expect that a discussion of data analytic strategies is helped by the precision obtained by casting strategies in terms of computational frameworks.

We would like to thank the authors for a stimulating paper and hope that this is not the end but the beginning of a discussion.

ADDITIONAL REFERENCES

BUJA, A. and ASIMOV, D. (1985). Grand tour methods: An outline. In *Computer Science and Statistics: Proc. of the 17th Symposium on the Interface* 63–67. North-Holland, Amsterdam.

BUJA, A., ASIMOV, D. and HURLEY, C. (1989). Methods for subspace interpolation in dynamic graphics. Technical Memorandum, Bellcore, Morristown, N.J.

BUJA, A., HURLEY, C. and MCDONALD, J. A. (1986). A data viewer for multivariate data. In *Computer Science and Statistics: Proc. of the 18th Symposium on the Interface* 171–174. North-Holland, Amsterdam.

HURLEY, C. (1987). The data viewer: A program for graphical data analysis. Ph.D. dissertation and Technical Report, Dept. Statistics, Univ. Washington.

HURLEY, C. and BUJA, A. (1990). Analyzing high dimensional data with motion graphics. *SIAM J. Sci. Statist. Comput.* To appear.

MCDONALD, J. A. and PEDERSEN, J. (1988). Computing environments for data analysis. III: Programming environments. *SIAM J. Sci. Statist. Comput.* **9** 380–400.

TIERNEY, L. (1990). *LISP-STAT*. Wiley, New York. To appear.

Comment

Frank Critchley

It is a pleasure to welcome this paper by Weihs and Schmidli with its emphasis on the practical benefits which derive from combining classical dimensionality reduction methods with recent advances in interactive, dynamic graphics in a single integrated computing environment. At the same time, however pressing the practical need, asking for "a fairly general *single* routine strategy" (Section 1.1) for multivariate exploratory analysis seems, to me at least, to be asking for the moon. A more realistic objective might be to establish a framework of methods through which the user is guided by an expert system. We elaborate a little on this possibility below.

With one exception, my comments are of two types: possible extensions and remarks on the example. The exception is a detail which we dispose of first. In the context of resampling and Procrustes transformation (Section 3.7), the authors suggest that "it may be worth looking for analytic expressions derived from data disturbances analogously to Sibson (1979)." At least for PCA-COV and PCA-COR, some relevant formulae are given in Sections 3.6.2 and 6.3 of Critchley (1985). Note that the covariance matrix used there has divisor n. Trivial modifications apply when the divisor is $(n - 1)$. The formulae given are essentially expansions in inverse powers of $(n - 1)$. In practice, these expansions are usually truncated to obtain approximations. In this case, greater accuracy can be achieved by renormalization of the eigenvalues to sum to the easily computed perturbed trace and of the eigenvectors to have unit length. Exact orthogonalization is also possible.

Frank Critchley is Chairman, Department of Statistics, University of Warwick, Coventry CV4 7AL, United Kingdom.

POSSIBLE EXTENSIONS

The following remarks are partly taken from the unpublished conference paper by Critchley (1987) on graphical data analysis. They relate principally to the dimensionality reduction methods employed.

1. In that paper I suggested that healthy progress requires constructive interaction between five ingredients: (a) important practical problems, (b) sufficient computing power, (c) a sound mathematical/statistical basis, (d) a good framework of methods, and (e) international cooperation. The present paper is an excellent example of the first three ingredients, while hopefully its publication in this format in this journal will encourage the last of these!

2. It is within the fourth ingredient that there is perhaps the greatest scope for fruitful extensions. The authors offer in Table 1 a classification of multivariate techniques in terms of two "dimensions": the preinformation required and the aspects of the data that are optimally represented. This framework of methods can be fruitfully extended by adding new methods (as the authors remark in Section 6) and also, we note here, by adding new "dimensions" to the classification of methods.

