canada is awarded to a statistical scientist of outstanding distinction who has contributed to the development of the statistical sciences in Canada", and as the statement from the Statistical Society of Canada itself says, this award is "for his deep research at the interface of Matrix Theory and Statistics; for his remarkable editorial work within Canada and beyond, his mentoring of graduate and postdoctoral students; and for innumerable other scholarly and professional contributions to the international statistical community".



Doctor *Honoris Causa*
University of Tampere
Finland (2000)

Honorary Member
Statistical Society of Canada
June, 2009

Figure 9 – George P. H. Styan honors

Family and friends

For George, we may say that maybe even more important than Mathematics, it is his family and his friends that play and have always played a central role in his life. How much George cheers Evelyn may be seen from the form he keeps her heavily guarded as it may be seen in Figure 10. Actually this is a magnification of a larger picture taken from Evelyn in front of the Queluz palace, near Lisbon, which is in Figure 11. In this Figure we may also see Evelyn, now guarded by a much better looking body-guard.

Figure 10 – Evelyn heavily guarded

Evelyn in front of Queluz palace *Evelyn and the author*

Figure 11 – Evelyn in Queluz palace, near Lisbon (2008)

And how much George cheers and enjoys his friends company may be easily seen from his looks when we find him around those he loves. Indeed even better looks than when he is enjoying good food together with a good wine, which are a must for an extremely well-educated wine drinker and appreciator. In Figures 12-15 we may see George enjoying the company of a number of his closer friends, being this the opportunity to apologize for all those other many who remained not depicted in any of these pictures.

The first picture in Figure 12 was taken by Soile Puntanen and it surely would make one the most beautiful stamps ever, not needing any further framing. We would say that only Soile is really missing there, but I think we may all easily imagine her with all care, love and enjoyment taking such a beautiful picture.

Simo Puntanen, George Styan and Evelyn at a restaurant in St. John's, Canada

Bernardete Ribeiro, Alexander Kovacec, Evelyn, George Styan and Soile Puntanen, Coimbra, Portugal, 2008

Figure 12 – George, Evelyn and their friends I.

In the second picture in Figure 12 we have a good take of Soile, but since we cannot have it all, now Simo Puntanen was taking the photo.
Simo is also the author of both pictures in Figure 13, depicting George and Evelyn Styan at two different dinner times in 2008.
In Figure 14 we have George together with a number of some of his friends.

T.W. Anderson, George Styan, Evelyn Styan and Dorothy Anderson, dinner at IWMS'08

the author, Evelyn Styan, Soile Puntanen and George Styan, dinner time in Lisbon, 2008

Figure 13 – George, Evelyn and their friends II.

Shuangzhe Liu, Augustyn Markiewicz, George Styan and Yogendra Chaubey at the time of IWMS'07, Windsor

Stephen Haslett, George Styan and Jeffrey Hunter heading for dinner (IWMS'08)

Figure 14 – George and some of his closer friends.

George P. H. Styan. A celebration of 75 years. A personal tribute.

Jeffrey Hunter

Auckland University of Technology, New Zealand

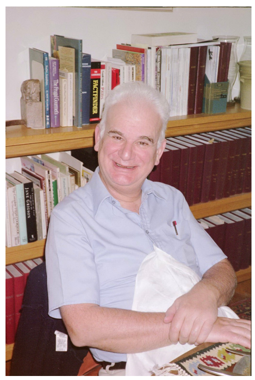


Fig. 1. GPHS at his residence at Vermont, 2001.

