FIFTH INTERNATIONAL WORKSHOP
ON
MATRIX METHODS FOR STATISTICS

THE GATEWAY CENTRE FOR
EDUCATION AND THE ARTS
SHREWSBURY
18–19 JULY 1996

PROGRAMME
LIST OF PARTICIPANTS
SOCIAL PROGRAMME
ABSTRACTS

Organising Committee:

R. W. Farebrother, Chairman, (University of Manchester, U.K.)

S. J. Puntanen (University of Tampere, Finland)

G. P. H. Styan (McGill University, Canada)

H. J. Werner (University of Bonn, Germany)

Administrative Assistant:

S. Farebrother (Royal Shrewsbury Hospital, U.K.)

Editorial Assistant:

H. C. Jessen (University of Manchester, U.K.)

1. The Kantorovich Inequality and Five Related Inequalities: 1914–1959

We consider the well-known Kantorovich Inequality (1948):

$$\frac{u'Au \cdot u'A^{-1}u}{(u'u)^2} \leq \frac{(\lambda_1 + \lambda_n)^2}{4\lambda_1\lambda_n}$$

where u is a real $n \times 1$ vector and A is a real $n \times n$ symmetric positive definite matrix, with λ_1 and λ_n, respectively, its (fixed) largest and smallest, necessarily positive, eigenvalues. We also consider five related inequalities due respectively to Schweitzer (1914), Pólya-Szegö (1925), Krasnosel'skiĭ-Kreĭn (1952), Cassels (1955), and Greub-Rheinboldt (1959), and show that these six inequalities are equivalent.

Gülhan ALPARGU and George P. H. STYAN, McGill University, Montréal

References

[1] G. Alpargu. The Kantorovich inequality, with some extensions and with some statistical applications. MSc thesis, Dept. of Mathematics and Statistics, McGill University, Montréal, in progress (1996).

[2] J. W. S. Cassels. Appendix: A proof of the inequality (3.2). In [10], pp. 340–341 (1955).

[3] W. Greub and W. Rheinboldt. On a generalization of an inequality of L. V. Kantorovich. *Proceedings of the American Mathematical Society* **10**, 407–415 (1959).

[4] L. V. Kantorovich. Functional analysis and applied mathematics [in Russian]. *Uspekhi Matematičeskiĭ Nauk, Novaya Seriya* **3**, no. 6 (28), 89–185 (1948). [Cf. pp. 142–144. Translated into English as [5].]

[5] L. V. Kantorovich. Functional Analysis and Applied Mathematics [in English]. Translated from the Russian [4] by Curtis D. Benster and edited by George E. Forsythe. Report no. 1509 (1101-10-5100), National Bureau of Standards, Los Angeles (1952). [Cf. pp. 106–109.]

[6] M. A. Krasnosel'skiĭ and S. G. Kreĭn. An iteration process with minimal residuals [in Russian]. *Matematičeskiĭ Sbornik, Novaya Seriya* **31** (73), 315–334 (1952).

[7] G. Pólya and G. Szegö. *Aufgaben und Lehrsätze aus der Analysis, Band I: Reihen, Integralrechnung, Funktiontheorie* [in German]. Fourth Edition. Heidelberger Taschenbücher, vol. 73. Springer-Verlag, New York (1970). [Cf. pp. xxx–yyy. Original version: 1925, Julius Springer, Berlin. Translated into English as [8].]

[8] G. Pólya and G. Szegö. *Problems and Theorems in Analysis, Volume I: Series, Integral Calculus, Theory of Functions* [in English]. Translated from the German [7] by Dorothee Aeppli (and revised and enlarged by the authors). Springer-Verlag, New York (1972). [Cf. pp. 71–72, 253–254.]

[9] P. Schweitzer. Egy egyenlötlenség az aritmetikai középértékről [in Hungarian: An inequality about the arithmetic mean]. 23. 257–261 (1914). [English translation by Levente T. Tolnai and Robert Vermes is Appendix A in [1].]

[10] G. S. Watson. Serial correlation in regression analysis. I. *Biometrika* **42**. 327–341 (1955).

