29^{th}

International Workshop on Matrices and Statistics

IWMS2023

BOOK OF ABSTRACTS

LINKÖPING UNIVERSITY

August 21-23, 2023
Linköping, Sweden

Contents

Preface

The purpose of the workshop is to stimulate research and, in an informal setting, foster the interaction of researchers using matrices in different areas of mathematics especially with a focus on statistics. The workshop will provide a forum through which mathematicians and statisticians may be better informed of the latest developments and newest techniques in linear algebra and matrix theory and may exchange ideas with researchers from a wide variety of countries. As well as range of invited speakers we are to strengthening the interactions between participants by organizing special sessions in various of areas.

The theme of the workshop target the theory and applications using matrices in different branches of science like: big data analytics, machine learning, computer and information science, biology, physics, economics, engineering, mathematics and statistics.

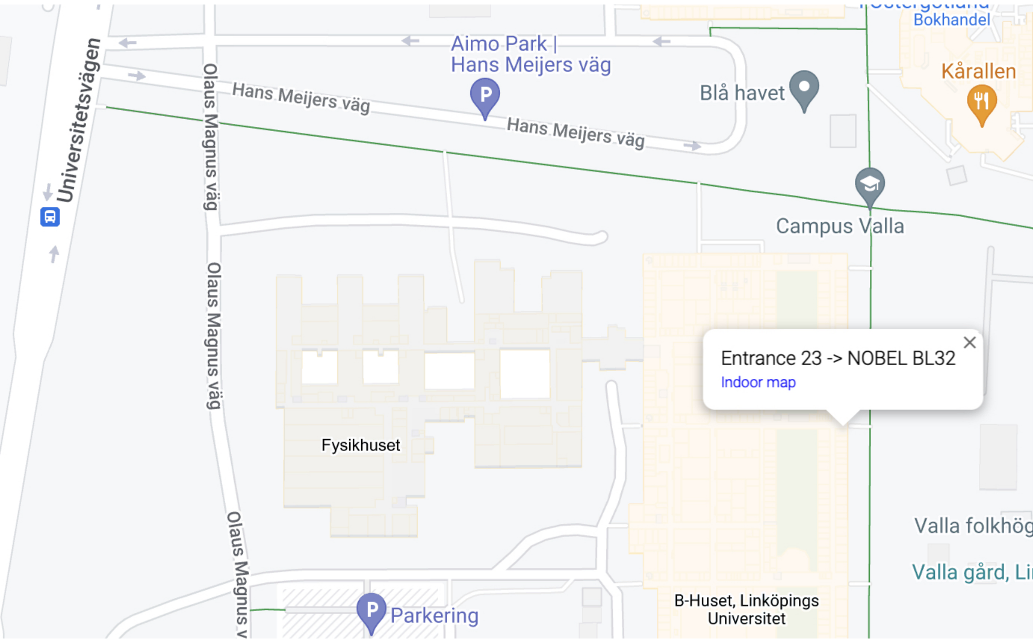
Enjoy the conference!

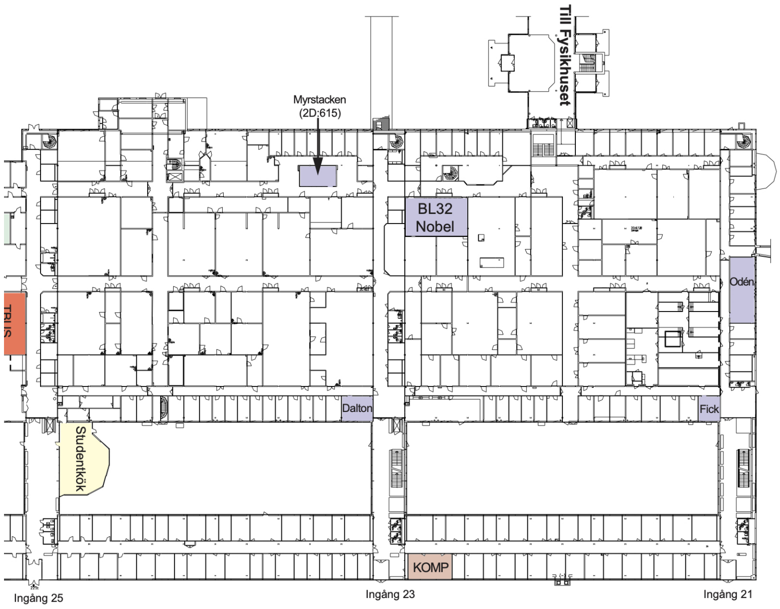
Martin Singull
Chair of Organizing Committee

Venue and Program

Venue

The venue for the conference will be lecture room BL32 (also known as *'Nobel'*), Campus Valla, B-building (entrance 23), Linköping University. Click the link to get an interactive map. All sessions will be in the same room.





Program

Monday 21/8		
9.25	Opening	
9.35-10.20	Taras Bodnar	Chair: Dietrich von Rosen
10.25-10.50	Raymond Kan	
10.50-11.15	Michel van de Velden	
HEALTH BREAK		
11.45-12.10	Jan Hauke	Chair: Kartarzyna Filipiak
12.10-12.35	Stephen Haslett	
12.35-13.00	Burkhard Schaffrin	
LUNCH		
14.30-15.15	Krzysztof Podgorski	Chair: Taras Bodnar
HEALTH BREAK		
15.45-16.10	Johannes Heiny	Chair: Stepan Mazur
16.10-16.35	Xiangfeng Yang	
16.35-17.00	Elias Erdtman	

Tuesday 22/8		
9.00-9.45	Göran Bergqvist	Chair: Berkant Savas
10.00-10.25	Nicola Loperfido	*Tensor methods in statistics*
10.25-10.50	Stepan Mazur	Chair: Nicola Loperfido
10.50-11.15	Cinzia Franceschini	
11.15-11.40	Dietrich von Rosen	
HEALTH BREAK		
12.10-12.35	Deepjyoti Deka	Chair: Malgorzata Bogdan
12.35-13.00	Shriram Srinivasan	
LUNCH		
14.30-15.15	Malgorzata Bogdan	Chair: Natalia Bebiano
HEALTH BREAK		
15.45-16.10	Oskar Baksalary	*Celebration of Götz Trenkler's 80th birthday*
16.10-16.35	Jarkko Isotalo	Chair: Simo Puntanen and Oskar Baksalary
16.35-17.00	Simo Puntanen	
17.00-17.25	George Styan (online)	
18.30-	Conference dinner	Chair: Martin Singull

Wednesday 23/8		
9.00-9.45	Natalia Bebiano	Chair: Göran Bergqvist (S10)
10.00-10.25	Hao Chi Kiang	Chair: Martin Singull (S11)
10.25-10.50	Katarzyna Filipiak	
10.50-11.15	Monika Mokrzycka	
HEALTH BREAK		
11.45-12.10	Martina Hancova	Chair: Jarkko Isotalo (S12)
12.10-12.35	Maryna Prus	
12.35-13.00	Francisco Diaz	
13.00-13.15	Closing	Chair: Dietrich von Rosen
LUNCH		

Abstract - Invited Speakers

The periodic pseudo-Jacobi inverse eigenvalue problem

Natalia Bebiano[1], Wei-Ru Xu, Yi Gong and Guo-Liang Chen
[1]Universidade de Coimbra, Coimbra, Portugal

Abstract

The problem of reconstructing a *periodic pseudo-Jacobi matrix*, which is derived from the discretization and truncation of Schrödinger equation, arises in non-Hermitian quantum mechanics. Also the reconstruction of the Hamiltonian system of an indefinite *Toda lattice* and the symmetry reduction of the *Wess-Zumino-Novikov-Witten* model in quantum field theory are problems deserving the attention of physicists and mathematicians. In mathematics, this problem is referred to as *periodic pseudo-Jacobi inverse eigenvalue problem* (hereafter **PPJIEP**), and concerns the reconstruction from assigned spectral data of a specified periodic pseudo-Jacobi matrix Inspired in a discrete version of Floquet theory in a space with indefinite metric [Math. Comp. 35 (1980) 1203-1220] and a van Moerbeke's idea [Invent. Math. 37 (1976) 45-81], the **PPJIEP** problem is solved. We use two methods to characterize the signature operator so that the solution exists.

This is joint work with Wei-Ru Xu, Yi Gong and Guo -Liang Chen (China). The problem of reconstructing a *periodic pseudo-Jacobi matrix*, which is derived from the discretization and truncation of Schrödinger equation, arises in non-Hermitian quantum mechanics. Also the reconstruction of the Hamiltonian system of an indefinite *Toda lattice* and the symmetry reduction of the *Wess-Zumino-Novikov-Witten* model in quantum field theory are problems deserving the attention of physicists and mathematicians. In mathematics, this problem is referred to as *periodic pseudo-Jacobi inverse eigenvalue problem* (hereafter **PPJIEP**), and concerns the reconstruction from assigned spectral data of a specified periodic pseudo-Jacobi matrix Inspired in a discrete version of Floquet theory in a space with indefinite metric [Math. Comp. 35 (1980) 1203-1220] and a van Moerbeke's idea [Invent. Math. 37 (1976) 45-81], the **PPJIEP** problem is solved. We use two methods to characterize the signature operator so that the solution exists.

Keywords: inverse problem, tridiagonal matrices

On tensor decompositions and rank

Göran Bergqvist
Linköping University, Linköping, Sweden

Abstract

We review some developments of tensor rank concepts and different ways of decomposing tensors, including results for random tensors and symmetric tensors. We also discuss the use of neural networks to determine tensor decompositions and ranks with some recent applications, and tensor decompositions as a tool to simplify neural networks.

Keywords: Tensors, Decompositions, Rank

Reviving pseudo-inverses: Asymptotic properties of large dimensional Moore-Penrose and ridge-type inverses with applications

Taras Bodnar[1] and Nestor Parolya[2]
[1]Stockholm University, Stockholm, Sweden
[2]Delft University of Technology, Delft, The Netherlands

Abstract

In this paper, we derive high-dimensional asymptotic properties of the Moore-Penrose inverse and the ridge-type inverse of the sample covariance matrix. In particular, the analytical expressions of the weighted sample trace moments are deduced for both generalized inverse matrices and are present by using the partial exponential Bell polynomials which can easily be computed in practice. The existent results are extended in several directions: (i) First, the population covariance matrix is not assumed to be a multiplier of the identity matrix; (ii) Second, the assumption of normality is not used in the derivation; (iii) Third, the asymptotic results are derived under the high-dimensional asymptotic regime. Our findings are used in the construction of improved shrinkage estimators of the precision matrix that minimizes the Frobenius norm. Also, shrinkage estimators for the coefficients of the high-dimensional regression model and the weights of the global minimum variance portfolio are obtained. Finally, the finite sample properties of the derived theoretical results are investigated via an extensive simulation study.

Keywords: Moore-Penrose inverse, Ridge-type inverse, Bell polynomials, Sample covariance matrix, Random matrix theory, High-dimensional asymptotics

References

[1] Bodnar, T., Dette, H., Parolya, N. (2016). Spectral analysis of the Moore-Penrose inverse of a large dimensional sample covariance matrix. *Journal of Multivariate Analysis*, 148:160-172.

[2] Imori, S., von Rosen, D. (2020). On the mean and dispersion of the Moore-Penrose generalized inverse of a Wishart matrix. *The Electronic Journal of Linear Algebra*, 36:124-133.

Sorted L-One Norm Penalized Estimator of the high dimensional precision matrices

Malgorzata Bogdan[1,2], Giovanni Bonaccolto[3], Ivan Hejny[1], Philipp Kremer[4], Sandra Paterlini[5], Riccardo Riccobello[5], Piotr Sobczyk[6], Jonas Wallin[1]
[1]Lund University, Lund, Sweden
[2]University of Wroclaw, Wroclaw, Poland
[3]Kore University of Enna, Enna, Italy
[4]EBS Universität für Wirtschaft und Recht, Wiesbaden, Germany
[5]University of Trento, Trento, Italy
[6]Wroclaw University of Science and Technology, Wroclaw, Poland

Abstract

Sparse graphical modeling has garnered significant attention in various academic disciplines. In this study, we introduce two novel approaches, namely Gslope and Tslope, for estimating sparse precision matrices. These approaches leverage Gaussian and T-student data, respectively, and achieve sparsity by using the sorted L_1-norm penalty on the elements the precision matrix. Unlike the widely-used graphical LASSO estimator, Gslope and Tslope allow for additional dimensionality reduction patterns by permitting some estimated elements of the precision matrix to be equal to each other.

To analyze the pattern recovery of Gslope and Tslope, we present original asymptotic results on the distribution of the SLOPE esimators of the precision matrices of elliptically contoured distributions. Furthermore, we propose tuning parameter selection strategies that guarantee control over the probability of including false edges between disconnected graph components and empirically control the False Discovery Rate for block diagonal covariance matrices.

To assess the performance of our proposed methods, we conduct extensive simulations and real-world analyses, comparing them against other state-of-the-art approaches in sparse graphical modeling. The results confirm that Gslope and Tslope serve as two effective tools for estimating precision matrices in scenarios involving Gaussian, T-student, and mixture data.

Keywords: Precision Matrix, Graphical Model, Penalized Estimation, Sorted L-One Norm

Efficient inversion of large sparse positive definite matrices

Krzysztof Podgórski, Michał Kos, and Hanqing Wu
Lund University, Lund, Sweden

Abstract

The decompositions of matrices such as triangular, Cholesky, singular value, etc, are often important tools for their inversion, which, in turn, is the basis for many statistical methods, to mention fitting regression models or estimation of covariance matrix with constraints for multivariate Gaussian data. Additionally to this, in many applications there is an increasing need for high-dimensional methods that often naturally lead to inverting high-dimensional sparse matrices. We start with an overview of the decompositions methods with the emphasis on an often not realized connection between the inversion and decomposition of a positive definite matrix. Then we argue that these decompositions are not suitable for the inversion of sparse matrices. Instead we propose an alternative method that is based on a dyadic algorithm that has been introduction in [2] for the purpose of efficient orthogonalization of the B-splines and used for the functional data analysis in [1]. The inversion method is made of two steps. The method requires first to gather non-zero terms close to the diagonal, the process that connects to the classical band-width minimization problem in computing. Then it uses the dyadic algorithm that compute inverse (and more) faster than any standard method. We discuss new contributions and generalizations to the both steps and show results of numerical studies involving the presented techniques.

