PSYCHOMETRIKA—VOL. 60, NO. 3, 447–448
SEPTEMBER 1995
REVIEWS

REVIEWS

Jos M.F. ten Berge (1993). *Least squares optimization in multivariate analysis*. Leiden, The Netherlands: DSWO Press, 88 pp, ISBN 90-6695-083-8, Dfl 33.gz.

This is a fascinating book. As soon as I received a copy, I started reading it, and I could not stop reading it until I read it through in a matter of two hours or so. As you can imagine, this is a compact book. Yet it serves well its focused objective of providing a unified treatment of constrained and unconstrained least squares (LS) problems. The constraints considered are linear constraints, rank constraints and orthogonality constraints. The author shows remarkable restraint to keep the material to absolute essentials, and yet succeeds in "painting a colorful picture rich in content".

An LS problem is a minimization problem; that of finding a minimum of a function. A standard technique for solution is to use calculus. The author of this monograph achieves the same goal without using calculus. Instead, he exploits known vector and matrix inequalities such as Schwarz and ten Berge's (1983). While this idea of using inequalities is not entirely new (e.g., Rao, 1980), it was done so remarkably well. I had never seen any better (simpler and clearer) proofs of various optimality properties of singular value decomposition (SVD) of a matrix than those given in this monograph.

Ten Berge's inequality plays a particularly important role. In the simplest case, it is an extension of von Neumann's (1937) theorem, which states that for any orthonormal matrix **G** and a diagonal matrix **C** (nnd), tr(**GC**) $\leq$ tr(**C**). Ten Berge (1983) generalized this inequality into tr(**GC**) $\leq \sum_i^r c_i$, where **G** is any *suborthogonal* matrix of rank r and $c_1 \geq \ldots \geq c_r \geq 0$ are their largest diagonal elements of **C**. (If you don't know what a suborthonormal matrix is, that is a good reason to read this monograph.) In the monograph the author effectively used this inequality to show important optimality properties of SVD used to solve certain LS problems. To find out exactly how effective the inequality is, you should read the monograph. You will enjoy reading it, just as I did, and discover the importance of the inequality, which is not yet so widely recognized as it should in the standard matrix algebra literature. In the second half of the monograph, the author demonstrates the use of his results (established in the first part) in nine concrete settings, multiple regression analysis, principal component analysis, simultaneous component analysis in two or more populations, MINRES factor analysis, canonical correlation analysis, redundancy analysis, PARAFAC, INDSCAL and homogeneity analysis.

Although I feel that the monograph is self-contained, I would like to have the following topics to be included, if any follow-up to it is ever to be written. They complement the basic ideas in the monograph and expand the scope of LS problems. (1) Metric matrices extend the LS analysis in important ways by capturing variance-covariance structures of rows and columns of a data matrix. (2) Projection operators (both with or without nonidentity metric matrices) play important roles in the theory of LS estimation with linear constraints. There are two alternative ways of incorporating them, as discussed by Takane, Yanai, & Mayekawa (1991). For example, in the problem of minimizing SS($\mathbf{y} - \mathbf{y}^*$) with respect to $\mathbf{y}^*$, the constraints, $\mathbf{y}^* = \mathbf{X}\boldsymbol{\beta}$ and $\mathbf{W}\mathbf{y}^* = \mathbf{0}$, are equivalent, if **X** and **W**′ span complementary subspaces. (3) Generalized SVD (or SVD under nonidentity metric matrices) is important in linearly constrained LS prob-

0033-3123/95/$00.75/0

lems with rank constraints with or without nonidentity metric matrices (Takane & Shibayama, 1991). While these notions can be built up from the basic building blocks given in the monograph, they are sort of macro language which takes you to your destination more quickly or enables you to reach a further destination.

MCGILL UNIVERSITY *Yoshio Takane*

References

Rao, C. R. (1980). Matric approximation and reduction of dimensionality in multivariate statistical analysis. In P. R. Krishnaiah (Ed.), *Multivariate analysis V* (pp. 3–22). Amsterdam: North Holland.

Takane, Y., & Shibayama, T. (1991). Principal component analysis with external information on both subjects and variables. *Psychometrika*, *56*, 97–120.

Takane, Y., Yanai, H., & Mayekawa, S. (1991). Relationships among several methods of linearly constrained correspondence analysis. *Psychometrika*, *56*, 667–684.

von Neumann, J. (1937). Some matrix inequalities and metrization of matrix space. *Tomsk University Review*, *1*, 286–300. (Reprinted in A. H. Taub (Ed.), (1962). *John von Neumann collected works* (Vol. IV). New York: Pergamon Press)

PSYCHOMETRIKA—VOL. 60, NO. 3, 449–450
SEPTEMBER 1995
REVIEWS

Ton Heinen (1993). Discrete Latent Variable Models. Tilburg University Press, ISBN 90-361-9753-8, $27.00.

Over the last 10–15 years the literature on categorical data with applications in education, psychology and sociology has been enriched by a number of books published by the University Press at a number of Dutch universities. In most cases these books are the Ph.D. dissertations of Dutch researchers. A recent addition to this list is Ton Heinen's dissertation, published by the Tilburg University Press. Unfortunately too few of these find their way to the major publishing Companies.