2. Making statistical quirks and anti–quirks

A statistical quirk is defined to occur in the multiple regression analysis of a variable to be explained when the coefficient of determination for the full multiple regression model exceeds the sum of the coefficients of determination obtained on all the separate bivariate regressions of the variable to be explained on each explanatory variable. Anti- quirks can be defined as the contrary case

Quirks and anti-quirks exhibit different difficulties of interpretation. Initial data analyses and descriptive statistics provide first impressions which are contrary to the analysis obtained on fitting the full model. High dimensional multiple correlations are, however, common features.

Matrix methods are used to construct design matrices and explained variables with quirk and with anti-quirk properties. Methods used include spectral decomposition, singular value decomposition, Gram-Schmidt orthonormalisation and generalised inverses.

Impressions provided by initial data analyses which are contrary to the conclusions reached on by fitting the full model are illustrated graphically where possible. Such examples dramatically illustrate how first impressions concerning data and simplistic data analysis can lead to incorrect conclusions.

Philip V Bertrand.
School of Mathematics and Statistics,
The University of Birmingham, Birmingham B15 2TT, UK

3. Statistical inference for random planes

A random p-plane F in n-space is characterized uniquely by the nxn perpen dicular projection matrix $\mathbf{X'X}$. where $\mathbf{X}$ is any p-frame (i.e. $\mathbf{XX} = \mathbf{I}$) spanning F. Let $\mathbf{T}$ denote the sum of a sample of $\mathbf{N}$ such projection matrices. Let $\mathbf{D}$ be diagonal with diagonal elements the eigenvalues of $\mathbf{T}$ in decreasin g order. and $\mathbf{M}_m$ orthogonal with columns the corresponding eigenvectors. Then we write, in accord with the principal axis theorem:

$$\begin{aligned} \mathbf{T} &= \sum_k \mathbf{X}_k\mathbf{X}'_k = \mathbf{M}_m\mathbf{D}\mathbf{M}'_m \quad \text{(principal form)} \\ \text{where } k &= 1, ..., \mathbf{N} \end{aligned}$$

Partition $\mathbf{M}_m$ into p– and q–frames: $\mathbf{M}_m = (\mathbf{M}_{mp}\mathbf{M}_{mq})$. where $n = p + q$. Span(M_{mp}) is taken to be the *sample mean p–plane* since $\mathbf{M}_{mp}$ corresponds to the large st p eigenvalues of $\mathbf{D}$.

The partition of $\mathbf{M}_m$ induces a partitioning of $M'_m X$ into the pxp matrix $\mathbf{M}'_{mp}\mathbf{X}$ over the qxp matrix $\mathbf{M}'_{mq}\mathbf{X}$. The closer span($\mathbf{X}$) to span($\mathbf{M}_{mp}$), the closer $M'_{mp}X$ to an orthogonal matrix and the closer $\mathbf{M}'_{mq}\mathbf{X}$ to the null matrix. We seek to measure the deviation of span($\mathbf{X}$) from span(M_{mp}) by a qxp "error matrix" that is the same for all p–frames X with the same span. sinc e only span($\mathbf{X}$) is relevant for this. $\mathbf{M}'_{mq}\mathbf{X}$ will not do because it depends on more than span($\mathbf{X}$): permuting the columns of $\mathbf{X}$, for example, doesn't change the span but results in different entries for $\mathbf{M}_{mq}\mathbf{X}$. More appropriate is the qxp *estimated error matrix* $\mathbf{E}_{mk}$ for $\mathbf{X}_k$, which depends only o n span($\mathbf{X}_k$) (Downs, in press). and is given by

$$\mathbf{E}_{mk} = \mathbf{M}'_{mq}\mathbf{X}_k\mathbf{R} \tag{3.1}$$

where $R = \left[\left(\mathbf{M}'_{mp}\mathbf{X}_k\right)'\left(\mathbf{M}'_{mp}\mathbf{X}_k\right)\right]^{-\frac{1}{2}}\left(\mathbf{M}'_{mp}\mathbf{X}_k\right)'$ is a pxp and orthogonal.