Keywords: Sparse positive definite matrices, Cholesky decomposition, Gauss elimination method, orthogonalization, band matrices.

References

[1] Basna, R., Nassar, H., and Podgórski, K. (2022). Data driven orthogonal basis selection for functional data analysis. *Journal of Multivariate Analysis*, 189:104868.

[2] Liu, X., Nassar, H., Podgórski, K. (2022). Dyadic diagonalization of positive definite band matrices and efficient b-spline orthogonalization. *Journal of Computational and Applied Mathematics*, 414: 114444.

Abstract - Contributed Speakers

Revisitation of matrix partial orderings

Oskar Maria Baksalary
Faculty of Physics, Adam Mickiewicz University, Poznań, Poland

Abstract

In 2008 the author delivered a talk at the *Fifth Autumn Symposium of the Research Training Group (Graduate School) "Statistical Modeling"*, in Bommerholz near Dortmund, Germany. The address, entitled *Along Route 66 with Götz Trenkler*, celebrated Professor Götz Trenkler's 65th birthday. At that time the set of joint publications of Professor Trenkler and the author consisted of four research articles and 7 problems/solutions published in *IMAGE – the Bulletin of the International Linear Algebra Society*. At present the set contains 41 joint articles and over 40 problems/solutions published in different mathematical journals.

The present talk will reach back to 2008 and lead through the 15 subsequent years of our joint research (and friendship). A subjective selection of the most valuable results obtained during this period will be presented. A particular attention will be paid to one of our recent papers dealing with the notion of a partial order, namely [O.M. Baksalary, G. Trenkler, A partial ordering approach to characterize properties of a pair of orthogonal projectors. Indian Journal of Pure and Applied Mathematics 52 (2021) 323-334]. Several new findings concerned with the topic will be indicated as well.

Keywords: Professor Götz Trenkler's 80th birthday on July 14, 2023.

Leveraging properties of Adjacency Matrices for topology learning in partially observed networks of physical flow

Deepjyoti Deka
Los Alamos National Laboratory, Los Alamos, USA

Abstract

Physical Flow networks model multiple large scale systems that are part of modern society, including power grids, gas networks and thermal-systems. Such networks, in both static or dynamic regime, follow flow conservation and are potential driven. For example, flows on lines/edges in power grids follow Kirchhoff's laws. Mathematically, the relation between network states θ and external disturbances p are represented as

$$p = H_B\theta \text{ where } H_B \text{ is the weighted adjacency matrix for network } \mathcal{G} = (\mathcal{N}, \mathcal{L}) \text{ with lines } \mathcal{L}. \quad (1)$$

Estimating the topology of physical flow networks has multiple applications in control, fault detection and cyber security of corresponding networks. Often measurements available for topology learning are limited to state measurements at a subset of nodes in the network. Further, as physical flow networks often pertain to critical infrastructure, learning algorithms for topology necessitate guarantees on correctness and convergence.

In this talk, we show that properties of H_B (weighted Laplacian matrix) provide a rich framework to arrive at tractable learning algorithms for physical flow networks, using only partially observed nodal states θ. We first show that for radial networks, in particular, Laplacian matrix leads to the development of efficient greedy algorithms for topology recovery using second order statistics of state values. Next we use properties of inverse Laplacian matrices on meshed networks within the framework of probabilistic graphical models to prove topology recovery using observations that scale polynomially with the size of the network [1, 2]

Keywords: Monotonicity, flow conservation, dynamical Systems, power grids, graphical models, concentration bounds.

References

[1] D. Deka, V. Kekatos, and G. Cavraro. "Learning Distribution Grid Topologies: A Tutorial." IEEE Transactions on Smart Grid (2023).

[2] H. Doddi, D. Deka, S. Talukdar, and M. Salapaka. "Efficient and passive learning of networked dynamical systems driven by non-white exogenous inputs." In International Conference on Artificial Intelligence and Statistics, pp. 9982-9997. PMLR, 2022.

Construction of the Design Matrix for Generalized Linear Mixed-Effects Models in the Context of Clinical Trials of Treatment Sequences

Francisco J. Diaz
Department of Biostatistics and Data Science, The University of Kansas
Medical Center, Kansas City, KS, United States

Abstract

The estimation of carry-over effects is a difficult problem in the design and analysis of clinical trials of treatment sequences including cross-over trials. Except for simple designs, carry-over effects are usually unidentifiable and therefore nonestimable. Solutions such as imposing parameter constraints are often unjustified and produce differing carry-over estimates depending on the constraint imposed. Generalized inverses or treatment-balancing often allow estimating main treatment effects, but the problem of estimating the carry-over contribution of a treatment sequence remains open in these approaches. Moreover, washout periods are not always feasible or ethical. A common feature of designs with unidentifiable parameters is that they do not have design matrices of full rank. Thus, we propose approaches to the construction of design matrices of full rank, without imposing artificial constraints on the carry-over effects. Our approaches are applicable within the framework of generalized linear mixed-effects models. We present a new model for the design and analysis of clinical trials of treatment sequences, called Antichronic System, and introduce some special sequences called Skip Sequences. We show that carry-over effects are identifiable only if appropriate Skip Sequences are used in the design and/or data analysis of the clinical trial. We explain how Skip Sequences can be implemented in practice, and present a method of computing the appropriate Skip Sequences. We show applications to the design of a cross-over study with 3 treatments and 3 periods, and to the data analysis of the STAR*D study of sequences of treatments for depression ([1]).

Keywords: carry-over effects, cross-over design, estimability, generalized inverses, generalized least squares, identifiability, random effects linear models

References

[1] Diaz, F. J. (2018). Construction of the Design Matrix for Generalized Linear Mixed-Effects Models in the Context of Clinical Trials of Treatment Sequences. *Revista Colombiana de Estadística*,41:191-233.

Variance change point detection with a Binomial method

Elias Erdtman[1], Dietrich von Rosen[1,2] and Martin Singull[1]
[1]Department of Mathematics, Linköping University, Linköping, Sweden
[2]Energy & Technology, Swedish University of Agricultural Sciences, Uppsala, Sweden

Abstract

We consider a simple method for detecting a change with respect to the variance in a sequence of independent normally distributed observations with a constant mean. The method filters out extreme values and divides the sequence into equally large subsequences. For each subsequence, the count of extreme values is translated as a binomial random variable which is tested toward the expected number of extremes. The expected number of extremes comes from prior knowledge of the sequence and a specified probability of how common an extreme value should be. Then specifying the significance level of the goodness of fit test gives how many extreme observations are needed to detect a change.

The approach is extended to a sequence of independent multivariate normally distributed observations by using the Mahalanobis distance to transform the sequence into a univariate sequence and apply the same approach.

Keywords: Change point, Variance

Comments on the estimation and testing under multivariate and matrix, normal and t-distributions

Katarzyna Filipiak
Institute of Mathematics, Poznan University of Technology, Poland

Abstract

In this talk we compare the estimators of unknown parameters under multivariate normal and multivariate t distribution as well as we show respective estimators expressed in terms of matrix normal and matrix t distribution. The comparison of likelihood ratio test for testing covariance structure under multivariate and matrix t distributions will be also given.

Keywords: multivariate normal distribution, matrix normal distribution, multivariate t-distribution, matrix t-distribution

Dimension Reduction for Clustering Italian Children According to their Attitude towards Food Consumption

Cinzia Franceschini[1] and Nicola Loperfido[2]
[1]Università di Scienze Gastronomiche di Pollenzo - Pollenzo (Bra), Italy
[2]Dipartimento di Economia, Società e Politica (DESP), Università degli Studi di Urbino "Carlo Bo", Urbino, Italy

Abstract

The University of Gastronomic Sciences (Pollenzo, Italy) conducted a survey aimed at investigating the Italian children's attitude towards food and its consumption at school. Data were collected from questionnaires administered to 1108 children all over Italy. First, we used model-based clustering and k-means clustering on the original data. Then we used principal component analysis to reduce the number of variables before clustering, a procedure often referred to as to either tandem clustering ([1]) or reduced k-means ([3]). Finally, we clustered the data projections onto the directions found by means of projection pursuit and invariant coordinate selection. Projection pursuit is a multivariate statistical technique aimed at finding interesting low-dimensional data projections. It addresses three major challenges of multivariate analysis: the curse of dimensionality, the presence of irrelevant features and the limitations of visual perception ([2]). Invariant coordinate selection is a multivariate statistical method aimed at detecting data structures by means of the simultaneous diagonalization of two scatter matrices. Statistical applications of invariant coordinate selection include cluster analysis, independent component analysis, outlier detection and regression analysis ([4]). Our statistical analysis makes a case for projecting the data onto lower dimensional subspaces before clustering, even when the number of variables is much smaller than the number of units. Also, choices of subspaces guided by either projection pursuit or invariant coordinate selection might lead to better clustering results than those guided by principal component analysis.

Keywords: K-means clustering, Invariant coordinate selection, Principal component analysis, Projection pursuit, Reduced K-means, Tandem analysis

References

[1] Arabie, P. and Hubert, L. (1994). Cluster analysis in marketing research. *In Advanced methods of marketing research*, (ed Bagozzi, R. P.), Blackwell, Oxford; 160-189.

[2] Sun, J. (2006). Projection Pursuit. *Encyclopedia of Statistical Sciences*, 10.

[3] Terada, Y. (2014). Strong Consistency of Reduced K-means Clustering. *Scandinavian Journal of Statistics*, 41:913 -931.

[4] Tyler, D. and Critchley, F. and Dümbgen L. and Oja, H. (2009). Invariant co-ordinate selection (with discussion). *J. R. Statist. Soc. B*, 71:549-592.

Convex optimization geometry of nonnegative variance estimators in time series kriging

Martina Hančová[1] and Jozef Hanč[2]
[1]Institute of Mathematics, P. J. Šafárik University, Košice, Slovakia
[2]Institute of Physics, P. J. Šafárik University, Košice, Slovakia

Abstract

Within the framework of convex optimization geometry [1], we investigate the theoretical properties and computational algorithms for nonnegative variance estimators as applied in time series predictions. Our forecasting approach, generally referred to as time series kriging [2], is founded on linear time series models where observations are described by a linear mixed model. Employing such a geometrical perspective not only fosters an intuitive and natural comprehension of the given estimators but also serves as a very valuable guide for accelerating and reducing computational complexity in numerical calculations. We use the convex geometry to improve existing algorithms for nonnegative estimators based on least squares or maximum likelihood [2].

These improvements are crucial in computational research as a third paradigm in controlling and developing mathematical theory and practice. Monte Carlo and bootstrap simulations, implementing our efficient algorithms, extend beyond time series econometrics and finance, proving their importance also for fast data analysis based on exact probability distributions in areas like multidimensional statistics or measurement uncertainty analysis in metrology. We illustrate our results through a systematic design of simulation experiments [3] using high-performance computing and open data science tools [4].

Keywords: time series prediction, linear mixed models, matrix analysis, high-performance computing

Funding: This work was supported by the Slovak Research and Development Agency under the Contract no. APVV-21-0216 and APVV-21-0369.

References

[1] Dattorro, J. (2019). *Convex Optimization Euclidean Distance Geometry 2e.* Meboo Publishing: Palo Alto.

[2] Hančová, M., Gajdoš, A., Hanč, J., and Vozáriková, G. (2021). Estimating variances in time series kriging using convex optimization and empirical BLUPs. *Stat Papers 62*, 4:1899–1938.

[3] Lorscheid, I., Heine, B. O., and Meyer, M. (2012). Opening the 'black box' of simulations: increased transparency and effective communication through the systematic design of experiments. *Comput Math Organ Theory 18*, 1:22–62.

[4] Hančová, M., Gajdoš, A., and Hanč, J. (2022). A practical, effective calculation of gamma difference distributions with open data science tools. *J. Stat. Comput. Simul. 92*, 11:2205–2232.

Equality of BLUEs and their covariances under error covariance change for a linear model and its submodels, with links to data confidentiality and encryption

Stephen Haslett
School of Mathematical and Computational Sciences & Research Centre for Hauora and Health, and Environmental Health Intelligence NZ, Massey University, Palmerston North & Wellington, Aoteoroa/New Zealand; Research School of Finance, Actuarial Studies and Statistics, The Australian National University, Canberra ACT 2601 Australia; NIASRA, Faculty of Engineering and Information Sciences, University of Wollongong, NSW 2522, Australia.

Abstract

The necessary and sufficient condition for BLUEs of estimable functions of parameters in a linear fixed effect model being un-altered by a change in error covariance structure is due to [1]. When the original full linear model is made smaller by reducing the number of regressors (which may include interactions of any order), block diagonal or diagonal matrices also provide insight into conditions for the entire set of full, small, and intermediate models each to retain their own BLUEs ([3]; [2]). The role that such changes in error covariance structure can play in data confidentiality and data encryption is outlined, especially when the covariance of the BLUEs is also retained.

This talk is based on joint work with Jarkko Isotalo, Augustyn Markiewicz and Simo Puntanen.

Keywords: BLUE, BLUP, Confidentialised unit record files, Covariance, Data cloning, Data confidentiality, Encryption, Linear model, Residuals.

References

[1] Rao, C. R. (1971) Unified theory of linear estimation (corr: 72v34 p 194; 72v34 p 477). *Sankhyä Ser A*, 33:371-394.

