A Derivation of the Algebraic Curve for Multi-dimensional Data using the Least-squares Distance

Masahiro Mizuta
Hokkaido University

Abstract

We propose a method for finding the algebraic curve that fits multi-dimensional data. An algebraic curve in n dimensional space is generally defined by n-1 polynomial expressions. The proposed method finds the n-1 polynomial expressions of the algebraic curve for n dimensional data. The sum of the squares' distances from the data point to the nearest point on the curve is minimum.

Utilizing MDPREF When the Number of Sources is Large

Akinori Okada and Junko Marumo
Rikkyo (St. Paul's) University, Tokyo

Abstract

The purpose of the present paper is to introduce a procedure to utilize MDPREF(Carroll, 1972, 1980) when the number of sources is large. The procedure, which is based on the external MDPREF analysis (Okada & Marumo, 1992), uses cluster analysis and INDSCAL (Arabie, Carroll, & DeSarbo, 1987; Carroll, 1972). When the number of sources becomes large, the difficulty in graphically representing the result obtained by MDPREF increases. Thus the conventional ways of utilizing MDPREF when the number of sources is large are (a) to analyze all sources simultaneously or (b) to clustering (grouping) homogeneous sources (by cluster analysis or other similar procedures), to obtain averaged data in each cluster (group), and to analyze average data as if each cluster is a source. These two conventional ways have shortcomings such that it is practically impossible to disclose differences among sources, differences among sources within the same cluster (group) or differences among clusters (groups).

The present procedure consists of (a) to find clusters of homogeneous sources by cluster analysis, (b) to compute inner product matrix among objects in each cluster, (c) to analyze inner product matrices by INDSCAL to obtain a common object configuration and a weight configuration, and (d) to derive a source configuration (vectors) by external MDPREF analysis using the common object configuration resulted from INDSCAL. The present procedure was applied to the data on eating habits.

References

[1]Arabie, P., Carroll, J.D., & DeSarbo, W.S. (1987). Three-way scaling and clustering. Newbury Park, CA: Sage.

[2]Carroll, J. D. (1972). Individual differences and multidimensional scaling. In R. N. Shepard, A. K. Romney, & S. B. Nerlove (Eds.), Multidimensional scaling: Theory and applications in the behavioral sciences. Vol. 1 (pp. 105-155). New York: Seminar Press.

[3]Carroll, J. D. (1980). Models and methods for multidimensional analysis of preferential choice (or other dominance) data. In E. D. Lantermann & H. Feger (Eds.), Similarity and choice (pp. 234-289), Bern: Hans Huber.

[4]Carroll, J. D., & Chang, J. J. (1970). Analysis of individual differences in multidimensional scaling via N-way generalization of "Eckart-Young" decomposition. Psychometrika, 35, 283-319.

[5]Okada, A., & Marumo, J. (1992). External MDPREF analysis based on INDSCAL: An application to the survey data comparing eating habits among Tokyo, Paris, and New York. Proceedings of the 1992 Fall National Conference of the Operations Research Society of Japan, pp. 252-253.

(eds. Y.Dodge and et al.), COMPSTAT '92, Switzerland, Computational Statistics, Volume 2, PP 43-48, Physica-Verlag Heidelberg.

[11] Ohsumi, N.(1992), An Experimental System for Navigating Statistical Meta-Information -- The Meta-Stat Navigator, (eds. Y.Dodge and et al.), COMPSTAT '92, Switzerland, Computational Statistics, Volume 2, PP 375-380, Physica-Verlag Heidelberg.

[12] Nikkei BP Publishing Company(1992), Existing state of Multimedia, Nikkei Byte, No.105, PP 284-289, (in Japanese).

[13] AuthorWare(1992), User's Manual of AuthorWare Star, ASCII, (in Japanese).

A Consultation System for Statistical Analysis on Multimedia Software

Atsuhiro HAYASHI
Kawasaki Medical School

Abstract

The development and popularization of personal computers have had a strong influence on the field of statistics during the last ten years. The analysis of data which was once carried out only on main frame computers can now even at home with a personal computer.

