Computers Ops Res. Vol. 21, No. 2, pp. 169–183, 1994
Printed in Great Britain. All rights reserved
0305-0548/94 $6.00 + 0.00

LINEAR METHODS IN MULTIMODE DATA ANALYSIS FOR DECISION MAKING

S. LIPOVETSKY† and A. TISHLER‡

Faculty of Management, Tel Aviv University, Ramat Aviv, Tel Aviv 69978, Israel

(Received January 1992; in revised form January 1993)

Scope and Purpose—Many-way matrices represent data received from many recipients, by a variety of scales, moments of time, different objects, situations, etc., depending on the specifics of the system under consideration. It is usually possible to evaluate such multimode data matrices by non-linear procedures of factor analysis, which are complicated methods that may not yield a satisfactory and/or reasonable solution. In this work, several simple linear approaches are considered that give solutions for very large matrices. We apply these methods to a problem of choice of university by applicants; that is, we evaluate a 3-way matrix using real data about university choice according to several attributes by high school graduates in Israel.

Abstract—This paper presents and analyses several methods for the evaluation of information given in the form of many-way matrices. These methods are based on the least squares approximation of a matrix by a many-vector product which can be represented as a nonlinear eigenvector problem. Using real data about university choice by high school graduates in Israel, we develop and compare the following three families of methods: parallel proportional profiles, various types of methods based on the use of cyclic matrices (canonical correlations, principal components, and planes' approximation), and minimization of relative deviations.

1. INTRODUCTION

Considerable attention is being focused today in the areas of economics, biology, political science, management, and other social and behavioral sciences to the evaluation of data represented by tables with many entries, or inputs, that is, data organized in a multi-facet matrix. This is due to the great complexity of the systems being investigated in these branches of human interest, and the development of computer capability which makes it possible to work with such information. Multi-facet matrices of data have been analysed in statistics as early as 1904, when Pearson studied complex correlation tables for multivariate distributions ([13], Chap. 33). Such tables are utilized in the processing of categorical data, multifactor cross-variance analysis, design of experiments, and sample surveys ([14], Chaps. 35, 38, 39). In factor analysis, many-facet matrices usually represent measurements, or observations, of various attributes of objects in different periods of time [11]. Different directions in multidimensional tables may correspond to scales, features, experts, proposals, concepts, situations, etc.—depending on the specific problem under consideration. Development of methods for the analysis of multi-facet data matrices (using all facets simultaneously) was originated by Tucker [27–29] in factor analysis for semantic differential problems, and in studies of the multivariate scaling problem [3, 4]. Various approaches to the evaluation of multi-dimensional data and their applications are found, among others, in [5, 12, 15, 16, 21, 24, 25]. These approaches are, usually, quite difficult to use in moderate and large-scale problems since they require the use of non-linear optimization methods.

†Stanislav Lipovetsky holds an M.Sc. in atomic physics and a Ph.D. in mathematical methods in economics from Moscow State University. Since 1990 Dr Lipovetsky has been living in Israel. Presently he is affiliated with the Faculty of Management at Tel Aviv University. His primary areas of research are in multivariate statistics, multicriteria decision models, analytic hierarchy processes, econometrics, microeconomics.

‡Asher Tishler received his B.A. in Economics and Statistics from the Hebrew University, Israel (1971), and his Ph.D. in Economics from the University of Pennsylvania (1976). He is currently at the Faculty of Management at Tel Aviv University. His publications have appeared in the *Journal of Econometrics, Review of Economics and Statistics, European Economic Review, The Energy Journal, Operations Research, Management Science, Journal of Optimization Theory and Applications* and *The Journal of the American Statistical Association.*

In this paper we present and analyse several approaches to the evaluation of multimode data that allow the use of analytical solutions. These solutions can be reduced to simple linear algebraic problems that are easy to use and interpret. For ease of exposition, we analyse in this paper a three-way matrix of attributes $i = 1, 2, \ldots, n$ which are used for considering objects $j = 1, 2, \ldots, m$ by individuals $k = 1, 2, \ldots, p$. Each element of the matrix Q_{ijk} is an evaluation (estimation) given to the ith attribute of the jth object by the kth individual. These evaluations (estimations) are given as numerical values. Our purpose here is to derive from these evaluations, Q_{ijk}, some kind of aggregated evaluations of weights $a_1, a_2, \ldots, a_n$ for attributes (which may be compared), weights $b_1, b_2, \ldots, b_m$ for objects in order to rank them in a meaningful way, and of weights $c_1, c_2, \ldots, c_p$ for individuals which will allow them to be classified in groups with similar opinions.