[2] Haslett, S. J., Isotalo, J., Markiewicz, A., Puntanen, S. (2023) Permissible covariance structures for simultaneous retention of BLUEs in small and big linear models. Chapter 11 in *Springer proceedings of the conferences in honor of C R Rao & A K Lal: Applied Linear Algebra, Probability and Statistics (ALAPS).*

[3] Haslett, S. J., Puntanen, S. (2023) Equality of BLUEs for full, small, and intermediate linear models under covariance change, with links to data confidentiality and encryption. Chapter 14 in *Springer proceedings of the conferences in honor of C R Rao & A K Lal: Applied Linear Algebra, Probability and Statistics (ALAPS).*

Spatial weight matrices and their properties

Jan Hauke[1] and Martin Singull[2]
[1]Adam Mickiewicz University, Poznan, Poland
[2]Linköping University, Linköping, Sweden

Abstract

Spatial econometrics models are inextricably linked with the consideration of spatial effects, the importance of which has been proven both theoretically and empirically. Therefore, an important element of this modeling is to determine the type and scope of existing spatial relationships between areas, by constructing spatial weights to reflect spatial interactions. The starting point is the consideration of two types of spatial relationships reflecting side effects modeled by incorporating a (spatially weighted) variable directly into the model (spatial lag model) or by incorporating a spatial relationship in the (spatially weighted) error term in the model (spatial error). One of the simplest ways to attain identification in spatial models, as is common in applied literature, is specifying exogenously a weighting matrix **W** (it has an arbitrary decided formula). The exogenous approach is by far the most common. It includes, i.a., the use of a binary contiguity criterion, k-nearest neighbours, and kernel functions based on distance. This approach, in addition to modeling, allows for quantitative estimation of the strength of spatial influence (based on the **W** matrix) on the change in the value of the observed features. One of the indicators used for this purpose is the Morans I coefficient. The second approach is estimating **W** from data. This approach has some drawbacks, as discussed by [2]. Regardless of the method of introducing the **W** matrix into the model, it plays a key role in spatial econometrics, see [3], [1], and [4]. The properties of the weighting matrix **W**, taking into account various aspects and models, will be discussed in this presentation.

Keywords: Spatial models, spatial weight matrices

References

[1] Bauman, D., Drouet, T., Fortin, M. J., and Dray, S. (2018). Optimizing the choice of a spatial weighting matrix in eigenvector-based methods. *Ecology*, 99: 2159-2166.

[2] Harris, R., Moffat, J., and Kravtsova, V. (2011). In search of *"W"*. *Spatial Economic Analysis*, 6: 249-270.

[3] Hays, J. C., Kachi, A., and Franzese Jr, R. J. (2010). A spatial model incorporating dynamic, endogenous network interdependence: A political science application. *Statistical Methodology*, 7: 406-428.

[4] Lam, C., and Souza, P. C. (2020). Estimation and selection of spatial weight matrix in a spatial lag model. *Journal of Business and Economic Statistics*, 38: 693-710.

Eigenvalues of large sample correlation matrices

Johannes Heiny
Stockholm University, Stockholm, Sweden

Abstract

Many fields of modern sciences are faced with high-dimensional data sets. In this talk, we investigate the spectral properties of a large sample correlation matrix R. Results for the spectral distribution, extreme eigenvalues and functionals of the eigenvalues of R are presented in both light- and heavy-tailed cases. The findings are applied to independence testing and to the volume of random simplices.

Keywords: Random matrix, high dimension, correlation matrix, eigenvalues, spectral statistics.

Prediction and Testing of Random Effects in Linear Mixed Models

Jarkko Isotalo[1], Stephen Haslett[2] and Simo Puntanen[1]
[1]Tampere University, Tampere, Finland
[2]Massey University, Palmerston North & Wellington, New Zealand

Abstract

We consider the linear mixed effects model

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \boldsymbol{\varepsilon},$$

where $\mathbf{y}$ is an $n \times 1$ observable random vector, $\mathbf{X}$ and $\mathbf{Z}$ are known model matrices, $\boldsymbol{\beta}$ is a $p \times 1$ vector of unknown fixed parameters, $\mathbf{u}$ is a $q \times 1$ vector of unobservable random effects, and $\boldsymbol{\varepsilon}$ is an $n \times 1$ unobservable random error vector. We further assume that random vectors $\mathbf{u}$ and $\boldsymbol{\varepsilon}$ are uncorrelated and are normally distributed, i.e., $\mathbf{u} \sim N(\mathbf{0}, \sigma^2\mathbf{G}_{\boldsymbol{\theta}}), \boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2\mathbf{R}), Cov(\mathbf{u}, \boldsymbol{\varepsilon}) = \mathbf{0}$, where σ^2 is a positive unknown scalar and $\boldsymbol{\theta}$ is an unknown parameter vector.

In this setup, we consider different methods for predicting the conditional mean $\mathbf{x}_*'\boldsymbol{\beta} + \mathbf{z}_*'\mathbf{u}$, when $\mathbf{x}_*$ and $\mathbf{z}_*$ are known given vectors. Particularly, we give conditions when the BLUP of $\mathbf{x}_*'\boldsymbol{\beta} + \mathbf{z}_*'\mathbf{u}$ is equal to the BLUE of the conditional mean, see [1]. Obtained results on prediction are then applied to the problem of testing hypotheses set on the conditional mean. In linear mixed models, hypothesis testing related to the random effects are often done by defining the structure of the covariance matrix $\mathbf{G}_{\boldsymbol{\theta}}$ in competing hypotheses and then testing them by using the likelihood ratio statistic or by some other suitable test statistic, see, e.g.,[2]. In this research, we set hypotheses on the conditional mean $\mathbf{x}_*'\boldsymbol{\beta} + \mathbf{z}_*'\mathbf{u}$ and then consider different approaches for testing them. Specifically, we show that testing statistic obtained by the interval prediction with use of the BLUP is equivalent to the extended likelihood ratio test statistic.

Keywords: BLUE, BLUP, Likelihood-ratio test.

References

[1] Puntanen, S., Isotalo, J. and Haslett, S. J., (2021). Equalities between the BLUEs and BLUPs under the partitioned linear fixed model and the corresponding mixed model, *Acta et Commentationes Universitatis Tartuensis de Mathematica*, 5, 2, 239-257.

[2] Hui, F. K. C., Muller, S. M. and Welsh, A. H., (2019). Testing random effects in linear mixed models: another look at the F-test (with discussion), *Australian and New Zealand Journal of Statistics*. 61(1), 61-84.

On the Expectations of Equivariant Matrix-valued Functions of Wishart and Inverse Wishart Matrices

Grant Hillier[1] and Raymond Kan[2]
[1]University of Southampton, Southampton, United Kingdom
[2]University of Toronto, Toronto, Canada

Abstract

Many matrix-valued functions of an $m \times m$ Wishart matrix W, $F_k(W)$, say, are homogeneous of degree k in W, and are equivariant under the conjugate action of the orthogonal group $\mathcal{O}(m)$, i.e., $F_k(HWH') = HF_k(W)H'$, $H \in \mathcal{O}(m)$. It is easy to see that the expectation of such a function is itself homogeneous of degree k in Σ, the covariance matrix, and are also equivariant under the action of $\mathcal{O}(m)$ on Σ. The space of such homogeneous, equivariant, matrix-valued functions is spanned by elements of the type $W^r p_\lambda(W)$, where $r \in \{0, \dots, k\}$ and, for each r, λ varies over the partitions of $k - r$. Here, $p_\lambda(W)$ denotes the power-sum symmetric function indexed by λ. In the analogous case where W is replaced by W^{-1}, these elements are replaced by $W^{-r} p_\lambda(W^{-1})$. In this paper we derive recurrence relations and analytical expressions for the expectations of such functions. Our results provide highly efficient methods for analysing the properties of, and the computation of, all such moments, even those of very high order k. We thus provide a complete toolbox for analysing the properties of any matrix-valued function in this class.

Keywords: Wishart distribution, Inverse Wishart distribution, Equivariant function, Recursive relation

MLEs and Hotelling's T^2 Statistics of a class of Matrix Normal Distribution Arisen from Phylogenetic Problems and Their Properties

Hao Chi Kiang
Linköping University, Linköping, Sweden

Abstract

Brownian Motion along phylogeny is commonly used to model trait evolution along a given phylogenetic tree [1]. This phylogenetic Brownian Motion model can be reduced to $\text{vec}X \sim N(\mu \otimes 1, \Sigma \otimes C)$ where X is an $n \times p$ matrix; C is an $n \times n$ given (constant) positive definite matrix and μ would be the trait value of the root of the tree.

Maximum likelihood estimators $\hat{\mu}$ and $\hat{\Sigma}$ of this model is well-known, but their exact distribution properties are much less discussed nor utilized. My work shows that 1) μ and $\frac{n}{n-1}\hat{\Sigma}$ are in fact unbiased; 2) that $\frac{n}{n-1}\hat{\Sigma}$ is Wishart-distributed; 3) that there is independence between $\hat{\mu}$ and $\hat{\Sigma}$; 4) that there is a Hotelling's T^2 distribution around μ, which can provide an exact confidence region; 5) a useful sufficient condition for the XAX^T to be independent of BXD for any suitably-sized matrices A, B, and D when the data matrix X were generated from this Brownian Motion model.

Keywords: matrix normal distribution, kronecker product covariance, confidence region, maximum likelihood estimator

References

[1] J. Felsenstein, 'Phylogenies and the Comparative Method', The American Naturalist, vol. 125, no. 1, pp. 1–15, 1985.

Matrix Operations for Tensor Algebra, with Statistical Applications

Nicola Loperfido
Dipartimento di Economia, Società e Politica (DESP), Università degli Studi di Urbino "Carlo Bo", Urbino, Italy

Abstract

The star product [1] and the matrix reshaping [2] are two matrix operations acting on block matrices. The star product of two matrices is the matrix arranging linear combinations of the block of the second matrix, where the combinations' coefficients are the elements of the first matrix. The reshaping of a block matrix vectorizes, transposes and stacks on top of each other the blocks of the matrix itself. Both the star product and the matrix reshaping connect matrices and tensors, since block matrices might be regarded as the unfoldings of third and fourth order tensors. Firstly, we investigate the mathematical properties of the two operations with respect to other matrix operations, as for example matrix multiplication, matrix transposition and tensor product. Secondly, we investigate the relationships between star product, tensor unfolding and tensor contraction. Thirdly, we use the matrix transposition, the commutation matrix and the matrix reshaping for tensor symmetrization and symmetric tensor decompositions. As a first statistical application, we consider multivariate density approximation, where a normal, finite location mixture is used to approximate a multivariate distribution with the same first three cumulants. As a second statistical application, we use tensor algebra to deal with random samples from random matrices.

Keywords: Block matrix, Density approximation, Finite mixture, Higher-order moments, Subtensor, Tensor contraction, Tensor unfolding.

References

[1] MacRae, E. (1974). Matrix derivatives with an application to an adaptive linear decision problem. *Annals of Statistics* 2:337-346.

[2] Van Loan, C. and Pitsianis, N. P. (1992). Approximation with Kronecker products. In: M. S. Moonen, G. H. Golub (Eds.), *Linear Algebra for Large Scale and Real Time Applications*, Kluwer Publications, Dordrecht, 293-314.

Fourth cumulant for the random sum of random vectors

Stepan Mazur[1,2], Farrukh Javed[3] and Nicola Loperfido[4]
[1]Örebro University, Örebro, Sweden
[2]Linnaeus University, Växjö, Sweden
[3]Lund University, Lund, Sweden
[4]Università degli Studi di Urbino "Carlo Bo", Urbino (PU), Italy

Abstract

The fourth cumulant for the random sum of random vectors is considered. A formula is presented for the general case when the aggregating variable is independent of the random vectors. Two important special cases are considered. In the first one, multivariate skew-normal random vectors are considered that are aggregated by a Poisson variable. The second case deals with multivariate asymmetric generalized Laplace random vectors and aggregation is made by a negative binomial variable. There is a well-established relation between asymmetric Laplace motion and negative binomial process that corresponds to the invariance principle of the random sum of random vectors for the generalized asymmetric Laplace distribution. We explore this relation and provide a multivariate continuous time version of the results.

Keywords: Fourth cumulant, Random sum of random vectors, Skew-normal, Laplace motion

Applications of partial and block trace operators

Monika Mokrzycka[1], Katarzyna Filipiak[2] and Daniel Klein[3]
[1]Institute of Plant Genetics, Polish Academy of Sciences, Poznań, Poland
[2]Institute of Mathematics, Poznań University of Technology, Poland
[3]Institute of Mathematics, Faculty of Science, P. J. Šafárik University, Košice, Slovakia

Abstract

Let us consider an experiment regarding some independent objects, for which we measure few characteristics at few time or location points. For two such sources of variability, the separability is the natural assumption for the covariance matrix structure, and means that the time or location points are correlated "independently" of characteristics and characteristics are correlated "independently" of time or location points. In the literature maximum likelihood estimators of the positive definite covariance matrix having the separable structure have been proposed, however, the explicit form is not available. It means, that to get the MLEs one should solve numerically respective system of equations. The aim of the talk is to compare the available approaches and to show, that the one based on the partial and block trace operators is the most efficient.

Since similar numerical problems appear in the estimation of separable covariance matrix via Frobenius norm or entropy loss function, we also show that the method based on partial and block trace operators outperforms the existing approaches.

Keywords: covariance matrix, separable structure, estimation, partial trace operator, block trace operator

Optimizing the Allocation of Trials to Sub-Regions in Multi-Environment Crop Variety Testing for Multi-Annual Experiments

Maryna Prus and Hans-Peter Piepho
University of Hohenheim, Stuttgart, Germany

Abstract

New crop varieties are usually evaluated for their performance in a target population of environments (TPE). This evaluation requires conducting randomized field trials at several environments sampled from the TPE. Such trials are called multi-environment trials (MET). If the TPE is large and can be suitably stratified along geographical borders or agro-ecological zonations, it may be advantageous to subdivide the TPE into sub-regions. If the same set of genotypes is tested at a number of locations in each of the sub-regions, a linear mixed model may be fitted with random genotype-within-sub-region effects. The first analytical results to optimizing allocation of trials to sub-regions have been obtained in [1]. That paper considers only a single year of trials. However, in practice the responses are usually being observed during several years. In this work we consider the extended linear mixed model that incorporates the influence of the years. We propose an analytical solution for optimal allocations of trials and illustrate the obtained results by a real data example.