Error matrices can be used to test hypotheses about the mean p–plane for the Bingham distribution of random p-planes in n–space. The Bingham distribution has the forms:

$$\mathbf{f}\left(\mathbf{X}_k\mathbf{X}'_k\right) = \mathbf{c}\left(\lambda\right)\mathbf{etr}\left(\mathbf{X}_k\mathbf{X}'_k\mathbf{G}\right) = \mathbf{c}\left(\boldsymbol{\lambda}\right)\exp\left[-\frac{1}{2}\sum_{ij}\left(2k_{ij}\mathbf{e}^2_{ijk}\right)\right]$$

where $c\left(\lambda\right)$ is a norming constant. the symmetric parameter matrix $\mathbf{G} = \mathbf{M}\boldsymbol{\lambda}\mathbf{M}'$ (principal form). the first p elements on the diagonal λ of sum to zero. the $\mathbf{k}_{ij}$ are functions óf $\boldsymbol{\lambda}$. and the e_{ijk} are elements of the *error matrix* E_k obtained from (2) by replacing $\mathbf{M}_m$= ($\mathbf{M}_{mp}\mathbf{M}_{mq}$) therein with $\mathbf{M}$ = ($\mathbf{M}_p\mathbf{M}_q$). When the$\mathbf{k}_{ij}$ are large the e_{ijk} are approximately independent and normall y distributed with null means and with variances the reciprocals of the $2\mathbf{k}_{ij}$ (Downs, in press). To construct a test of the null hypothesis that span($\mathbf{M}_p$)=span($\mathbf{M}_{0p}$). first estimate $\mathbf{M}$ under the null hypothesis by $\mathbf{M}_0$. par titioned as $\mathbf{M}_0$=($\mathbf{M}_{0p}\mathbf{R}_{0p}$ $\mathbf{M}_{0q}\mathbf{R}_{0q}$). where

$$\begin{aligned} R_{0p} &= [(M'_{mp}M_{0p})'(M_{mp}M_{0p})]^{-\frac{1}{2}}(M'_{mp}M_{0p})', \\ R_{0q} &= [(M'_{mq}M_{0q})'(M'_{mq}M_{0q})]^{-\frac{1}{2}}(M'_{mq}M_{0q})' \end{aligned}$$

and $\mathbf{M}_{0q}$ is any q–frame orthogonal to $\mathbf{M}_{0p}$. Then get the *null error matrices* $\mathbf{E}_{0k}$ (k=1,....n) from (2) by using $\mathbf{M}_0$ for $\mathbf{M}_m$. Form N pqx1 supervectors $\mathbf{e}_{0k}$ from the N $\mathbf{E}_{0k}$, and get their sample mean vector and dispersion matrix. Then test the null hypothesis that span($\mathbf{M}_p$)=span($\mathbf{M}_{0p}$) by a standard Hotelling-type F-test of the hypothesis that E($\mathbf{e}_{0k}$)=0.

Tom Downs.
The University of Texas Health Science Center,
School of Public Health.
P.O. Box 20186.
Houston. Texas

4. A class of statistical estimators related to principal components

In this paper we investigate the algebraic relationships between some of the more familiar estimation and testing procedures employed in multivariate econometrics and the principal components and continuum regression techniques of multivariate statistics.

Dr R.W.Farebrother. Department of Economic Studies, Victoria University of Manchester, Manchester M13 9PL, United Kingdom,
R.W.Farebrother@man.ac.uk

5. Some properties of shrinkage parameters of operational ridge regression estimators

It is a known fact that the Ordinary Least Squares estimator does not provide satisfactory estimates when the regression equation is affected by multicollinearity. In this case it may be preferrable to use Ridge Regression (RR) which has been usefully applied in diverse areas such as chemistry, criminology and economics. However these estimators depend upon some shrinkage parameters whose optimal values are unknown. Obviously the properties of the RR estimators will depend upon how well the shrinkage parameters have been estimated. For this reason several devices have been proposed to estimate these parameters, and their behaviour should help to explain the behaviour of the different operational RR estimators. It is the purpose of this paper to derive the distribution and moments of order one and two of some commonly used estimators of the shrinkage parameters. These estimators are in fact ratios of quadratic forms of non-central normal random variables. Given the complexity of the results, some numerical calculations will be presented. The expected contribution of this work is to be able to determine conditions under which an operational RR estimator would be preferable to the others.