3. The methods currently considered can be characterized as corresponding to one of several possibilities on each of a (nonexhaustive) number of additional

dimensions. For the most part this is the first-named possibility in each of the following lists. Used in concert these dimensions provide a rather rich framework of methods:

(a) Continuous, discrete, mixed data.
(b) Primal ($n \times n$), dual ($p \times p$), or both ($n \times p$).
(c) Spatial, discrete, and hybrid representations.
(d) Single, replicated, and structured data set(s).
(e) Two-way, three-way, and multiway data.
(f) Second, third, fourth and higher degree in the configuration coordinates.
(g) Inner-products, distances, and squared distances as the selected features of data and/or representation.
(h) (Generalized) least squares or any other measure of the accuracy of the representation of the data.
(i) Objects or variables compared two-, three-, or more at a time.
(j) Two, three, or more groups of variables considered (in a canonical analysis context).
(k) Euclidean, hyperbolic, or other geometries.
(l) Raw, standardized, filtered, or derived data.

4. In particular, in the same lettering as (3) above, we offer the following remarks or references:

(b) "Both" methods include not only the biplot which the authors mention, but also the considerable literature on correspondence analysis. See, for example, Greenacre (1984) and Lebart, Morineau and Warwick (1984).
(c) "Discrete" representations include, for example, those based on graphs (in the mathematical sense). The dendrograms of hierarchical cluster analysis are a leading case of this. Hybrid representations try to combine such discrete structures with the continuous structure of (typically) Euclidean space.
(e) Coppi and Bolasco (1989) offer a recent overview of multiway data analysis.
(f) This important dimension distinguishes essentially all classical (second degree) methods from several interesting newcomers. The lead example of these latter is perhaps the version of projection pursuit in which the criterion of "interestingness" selected is that combination of third and fourth moments which, in a certain sense, optimally detects departures from multivariate normality. At the same time, it is worth making the obvious remark that higher degree methods are intrinsically less robust than lower degree ones.
(g) (Squared) distance methods embrace that wide field known as multidimensional scaling. There is a natural duality between an inner-product based method and its corresponding squared-distance method. The algebra of this duality is partially explored in Critchley (1988).
(h) An appropriate measure to use depends on the context which is defined (at least in part) by the framework currently under discussion. For example, least-squares is often appropriate for spatial representations but not always (essentially for reasons of nonconvexity) for discrete representations such as dendrograms. Here alternative criteria, such as subdominance, have been developed. In other words, the menu of appropriate measures depends amongst other things on the coordinates of the other "dimensions" of the method.
(j) Van der Burg (1988) offers a recent review of and original contributions to (among other things) nonlinear canonical correlation analysis with two *or more* groups of variables.
(k) Non-Euclidean geometries applied to statistics have both a long history and form a research area of much current activity. See for example the work on differential geometry reported in Amari (1985), Amari, Barndorff-Nielsen, Kass, Lauritzen and Rao (1987), Barndorff-Nielsen (1988) and Dodson (1987). In our view, the advent of powerful graphics workstations, with their facilities for displaying, manipulating and updating curves and surfaces, will prove to be a decisive factor in the implementation of practical procedures based on recent primarily theoretical advances in this domain.
(l) We use "filtering" to cover any preprocessing of the data of which standardization is just one possibility. Model-based filters are described in, for example, Van der Heijden, de Falguerolles and De Leeuw (1989). The term "derived data" covers not only correlations but also any measure of association or dissimilarity.

5. Transformation of data is discussed by the authors. Transformation of the representation is also possible of course. The optimal transformation methodology of Gifi referred to by the authors is developed in a distance context in Meulman (1986). Monotone spline transformations are reviewed in Ramsay (1988).

In summary, there appears to be scope for intelligent, structured elaboration of the framework of methods offered. The user could potentially be guided through this framework of methods by an expert system which presents a sequence of menus. The user's responses then determine a set of appropriate, sensible analyses. Within this set it may well be a good strategy to select several methods which are "orthogonal" in those dimensions for which there is not a unique response. For example, a second degree, distance-

based, multidimensional scaling method may be selected along with a higher degree, inner-product based, projection pursuit method. If the same qualitative features are present in such "orthogonal" analyses, the user can be more sure that the corresponding effects are real ones and not just an artifact of a particular method employed.