George Peter Hansbenno Styan was born 75 years ago on 10 September 1937, in Hendon, a suburb of Greater London in England, U.K. Following a BSc (Hons) degree in Mathematics (1959) from the University of Birmingham and a Certificate in Statistics (1960) from Wadham College, Oxford he pursued post graduate studies at Columbia University with an MA in Mathematical Statistics in 1964, with a thesis on selected topics in Markov Chains with Lajos Tackacs as his supervisor, and a PhD in Mathematical Statistics in 1969 on "Multivariate Normal Inference with Correlation Structure" under the supervision of T.W. Anderson with whom he established a life long friendship. My association with George goes back to 1973 when I attended an Institute of Mathematical Statistics meeting at Ithaca College to hear George talk on some research on Markov chains that included reference to generalized inverses. This was of much interest to me as I had published a paper in 1969 identifying Kemeny and Snell's fundamental matrix of Markov chains as a generalized inverse. Starting from that meeting our subsequent association has spanned the globe with George visiting the University of Auckland over the period July 1984 to June 1985 on a sabbatical leave to spend time primarily with George Seber and Alastair Scott. At that time I was a member of the Department of Mathematics and Statistics at the University of Auckland. I followed up his visit with me visiting McGill University for a month in May 1988 and again visiting McGill in June 2001 (when I was based at Massey University). Both of these latter visits occurred while I was on a sabbatical leave. In 1988 George tried to discourage me from pursuing any further activity on generalized inverses but not all was known about their properties when associated with Markov chains so that I failed to take his advice!

However, it is the series of International Workshops on Matrices and Statistics that we owe a debt of gratitude to George for promoting and maintaining the impetus that has seen this annual series of meetings continue to flourish. I think that initially it was an opportunity for George to maintain links with his by now very extensive group of friends and research colleagues but the workshops continue to attract much interest. I was a member of the organising committee of the second such workshop that was held 4-5 December 1992 in Auckland immediately preceding the International Biometrics Conference that was held in Hamilton, New Zealand. Some of those participating in that workshop meeting – Alastair Scott, Simo Puntanen, Bill Farebrother, Thomas Mathew, Chris Paige, Shayle Searle have also maintained their association with the workshops and George over the years. It was following my visit with George in Montreal in 2001 that I was persuaded to Chair the Local Organising Committee of IWMS in Auckland in 2005. As a precursor to that I renewed my association with the workshop series in 2003 at Dortmund and 2004 at Bedlewo. After the Auckland meeting we met again in 2006 (Uppsala), 2007 (Windsor), 2008 (Tomar), 2009 (Smolenice) and 2011 (Tartu). The expansion of the International Organising Committee in 2007 saw George taking on the Honorary Chair role with a rotation of Chairs of the IOC being established. I chaired the IOC in 2010 for the very large meeting that we held in Shanghai. However with a bereavement in Evelyn's family he was not able to attend that meeting – the first held in China.

George has been a very hospitable host over the years of our association – we have progressed from cask wines served in 1984 to more sophisticated tastes. George however at that time had a liking of Montana Fairhall River Claret on our weekly Friday evening drinks at the Senior Common at the University of Auckland. He also has a passion for New Zealand Bluff Oysters requesting that they be on the menu at a dinner at our home in 2005 following the IWMS meeting in Auckland. George's gastronomic tastes are legendary with some wonderful meals invariably scheduled whenever he has a group of friends to partake of the opportunities!

George has been very generous of his time and commitments that he has made with these workshops in many cases bringing with him his graduate students so that they could share in the experiences.

George's research interests are wide ranging including matrices and statistics, with particular emphasis on canonical correlations, canonical efficiency factors in experimental design, efficiency and optimality of ordinary least squares, generalized inverses, Hadamard products, matrix inequalities, matrix partial orderings, matrix rank additivity and subtractivity, rank equalities and inequalities, Schur complements. Applications to electrical networks. Bibliography and biography. Experimental designs involving Graeco-Latin squares, Latin squares, Youden squares, and magic squares, Bibliography, biography and history. Postage stamps, playing cards and other artefacts associated with statistics and mathematics. Because of these extensive interests he has developed a very broad academic community with which he collaborates and interacts with. For fear of omitting any names I refrain from naming individuals but his range activity is extensive scholarly and noteworthy.

George has a wealth experience and knowledge in matrices and statistics and his recent book with Simo Puntanen and Jarkko Isotalo was a wonderful opportunity for them to share with the scientific community many of the "tricks" that they have developed and fostered over a number of years of active and supportive collaboration.