Dr L Firinguetti. Departamento de Matematica, y C.C. Universidad de Santiago de Chile Casilla 307 - Correo 2 Santiago Chile,
LFIRINGU@USACHVMI.CL

6. Second order methods for computing linear minimax estimators

In a linear statistical model with restricted parameter set, alternatives to the Least Squares Estimator are of interest. One alternative is the linear minimax estimator, minimizing among the linear estimators the maximum mean squared error (the maximum being taken over the parameter set).

We address to a situation which has been studied fairly extensively, namely the case that the parameter comes from an ellipsoid, the covariance model matrix is known and positive definite, and the mean value model matrix has full column rank. Then the linear minimax estimator is obtainable from solutions M^* to the 'dual' problem

$$\text{minimize } \Phi(M) := \mathit{trace}[C(D+M)^{-1}] \text{ over } M \in \mathcal{M}, \tag{6.1}$$

where C and D are certain positive semidefinite and positive definite $(k \times k)$-matrices, resp., and $\mathcal{M}$ is the set of positive definite $(k \times k)$-matrices M with $\mathit{trace}(M) = 1$. Moreover, it is known that if M_n is a sequence of feasible matrices for (1) converging to the optimum, then the corresponding sequence of linear estimators tends to the minimax estimator.

For solving (1) numerically, we propose a Newton type method, described in Gaffke & Heiligers (1996) (*Optimization* 36, pp. 41–57): The gradient G_n and the Hessian H_n

(viewed as self-adjoint linear operator from $Sym(k)$ to $Sym(k)$) of Φ at each point M_n of the iteration are available, and thus a search direction for the next Newton step can be obtained by solving the quadratic subproblem

$$\text{minimize } trace[G_n(M - M_n) + \frac{1}{2}(M - M_n)H_n((M - M_n))] \text{ over } M \in \mathcal{M}. \quad (6.2)$$

Instead of solving (2), we replace the feasibility set $\mathcal{M}$ by

$$\mathcal{M}_n = conv(\{x_1x_1', \ldots, x_sx_s', y_1y_1', \ldots, y_ky_k'\}).$$

where $x_1, \ldots, x_s$ and $y_1, \ldots, y_k$ are orthonormal systems of eigenvectors of G_n and M_n, resp., corresponding to the nonzero eigenvalues. For this modified problem the Higgins & Polak method finds in an acceptable time a solution accurate enough to lead to a good Newton step in the overall algorithm.

Combining the obtained search directions with suitable steplengths, theoretic convergence of the resulting sequence M_n to the optimum can be proven. Moreover, in numerous, randomly generated examples the procedure also worked well, as for problem sizes up to $k = 30$ in the very most cases not more than 100 iterations were required to obtain linear estimators with guaranteed 'minimax' efficiency $1 - 10^{-6}$.

Robert Gaffke and Berthold Heiligers
Berthold.Heiligers@Mathematik.Uni-Magdeburg.De

7. Approximating a matrix of nominal values

The rank r least squares approximation $\mathbf{X}_r$ to a matrix $\mathbf{X}$ of real numbers is a fundamental result of matrix approximation with a well-known solution expressed in terms of the simgular value decomposition of $\mathbf{X}$ (Eckart and Young, 1936). When $\mathbf{X}$ is a matrix of nominal (categorical, qualitative) values it is usual to replace the category names by numerical scores (uanitification) estimated by optimising a ratio of two sums of squares leading to two-sided eigenvalue probl ms as in Multiple Correspondence Analysis and similar methods. Unfortunately, this gives a numerical $\mathbf{X}_r$ wheras it may be more appropriate to approximate $\mathbf{X}$ by another nominal matrix . The paper will show how the rank of a nominal matrix (ominal rank) may be defined in terms of an r-dimensional geometrical representation of a nominal matrix in a similar manner to how a rank r real matrix may be regarded as a set of points in R^r. The definition may also be adapted to cate for ordered categorical variables (rdinal rank) . Then the best approximation may be defined as the nominal matrix $\mathbf{X}_r$ which best matches $\mathbf{X}$ in terms of the number of agreements in their corresponding nominal values. It will be shown that this may be expressed as a constrained least squares approximation to an indicator matrix. Algorithms for maximising the number of agreements would lead to interesting new forms of multidimensional scaling but remain to be developed. Even without an ptimal algorithm . the new definition of nominal rank permits a nominal $\mathbf{X}_r$ to be derived from quantified $\mathbf{X}_r$ and hence the relative performances of existing (sub-optimal) quantification methods may be assessed. Examples will illustrate these ideas.