Finally, I wonder to what extent the OMEGA system could fruitfully be developed along the general lines very briefly sketched in my published discussion of Van der Heijden, de Falguerolles and De Leeuw (1989, page 275). The thrust of those remarks was in favor of a general constructive interplay between two broad approaches to data analysis: the exploratory, graphical approach and the confirmatory, modeling approach. Could OMEGA benefit from blending with the second of these? Some particular possibilities that come to mind are: brushing points that are influential for particular aspects of the analysis; examining the robustness of the methods proposed; borrowing ideas from the *model* choice literature in the present *method* choice context; and filtering to remove uninteresting model effects to see more clearly what remains (the thrust of the original paper).

REMARKS ON THE EXAMPLE

The following remarks concern "color strength: unexpected nonpredictability" (Section 5.2):

To what extent is the reduction from 29 to 5 variables in the PCA-COV analysis a reflection of dominant variation of these variables compared to the rest? Recalling the discussion in Section 3.1, it would be helpful to know to what extent the results go through in a PCA-COR analysis.

The (3, 5) and (4, 5) scatterplots in Figures 6 and 7 seem to reveal an outlier with low STRVI and STRREM values for its STRTRA figure.

The authors note two oddly placed batches in Figure 8: numbers 84 and 93. Could it be that these are ill-fitting points in the dominant PCA plane (perhaps with high loadings on a particular minor component)?

ADDITIONAL REFERENCES

AMARI, S.-I. (1985). *Differential Geometric Methods in Statistics. Lecture Notes in Statist.* **28.** Springer, New York.

AMARI, S.-I., BARNDORFF-NIELSEN, O. E., KASS, R. E., LAURITZEN, S. L. and RAO, C. R. (1987). *Differential Geometry in Statistical Inference.* IMS, Hayward, Calif.

BARNDORFF-NIELSEN, O. E. (1988). *Parametric Statistical Models and Likelihood. Lecture Notes in Statist.* **50.** Springer, New York.

COPPI, R. and BOLASCO, S. (1989). *Multiway Data Analysis.* North-Holland, Amsterdam.

CRITCHLEY, F. (1985). Influence in principal components analysis. *Biometrika* **72** 627–636.

CRITCHLEY, F. (1987). Graphical presentation of data: A review and some recent developments. Presented at the 17th European Meeting of Statisticians, Thessaloniki.

CRITCHLEY, F. (1988). On certain linear mappings between inner-product and squared-distance matrices. *Linear Algebra Appl.* **105** 91–107.

DODSON, C. T. J. (1987). *Geometrisation of Statistical Theory.* ULDM Publications, Univ. Lancaster, England.

GREENACRE, M. J. (1984). *Theory and Applications of Correspondence Analysis.* Academic, London.

LEBART, L., MORINEAU, A. and WARWICK, K. M. (1984). *Multivariate Descriptive Statistical Analysis: Correspondence Analysis and Related Techniques for Large Matrices.* Wiley, New York.

MEULMAN, J. (1986). *A Distance Approach to Nonlinear Multivariate Analysis.* DSWO, Leiden.

RAMSAY, J. O. (1988). Monotone regression splines in action (with discussion). *Statist. Sci.* **3** 425–461.

VAN DER BURG, E. (1988). *Nonlinear Canonical Correlation and Some Related Techniques.* DSWO, Leiden.

VAN DER HEIJDEN, P. G. M., DE FALGUEROLLES, A. and DE LEEUW, J. (1989). A combined approach to contingency table analysis using correspondence analysis and log-linear analysis (with discussion). *Appl. Statist.* **38** 249–292.

Comment

N. I. Fisher

I am grateful for the opportunity to comment on this interesting piece of work. I regret that the rude interjection of the Australian holiday season has prevented me from giving the paper the attention it deserves, so I shall confine my remarks to a couple of specific aspects relating to graphical testing and estimation.

The authors are confronted by a common problem: the sheer volume of data sets being presented to the in-house statisticians means that the treatment of all but a very small number of sets must necessarily be

N. I. Fisher is Program Leader, Applied and Industrial Statistics, in the CSIRO Division of Mathematics and Statistics. His mailing address is CSIRO DMS, PO Box 218, Lindfield NSW 2070, Australia.