When he retired from McGill in 2005 ago he was honoured with Professor Emeritus status, a recognition that he was very proud of. He has also been honoured with a variety of awards based upon his contributions to numerous professional associations. His honorary doctorate "For his great scientific contributions and merits in mathematics and statistics, and in the promotion of research in the University of Tampere" in May 2000, was a significant event in his career.

George and Götz Trenkler also honoured my retirement with an article on a philatelic excursion in probability and matrix theory with references to mathematicians featured in my books, and published in a festschrift devoted to my research activities. I appreciated that recognition from such valued colleagues.

My wife Hazel has been able to share some times with both Evelyn and George be it in our respective homes in Auckland or at the VV (Villa Vermont) or at their apartment, initially in Wilderton Ave and later on Nuns Island in Montreal.

There are many personal anecdotes that I could share with you all but I think that this is best done through a series of photographs taken mainly by me over the years and appended below.

It is with regret that due to the timing of the Workshop I will not be able to be with you in Bedlewo this year but I do wish to convey to George my very best wishes for the occasion of celebrating this milestone birthday. George, you are a larger than life personality, may you have many more years of active and enjoyable pursuit of things mathematical and statistical. I wish you all the best for the future. Celebrate well!

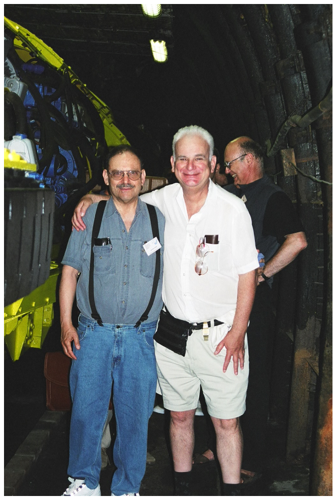


Fig. 2. Adi Ben-Israel with George on the excursion at IWMS'2003 in Dortmund

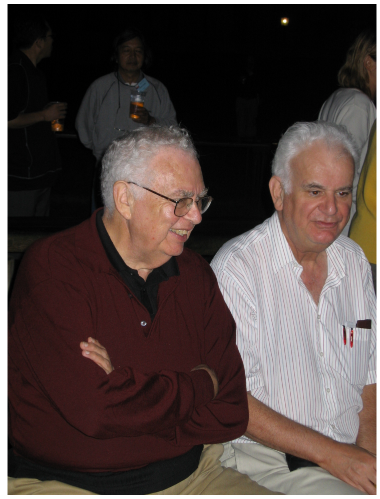


Fig. 3. Gene Golub and George at a Barbecue at IWMS'2004 at Będlewo

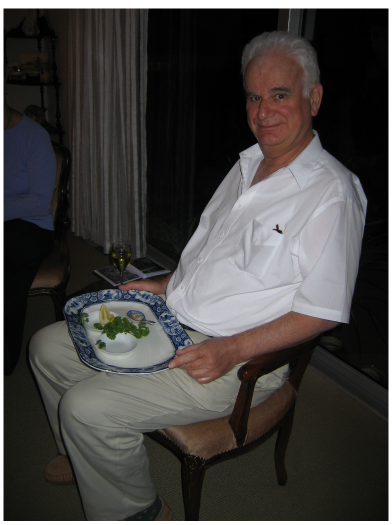


Fig. 4. George with a partially consumed plate of oysters, April 2005

Fig. 5. George at IWMS'2006, Uppsala

Fig. 6. Simo, Jochen, Jeff, George, Götz at IWMS'2005, Auckland

Fig. 7. George, Jeff and Simo at Jeff's home, April 2005

Fig. 8. George and C.R. Rao, with Simo, Jarkko Isotalo & Götz in the background, IWMS'2005, Auckland

Fig. 9. George honouring Jerzy Baksalary at a special session at IWMS'2005

Fig. 10. Shuangzhe Liu, Augustyn, George, and Yogendra Chaubey at IWMS'2007, Windsor, Canada