Dr J.C.Gower, Faculty of Mathematics, Open University Milton Keynes MK, United Kingdom, J.C.Gower@open.ac.uk

8. Iterative majorization for minimizing complicated functions

Several standard methods exist for minimizing a (multivariable) function. Here we focus on the relatively unknown method of iterative majorization (Ortega & Rheinboldt, 1970; De Leeuw, 1977, 1988, 1993; Heiser, 1995). This method has been successfully applied to minimize several loss functions. One of the key features of iterative majorization is that the sequence of function values decreases monotonically. For functions that are bounded below or sufficiently constrained, this property implies that usually a local minimum is found. Two particularly useful forms of majorization are distinguished: linear and quadratic majorization (De Leeuw, 1993; Groenen & Heiser, in press). Several majorization inequalities from the literature are discussed and categorized according this classification.

Majorization is illustrated in two applications. First, we discuss the tunneling method for global optimization (Montalvo, 1979; Groenen & Heiser, in press). This technique searches for increasingly better local minima. It consists of a local phase in which a local minimum is obtained, and a tunneling phase in which we search for a different configuration with the same loss function value. The second application comes from the area of distance-based multivariate analysis (Meulman, 1986, 1992), which has a strong relation with multidimensional scaling. Here, majorization is used to update the configuration and to find optimal nonlinear transformations of the variables.

Patrick J. F. Groenen, Department of Data Theory, PO Box 9555, 2300 RB Leiden, The Netherlands, groenen@ruls30.fsw.leidenuniv.nl. *Supported by The Netherlands Organization for Scientific Research (NWO) by grant nr. 575-67-053 for the 'PIONEER' project 'Subject Oriented Multivariate Analysis'.*

9. Solving a constrained nonnegative least squares problem within an iterative majorization scheme

Monotone spline transformations are an appealing alternative to general monotone transformations in nonlinear multivariate analysis (MVA) and distance-based MVA (Meulman, 1992). In this context, fitting monotone spline-transformations amounts to repeatedly solving a least-squares problem (LS), subjected to both nonnegativity constraints on the parameters and a length constraint on the transformed variable. Since the overall minimization problem in distance-based MVA constitutes a major computational task, it is necessary to find the optimal transformations as efficiently as possible. It will be shown that within the larger iterative scheme for distance-based MVA, the constrained LS problem basically can be resolved using two approaches. First, by using an analytic procedure to solve the LS problem subject to the nonnegativity constraints (Lawson & Hanson, 1992) and afterwards applying the length constraint. Second, by using an iterative alternating least squares approach that satisfies both types of constraints in each iteration. Using the results of a Monte-Carlo study, the efficiency of both approaches will be compared in terms of the quality of the found minimum and the convergence speed of the overall algorithm.

P.J.F. Groenen, B.J. van Os[1] *and J.J. Meulman*

Department of Data Theory, Leiden University, P.O. Box 9555, 2300 RB Leiden, The Netherlands, e-mail: groenen@ruls30.fsw.leidenuniv.nl
Meulman, J.J. (1992). The integration of multidimensional scaling and multivariate analysis with optimal transformations. *Psychometrika*, 57,539-565.
Lawson, L. and Hanson, R.J. (1974). *Solving least squares problems.* Englewood Cliffs, NJ: Prentice-Hall.

10. Issues involved in star orderings and their statistical applications

The star ordering introduced by Drazin (1978) is a very restrictive ordering. We analyze matrix partial orderings which are less restrictive. Examples of such orderings are the left star partial ordering, the right star partial ordering (introduced by Baksalary and Mitra, LAA, 149,1991), and σ-minus partial ordering (introduced by Baksalary and Hauke, LAA, 96, 1987). These orderings imply the minus partial ordering. There exist other two orderings extending the star partial ordering which do not imply the minus partial ordering. The first was introduced by Baksalary and Hauke, LAA, 127, (1990) and it coincides with the minus partial ordering of the Moore-Penrose inverses of ordered matrices. To define the second one let us observe that matrices ordered with respect to the star ordering belong to a family $\mathcal{S}$ of simultaneously singular-value decomposable matrices from the set of complex rectangular matrices, i.e. $\mathbf{A}, \mathbf{B} \in \mathcal{S}$ if matrices $\mathbf{AB}^*$ and $\mathbf{A}^*\mathbf{B}$ are positive semi-definite. In the family we can define the following ordering