But, many users who have data sets and purposes for analyses are worried about how to select statistical methods and combine them, or their possible misuse of methods, because they lack knowledge of and experience with statistical analysis. Furthermore it is difficult that how to select software which are fit to their purposes, and how to use them.

In the field of Artificial Intelligence(AI), there are some representational models of knowledge. The semantic network model is an effective model for the hierarchical knowledge, and the knowledge may be easily modified. The hypertool has the same structure as the semantic network model.

To deal with the above mention problems, we have developed a consultation system for statistical analysis based on hypertool. This system advises to the user one appropriate statistical method based on our knowledge base linking of the knowledge and programs of statistics. And this system utilized the feature of multimedia for user interface. It offers the representation of consulting things easily with figures.

The system derives a statistical method to fit the user's purpose by means of a dialogue between the user and this system. The system is designed for data analysts and students without statistical knowledge or experience.

In this paper, we describe a statistical consultation system that provides assistance to non-statisticians. There are many kinds of statistical software programs, but most of them require some knowledge of and experience with statistics. Therefore, the development of a consultation system that can provide knowledge of data analyses is eagerly desired by many users.

We developed the consultation system based on hypertool. This system is expandable and has the advantage of being able to add/modify statistical knowledge. The system derives a statistical method fitting user's purpose from a dialogue between the user and this system. The system aims at data analysts and students without statistical knowledge or experience.

Key Words : expert system; consultation system; hypertool; multimedia; semantic network

References
[1] Gale, W.A.(1986), Artificial Intelligence and Statistics, Addison Wesley.
[2] Ueno, H. and Ishizuka, M.(1987), Usage of Knowledge and it's Representation, Ohm Inc., (in Japanese).
[3] Tarumi, T. and Hayashi, A.(1988), Statistical Analysis Software --- Seto/B, Kyoritsu Publishing Company, (in Japanese).
[4] Hayashi, A. and Tarumi, T.(1988), A Consultation System for Method Selection on Data Analysis, The Proceedings of the 5th Annual JCS Meeting, PP 50-51, (in Japanese).
[5] Christopher F.Chabris(1989), Artificial Intelligence and Turbo C, Multiscience Press, Inc.
[6] MaxThink Inc.(1989), User's Manual of HyperLink, (in Japanese).
[7] Afifi, A.A. and Clark, V.(1990), Computer-aided multivariate analysis 2nd ed., Van Nostrand Reinhold Company.
[8] Van den Berg G.M. and Visser R.A.(1990), Knowledge Modelling for Statistical Consultation System; Two Empirical Studies, COMPSTAT '90, Yugoslavia, Physica-Verlag.
[9] Gebhardt F.(1990), An Expert System Strategy for Selecting Interesting Results, COMPSTAT '90, Yugoslavia, Physica-Verlag.
[10] Hayashi,A. and Tarumi,T.(1992), A Consultation System for Statistical Analysis on Hypertool,

A Latent Structure in College Entrance Data

Shoichi UEDA
Ryukoku University

Abstract

Due to an excessive preparation for the race to university entrance, students are sliced by 'hensachi' scores of preliminary achievement tests and guided to a route for entrance with higher passing possibility. Furthermore, this deduce a hierarchical ranking for high schools and universities regardless their identical tradition or educational target. This hierarchical structure is clearly found in high-school-colledge matrix data in Kansai area. Distributions of passing university are compared and two principal components are derived. These are

   1: two dimensional scores reflecting achievement ranking
   2: score reflecting commuting distance

with 47% and 16% contribution respectively. Using 1991 and 1992 data, facter scores are computed for each high school and university, and then two score-maps are compared. Quite stable configurations are found and demonstrates that the 'hensachi-slicing' is unchanged in this period. This suggests a drastic change could follow in comming years when 18-years-old populatin decrease but university capacity keep unchange. To find any sign of future change, repeated analysis should be made as soon as 1993 years data become available.

Contents

Bulletin

of

The Japanese Classification Society