Fig. 11. George at IWMS'2008, Tomar, Portugal

Fig. 12. George, Fikri Akdenis & Jochen at IWMS'2009, Smolenice, Slovakia

Fig. 13. George and Jeff with Ingram Olkin seated at the right at IWMS'2011, Tartu, Estonia

Part VIII

List of Participants

Participants

1. **Haftom Abebe**
   Department of Methodology and Statistics, Maastricht University, Maastricht, The Netherlands, *haftom.temesgen@maastrichtuniversity.nl*
2. **Nihan Acar**
   Department of Statistics, Mimar Sinan Fine Arts University, Istanbul, Turkey, *nihan.acar@msgsu.edu.tr*
3. **S. Ejaz Ahmed**
   Faculty of Mathematics and Science, Brock University, St. Catharines, Ontario, Canada, *seahmed12@gmail.com*
4. **Kadri Ulaş Akay**
   Department of Mathematics, University of Istanbul, Istanbul, Turkey, *kulas@istanbul.edu.tr*
5. **Aylin Alin**
   Department of Statistics, Dokuz Eylül University, Izmir, Turkey, *aylin.alin@deu.edu.tr*
6. **Barbora Arendacká**
   Physikalisch-Technische Bundesanstalt, Berlin, Germany, *barendacka@gmail.com*
7. **Baris Asikgil**
   Department of Statistics, Mimar Sinan Fine Arts University, Istanbul, Turkey, *basikgil@msgsu.edu.tr*
8. **Anthony C. Atkinson**
   The London School of Economics, London, UK, *a.c.atkinson@lse.ac.uk*
9. **Alena Bachratá**
   Department of Applied Mathematics and Statistics, Comenius University Bratislava, Slovakia, *Alena.Bachrata@fmph.uniba.sk*
10. **Rosemary A. Bailey**
    Queen Mary, University of London, London, UK, *r.a.bailey@qmul.ac.uk*
11. **Oskar M. Baksalary**
    Faculty of Physics, Adam Mickiewicz University, Poznań, Poland, *OBaksalary@gmail.com*
12. **Narayanaswamy Balakrishnan**
    McMaster University, Ontario, Canada, *bala@mcmaster.ca*
13. **Krzysztof Bartoszek**
    Department of Mathematical Statistics, Chalmers University of Technology and University of Gothenburg, Gothenburg, Sweden, *krzbar@chalmers.se*

14. **Silvie Bělašková**
Department of Mathematics, Tomas Bata University in Zlín, Zlín, Czech Republic, *belaskova@fai.utb.cz*

15. **Philip Bertrand**
Solihull, UK, *researchall@blueyonder.co.uk*

16. **Ufuk Beyaztaş**
Department of Statistics, Dokuz Eylül University, Izmir, Turkey, *ufuk.beyaztas@deu.edu.tr*

17. **Rajendra Bhatia**
Indian Statistical Institute, New Delhi, India,*rajenbhatia@gmail.com*

18. **Olivia Bluder**
Department of Statistics, KAI-Kompetenzzentrum fur Automobil und Industrieelektronik GmbH, Villach, Austria
Alpen-Adria University of Klagenfurt, Klagenfurt, Austria, *olivia.bluder@k-ai.at*

19. **Barbara Bogacka**
School of Mathematical Sciences, Queen Mary, University of London, London, UK, *b.bogacka@qmul.ac.uk*

20. **Tadeusz Caliński**
Department of Mathematical and Statistical Methods, Poznań University of Life Sciences, Poznań, Poland,*calinski@up.poznan.pl*

21. **Francisco Carvalho**
Unidade Departamental de Matemática e Física, Instituto Politécnico de Tomar, Tomar, Portugal
CMA - Centro de Matemática e Aplicaçoes, Universidade Nova de Lisboa, Lisboa, Portugal, *fpcarvalho@ipt.pt*