$$\mathbf{A} \preceq_G^* \mathbf{B} \text{ if } \mathbf{A}, \mathbf{B} \in \mathcal{S} \text{ and } \mathbf{AB}^* - \mathbf{AA}^* \text{ is positive semi-definite.}$$

Above orderings are used in some statistcal problems connected with estimation in the Gauss-markov model and with distributional properties of quadratic forms.

Dr Jan Hauke Adam Mickiewicz, University University
Institute of Socio-Economic Geography and Spatial Planning
Fredry 10
61-701 Poznań, Poland
Tel: (+48-61) 526383
Fax: (+48-61) 524336
e-mail: jhauke@hum.amu.edu.pl

11. Linear Sufficiency and Admissibility in Restricted Linear Models

The linearly sufficient and admissible estimators of parameters vector in linear regression models are found in Markiewicz (1996) to be precisely the general rigde estimators.

[1] Department of Data Theory, PO Box 9555, 2300 RB Leiden, The Netherlands, os@ruls30.fsw.leidenuniv.nl. Supported by The Netherlands Organization for Scientific Research (NWO) by grant nr. 575-67-053 for the 'PIONEER' project 'Subject Oriented Multivariate Analysis'.

Linear sufficiency and admissibility is a desirable property especially in restricted linear models in which unbiased estimators are inadmissible. In such models the linearly sufficient and admissible estimators are identified as a special subclass of the class of general ridge estimators.

This characterization is based on a decomposition result for the admissible estimators associated with the subclass of estimators with bounded mean squared error. Following Gaffke and Heiligers (1989), and Heiligers (1993) we describe them as sums of the only admissible estimator of some unrestricted subparameter with bounded mean squared error, and of admissible estimators of certain completely restricted subparameter.

Gaffke, N. and B. Heiligers (1989). Bayes-, admissible, and minimax linear estimators in linear models with restricted parameter space. Statistics 20, 487–508.

Heiligers, B. (1993). Linear Bayes and minimax estimation in linear models with partially restricted parameter space. J. Statist. Plann. Inference 36, 175–184.

Markiewicz, A. (1996). Characterization of general ridge extimators. Statistics and Probability Letters 27, 145–148.

Berthold Heiligers, Fakultät für Mathematik Institut für Mathematische Stochastik, Universität Magdeburg Postfach 4120 D–39106 Magdeburg, Germany,
e-mail: heiliger@imst.math.uni-magdeburg.de
Augustyn Markiewicz Department of Mathematical and Statistical Methods, Agricultural University of Poznań Wojska Polskiego 28 PL–60637 Poznań, Poland e-mail: amark@owl.au.poz*

12. Narrower Bounds for Characteristic Roots of Hadamard Product

J. Schur (1911) established the upper and lower bounds for characteristic roots of Hadamard product, which are discussed by G. Styan (1973) and R. Horn (1990) in their respective survey articles on Hadamard product. However, the bounds as established by Schur are global bounds for all characteristic roots, hence may be of a limited use.

In this paper, we establish a theorem which defines the bounds narrower and specific to each characteristic root: Define a Hadamard product as $H : A \odot B$ where A and B are positive semi-definite matrices of order n. Then, $\lambda_i(\cdot)$ and $\lambda_i(\cdot)$ denoting, respectively, characteristic roots and diagonal elements of the argument matrix both in descending order, $\lambda_n(\mathbf{A})\,\alpha_i(\mathbf{B}) \leq \lambda_i(H) \leq \lambda_1(\mathbf{A})\,\alpha_1(\mathbf{B})$with A and B interchangeable.