Keywords: Target population of environments, Multi-environment trials, Linear mixed model, Prediction, Optimal design

References

[1] Prus, M. and Piepho, H.-P. (2021). Optimizing the allocation of trials to sub-regions in multi-environment crop variety testing. *Journal of Agricultural, Biological and Environmental Statistics*, 26: 267-288.

Personal Photographic Glimpses of Professor Götz Trenkler

Simo Puntanen
Tampere University, Tampere, Finland

Abstract

In addition to  more photos, since 1983, will be shown up.

Keywords: Dortmund, Tampere, Calcutta, Montreal, Shanghai, Windsor, Bedlewo, Tomar, Auckland, Uppsala, Lyngby, Hyderabad, Fort Lauderdale, . . . , Manipal, you name it!

Applications of the Kronecker product when solving matrix equations and constructing matrix derivatives

Dietrich von Rosen
Energy & Technology, Swedish University of Agricultural Sciences, Uppsala, Sweden, and Department of Mathematics, Linköping University, Linköping, Sweden

Abstract

The Kronecker product of two matrices generates a tensor space and an orthogonal complement to this space can be obtained. The result is useful when, for example, the matrix equation in $\boldsymbol{X}$, say $\boldsymbol{AXB} = \boldsymbol{0}$, with $\boldsymbol{A}$ and $\boldsymbol{B}$ known, has to be solved. The approach can be extended to cover the equation $\boldsymbol{A}_i\boldsymbol{X}\boldsymbol{B}_i = \boldsymbol{0}$, $i \in \{1, 2\}$, but it cannot handle $\boldsymbol{A}_i\boldsymbol{X}\boldsymbol{B}_i = \boldsymbol{0}$, $i \in \{1, 2, 3\}$.

Concerning matrix derivatives the Kronecker product can be used to define different types of derivatives. Let $\boldsymbol{X} = \sum_{ij} x_{ij}\boldsymbol{e}_i\boldsymbol{d}_j^\top$ and $\boldsymbol{Y} = \sum_{k\ell} y_{k\ell}\boldsymbol{g}_k\boldsymbol{f}_\ell^\top$, where $\boldsymbol{d}_j, \boldsymbol{e}_i, \boldsymbol{g}_k, \boldsymbol{f}_\ell$ are unit basis vectors. Then

$$\frac{d\boldsymbol{X}}{d\boldsymbol{Y}} = \sum_{i,j,k,\ell} \frac{\partial x_{ij}}{\partial y_{k\ell}} \boldsymbol{e}_i\boldsymbol{d}_j^\top \otimes \boldsymbol{g}_k\boldsymbol{f}_\ell.$$

There exist several alternative derivatives, for example,

$$\frac{d\boldsymbol{X}}{d\boldsymbol{Y}} = \sum_{i,j,k,\ell} \frac{\partial x_{ij}}{\partial y_{k\ell}} (\boldsymbol{d}_j \otimes \boldsymbol{e}_i)(\boldsymbol{f}_\ell \otimes \boldsymbol{g}_k)^\top.$$

How to Treat Partial Errors-In-Variables Models Efficiently

Burkhard Schaffrin[1] and Kyle Snow[2]*
[1]The Ohio State University, Division of Geodetic Science, School of Earth Sciences, Columbus, Ohio, USA
[2]Polaris Geospatial Services, Westerville, Ohio, USA

* With computational assistance from Shahram Jazaeri.

Abstract

The standard Errors-In-Variables (EIV) Model arises when all the entries of the coefficient matrix that describes the relationships between the unknown parameters and a collected data set stem from observations themselves and must be considered random. When some of these elements are, however, fixed, then one would face a Partial EIV-Model for which Xu et al. [4] proposed a split approach in which the parameters with fixed coefficients are determined by a classical Least-Squares approach and those with random coefficients by a (weighted) Total Least-Squares (TLS) procedure. In order to reduce the necessary bookkeeping for two sorts of coefficients, it is here proposed to consider the fixed coefficients as "random with zero variance" and, as a consequence, allow *singular dispersion matrices* for the original EIV-Model. Snow [3] created the first generation of algorithms for such generalized EIV-Models, and here the next generation will be presented which shows higher efficiency, in particular when the dispersion matrix of the coefficients exhibits a Kronecker-product structure as exploited by Schaffrin and Wieser [2].

Keywords: Errors-In-Variables Model, Total Least-Squares, singular covariance matrices

References

[1] Jazaeri S., Schaffrin, B., and Snow K. (2022). A simple TLS-treatment of the partial EIV-Model as one with singular cofactor matrices I: The case of a Kronecker product for $Q_A = Q_0 \otimes Q_x$, *Springer International Publishing, IAG Symp*, (under review).

[2] Schaffrin, B. and Wieser A. (2008). On weighted total least-squares adjustment for linear regression. *Journal of Geodesy*, 82:415–421.

[3] Snow, K. (2012). Topics in Total Least-Squares adjustment within the Errors-In-Variables Model: Singular cofactor matrices and prior information. *Div. of Geodetic Science, School of Earth Sciences, The Ohio State University*, Report No. 502.

[4] Xu, P., Liu J., and Shi C. (2012). Total least-squares adjustment in partial Errors-In-Variables models: algorithm and statistical analysis. *Journal of Geodesy*, 86:661–675.

What is the gradient of a scalar function defined on a subspace of square matrices?

Shriram Srinivasan and Nishant Panda
Los Alamos National Laboratory, Los Alamos, United States

Abstract

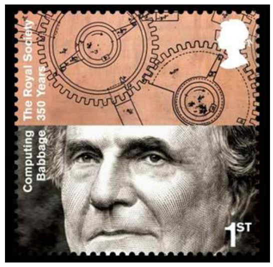
We illustrate a technique to calculate the gradient of scalar functions defined on *any* arbitrary matrix subspace. It generalizes our earlier work titled "What is the gradient of a scalar function of a symmetric matrix?", in which we considered the special case of the subspace of symmetric matrices. Extant methods to calculate the gradient in such cases have an inherent flaw that leads to spurious results which populate several publications, as well as respected textbooks and handbooks on matrix calculus. One of our important contributions has been to examine these sources and reproduce the spurious results in a rigorous and concrete mathematical setting of a finite-dimensional inner-product space. In this process, we discover the inherent flaw and also a remedy. We demonstrate two ways to calculate the derivative/gradient and second derivative for scalar functions of matrices defined over an arbitrary matrix subspace; the first method is by considering *any* (differentiable) extension to the space of square matrices and projection of its gradient onto the given subspace. The second method utilizes an ordered basis and computes each component of the gradient through evaluation of the directional derivative. All the ideas presented are illustrated by non-trivial examples. Moreover, the presentation of matrix calculus in the language of calculus on inner-product spaces will be significant and meaningful for engineers and researchers working in inter-disciplinary fields to avoid the conceptual pitfalls that exist.

Keywords: matrix calculus, patterned matrix, gradient, matrix functional

A philatelic introduction to Ada Lovelace, Lord Byron & Charles Babbage

George P. H. Styan
McGill University, Montréal (Québec), Canada

Abstract



Augusta Ada King, countess of Lovelace (1815–1852) was a daughter of the poet Lord George Gordon Byron (1788–1824). Ada Lovelace became interested in Babbage's analytic engine and described how it could be programmed. Charles Babbage (1791–1871) originated the modern analytic computer. He invented the principle of the analytical engine, the forerunner of the modern electronic computer. [*MacTutor*]

Keywords: Ada Lovelace, Lord Byron, Charles Babbage, mathematical philately, analytical engine, modern electronic computer, postage stamp, personalized stamp, stamps from Ukraine, Jeff Miller's website, `delcampe`, *MacTutor*, *Wikipedia*.

References

[1] A philatelic introduction to Ada Lovelace, Lord Byron & Charles Babbage in celebration of the 80th birthday of Götz Trenkler, beamer file by Oskar Maria Baksalary, Ka Lok Chu, Simo Puntanen & George P. H. Styan, to be presented at the 29th International Workshop on Matrices & Statistics (IWMS), Linköping, Sweden (21–23 August 2023), 20 pp.

On the Uniqueness of Correspondence Analysis Solutions

Michel van de Velden, Rick Willemsen, and Wilco van den Heuvel
Econometric Institute, Erasmus University Rotterdam, The Netherlands

Abstract

In correspondence analysis (CA), rows and columns of a contingency table are optimally represented by a k-dimensional approximation. As CA is a dimension reduction technique, one expects the k-dimensional approximation to be non-unique. That is, we expect that there are different contingency tables that lead to the same k-dimensional CA approximation. However, [1] find in computational experiments that, for the case where $k = 3$ (which is commonly used in CA applications) only one contingency table exists corresponding to the low dimensional CA solution. In this paper, we tackle this problem from a theoretical perspective. We show that k-dimensional CA solutions are not necessarily unique. That is, two distinct contingency tables may have the same k-dimensional approximation. We present necessary and sufficient conditions for the non-uniqueness of CA solutions. Furthermore, based on the sufficient conditions, we present a procedure to generate contingency tables with non-unique k-dimensional CA solutions. Finally, we note that the necessary conditions are rather restrictive and unlikely to be satisfied by empirical data. Hence, in practice, a CA solution most likely only corresponds to one contingency table.

Keywords: Correspondence analysis, inverse problems, uniqueness

References

[1] Van de Velden, M., Van den Heuvel, W., Galy, H. and Groenen, P. J. F. (2020). Retrieving a contingency table from a correspondence analysis solution. *European Journal of Operational Research*, 283:541-548.

Large deviations of extremal eigenvalues of sample covariance matrices

Martin Singull[1], Denise Uwamariya[2] and Xiangfeng Yang[1]
[1]Department of Mathematics, Linköping University, Linköping, Sweden
[2]Department of Mathematics, University of Rwanda, Rwanda

Abstract

Let $\mathbf{X}$ be a $p \times n$ random matrix whose entries are independent and identically distributed real sub-Gaussian random variables with zero mean and unit variance. (i) Large deviations of the largest and smallest eigenvalues of $\mathbf{X}\mathbf{X}^T/n$ are discussed in this talk, under the assumption that both the dimension size p and the sample size n tend to infinity with $p(n) = o(n)$. This study generalizes one result obtained in [1] and [2]. (ii) Large deviations of the 2-norm condition number of $\mathbf{X}$ are also discussed in the talk when p is either fixed or $p = p(n) \to \infty$ with $p(n) = o(n)$.

Keywords: Large deviations; sample covariance matrices; extremal eigenvalues; condition numbers

References

[1] Fey, A., van der Hofstad, R. and Klok, M. (2008). Large deviations for eigenvalues of sample covariance matrices, with applications to mobile communication systems. *Adv. in Appl. Probab.* **40** 1048–1071.

[2] Jiang, T. and Li, D. (2015). Approximation of rectangular beta-Laguerre ensembles and large deviations. *J. Theoret. Probab.* **28** 804–847.

Special Session - Tensor methods in Statistics

Organizer: Nicola Loperfido
Dipartimento di Economia, Società e Politica (DESP), Università degli Studi di Urbino "Carlo Bo", Urbino, Italy

Tensor methods are becoming increasingly popular in Statistics. They are particularly useful when dealing with higher moments of multivariate distributions or with data generated from matrix-variate distributions. This session presents some of the many applications of tensor methods in Statistics, and motivates them with both real data and theoretical arguments.

The session will consist of the following talks:

- Cinzia Franceschini – Dimension Reduction for Clustering Italian Children According to their Attitude towards Food Consumption
- Nicola Loperfido – Matrix Operations for Tensor Algebra, with Statistical Applications
- Stepan Mazur – Fourth cumulant for the random sum of random vectors
- Dietrich von Rosen – Applications of the Kronecker product when solving matrix equations and constructing matrix derivatives

Celebration of Götz Trenkler's 80th birthday

Special Session in celebration of Götz Trenkler's 80th birthday – Matrix analysis as a basis of applied sciences

Organizers: Oskar Maria Baksalary[1], Martin Singull[2], and Simo Puntanen[3]
[1]Adam Mickiewicz University, Poznań, Poland
[2]Linköping university, Linköping, Sweden
[3]Tampere University, Tampere, Finland

The session celebrates the 80th birthday of Professor Götz Trenkler, which took place on July 14, 2023. Professor Trenkler is at present Professor Emeritus at the Dortmund University of Technology, where he has spent last forty years of his scientific career, since he was appointed Full Professor of Statistics and Econometrics in 1983. Scientific interests of Professor Trenkler spread over several research areas, such as: applications of mathematical methods (e.g., in econometrics, physics, statistics), pure and applied linear algebra, statistical inference, and mathematical education. Professor Trenkler is an author or coauthor of eight monographs, around 200 scientific articles, and about 200 other contributions published in scientific journals. He has supervised 25 Ph.D. students. A light was shed on selected (and then up-to-dated) scientific achievements of Professor Trenkler in a Festschrift dedicated to him on the occasion of his 65th birthday published in 2009 [*Statistical Inference, Econometric Analysis and Matrix Algebra – Festschrift in Honour of Götz Trenkler* (B. Schipp, W. Krämer, eds.), Springer, Heidelberg, 2009, DOI.

Professor Trenkler has been a frequent participant of the meetings within the series of International Workshops on Matrices and Statistics (IWMS), taking part already in the very first of them organized in Tampere, Finland, in 1990 (under the name International Workshop on Linear Models, Experimental Designs, and Related Matrix Theory). Since then Professor Trenkler has taken part in 14 workshops bearing the logo of IWMS, for several years acting as a member of their International Organizing Committee. In 2003 Professor Trenkler was the chair of the Local Organizing Committee of The Twelfth International Workshop on Matrices and Statistics held at the University of Dortmund.

The session will consist of the following talks:

- Oskar Maria Baksalary – Revisitation of matrix partial orderings,
- Jarkko Isotalo – Prediction and Testing of Random Effects in Linear Mixed Models,
- Simo Puntanen – Personal photographic glimpses of Professor Götz Trenkler,
- George P. H. Styan, – A philatelic introduction to Ada Lovelace, Lord Byron & Charles Babbage (online).

Further contributions to the session by Professor Trenkler's collaborators, colleagues, friends, and scientific followers are foreseen and welcome.