22. **Garry Ka Lok Chu**
Mathematical Department, Dawson College, Montreal, Canada, *prof_ chu@yahoo.ca*

23. **Carlos A. Coelho**
Departamento de Matemática e Centro de Matemática e Aplicações, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Lisboa, Portugal, *cmac@fct.unl.pt*

24. **Knut Conradsen**
Department of Informatics and Mathematical Modeling, Technical University of Denmark, Lyngby, Denmark, *kc@imm.dtu.dk*

25. **Ricardo Covas**
Unidade Departamental de Matemática e Física, Instituto Politécnico de Tomar, Tomar, Portugal
CMA - Centro de Matemática e Aplicaçoes, Universidade Nova de Lisboa, Lisboa, Portugal, *ricardocovas@gmail.com*

26. **Carlos Cuevas-Covarrubias**
Center for Research in Statistics and Applied Mathematics, Acturial Sciences School, Anahuac, Mexico, *ccuevas@anahuac.mx*

27. **Somnath Datta**
Department of Bioinformatics and Biostatistics, School of Public Health and Information Sciences, University of Louisville, Louisville, USA, *somnath.datta@louisville.edu*

28. **Susmita Datta**
Department of Bioinformatics and Biostatistics, School of Public Health and Information Sciences, University of Louisville, Louisville, USA, *susmita.datta@louisville.edu*

29. **Nino Demetrashvili**
Unit of Medical Statistics, Department of Epidemiology, University of Groningen, and University Medical Center Groningen, Groningen, The Netherlands, *n.demetrashvili@umcg.nl*, *ninobiostat@gmail.com*

30. **Sandra Donevská**
Department of Mathematical Analysis and Applications of Mathematics, Palacký University Olomouc, Olomouc, Czech Republic, *sdonevska@seznam.cz*

31. **Hilmar Drygas**
Kassel University, Kassel, Germany, *hilmar.drygas@onlinehome.de*

32. **Rie Enomoto**
Department of Mathematical Information Science, Tokyo University of Science, Tokyo, Japan, *j1410701@ed.tus.ac.jp*

33. **Birsen Eygi Erdogan**
Department of Statistics, Marmara University, Istanbul, Turkey, *birsene@marmara.edu.tr*

34. **Ali Erkoç**
Department of Statistics, Mimar Sinan University, Istanbul, Turkey, *erkoc_a@hotmail.com*

35. **Célia Fernandes**
Área Departamental de Matemática, Instituto Politécnico de Lisboa, Lisboa, Portugal, *cfernandes@adm.isel.pt*

36. **Katarzyna Filipiak**
Department of Mathematical and Statistical Methods, Poznań University of Life Sciences, Poznań, Poland, *kasfil@up.poznan.pl*

37. **Lenka Filová**
Department of Applied Mathematics and Statistics, Comenius University Bratislava, Bratislava, Slovakia, *filova@fmph.uniba.sk*

38. **Eva Fišerová**
Department of Mathematical Analysis and Applications of Mathematics, Palacký University Olomouc, Olomouc, Czech Republic, *fiserova@upol.cz*

39. **Miguel Fonseca**
Centro de Matemática e Aplicações, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Lisboa, Portugal, *fmig@fct.unl.pt*

40. **Stelios D. Georgiou**
Department of Statistics and Actuarial-Financial Mathematics, University of the Aegean, Karlovassi, Samos, Greece, *stgeorgiou@aegean.gr*

41. **Steven Gilmour**
Southampton Statistical Sciences Research Institute, University of Southampton, Southampton, UK, *S.Gilmour@soton.ac.uk*

42. **Atilla Göktaş**
Department of Statistics, University of Muğla, Muğla, Turkey, *gatilla@mu.edu.tr*

43. **Tomasz Górecki**
Faculty of Mathematics and Computer Science, Adam Mickiewicz University, Poznań, Poland, *tomasz.gorecki@amu.edu.pl*

44. **Małgorzata Graczyk**
Department of Mathematical and Statistical Methods, Poznań University of Life Sciences, Poznań, Poland, *magra@up.poznan.pl*