The implications of the theorem herein established, from which Schur's Theorem readily follows as a corollary, may be summarized as: (1) the smaller the range of characteristic roots of one matrix in a Hadamard product of two positive semi-definite matrices, the closer each characteristic root of the Hadamard product to the correspondingly large diagonal element of the other matrix; (2) if characteristic roots of a matrix in the Hadamard product are finite, then the order of each characteristic root of the Hadamard product has the same order of the correspondingly large diagonal element of the other matrix; (3) because of interchangeability of the two constituent matrices, there are two sets of bounds defined, not necessarily identical, one can choose from. The narrower bounds herein established may be more useful if constituent matrices are both positive definite since otherwise lower bounds are uniformly unity. One may well find these properties of the characteristic roots to be of value in the context of theoretical work in statistics and elsewhere.

Eric Iksoon Im
Department of Economics, University of Hawaii at Hilo
200 W. Kawili Street, Hilo, Hawaii 96720-4091, USA
E-mail: eim@uhunix.uhcc.hawaii.edu
fax: (808)933-3685

13. Matrix methods in demographic analysis

Consider a female population ($P_{x,t}$) consisting of n age-groups over discrete points of time. Let f_x be the probability that a female from this population in age–group x will give birth to a daughter (only single births being considered) during the interval to t+1, and this daughter then survives to t+1. Let s_x be the probability that female in the age-group x at time t will survive to reach the next age-group at time t+1.

A matrix representation of the fertility and mortality process can be achieved through the definition of a Fertility-Survivorship Matrix (or Leslie Matrix) denoted by **L** and the population projection process may be represented as:

$$\mathbf{P}^{(t+1)} = \begin{bmatrix} 0 & 0 & . & . & f_1 & f_2 & . & . & f_k & 0 & . & . & 0 & 0 \\ s_1 & 0 & 0 & . & . & . & . & . & . & . & . & . & 0 & 0 \\ 0 & s_2 & 0 & . & . & . & . & . & . & . & . & . & 0 & 0 \\ . & . & . & . & . & . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . & . & . & . & . & . \\ 0 & 0 & 0 & . & . & . & . & . & . & . & . & 0 & s_{n-1} & 0 \end{bmatrix} \begin{bmatrix} P_1 \\ P_1 \\ P_1 \\ . \\ . \\ . \\ P_n \end{bmatrix} = \mathbf{LP}^{(t)}$$

The first row of the L matrix is filled by non-zero elements in the childbearing age-groups to represent fertility that occurred during the projection period. The non-zero elements of the subsequent rows represent the survival probabilities. Assuming an unchanging fertility and mortality regime (ie **L** is constant), then the projection process may be represented by thus:

$$\mathbf{P}^{(t+2)} = \mathbf{LP}^{(t+1)} = \mathbf{L}^2\mathbf{P}^{(t)} \Longrightarrow \mathbf{P}^{(t+n)} = \mathbf{L}^n\mathbf{P}^{(t)}$$

To this basic model, the following extensions are introduced:

1. The number of females in age group x at time t, denoted by $P_{x,t}$ and the number of births that survive to t+1 are assumed to be random variables.

2. A spatial disaggregation of the population by regions is undertaken, thereby requiring the incorporation of inter-regional migration into the projection process.

3. An examination of the stable properties of this projection model is undertaken in the context of some well-known results from demographic theory.

George Joseph, Department of Economic Studies, Victoria University of Manchester, Manchester M13 9PL, United Kingdom

14. Derivative of the Eigenvectors of a Symmetric matrix

The authors supply the derivative of an orthogonal matrix of eigenvectors of a real symmetric matrix. To illustrate the applicability of their result they consider a real symmetric random matrix for which a more or less standard convergence in distribution is assumed to hold. The well-known delta method is then used to get the asymptotic distribution of the orthogonal eigenmatrix of the random matrix.

Tonu Kollo and Heinz Neudecker, School of Actuarial Sciences and Econometrics, University of Amsterdam, Roetersstraat 11 1018 WB Amsterdam, The Netherlands

15. Stability conditions for linear stochastic models: a survey

This paper presents implementations, in Pascal, of the Routh–Hurwitz conditions for determining the stability of linear differential equations and of the Schur–Cohn conditions for determining the stability of linear difference equations. The two sets of conditions are intimately related. The Schur–Cohn conditions can be used to determine the stationarity and invertibility of linear stochastic models of the ARMA variety. The Schur–Cohn conditions are to be found in time-series analysis in a variety of disguises; and, in this paper, we also demonstrate some surprising relationships between the stability conditions and some other conditions which seem, at first, to be quite unconnected to them.