A philatelic introduction to Alfred Nobel, Sir Roger Penrose & Zu Chongzhi

Oskar Maria Baksalary[1] and George P. H. Styan[2]
[1]Faculty of Physics, Adam Mickiewicz University, Poznań, Poland
[2]McGill University, Montréal (Québec), Canada

Abstract

Alfred Bernhard Nobel (1833–1896) was a Swedish chemist, inventor, and philanthropist, well known for having bequeathed his fortune to establish the Nobel Prize. [*MacTutor*] "Sir Roger Penrose (b. 1931) was awarded one half of the 2020 Nobel Prize in Physics for the discovery that **black hole** formation is a robust prediction of the general theory of relativity". [*Wikipedia*] Zu Chongzhi (429–500) was a "Chinese astronomer and mathematician most notable for calculating π as between 3.1415926 and 3.1415927". [*Wikipedia*]

Keywords: Alfred Bernhard Nobel, Nobel Prize, Sir Roger Penrose, Zu Chongzhi, mathematical philately, postage stamp, personalized stamp, stamps from Ukraine, approximations to π, Jeff Miller's website, `delcampe`, *MacTutor*, *Wikipedia*.

References

[1] A philatelic introduction to Alfred Nobel, Sir Roger Penrose & Zu Chongzhi in celebration of the 80th birthday of Götz Trenkler, beamer file by Oskar Maria Baksalary, Ka Lok Chu, Simo Puntanen & George P. H. Styan, poster to be presented at the 29th International Workshop on Matrices & Statistics (IWMS), Linköping, Sweden (21–23 August 2023), 37 pp.

A philatelic introduction to Stefan Banach, Stanisław Ulam & John von Neumann

Oskar Maria Baksalary[1] and George P. H. Styan[2]
[1]Faculty of Physics, Adam Mickiewicz University, Poznań, Poland
[2]McGill University, Montréal (Québec), Canada

Abstract

Stefan Banach (1892–1945) is generally considered one of the 20th-century's most important and influential mathematicians and was an original member of the Lwów School of Mathematics, a group of Polish mathematicians who worked in the interwar period in Lwów, Poland (since 1945 Lviv, Ukraine). Stanisław Ulam (1909–1984) devised the "Monte-Carlo method" widely used in solving mathematical problems using statistical sampling. John von Neumann (1903–1957) was regarded as having perhaps the widest coverage of any mathematician of his time and was said to have been "the last representative of the great mathematicians who were equally at home in both pure and applied mathematics".

Keywords: Stefan Banach, Stanisław Ulam, John von Neumann, mathematical philately, postage stamp, personalized stamp, stamps from Ukraine, Lwów School of Mathematics, *The Scottish Book,* Jeff Miller's website, `delcampe`, *MacTutor*, *Wikipedia.*

References

[1] A philatelic introduction to Stefan Banach, Stanisław Ulam & John von Neumann, in celebration of the 80th birthday of Götz Trenkler, beamer file by Oskar Maria Baksalary, Ka Lok Chu, Simo Puntanen & George P. H. Styan, poster to be presented at the 29th International Workshop on Matrices & Statistics (IWMS), Linköping, Sweden (21–23 August 2023), 15 pp.

A philatelic introduction to perfect numbers & Mersenne primes

Oskar Maria Baksalary[1] and George P. H. Styan[2]
[1]Faculty of Physics, Adam Mickiewicz University, Poznań, Poland
[2]McGill University, Montréal (Québec), Canada

Abstract

A "perfect number" is a positive integer that is equal to the sum of its positive divisors, excluding the number itself. It is not known when perfect numbers were first studied and indeed the first studies may go back to the earliest times when numbers first aroused curiosity. This definition is ancient, appearing as early as Euclid's *Elements* (VII.22). Euclid also proved a formation rule (IX.36) whereby $q(q+1)/2$ is an even perfect number whenever q is a prime of the form $2^p - 1$ for positive integer p, which is now called a "Mersenne prime". Two millennia later, Leonhard Euler proved that all even perfect numbers are of this form. This is now known as the Euclid–Euler theorem. The stamps displayed here, from left to right, depict Euclid, Euler and Mersenne, and were issued by Ukraine (2019), Guinea-Bissau (2009) and Ukraine (2022).

Keywords: perfect numbers, Mersenne primes, Jonathan Pace, 51st largest known prime number, Great Internet Mersenne Prime Search (GIMPS), church computer in Memphis suburb, Euclid–Euler theorem, Euler–Mascheroni constant, stamps from Ukraine, mathematical philately, postage stamps, personalized stamps, Jeff Miller's website, `delcampe`, *MacTutor*, *Wikipedia*.

References

[1] A philatelic introduction to perfect numbers & Mersenne primes, in celebration of the 80th birthday of Götz Trenkler, beamer file by Oskar Maria Baksalary, Ka Lok Chu, Simo Puntanen & George P. H. Styan, poster to be presented at the 29th International Workshop on Matrices & Statistics (IWMS), Linköping, Sweden (21–23 August 2023), 15 pp.

Matrix derivatives and Kronecker products for the core and generalised core inverses

Shuangzhe Liu[1]*, **Hongxing Wang**[2], **Yonghui Liu**[3], **Conan Liu**[4]

[1]Faculty of Science and Technology, University of Canberra, Australia

[2]College of Mathematics and Physics, Guangxi Minzu University, China

[3]School of Statistics and Information, Shanghai University of International Business and Economics, China

[4]Business School, University of New South Wales, Australia

This paper is dedicated to Professor Götz Trenkler on his 80th birthday.

Abstract

We first present the Moore-Penrose, Drazin, core and some generalised core inverses, and then use the matrix differential calculus to establish matrix derivatives for some of the important matrix functions involving these gereralised inverses. We include the generalised inverses of the Kronecker product of two matrices as well.

Keywords: Moore-Penrose inverse, Drazin inverse, group inverse, core inverse, generalised core inverse, matrix differential calculus, Kronecker product

1 Introduction

On the occasion of Professor Götz Trenkler's 80th birthday, we would like to honour his exceptional contributions to the field of mathematics and statistics. His work encompasses a broad range of topics, including non-negative definite matrices, Löwner ordering, matrix inequalities, partitioned matrices, matrix and cross products, the Moore-Penrose inverse, core and generalized core inverses of matrices and operators, the ordinary least squares estimator, best linear unbiased estimator, estimators in restricted regression models, biased estimators, mean square errors, and other topics with applications ranging from mathematics, econometrics to nonparametric statistics. Through his expertise, Professor Trenkler has inspired and influenced numerous researchers within our community, enriching our collective knowledge.

With deep gratitude and respect, we proudly dedicate this paper to Professor Götz Trenkler in celebration of his 80th birthday and his outstanding achievements in academia and life. The focus of this paper lies on several matrix functions involving some matrix inverses, a subject that Professor Trenkler has devoted considerable time and expertise to studying throughout his illustrious career.

*Corresponding author: S. Liu. Email: shuangzhe.liu@canberra.edu.au

Solvability of a system of linear equations – an approach based on the generalized inverses determined by the Penrose equations[1]

Oskar Maria Baksalary

Faculty of Physics, Adam Mickiewicz University,
ul. Uniwersytetu Poznańskiego 2, PL 61-614 Poznań, Poland

Abstract. The paper aims to play an expository role, providing a tailored introduction to the theory of generalized inverses determined by the so-called Penrose equations, with the Moore–Penrose inverse as the jewel in the crown. A particular attention is paid to the applications of the inverses to the solvability of a system of linear equations, which covers inter alia the least squares method. Various links between the generalized inverses and the theory of projectors are also discussed, indicating issues which are relevant from the point of view of physics. In fact, the paper can be viewed as a sequel of [Baksalary, O.M. and Trenkler, G., The Moore–Penrose inverse – a hundred years on a frontline of physics research, The European Physical Journal H, 46 (2021) 9], the paper prepared to celebrate the 100th anniversary of the first definition of the Moore–Penrose inverse, which shades a spotlight on the role which the inverse plays in physics.

Most of the results given in the paper are known in the literature, though are scattered among various sources. Some results were so far likely not explicitly expressed in the literature, but could be derived by combining different known facts (again distributed among diverse sources). It is believed that the article will prove to provide a compendious, though fully-fledged guidance to the researchers utilizing the least squares methods as well as to those looking for tools rooted in the theory of matrix generalized inverses to cope with the problems they face in their investigations.

Keywords: applications of generalized inverses; Moore–Penrose inverse; least squares method; matrix equations

E-mail address: OBaksalary@gmail.com
ORCID: 0000-0002-5172-2068

[1] *The paper is dedicated to Professor Götz Trenkler on the occasion of his eighties birthday on July 14, 2023. I am sincerely thankful for over twenty years of our friendship and collaboration! Many happy returns, my Best Man!*

Clockwise and Counter-Clockwise Transformations with Applications

Richard William Farebrother

The paper is dedicated to Professor Götz Trenkler on the occasion of his 80th birthday on July 14, 2023. The author thankfully acknowledges inspiring contributions of Professor Trenkler to the theory of magic squares, reflected, inter alia, by the publications:

O.M. Baksalary, D. Trenkler, G. Trenkler, On most perfect magic squares of order four, Linear and Multilinear Algebra 68 (2020) 1411-1423.

J. Groß, S.-O. Troschke, G. Trenkler, On properties of 3×3 semi-magic squares, International Journal of Mathematical Education in Science and Technology 30 (1999) 861-865.

K. Schmidt, G. Trenkler, The Moore–Penrose inverse of a semi-magic square is semi-magic, International Journal of Mathematical Education in Science and Technology 32 (2001) 624-629.

D. Trenkler, G. Trenkler, Magic squares, Melancholy and the Moore–Penrose inverse, IMAGE: The Bulletin of the International Linear Algebra Society 27 (2001) 3-10.

D. Trenkler, G. Trenkler, Most-perfect pandiagonal magic squares and their Moore–Penrose inverse, International Journal of Mathematical Education in Science and Technology 35 (2004) 697-701.

G. Trenkler, Singular magic squares, International Journal of Mathematical Education in Science and Technology 25 (1994) 595-597.

G. Trenkler, Another look at cubes and inverses of magic squares, The Mathematical Gazette 82 (1998) 288-289.

G. Trenkler, On the eigenvalues of 3×3 magic squares, International Journal of Mathematical Education in Science and Technology 30 (1999) 307-308.

Clockwise and Counter-Clockwise Transformations with Applications

Richard William Farebrother

11 Castle Road, Bayston Hill, Shrewsbury SY3 0NF, United Kingdom.
R.W.Farebrother@Hotmail.com

Abstract: In this paper we introduce a class of clockwise and counter-clockwise transformations of square matrices and discuss their application to Graeco-Latin squares, magic squares and the solution of a family of generalised Sudoku problems.

Key words: Graeco-Latin square; Latin square; magic square; permutation matrix; Sudoku problem.

1. Introduction

Given two$n \times n$ matrices with typical elements (f, g) and (i, j) respectively, then we may associate the n^2 cells of the $n \times n$ source matrix with the n^2 cells of the $n \times n$ target matrix to obtain a one-to-one transformation which carries the (f, g)th element of the source matrix into the (i, j)th element of the target matrix. We represent the effect of this transformation by inserting the symbol (i, j) in the (f, g)th cell of a third $n \times n$ matrix which serves to represent the transformation. In other words, for $i, j = 1, 2, ..., n$, we have to allocate the n^2 symbols (i, j) to the n^2 elements of a third $n \times n$ matrix

Let $\mathbf{A}$ be an $n \times n$ matrix and let h be an integer in the range $1 \leq h \leq n$, then we may define the hth row of $\mathbf{A}$ by $\{a_{hj} \quad j = 1, 2, ..., n\}$, the hth column of $\mathbf{A}$ by $\{a_{ih} \quad i = 1, 2, ..., n\}$, the hth primary cyclic-diagonal of $\mathbf{A}$ by $\{a_{ij} \ : \ i = h + j \pmod{n},\ j = 1, 2, ..., n\}$, and the hth secondary cyclic-diagonal of $\mathbf{A}$ by $\{a_{ij} \ : \ i = h + 1 - j \pmod{n},\ j = 1, 2, ..., n\}$. In each case, the nth primary or secondary cyclic-diagonal of $\mathbf{A}$ may be named the principal primary or secondary diagonal as they are also the familiar non-cyclical diagonals of the matrix. [Henceforth we shall suppress the 'cyclic' prefix.]

For example, omitting the matrix identifier a in the 3×3 case, we find that the element-indicators in the first, second and third rows are $\{(1, 1), (1, 2), (1, 3)\}$, $\{(2, 1), (2, 2), (2, 3)\}$, $\{(3, 1), (3, 2), (3, 3)\}$, the indicators in The first second and third columns are $\{(1, 1), (2, 1), (3, 1)\}$, $\{(1, 2), (2, 2), (3, 2)\}$, $\{(1, 3), (2, 3), (3, 3)\}$, the indicators in The first, second and third primary diagonals are $\{(1, 2), (2, 3), (3, 1)\}$, $\{(1, 3), (2, 1), (3, 2)\}$, $\{(1, 1), (2, 2), (3, 3)\}$, and the indicators in the first second and third secondary diagonals are $\{(1, 1), (2, 3), (3, 2)\}$, $\{(1, 2), (2, 1), (3, 3)\}$, $\{(1, 3), (2, 2), (3, 1)\}$.