45. **Romesh Gupta**
Department of Statistics, Jammu University, Jammu, India, *romesh_68@yahoo.com*

46. **Nesrin Gürel**
Department of Statistics, Sakarya University, Sakarya, Turkey *nesring@sakarya.edu.tr*

47. **Duygu Haki**
Institute of Science, Marmara University, Istanbul, Turkey, *duygukocaman@ymail.com*

48. **Zofia Hanusz**
University of Life Sciences in Lublin, Lublin, Poland, *zofia.hanusz@up.lublin.pl*

49. **Chengcheng Hao**
Department of Statistics, Stockholm University, Stockholm, Sweden, *chengcheng.hao@stat.su.se*

50. **Stephen J. Haslett**
Department of Statistics, Massey University, Palmerston North, New Zealand, *s.j.haslett@massey.ac.nz*

51. **Jan Hauke**
Faculty of Geography and Geology, Adam Mickiewicz University, Poznañ, Poland, *jhauke@amu.edu.pl*

52. **Deniz İnan**
Department of Methodology and Statistics, Marmara University, Istanbul, Turkey, *denizlukuslu@marmara.edu.tr*

53. **Krystyna Katulska**
Faculty of Mathematics and Computer Science, Adam Mickiewicz University, Poznań, Poland, *krakat@amu.edu.pl*

54. **Dariusz Kayzer**
Department of Mathematical and Statistical Methods, Poznań University of Life Sciences, Poznań, Poland, *dkayzer@up.poznan.pl*

55. **Abbas Khalili**
Department of Mathematics and Statistics, McGill University, Montreal, Canada, *khalili@math.mcgill.ca*

56. **Daniel Klein**
Institute of Mathematics, University of P.J. Šafárik, Košice, Slovakia, *daniel.klein@upjs.sk*

57. **Daniel Kosiorowski**
Department of Statistics, Cracow University of Economics, Cracow, Poland, *daniel.kosiorowski@uek.krakow.pl*

58. **Tomasz Kossowski**
Faculty of Geography and Geology, Adam Mickiewicz University, Poznañ, Poland, *tkoss@amu.edu.pl*

59. **Mirosław Krzyśko**
Faculty of Mathematics and Computer Science, Adam Mickiewicz University, Poznań, Poland, *mkrzysko@amu.edu.pl*

60. **Joachim Kunert**
Department of Statistics, Technical University of Dortmund, Dortmund, Germany, *joachim.kunert@udo.edu*

61. **Fatma Sevinç Kurnaz**
Department of Statistics and Computer Sciences, Karadeniz Technical University, Trabzon, Turkey, *fskurnaz@ktu.edu.tr*

62. **Agnieszka Łacka**
Department of Mathematical and Statistical Methods, Poznań University of Life Sciences, Poznań, Poland, *aga@riders.pl*

63. **Lynn R. LaMotte**
Department of Biostatistics, Louisiana State University, Health Sciences Center, New Orleans, USA, *llamot@lsuhsc.edu*

64. **Alan Lee**
Department of Statistics, University of Auckland, Auckland, New Zealand,*lee@stat.auckland.ac.nz*

65. **Yuli Liang**
Department of Statistics, Stockholm University, Stockholm, Sweden, *yuli.liang@stat.su.se*

66. **Pen-Hwang Liau**
Department of Mathematics, National Kaohsiung Normal University, Kaohsiung, Taiwan, R.O.C., *phliau@nknucc.nknu.edu.tw*

67. **Erkki P. Liski**
Department of Mathematics, Statistics and Philosophy, University of Tampere, Tampere, Finland, *epl@uta.fi*

68. **Augustyn Markiewicz**
Department of Mathematical and Statistical Methods, Poznań University of Life Sciences, Poznań, Poland, *amark@up.poznan.pl*

69. **Janko Marovt**
Institute of Mathematics, Physics and Mechanics, University of Maribor, Maribor, Slovenia, *janko.marovt@uni-mb.si*