D.S.G. Pollock Department of Economics Queen Mary College University of London Mile End Road London E1 4 NS Tel : +44-71-975-5096 Fax : +44-71-975-5500

16. Power additivity and orthogonality

Suppose we are given a family $\{A_i\}_{i=1}^{i=s}$ of complex $n \times n$ matrices, and define $A = A_1 + \ldots + A_s$. We say that the family is power additive if

$$A^k = A_1^k + \ldots + A_s^k \quad \textit{for all } k = 1, 2, \ldots$$

and that the family is orthogonal if

$$A_i A_j = 0 \quad \textit{for } i \neq j.$$

It is clear that the orthogonality implies power additivity. The converse is not true as the following example shows

$$A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix}.$$

Our aim in this talk is to obtain structural results for orthogonal families of matrices and for power additive families of matrices. In particular, we will show that in certain cases these concepts coincide.

P. Šemrl and R.E. Hartwig Faculty of Mechanical Engineering, University of Maribor, Smetanova 17, 2000 Maribor, Slovenia

17. On semi-orthogonality and a special class of matrices

The concept of semi-orthogonality of two complex vectors is introduced. As a consequence, a generalization of the class of orthogonal projectors is investigated and the distribution of a quadratic form is considered.

Götz Trenkler
Jürgen Gross
Sven-Oliver Troschke
Department of Statistics
University of Dortmund
Vogelpothsweg 87
D-44221 Dortmund, Germany
e-mail: trenkler@amadeus.statistik.uni-dortmund.de

18. Lowner ordering of Matrices with Constraints and Matrices Extensions of Cauchy–Schwarz Inequality

We consider Lowner ordering of matrices under linear constraints. The restrictions are given in the form of linear matrix equations,which involve two matrices to be chosen. Inequalites are expressed as Lowner ordering of several matrices some of which are given and others to be chosen. Since the choice of matrices in the constractions is flexible, this approach provides an efficient method to generate Lowner orderings of matrices which are generalizations of the Cauchy–Schwarz inequality. Using this technique we have derived a plenty of inequalites: some of them are new,but many of them have recently appeared in the literature.

Song-Gui Wang
Beijing Polytechnic University, Beijing,China 100022
and
Erkki P. Liski
University of Tampere, Finland
e-mail: epl@uta.fi

19. On inequality constrained generalized least squares selections in the general possibly singular Gauss–Markov model

This paper deals with the general possibly singular linear model. It is assumed that in addition to the sample information we have some nonstochastic prior information concerning the unknown regression coefficients that can be expressed in form of linear independent inequality constraints. Since these constraints are part and parcel of the model the inequality constrained generalized least squares (ICGLS) problem arises that contains some unknown aspects up to now. Based on a projector theoretical approach we show in this paper how the set of ICGLS selections under the constrained model is related to the set of GLS selections under the associated unconstrained model. As a by-product we obtain an interesting method for determining an ICGLS selection from a GLS selection. Certain

special model cases are also considered. Some of the results discussed in Werner (1990) and Firoozi (1990) are reobtained.

Hans Joachim Werner
Institute for Econometrics and Operations Research
Econometrics Unit, University of Bonn
Adenauerallee 24-42, D-53113 Bonn, Germany
or470@unitas.or.uni-bonn.de; na.werner@na-net.ornl.gov

Index of Participating Authors

P. Bertrand, *section 2*

J. S. Chipman, *no abstract*

T. Downs, *section 3*

R. W. Farebrother, *section 4*

L. Firinguetti, *section 5*

J. C. Gower, *section 7*

P. Groenen, *section 8 and 9*

J. Hauke, *section 10*

B. Heiliger, *section 6 and 11*

E. I. Im, *section 12*

G. G. Joseph, *section 13*

A. Markiewicz, *section 10*

H. Neudecker, *section 14*

D. S.G. Pollock, *section 15*

P. Semrl, *section 16*

G. P. H. Styan, *section 1*

G. Trenkler, *section 17*

S. O. Troschke, *section 17*

B. J. van Os, *section 9*

S. Wang, *section 18*

H. J. Werner, *section 19*