The typical square matrix have six classes of transformation of particular interest to us in this paper: a: rotating the given matrix about its middlemost row or column carries rows into rows and columns into columns but primary diagonals into secondary diagonals and *vice versa*. b: rotating the given matrix about its principal primary or secondary diagonal carries primary diagonals into primary diagonals and secondary diagonals into secondary diagonals but rows into columns and *vice versa*. c: rotating the given matrix about its middlemost row or column and about its principal primary or secondary diagonal carries rows into columns and *vice versa* and primary diagonals into secondary diagonals and *vice versa*. d: What we shall call a clockwise transformation carries rows into primary diagonals, primary diagonals into columns, columns into secondary diagonals, and secondary diagonals into rows. e: What we shall call a counter-clockwise transformation carries rows into secondary diagonals, secondary diagonals into columns, columns into primary diagonals, and primary diagonals into rows. f: Two applications of a clockwise or a counter-clockwise transformation will carry rows into columns and *vice versa* and primary diagonals into secondary diagonals and *vice versa*. These results are summarised in the following table where R, Q, P, and S signify rows, columns, primary diagonals and secondary diagonals respectively:

Table 1

Transformation	R	Q	P	S
mid row/column	R	Q	S	P
Principal P/S Diagonal	Q	R	P	S
R/Q + P/S Combination	Q	R	S	P
Clockwise	P	S	Q	R
Counter Clockwise	S	P	R	Q
Double Clockwise	Q	R	S	P

Note that an application of the double transformations (c) or (f) with the single transformations (d) or (e) will convert a clockwise transformation (d) into a counter-clockwise transformation (e) and *vice versa*, see Section 5 below.

If $n = 2m + 1$ is odd then each of the n rows of $\mathbf{A}$ has a single intersection with each of its n columns, each of its n primary diagonals, and each of its n secondary diagonals. Similarly, each of the n primary diagonals of $\mathbf{A}$has a single intersection with each of its n secondary diagonals, each of its n rows and each of its n columns, and so on. [In this context, the principal diagonals may be associated with the middlemost value $h = m + 1$ of h by replacing $i = h + j$ by $i = m + h + j$ and $i = h + 1 - j$ by $i = h - m - j$ in the above definitions of the hth primary and secondary cyclic-diagonals where $m = (n - 1)/2$.]

In this paper, we shall restrict our attention to odd values of n, and shall seek to identify a family of "counter-clockwise" transformations which carries each column into a primary diagonal, each primary diagonal into a row, each row into a secondary diagonal, and each secondary diagonal into a column. [The corresponding reverse "clockwise" transformation will be discussed in Section 5 below.

For this purpose,we seek a rule for determining an $n \times n$ matrix of quadruplets in which four distinct sequences of values $1, 2, ..., n$ are assigned to the n rows, to the n columns, to the n primary diagonals, and to the n secondary diagonals of the target matrix. In this context, a typical quadruplet takes the form $(ijkl)$ where the First element i is constant on each primary diagonal, the Second element j is constant on each secondary Diagonal, the Third Element k is constant on each column, and the Fourth Element l is constant on each row of the target matrix, where, for consistency, the parameters i, j, k and l must satisfy the conditions $k = j - i \pmod n$ and $l = i + j \pmod n$, whence $2i = l - k \pmod n$ and $2j = k + l \pmod n$. Given any three entities from the n rows, the n columns, the n primary diagonals, and the n secondary diagonals, not all from the same class (so that there is at least one intersection) and not all meeting in a single point (so that there are at least two intersections), and assuming that the first two choices are from distinct classes, we may assign any one of $n$values to the parameter associated with the first choice, any one of n values to the parameter associated with the second choice, and any one of $n - 1$ values to the parameter associated with the third choice. (Note that the value associated with the third choice must be distinct from that associated with the member of the third class implied by the intersection of the first two choices). It is readily apparent from the examples given below that we may obtain a full set of n^2 quadruplets from any such selection. Thus, somewhat surprisingly, we find that we only have three degrees of freedom in this problem whatever the value of n. and hence that we have $n^2(n - 1)$ possible arrangements of the elements in an $n \times n$counter-clockwise transformation matrix.

2. The 3×3 Case

As a specific example of this procedure, in the 3×3 case, we may set the fourth element of all quadruplets in the second row equal to 2 and the third element of all quadruplets in the second column equal to 3. These lines intersect in the middlemost cell and we may identify the corresponding quadruplet as (1132). we may thus associate the value 1 with the principal primary diagonal through this cell and the value 1 with the principal secondary diagonal through the same cell.

Altermatively, we could have set the first element of all quadruplets on the principal primary diagonal equal to 1 and also the second element of all quadruplets on the principal secondary diagonal equal to 1. Again we have to associate the quadruplet (1132) with the middlemost cell, and thus the value 2 with the second row and 3 with the second column.

In either case, we have the partially completed matrix:

$$\begin{bmatrix} (1***) & (**3*) & (*1**) \\ (***2) & (1132) & (***2) \\ (*1**) & (**3*) & (1***) \end{bmatrix}$$

We now have to associate any unused value with any row, column, primary diagonal or secondary diagonal not already in use. for example, we may associate the value 3 with the fourth element of all quadruplets in the first row (and the unused value 1 with the fourth element of all quadruplets in the third row). We then have:

$$\begin{bmatrix} (1**3) & (**33) & (*1*3) \\ (***2) & (1132) & (***2) \\ (*1*1) & (**31) & (1**1) \end{bmatrix}$$

whence

$$\begin{bmatrix} (1213) & (3333) & (2123) \\ (2312) & (1132) & (3222) \\ (3111) & (2231) & (1321) \end{bmatrix}$$

In fact, the first two parameters suffice to identify any quadruplet and we delete the third and fourth element of each quadruplet to obtain the required transformation:

$$\begin{bmatrix} (1,2) & (3,3) & (2,1) \\ (2,3) & (1,1) & (3,2) \\ (3,1) & (2,2) & (1,3) \end{bmatrix}$$

Examining in turn the columns, the primary diagonals, the rows, and the secondary diagonals of this matrix, we find that the transformation represented by this matrix carries The columns $\{(1,1),(2,1),(3,1)\}$, $\{(1,2),(2,2),(3,2)\}$, $\{(1,3),(2,3),(3,3)\}$ into the primary diagonals $\{(1,2),(2,3),(3,1)\}$, $\{(3,3),(1,1),(2,2)\}$, $\{(2,1),(3,2),(1,3)\}$.

It carries the primary diagonals $\{(1,1),(2,2),(3,3)\}$, $\{(1,2),(2,3),(3,1)\}$, $\{(1,3),(2,1),(3,2)\}$ into the rows $\{(1,2),(1,1),(1,3)\}$, $\{(3,3),(3,2),(3,1)\}$, $\{(2,1),(2,3),(2,2)\}$. it carries the rows {(1, 1), (1, 2), (1, 3)}, {(2, 1), (2, 2), (2, 3)}, {(3, 1), (3, 2), (3, 3)} into the secondary diagonals {(1, 2), (3, 3), (2, 1)}, {(2, 3), (1, 1), (3, 2)} {(3, 1), (2, 2), (1, 3)}. And it carries the secondary diagonals {(1, 1), (2, 3), (3, 2)}, {(1, 2) , (2, 1) , (3, 3)}, {(1, 3), (2, 2), (3, 1)} into the columns {(1, 2), (3, 2), (2, 2)}, {(3, 3), (2, 3), (1, 3)}, {(2, 1), (1, 1),(3, 1)}.

This transformation clearly exhibits a fixed point at $(3,1)$ combined with an eight period cycle that carries $(1,1)$ into $(1,2)$, $(1,2)$ into $(3,3)$, $(3,3)$ into $(1,3)$, $(1,3)$ into $(2,1)$, $(2,1)$ into $(2,3)$, $(2,3)$ into $(3,2)$, $(3,2)$ into $(2,2)$, $(2,2)$ into $(1,1)$, and so on. Thus, whatever the starting value, it is clear that eight applications of this transformation carries every element back to its original position. Moreover, on examining the second and fourth successors in this sequence, we find that two applications of this transformation Carries Rows 1, 2 and 3 into columns 3, 2 and 1, and columns 1, 2 and 3 into rows 3, 1 and 2whilst four applications Carries Rows 1, 2 and 3 into Rows 2, 1 and 3, and columns 1, 2 and 3 into Columns 1, 3 and 2 in a two period block cycle, see Table 2.

Table 2

(1, 1)	(1, 2)	(3, 3)	(1, 3)	(2, 1)
(1, 2)	(3, 3)	(1, 3)	(2, 1)	(2, 3)
(1, 3)	(2, 1)	(2, 3)	(3, 2)	(2, 2)
(2, 1)	(2, 3)	(3, 2)	(2, 2)	(1, 1)
(2, 2)	(1, 1)	(1, 2)	(3, 3)	(1, 3)
(2, 3)	(3, 2)	(2, 2)	(1, 1)	(1, 2)
(3, 1)	(3, 1)	(3, 1)	(3, 1)	(3, 1)
(3, 2)	(2, 2)	(1, 1)	(1, 2)	(3, 3)
(3, 3)	(1, 3)	(2, 1)	(2, 3)	(3, 2)

3. A Simple General Procedure

Rather than applying an *ad hoc* procedure to each new problem as it arises, Farebrother (2007-08) has suggested the following simple procedure for generating a transformation which carries each column into a primary diagonal, each primary diagonal into a row, each row into a secondary diagonal, and each secondary diagonal into a column when n is odd.

Given n odd and an empty $n \times n$ matrix, we suppose that the transformation carries a typical element (associated with the quadruplet $(ijkl)$) into the ith row, the jth column, the kth primary diagonal, and the lth secondary diagonal, where $k = j - i \mod n$ and $l = i + j \mod n$ and hence $2i = l - k \mod n$ and $2j = k + l \mod n$. A specific transformation may now be obtained by attaching consistent values to any three rows, columns, primary diagonals or secondary diagonals that intersect in two or three cells. In particular, we note the following simple, if somewhat inefficient, procedure. Attaching any one of n values to the fourth element of all quadruplets in the first row and any one of n values to the third element of all quadruplets in the first column. Then the intersection of the first row with the first column at $(1, 1)$ identifies the value of the first element of all quadruplets in the principal primary diagonal which passes through the points $(1, 1)$, $(2, 2)$ and (n, n). Attaching any one of $n - 1$ values to the fourth element of all quadruplets in the last row, then the intersection between this row and the principal primary diagonal at (n, n) identifies the value of the third element of all quadruplets in the last column. Now the intersection between the nth row and the first column at $(n, 1)$ identifies the value of the first element of all quadruplets in the primary diagonal that passes through this point and through $(1, 2)$, $(2, 3)$, and $(n - 1, n)$ whilst the intersection between the first row and the nth column at $(1, n - 1)$ identifies the value of all quadruplets in the primary diagonal that passes through this point and through $(2, 1)$, $(3, 2)$, and $(n, n - 1)$. The intersections between these two primary diagonals and the first and last rows and columns at the points $(1, 2)$, $(n - 1, n)$, $(2, 1)$, and $(n, n - 1)$ identifies the values to be associated with the second and $(n - 1)$th rows and columns. Missing the opportunity of defining four more primary diagonals at this stage, we simply note that the intersections between the two cyclic-diagonals immediately above and below the principal primary diagonal and the second and $(n - 1)$th rows and columns define the values to be associated with the third and $(n - 2)$th rows and columns; and so on to completion.

4. The 5×5 case

In order to obtain a sufficiently explicit example of this general procedure, we consider the case of a 5×5 matrix. Setting $n = 5$, and arbitrarily setting the fourth element of all quadruplets in the first row of the empty 5×5 matrix equal to 4, and the third element of all quadruplets in the first column equal to 2, we find that the principal primary diagonal is associated with a value of $1 = (4 - 2)/2$ and set the first element of all quadruplets in this diagonal equal to 1. Again, arbitrarily setting the fourth element of all quadruplets in the fifth row equal to 5, we find that the last column is associated with a value of

$3 = 5 - 2$, and we have to set the third element of all quadruplets in this column equal to 3. In this way we obtain the matrix:

$$\begin{bmatrix} (1*24) & (***4) & (***4) & (***4) & (**34) \\ (**2*) & (1***) & (****) & (****) & (**3*) \\ (**2*) & (****) & (1***) & (****) & (**3*) \\ (**2*) & (****) & (****) & (1***) & (**3*) \\ (**25) & (***5) & (***5) & (***5) & (1*35) \end{bmatrix}$$

From this matrix we may establish the values associated with the first primary diagonals immediately above and below the principal primary diagonal, and hence the values associated with the second and fourth rows and the second and fourth columns, we thus have:

$$\begin{bmatrix} (1*24) & (4*14) & (**44) & (***4) & (3*34) \\ (3*23) & (1*13) & (4**3) & (**43) & (**33) \\ (**2*) & (3*1*) & (1***) & (4*4*) & (**3*) \\ (**21) & (**11) & (3**1) & (1*41) & (4*31) \\ (4*25) & (**15) & (***5) & (3*45) & (1*35) \end{bmatrix}$$

From which it is easy to obtain the full matrix:

$$\begin{bmatrix} (1324) & (4514) & (2254) & (5444) & (3134) \\ (3523) & (1213) & (4453) & (2143) & (5333) \\ (5222) & (3412) & (1152) & (4342) & (2532) \\ (2421) & (5111) & (3351) & (1541) & (4231) \\ (4125) & (2315) & (5555) & (3245) & (1435) \end{bmatrix}$$

Once again, the first two elements suffice to identify the cells of this matrix, and we delete the third and fourth elements to obtain:

$$\begin{bmatrix} (1,3) & (4,5) & (2,2) & (5,4) & (3,1) \\ (3,5) & (1,2) & (4,4) & (2,1) & (5,3) \\ (5,2) & (3,4) & (1,1) & (4,3) & (2,5) \\ (2,4) & (5,1) & (3,3) & (1,5) & (4,2) \\ (4,1) & (2,3) & (5,5) & (3,2) & (1,4) \end{bmatrix}$$

Clearly, the transformation represented by this matrix carries each column into a primary diagonal, each primary diagonal into a row, each row into a secondary diagonal, and each secondary diagonal into a column. This transformation clearly exhibits a twenty-period cycle that carries $(1, 1)$ into $(1, 3)$, $(1, 3)$ into $(2, 2)$, $(2, 2)$ into $(1, 2)$, $(1, 2)$ into $(4, 5)$, $(4, 5)$ into $(4, 2)$, $(4, 2)$ into $(5, 1)$, $(5, 1)$ into $(4, 1)$, $(4, 1)$ into $(2, 4)$, $(2, 4)$ into $(2, 1)$, $(2, 1)$ into $(3, 5)$, $(3, 5)$ into $(2, 5)$, $(2, 5)$ into $(5, 3)$, $(5, 3)$ into $(5, 5)$, $(5, 5)$ into $(1, 4)$, $(1, 4)$ into $(5, 4)$, $(5, 4)$ into $(3, 2)$, $(3, 2)$ into $(3, 4)$, $(3, 4)$ into $(4, 3)$, $(4, 3)$ into $(3, 3)$, $(3, 3)$ into $(1, 1)$. It also exhibits a five period cycle that carries $(4, 4)$ into $(1, 5)$, $(1, 5)$ into $(3, 1)$, $(3, 1)$ into $(5, 2)$, $(5, 2)$ into $(2, 3)$, $(2, 3)$ into $(4, 4)$. However, whatever the starting value, it is clear that twenty applications of this transformation carries every element back to its original position.