70. **Jean-Pierre Masson**
INRA BIO3P Rennes, Agrocampus Ouest, Rennes, France, *jeanpierre.masson@me.com*

71. **Thomas Mathew**
Department of Mathematics and Statistics, University of Maryland, Baltimore, USA, *mathew@math.umbc.edu*

72. **Caterina May**
University of Eastern Piedmont, Novara, Italy, *caterina.may@eco.unipmn.it*

73. **France Mentré**
School of Medicine, University Paris Diderot, Paris, France, *france.mentre@inserm.fr*

74. **João T. Mexia**
Centro de Matemática e Aplicações, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Lisboa, Portugal, *jtm@fct.unl.pt*

75. **Andrzej Michalski**
Department of Mathematics, Wrocław University of Environmental and Life Sciences, Wrocław, Poland, *apm.mich@gmail.com*

76. **John P. Morgan**
Department of Statistics, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, USA, *jpmorgan@vt.edu*

77. **Jamal Najim**
Department of Statistics, Télécom Paristech and CNRS, Paris, France, *najim@telecom-paristech.fr*

78. **Konrad Nosek**
Department of Mathematical Analysis, Computational Mathematics and Probability Methods, AGH University of Science and Technology, Cracow, Poland *konosek@agh.edu.pl*

79. **Naoya Okamoto**
Faculty of Health and Nutrition, Tokyo Seiei College, Tokyo, Japan, *n_okamoto@auone.jp*

80. **Paulo Eduardo Oliveira**
Department of Mathematics, University of Coimbra, Coimbra, Portugal, *paulo@mat.uc.pt*

81. **Mizuki Onozawa**
Department of Mathematical Information Science, Tokyo University of Science, Tokyo, Japan, *mizuki.onozawa@hotmail.co.jp*

82. **Dulce G. Pereira**
Department of Mathematics, University of Évora, Évora, Portugal, *dgsp@uevora.pt*

83. **Domenico Perrotta**
European Commission, Joint Research Centre, Ispra, Italy, *Domenico.Perrotta@ec.europa.eu*

84. **Jolanta Pielaszkiewicz**
Department of Mathematics, Linköping University, Linköping, Sweden, *jolanta.pielaszkiewicz@liu.se*

85. **Maryna Prus**
Faculty of Mathematics, Otto-von-Guericke University, Magdeburg, Germany,*maryna.prus@ovgu.de*

86. **Marcin Przystalski**
The Research Center for Cultivar Testing, Słupia Wielka, Poland, *marprzyst@gmail.com*

87. **K. Manjunatha Prasad**
Department of Statistics, Manipal University, Manipal, India, *kmprasad63@gmail.com*

88. **Simo Puntanen**
Department of Mathematics Statistics and Philosophy, University of Tampere, Tampere, Finland, *simo.puntanen@uta.fi*

89. **Sonja Radosavljevic**
MAI, Linköping University, Linköping, Sweden, *sonja.radosavljevic@liu.se*

90. **Paulo Ramos**
Área Departamental de Matemática, Instituto Politécnico de Lisboa, Lisboa, Portugal, *pramos@adm.isel.pt*

91. **Paulo Canas Rodrigues**
Centro de Matemática e Aplicações, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Lisboa, Portugal, *paulocanas@gmail.com*

92. **Dietrich von Rosen**
Swedish University of Agricultural Sciences, Uppsala, Sweden
Linköping University, Linköping, Sweden, *dietrich.von.rosen@et.slu.se*

93. **Anuradha Roy**
Department of Management Science and Statistics, University of Texas at Santo Antonio, Santo Antonio, USA, *anuradha.roy@utsa.edu*

94. **Burkhard Schaffrin**
School of Earth Sciences, The Ohio State University, Columbus, Ohio, USA, *schaffrin.1@osu.edu*

95. **Alastair Scott**
Department of Statistics, University of Auckland, Auckland, New Zealand, *a.scott@auckland.ac.nz*