The present book has many merits and should be a standard book on the bookshelf for many researchers in the social sciences, who works with data analysis. During my reading, I found some weaknesses, though, to which I shall return later.

The organization of the book is as follows. Chapter 1 is an overview of the models to be discussed in the book. Table 1.1 contains a typology of latent structure models. This typology has with variations been presented before, but is useful as a reference. I did not particularly like the parentheses with names. Several names, who have made valuable contributions, and who are not mentioned, came to my mind, most notable Paul Lazarsfeld and Georg Rasch.

Latent class models are discussed in chapter 2. Their connections to loglinear models are carefully explained and estimation problems as well as testing procedures are discussed in great details. This chapter also contains a relatively complete listing of the various latent class models obtained by putting restrictions or added assumptions on the parameters of the models.

Chapter 3 is devoted to a similar discussion of latent trait models, that is, latent structure models with a latent parameter, having a continuous variation. Much of the chapter is devoted to a description and a comparison of different latent trait models especially for the case of polytomous manifest variables. As a tool for the comparison, the author uses a number of tables showing parameter estimates and test quantities. Some readers, not familiar with the general literature, may find some difficulties following the presentation here. A full discussion of estimation procedures is postponed to chapter 4 (with a very brief outline in section 3.2.3). As regards the test statistics G^2 and Pearson's χ^2, they are not discussed in relation to the latent trait model, and I did not find any reference to section 2.2.2, where they are introduced. But in general I found the chapter interesting reading and it certainly gave me many useful insights in the potential usefulness of the various models.

As mentioned, estimation in latent trait models is discussed in more detail in Chapter 4. This is difficult stuff, but I felt during the reading, that the author has done a good job of explaining and evaluating the various procedures. Section 4.3 on goodness of fit tests is not so well written and suffers especially from a lack of illustrations with data.

The final chapter 5 deals with two types of extensions: Models with multidimensional latent variables and models with explanatory variables.

Now for the weaknesses I found. In most chapters the author concludes with a section titled: "Evaluation". I did not find these sections very valuable. To a large extent, they contain the authors own personal views and preferences. As a summary of

0033-3123/95/$00.75/0

the chapter, they are not clear and well-structured enough to my taste. A second point is the fact that only one data set is used for all illustrations. This makes, of course, comparisons possible between various models, but one data set cannot be expected to illustrate all potentials and drawbacks of the various models. In addition, it is a very limited data set. Only five items with three response categories. Such a limited data set cannot, I suspect, illustrate weaknesses of estimation and test procedures. On the positive side is that the complete data set is shown in Appendix B. A third point of critisism is that cross references between various parts of the book have not been carefully edited. Some examples: It would have helped the reader, if, in the short section 3.2.3, there had been appropriate references to more detailed discussions in chapter 4. On page 74 there is a reference to Table 2.4, which one (after some searching) can find on page 60. In fact an index or at least a list of the location of major tables would help the reader.

It is clear that the author is more concerned with estimation than with testing goodness of fit. This is to some extent reflected in the tables showing values of test statistics. One example is Table 3.8, page 159. The text is primarily concerned with comparing the five models, for which the G^2-value is shown. These comparisons takes the form of evaluations of differences between G^2-values. Why then are the differences and their appropriate numbers of degrees of freedom not shown in Table 3.8?

Two final minor points need to be mentioned: Why are the Haberman-conditions for consistency of the JML-procedure only mentioned in a footnote (on page 196)? Actually the conditions are very simple and very illustrative of the consistency problem, and I think the author should have stated them. The second minor point is a comment on page 237: "for polytomous data CML and MML only rarely lead to the same estimated item parameters". I wonder in what sense. My own experiences over the years with many data sets are that the item parameter estimates by CML and MML (with a normal latent density) seldom differ apart from the second or third decimal point. For the author's two basic data sets, CATANA produced an average numerical difference between the two sets of estimates equal to 0.023, with a range of estimated item parameters from +2.573 to −2.944. Only one difference was on the first decimal point (0.141), all other differences were on the second or third decimal point. As far as I could tell the comment on page 237 refers to Table 4.8, where two sets of parameters are estimated simultaneously. One may in fact note some strange facts in Table 4.8: Why are the MML and CML estimates equal for Category 2, when $T = 3$ but unequal for Category 2, when $T = 4$? In fact Table 4.4 is more in line with my general experiences.

In summary, I can whole-heartedly recommend the book to researchers using latent structure models for their data analyses. It offers many useful insights and stimulating discussions of the various models. As a textbook for an advanced course in latent structure models I can only recommend it as supplementary reading. A textbook version would need more careful editing, an index and more illustrative examples.

UNIVERSITY OF COPENHAGEN *E.B. Andersen*

Editor's note: The original edition of the book is (nearly) sold out, but a shortened, updated version will probably be published by Sage Publications, Thousand Oaks CA. Information can be obtained from: T.U.P, P.O. Box 90153, 5000 LE Tilburg, The Netherlands.