Further, on examining fourth successors in this sequence, we find that four applications of this transformation carries rows $1,\ 2,\ 3,\ 4$ and 5 into rows $4,\ 5,\ 1,\ 2$ and 3 and columns $1,\ 2,\ 3,\ 4$ and 5 into columns $5,\ 1,\ 2,\ 3$ and 4 in a five period block cycle.

here we have shown implicitly that there are $n^2(n-1)$ transformations of the required form. There are also $n^2(n-1)$ reverse transformations that carry columns into secondary diagonals, secondary diagonals into rows, rows into primary diagonals, and primary diagonals into columns.

5. Rotational solutions

Having obtained a counter-clockwise transformation which carries columns into primary diagonals, primary diagonals into rows, rows into secondary diagonals, and secondary diagonals into columns, we may readily obtain the reverse of this transformation by reversing each of the individual cell-to-cell operations before arranging the results in an $n \times n$ matrix. Applying this idea to the transformation in Section 2, we find that the corresponding reverse transformation carries $(1, 1)$ into $(2, 2)$, $(1, 2)$ into $(1, 1)$, $(1, 3)$ into $(3, 3)$, $(2, 1)$ into $(1, 3)$, $(2, 2)$ into $(3, 2)$, $(2, 3)$ into $(2, 1)$, $(3, 1)$ into $(3, 1)$, $(3, 2)$ into $(2, 3)$, and $(3, 3)$ into $(1, 2)$. Thus, the reverse of the transformation described in section 2 may be represented by the 3×3 matrix:

$$\begin{bmatrix} (2,2) & (1,1) & (3,3) \\ (1,3) & (3,2) & (2,1) \\ (3,1) & (2,3) & (1,2) \end{bmatrix}$$

[Note that I have use the term ‘reverse’ in this context as the more usual ‘inverse’ might suggest that a matrix inverse is intended.]

On the other hand, if we are not concerned with obtaining the reverse of a specific counter-clockwise transformation but only a clockwise transformation that carries rows into primary diagonals, primary diagonals into columns, columns into secondary diagonals, and secondary diagonals into rows, then we may do so by reversing the order of the rows or columns, and transposing the resulting matrix about its principal primary or secondary diagonal or *vice versa*. As indicated in Table 1, Reversing the order of the rows or columns automatically interchanges the roles of the first and second indicators (relating to the primary and secondary diagonals of the target matrix) and the transposition of the matrix about either principal diagonal automatically interchanges the roles of the third and fourth indicators (relating to the rows and columns of the target matrix). [Alternatively, two applications of a clockwise or a counter-clockwise transformation interchanges the roles of both pairs of operators.]

To illustrate this point, we again consider the 3×3 transformation matrix identified in Section 2:

$$\begin{bmatrix} (1,2) & (3,3) & (2,1) \\ (2,3) & (1,1) & (3,2) \\ (3,1) & (2,2) & (1,3) \end{bmatrix}$$

Reversing the order of the rows, we have:

$$\begin{bmatrix} (3,1) & (2,2) & (1,3) \\ (2,3) & (1,1) & (3,2) \\ (1,2) & (3,3) & (2,1) \end{bmatrix}$$

Then transposing this matrix about its principal primary diagonal, we have:

$$\begin{bmatrix} (3,1) & (2,3) & (1,2) \\ (2,2) & (1,1) & (3,3) \\ (1,3) & (3,2) & (2,1) \end{bmatrix}$$

As required, this double operation produces a clockwise transformation that carries rows into primary diagonals, primary diagonals into columns, columns into secondary diagonals, and secondary diagonals into rows. Note that the fixed point of this transformation has moved to the $(3,2)$ position.

Alternatively, we can generate an arbitrary clockwise transformation by means of three applications of the counter-clockwise transformation of Section 2. As indicated in Table 2, this procedure yields a clockwise transformation which carries $(1,1)$ into $(1,3)$, $(1,2)$ into $(2,1)$, $(1,3)$ into $(3,2)$, $(2,1)$ into $(2,2)$, $(2,2)$ into $(3,3)$, $(2,3)$ into $(1,1)$, $(3,1)$ into $(3,1)$, $(3,2)$ into $(1,2)$, and $(3,3)$ into $(2,3)$, and hence the transformation matrix:

$$\begin{bmatrix} (1,3) & (2,1) & (3,2) \\ (2,2) & (3,3) & (1,1) \\ (3,1) & (1,2) & (2,3) \end{bmatrix}$$

6. Application to Graeco-Latin Squares

To identify a Graeco-Latin square generated by the 5×5 Transformation matrix of Section 4, we replace the numerical row indicators $1, 2, 3, 4, 5$ by their alphabetical equivalents A, B, C, D, E (in some order), and obtain:

$$\begin{bmatrix} A3 & D5 & B2 & E4 & C1 \\ C5 & A2 & D4 & B1 & E3 \\ E2 & C4 & A1 & D3 & B5 \\ B4 & E1 & C3 & A5 & D2 \\ D1 & B3 & E5 & C2 & A4 \end{bmatrix}$$

which contains all twenty-five combinations of the five letters and the five numerals. In the present context, this matrix would probably be called an alphanumeric square. But, for historical reasons, it is more usually known as a Graeco-Latin square.

Associated with each Graeco-Latin square we have two Latin squares with the same dimensions, the first is obtained by deleting all 25 numerals from the above Graeco-Latin square and the second by deleting all 25 letters. In this context, the given Graeco-Latin square may be regarded as being formed by superimposing one of these Latin squares on the other in a symbolic Hadamard product. Now, each row and each column of the alphabetical Latin square contains one of each letter; each row and each column of the numerical Latin square contains one of each numeral; and hence each row and each column of the Graeco-Latin square formed by superimposing one of these Latin squares on the other contains one of each letter and one of each numeral.

Moreover, we are not restricted to the Graeco-Latin square that arises naturally from the given $n \times n$ transformation as we may readily obtain $2(n!)^2$ variants of our original Graeco-Latin square by permuting the order of its rows, permuting the order of its columns, and transposing the result about its principal primary or secondary diagonal.

Thus, if it is thought inappropriate that the common letters in the above Graeco-Latin square should lie on its five primary diagonals and that the common numerals should lie on its five secondary diagonals then we can easily replace the 'bishop's move' structure of this Graeco-Latin square by a 'knights's move' structure by interleaving the last two rows of the matrix of Section 4 into its first three rows in the order 1, 4, 2, 5, 3 to obtain the 'knight's move' transformation matrix:

$$\begin{bmatrix} (1,3) & (4,5) & (2,2) & (5,4) & (3,1) \\ (2,4) & (5,1) & (3,3) & (1,5) & (4,2) \\ (3,5) & (1,2) & (4,4) & (2,1) & (5,3) \\ (4,1) & (2,3) & (5,5) & (3,2) & (1,4) \\ (5,2) & (3,4) & (1,1) & (4,3) & (2,5) \end{bmatrix}$$

before replacing the numerical row-indicators by letters to obtain the corresponding Graeco-Latin square:

$$\begin{bmatrix} A3 & D5 & B2 & E4 & C1 \\ B4 & E1 & C3 & A5 & D2 \\ C5 & A2 & D4 & B1 & E3 \\ D1 & B3 & E5 & C2 & A4 \\ E2 & C4 & A1 & D3 & B5 \end{bmatrix}$$

Note that each row, each column, each primary diagonal and each secondary diagonal of this (Knut Vik) variant of our basic Graeco-Latin square still contains one of each letter and one of each numeral.

$n \times n$ Latin and Graeco-Latin squares are primarily employed in the design of statistical experiments. but they also serve as the source of a class of $n \times n$ magic square defined by assigning the additive values $0,\ n,\ ...,\ n(n-1)$ (in some order) to one set of symbols and the values $1,\ 2,\ ...,\ n$ to the other. In particular, on setting $A = 0,\ B = 5,\ C = 10,\ D = 15$, and $E = 20$ in our 5×5 'knight's move' Graeco-Latin square, we have the magic square:

$$\begin{bmatrix} 3 & 20 & 7 & 24 & 11 \\ 9 & 21 & 13 & 5 & 17 \\ 15 & 2 & 19 & 6 & 23 \\ 16 & 8 & 25 & 12 & 4 \\ 22 & 14 & 1 & 18 & 10 \end{bmatrix}$$

which shows that the numbers $1,\ 2,\ ...,\ 25$ can be arranged in a 5×5 array in such a way that the elements in each of five rows, in each of five columns, in each of five primary diagonals, and in each of five secondary diagonals sum to 65, see Farebrother (2009).

7. A Simple Algebraic Procedure for Knut Vik Graeco-Latin Squares

Readers will be interested to learn that the above 5×5 knight's move matrix may be generalised to yield a family of $n \times n$ Graeco-Latin squares with special properties (Knut Vik designs) provided only that $n \geq 5$ is a positive integer that is not divisible by 2 or by 3. To see this, let h and k be arbitrary integers in the range $1 \leq h, k \leq n$ and let m be an integer in the range $2 \leq m \leq n-2$ such that the three functions $x_t = t(m-1)\ (modn),\ y_t = tm\ (modn)$ and $z_t = t(m+1)\ (modn)$ each take the values $0,\ 1,\ ...,\ n-1$ in some order as t takes the values $0,\ 1,\ ...,\ n-1$. then, our general algebraic Knut Vik square is defined by supposing that the (i,j)th cell of the array contains an ordered pair with first element $i + m(h-j)\ (mod\ n)$ and second element $i + m(j-k)\ (mod\ n)$.

Clearly, the parameters h and k serve to identify the columns $j = h$ and $j = k$ in which the first and second elements in this (i,j)th cell take the value $i\ (mod\ n)$. Moreover, for all values of i, j, the value of the first (second) element remains the same if i is increased by $f = gm\ (mod\ n)$ at the same time as j is increased (decreased) by $g = fp\ (mod\ n)$ where f and g are integers in the range $1 \leq f, g \leq n-1$ and p is an integer in the same range such that $mp = 1\ (mod\ n)$. In this context, it is clear that the 5×5 knight's move matrix given above corresponds to the values $n = 5,\ m = 2,\ p = 3,\ h = 1$ and $k = 5$.

Further, it may readily be established that the $n \times n$ arrays determined by this algebraic procedure have the property that the set of pairs in each row, each column, each primary cyclic-diagonal and each secondary cyclic-diagonal contain one element of each value in the first position and one element of each value in the second position. In addition, no ordered pair of elements is repeated in any two cells of the array.

8. Application to Generalised Sudoku problems

Let $e_1,\ e_2,\ ...,\ e_9$ be a set of nine distinct symbols, then the standard 9×9 Sudoku problem concerns the possibility of inserting nine copies of each of these nine symbols into a 9×9 array in such a way that the nine symbols in each of nine rows are distinct; the nine symbols in each of nine columns are distinct; and the nine symbols in each of nine contiguous 3×3 submatrices are distinct.

In the present section, we are concerned with a generalisation of this standard problem that requires that a further three (or seven) sets of conditions be satisfied. Namely, the nine symbols in each of nine corresponding locations in the nine contiguous 3×3 matrices are distinct; the nine symbols in the first, fourth and seventh 3×1 minicolumns of the first, second and third 3×9 row-blocks are distinct; as are the nine symbols in the second, fifth and eighth minicolumns and the nine symbols in the third, sixth and ninth minicolumns of these row-blocks. Similarly, the nine symbols in the first, fourth and seventh 1×3 minirows of the first, second and third 9×3 column-blockss are distinct; as are the nine symbols in the second, fifth and eighth minirows and the nine symbols in the third, sixth and ninth minirows of these column-blocks. Indeed, we shall have cause to mention a further generalisation in which four more sets of conditions based on combinations of primary and secondary minidiagonals are satisfied. These new sets of restrictions are more conveniently expressed in the form of Hadamard products, as we shall now show:

Let $\mathbf{E}_{ij}$ be a 3×3 matrix with unity in its ijth position and zeros elsewhere and let $\mathbf{F}$ be a 3×3 matrix of ones. Then we may define four sets of three 3×3 matrices:

The set of 3×3 minirow matrices $\mathbf{R_1}=\mathbf{E_{11}}+\mathbf{E_{12}}+\mathbf{E_{13}}$, $\mathbf{R_2}=\mathbf{E_{21}}+\mathbf{E_{22}}+\mathbf{E_{23}}$, $\mathbf{R_3}=\mathbf{E_{31}}+\mathbf{E_{32}}+\mathbf{E_{33}}$:

$$\mathbf{R}_1=\begin{bmatrix}1&1&1\\0&0&0\\0&0&0\end{bmatrix}\qquad \mathbf{R}_2=\begin{bmatrix}0&0&0\\1&1&1\\0&0&0\end{bmatrix}\qquad \mathbf{R}_3=\begin{bmatrix}0&0&0\\0&0&0\\1&1&1\end{bmatrix}$$

The set of 3×3 minicolumn matrices $\mathbf{Q_1}=\mathbf{E_{11}}+\mathbf{E_{21}}+\mathbf{E_{31}}$, $\mathbf{Q_2}=\mathbf{E_{12}}+\mathbf{E_{22}}+\mathbf{E_{32}}$, $\mathbf{Q_3}=\mathbf{E_{13}}+\mathbf{E_{23}}+\mathbf{E_{33}}$:

$$\mathbf{Q}_1=\begin{bmatrix}1&0&0\\1&0&0\\1&0&0\end{bmatrix}\qquad \mathbf{Q}_2=\begin{bmatrix}0&1&0\\0&1&0\\0&1&0\end{bmatrix}\qquad \mathbf{Q}_3=\begin{bmatrix}0&0&1\\0&0&1\\0&0&1\end{bmatrix}$$