96. **Peter Šemrl**
Department of Mathematics, University of Ljubljana, Ljubljana, Slovenia, *peter.semrl@fmf.uni-lj.si*

97. **Takashi Seo**
Department of Mathematical Information Science, Tokyo University of Science, Tokyo, Japan, *seo@rs.kagu.tus.ac.jp*

98. **Łukasz Smaga**
Faculty of Mathematics and Computer Science, Adam Mickiewicz University, Poznań, Poland, *ls@amu.edu.pl*

99. **George P. H. Styan**
Department of Mathematics and Statistics, McGill University, Montreal, Canada, *styan@math.mcgill.ca*

100. **Stella Stylianou**
Department of Statistics and Actuarial-Financial Mathematics, University of the Aegean, Karlovassi, Samos, Greece,*sstylian@aegean.gr*

101. **Reijo Sund**
Service Systems Research Unit, National Institute for Health and Welfare, Helsinki, Finland, *reijo.sund@thl.fi*

102. **Sho Takahashi**
Department of Mathematical Information Science, Tokyo University of Science, Tokyo, Japan, *j1410704@ed.tus.ac.jp*

103. **Müjgan Tez**
Department of Statistics, Marmara University, Istanbul, Turkey, *mtez@marmara.edu.tr*

104. **Secil Toprak**
Department of Mathematics, Dicle University, Diyarbakır, Turkey, *secilyalaz@gmail.com*

105. **Götz Trenkler**
Faculty of Statistics, Dortmund University of Technology, Dortmund, Germany, *trenkler@statistik.tu-dortmund.de*

106. **Tsung-Shan Tsou**
Institute of Statistics, National Central University, Jhongli City, Taiwan, *tsou@mx.stat.ncu.edu.tw*

107. **Michaela Tučková**
Department of Geoinformatics, Palacký University Olomouc, Olomouc, Czech Republic, *michaela.tuckova@upol.cz*

108. **Semra Türkan**
Department of Statistics, Hacettepe University, Ankara, Turkey,
sturkan@hacettepe.edu.tr

109. **Dariusz Uciński**
Faculty of Electrical Engineering, Computer Science and Telecommunications, University of Zielona Góra, Zielona Góra, Poland,
D.Ucinski@issi.uz.zgora.pl

110. **Florin Vaida**
Division of Biostatistics and Bioinformatics, University of California, San Diego, California, USA, *fvaida@ucsd.edu*

111. **Júlia Volaufová**
Department of Biostatistics, Louisiana State University, Health Sciences Center, New Orleans, USA, *jvolau@lsuhsc.edu*

112. **Łukasz Waszak**
Faculty of Mathematics and Computer Science, Adam Mickiewicz University, Poznań, Poland, *lwaszak@amu.edu.pl*

113. **Hans Joachim Werner**
Institute for Financial Economics and Statistics, University of Bonn, Bonn, Germany, *hjw.de@uni-bonn.de*

114. **Douglas Wiens**
Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, Alberta, Canada, *doug.wiens@ualberta.ca*

115. **Viktor Witkovský**
Institute of Measurement Science, Slovak Academy of Sciences, Bratislava, Slovakia, *witkovsky@savba.sk*

116. **Dominika Wojtera-Tyrakowska**
Faculty of Mathematics and Computer Science, Adam Mickiewicz University, Poznań, Poland, *dwt@amu.edu.pl*

117. **Waldemar Wołyński**
Faculty of Mathematics and Computer Science, Adam Mickiewicz University, Poznań, Poland, *wolynski@amu.edu.pl*

118. **Ivan Žežula**
Institute of Mathematics, P. J. Šafárik University in Košice, Košice, Slovakia, *ivan.zezula@upjs.sk*

119. **Roman Zmyślony**
Faculty of Mathematics, Computer Sciences and Econometrics, University of Zielona Góra, Zielona Góra and Institute of Mathematics and Informatics, University of Opole, Opole, Poland,
r.zmyslony@wmie.uz.zgora.pl

Index