The set of 3×3 primary mini-diagonal matrices $\mathbf{P_1}=\mathbf{E_{11}}+\mathbf{E_{22}}+\mathbf{E_{33}}$, $\mathbf{P_2}=\mathbf{E_{12}}+\mathbf{E_{23}}+\mathbf{E_{31}}$, $\mathbf{P_3}=\mathbf{E_{13}}+\mathbf{E_{21}}+\mathbf{E_{32}}$:

$$\mathbf{P}_1=\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}\qquad \mathbf{P}_2=\begin{bmatrix}0&1&0\\0&0&1\\1&0&0\end{bmatrix}\qquad \mathbf{P}_3=\begin{bmatrix}0&0&1\\1&0&0\\0&1&0\end{bmatrix}$$

And the set of 3×3 secondary mini-diagonal matrices $\mathbf{S_1}=\mathbf{E_{11}}+\mathbf{E_{23}}+\mathbf{E_{32}}$, $\mathbf{S_2}=\mathbf{E_{12}}+\mathbf{E_{21}}+\mathbf{E_{33}}$, $\mathbf{S_3}=\mathbf{E_{13}}+\mathbf{E_{22}}+\mathbf{E_{31}}$.

$$\mathbf{S}_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \qquad \mathbf{S}_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \mathbf{S}_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

Taking Kronecker products within the sets $\{\mathbf{E}_{11}, \mathbf{F}\}$, $\{\mathbf{E}_{12}, \mathbf{F}\}$, ..., $\{\mathbf{E}_{33}, \mathbf{F}\}$, and within the sets $\{\mathbf{R_1}, \mathbf{R_2}, \mathbf{R_3}, \mathbf{Q_1}, \mathbf{Q_2}, \mathbf{Q_3}\}$ and $\{\mathbf{P_1}, \mathbf{p_2}, \mathbf{P_3}, \mathbf{S_1}, \mathbf{S_2}, \mathbf{S_3}\}$, we have (Hadamard) masks for the nine 3×3 contiguous submatrices defined by $\mathbf{E}_{ij} \otimes \mathbf{F}$; masks for the nine location matrices defined by $\mathbf{F} \otimes \mathbf{E}_{ij}$; masks for the nine long row matrices defined by $\mathbf{R}_i \otimes \mathbf{R}_j$; masks for the nine long column matrices defined by $\mathbf{Q}_i \otimes \mathbf{Q}_j$; masks for the nine minirow $\otimes$ minicolumn (or broken column) matrices defined by $\mathbf{R}_i \otimes \mathbf{Q}_j$; masks for the nine minicolumn $\otimes$ minirow (or broken row) matrices defined by $\mathbf{Q}_i \otimes \mathbf{R}_j$; masksfor the nine primary $\otimes$ primary mini-diagonal matrices defined by $\mathbf{P}_i \otimes \mathbf{P}_j$; masks for the nine secondary $\otimes$ secondary mini-diagonal matrices defined by $\mathbf{S}_i \otimes \mathbf{S}_j$; masks for the nine primary $\otimes$ secondary mini-diagonal matrices defined by $\mathbf{P}_i \otimes \mathbf{S}_j$; and masks for the nine secondary $\otimes$ primary mini-diagonal matrices defined by $\mathbf{S}_i \otimes \mathbf{P}_j$.

The first six of these ten sets of nine masks were implicitly employed by Bailey, Cameron and Connelly (2008, p.390) in their definition of a 9×9 'symmetric' Sudoku matrix which requires that the nine nonzero symbols of the Hadamard product of each of these 54 matrices with the 9×9 candidate matrix should comprise a full set of the symbols e_1, e_2, ..., e_9.

Indeed, it may readily be established that this typical matrix also satisfies the conditions generated by the remaining four sets of nine Hadamard masks. Whence we may deduce that all 9×9 Sudoku matrices which can be obtained from this matrix by renumbering the symbols, by transposition or by pre-multiplication by a Kronecker product of two matrices from the set of 3×3 permutation matrices $\{\mathbf{P_1},\ \mathbf{P_2},\ \mathbf{P_3},\ \mathbf{S_1},\ \mathbf{S_2},\ \mathbf{S_3}\}$ and post-multiplication by a second Kronecker product of the same type will yield another Sudoku matrix with the same properties. Moreover,it is immediately apparent that these two sets of conditions are distinct as each of the first 54 masks is of rank one whilst each of the last 36 masks is nonsingular.

9. Solution of Generalised Sudoku Problems

A family of general solutions to the problem of Section 8 is easily found: Given any arrangement of the nine symbolss e_1, e_2, ..., e_9, we define the 3×3 seed matrix:

$$\mathbf{B} = \begin{bmatrix} e_1 & e_2 & e_3 \\ e_4 & e_5 & e_6 \\ e_7 & e_8 & e_9 \end{bmatrix}$$

In this context, the values of i and j in a typical element (i, j) from our chosen 3×3 clockwise or counter-clockwise transformation matrix serve to define the row and column permutations required to bring the $h = 3(i-1) + j$th element of $\mathbf{B}$ into the upper left corner of the transformed matrix $\mathbf{M}_h = \mathbf{P}_1^{i-1}\mathbf{B}\mathbf{P}_2^{j-1}$. Further, on substituting the 3×3 matrix $\mathbf{M}_h$ for the symbol (i, j) in our selective 3×3 transformation matrix from Section 2, we have the 9×9 matrix:

$$\begin{bmatrix} \mathbf{M}_2 & \mathbf{M}_9 & \mathbf{M}_4 \\ \mathbf{M}_6 & \mathbf{M}_1 & \mathbf{M}_8 \\ \mathbf{M}_7 & \mathbf{M}_5 & \mathbf{M}_3 \end{bmatrix}$$

or

e_2	e_3	e_1	e_9	e_7	e_8	e_4	e_5	e_6
e_5	e_6	e_4	e_3	e_1	e_2	e_7	e_8	e_9
e_8	e_9	e_7	e_6	e_4	e_5	e_1	e_2	e_3
e_6	e_4	e_5	e_1	e_2	e_3	e_8	e_9	e_7
e_9	e_7	e_8	e_4	e_5	e_6	e_2	e_3	e_1
e_3	e_1	e_2	e_7	e_8	e_9	e_5	e_6	e_4
e_7	e_8	e_9	e_5	e_6	e_4	e_3	e_1	e_2
e_1	e_2	e_3	e_8	e_9	e_7	e_6	e_4	e_5
e_4	e_5	e_6	e_2	e_3	e_1	e_9	e_7	e_8

It may readily be established that this 9×9 matrix, due to Farebrother (2006-07, p.30), satisfies all ten sets of conditions defining the generalised Sudoku problem of the previous Section.

Additional solutions to this problem may readily be obtained by making an alternative choice of a 3×3 transformation matrix $\mathbf{A}$ and/or by allocating the nine given symbols to the seed matrix $\mathbf{B}$ in a different order. For example, we may obtain a variant of the above general solution by replacing $\mathbf{B}$ by its transpose and the counter-clockwise transformation from Section 2 by the first arbitrary clockwise transformation from Section 5. Then, on setting $e_1 = 1,\ e_2 = 5,\ e_3 = 9,\ e_4 = 6,\ e_5 = 7,\ e_6 = 2,\ e_7 = 8,\ e_8 = 3,\ e_9 = 4$, we have the 9×9 typical Sudoku Latin square given by Bailey, Cameron and Connelly (2008, p.391):

8	1	6	2	4	9	5	7	3
3	5	7	6	8	1	9	2	4
4	9	2	7	3	5	1	6	8
7	3	5	1	6	8	4	9	2
2	4	9	5	7	3	8	1	6
6	8	1	9	2	4	3	5	7
9	2	4	3	5	7	6	8	1
1	6	8	4	9	2	7	3	5
5	7	3	8	1	6	2	4	9

Bailey, Cameron and Connelly (2008, p.390) have shown that there are just two equivalence classes of 9×9 Sudoku matrices, where

> "... two Sudoku solutions are equivalent if one can be obtained from the other by a combination of row and column permutations (and possibly transposition) which preserve all the relevant partitions, and re-numbering of the symbols."

For instance, Bailey, Cameron and Connelly's (2008, p.391) typical matrix can be obtained from a particular case of Farebrother's general solution by interchanging its first three rows with its last three rows and transposing the result about its principal primary diagonal

Besides the ten sets of nine conditions outlined in Section 7, these two general matrices have the property that they are 'orthogonal' to their transposes in the sense that the corresponding Graeco-Latin squares comprising eighty-one distinct combinations of nine letters with nine numerals may readily be obtained by superimposing the above matrices in numerical form on their transposes in alphabetical form, see Farebrother and Styan (2008) for details.

References

Rosemary A. Bailey, Peter J. Cameron, and Robert Connelly (2008). Sudoku , gerechte designs, revolutions, affine space, spreads, reguli, and Hamming codes. *The American Mathematical Monthly,* 115: 383–404.

Richard William Farebrother (2006-07). A general matrix transformation: problem and solution. *ILAS Image* No.37, Fall 2006, p.39 and No.39, Fall 2007, 28–30; online open access.

Richard William Farebrother (2007). Simple sudoku, *ILAS Image* No.38, Spring 2007, p.14; online open access.

Richard William Farebrother (2007-08). A simple procedure for a general matrix transformation: problem and solution. *ILAS Image* No.39, Fall 2007, p.32 and No.41, Fall 2008, 32–34; online open access.

Richard William Farebrother (2009). Symmetric Sudoku matrices and magic squares, *ILAS Image* No.43, Fall 2009, 23–24; online open access.

Richard William Farebrother and George P. H. Styan (2008). Double Sudoku Graeco-Latin squares. *ILAS Image* No.41, Fall 2008, 22–24; online open access.

List of Participants

Invited Speakers

1. **Natlia Bebiano**
   University of Coimbra, Portugal
   bebiano@mat.uc.pt

2. **Göran Bergqvist**
   Linköping University, Sweden
   goran.bergqvist@liu.se

3. **Taras Bodnar**
   Stockholm University, Sweden
   taras.bodnar@math.su.se

4. **Malgorzata Bogdan**
   Lund University, Sweden
   malgorzata.bogdan@stat.lu.se

5. **Krzysztof Podgorski**
   Lund University, Sweden
   krzysztof.podgorski@stat.lu.se

Contributed Speakers

6. **Oskar Maria Baksalary**
   Adam Mickiewicz University, Poland
   baxx@amu.edu.pl

7. **Deepjyoti Deka**
   Los Alamos National Laboratory, USA
   deepjyoti@lanl.gov

8. **Francisco Diaz**
   University of Kansas, USA
   fdiaz@kumc.edu

9. **Elias Erdtman**
   Linköping University, Sweden
   elias.erdtman@liu.se

10. **Katarzyna Filipiak**
    Poznan University of Technology, Poland
    katarzyna.filipiak@put.poznan.pl

11. **Cinzia Franceschini**
    University of Gastromic Sciences of Pollenzo, Italy
    c.franceschini@unisg.it

12. **Martina Hancova**
Pavol Jozef Šafárik University, Slovakia
martina.hancova@upjs.sk

13. **Stephen Haslett**
Massey University, New Zealand and
Australian National University and
University of Wollongong, Australia
s.j.haslett@massey.ac.nz

14. **Jan Hauke**
Adam Mickiewicz University, Poland
jhauke@amu.edu.pl

15. **Johannes Heiny**
Stockholm University, Sweden
johannes.heiny@math.su.se

16. **Jarkko Isotalo**
Tampere University, Finland
jarkko.isotalo@tuni.fi

17. **Raymond Kan**
University of Toronto, Canada
Raymond.Kan@rotman.utoronto.ca

18. **Hao Chi Kiang**
Linköping University, Sweden
hao.chi.kiang@liu.se

19. **Nicola Loperfido**
Università degli Studi di Urbino "Carlo Bo", Italy
nicola.loperfido@uniurb.it

20. **Stepan Mazur**
Örebro University and Linnaeus University, Sweden
stepan.mazur@oru.se

21. **Monika Mokrzycka**
Polish Academy of Sciences, Poland
mmok@igr.poznan.pl

22. **Maryna Prus**
University of Hohenheim, Germany
maryna.prus@uni-hohenheim.de

23. **Simo Puntanen**
Tampere University, Finland
simo.puntanen@tuni.fi

24. **Burkhard Schaffrin**
Ohio State University, USA
aschaffrin@earthlink.net

25. **Shriram Srinivasan**
Los Alamos National Laboratory, USA
shrirams@lanl.gov

26. **George Styan**
McGill University, Canada
geostyan@gmail.com

27. **Michel van de Velden**
    Erasmus University Rotterdam, The Netherlands
    vandevelden@ese.eur.nl

28. **Dietrich von Rosen**
    Swedish University of Agricultural Sciences and Linköping University, Sweden
    Dietrich.von.Rosen@slu.se

29. **Xiangfeng Yang**
    Linköping University, Sweden
    xiangfeng.yang@liu.se

Other participants

30. **Zahra Abbasalinejadkolaei**
    Linköping University, Sweden
    zahra.abbasalinejadkolaei@liu.se

31. **Cigdem Cengiz**
    Linköping University, Sweden
    cigdem.cengiz@liu.se

32. **Carl Ekblad**
    Linköping University, Sweden
    carek682@student.liu.se

33. **Jozef Hanc**
    Pavol Jozef Šafárik University, Slovakia
    jozef.hanc@upjs.sk

34. **Jolanta Pielaszkiewicz**
    Linköping University, Sweden
    jolanta.pielaszkiewicz@liu.se

35. **Berkant Savas**
    Linköping University, Sweden
    berkant.savas@liu.se

36. **Martin Singull**
    Linköping University, Sweden
    martin.singull@liu.se

37. **Götz Trenkler**
    Dortmund University of Technology, Germany
    trenkler@statistik.tu-dortmund.de

Index