1. Set up of the model: least squares approximation of a matrix by vectors

Consider a 2-dimensional matrix x with n rows and m columns. Denote the elements of x by x_{ij} and its approximation by the two vectors' external product

$$x = \lambda a \cdot b' + \varepsilon, \tag{1}$$

where a and b are column vectors with elements a_i and b_j, ε—matrix of residuals in approximation (1). The least squares (LS) procedure for determining the vectors a and b is known as the Eckart–Young approximation [9]:

$$S = \|\varepsilon\|^2 = \|x - \lambda ab'\|^2 \to \min, \tag{2}$$

which can be written as

$$S = \sum_{ij} \varepsilon_{ij}^2 = \sum_{i=1}^{n} \sum_{j=1}^{m} (x_{ij} - \lambda a_i b_j)^2 \to \min. \tag{3}$$

From the first-order conditions

$$\frac{\partial s}{\partial a_i} = 0, \qquad \frac{\partial s}{\partial b_j} = 0, \tag{4}$$

with the normalization restrictions

$$\sum_{i=1}^{n} a_i^2 = 1, \qquad \sum_{j=1}^{m} b_j^2 = 1, \tag{5}$$

we get a system of equations

$$\sum_j x_{ij} b_j = \lambda a_i, \qquad i = 1, \ldots, n \tag{6a}$$

$$\sum_i x_{ij} a_i = \lambda b_j, \qquad j = 1, \ldots, m \tag{6b}$$

or, in matrix form,

$$xb = \lambda a, \qquad x'a = \lambda b. \tag{7}$$

Substituting one of the equations in (7) into the other we obtain the usual eigenvalue problems

$$xx'a = \lambda^2 a, \tag{8a}$$

$$x'xb = \lambda^2 b. \tag{8b}$$

The eigenvectors a and b which solve problem (1–3) correspond to the maximal eigenvalue λ^2 in (8). If x is centered and normalized by the standard deviations of its rows, xx' becomes a correlation matrix. Thus, (8a) presents the usual principal component (PC) analysis for attributes, and (8b) is the dual principal component method for the objects. Relations (7) can be written as a combined eigenvalue problem as follows

$$\begin{pmatrix} 0 & x \\ x' & 0 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \end{pmatrix} = \lambda \begin{pmatrix} a \\ b \end{pmatrix}. \tag{9}$$

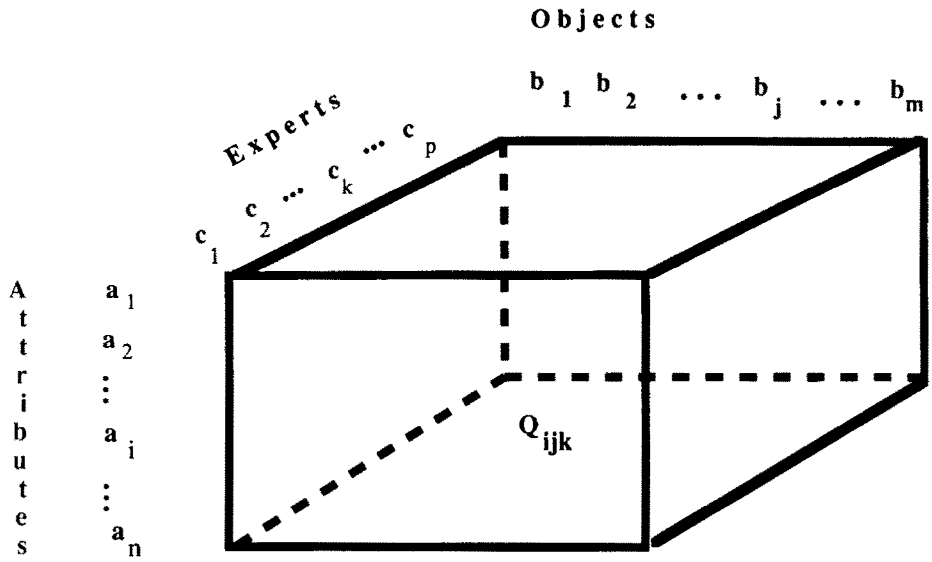

Fig. 1. 3-Way matrix of data.

The vectors a and b in (9) can be jointly obtained as the coordinate subspace of the right-hand side vector in (9). Multiplying (9) by the matrix in the left-hand side of (9) yields

$$\begin{pmatrix} xx' & 0 \\ 0 & x'x \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \lambda^2 \begin{pmatrix} a \\ b \end{pmatrix} \tag{10}$$

which is a combined representation for both equations in (8).

Procedure (1)–(10) was described for a 3-dimensional matrix Q_{ijk} (see Fig. 1) in [18]. Analogously to (1) we can approximate Q_{ijk} in Fig. 1 by the direct product (element by element) of three vectors

$$Q_{ijk} = \lambda a_i b_j c_k + \varepsilon_{ijk}, \tag{11}$$

where ε_{ijk} is the residual of the approximation of Q_{ijk} by the three vectors a, b, and c in Fig. 1. λ is a normalizing constant. LS minimization of the squared Euclidean norm of ε_{ijk} is given by

$$S = \sum_{ijk} \varepsilon_{ijk}^2 = \sum_{ijk} (Q_{ijk} - \lambda a_i b_j c_k)^2 \to \min. \tag{12}$$

The first-order conditions

$$\frac{\partial s}{\partial a_i} = 0, \qquad \frac{\partial s}{\partial b_j} = 0, \qquad \frac{\partial s}{\partial c_k} = 0, \qquad \frac{\partial s}{\partial \lambda} = 0, \tag{13}$$

and the restrictions of normalization

$$a'a = 1, \qquad b'b = 1, \qquad c'c = 1, \tag{14}$$

yield the following system of equations

$$\sum_{jk} Q_{ijk} b_j c_k = \lambda a_i, \qquad i = 1, \ldots, n \tag{15a}$$

$$\sum_{ik} Q_{ijk} a_i c_k = \lambda b_j, \qquad j = 1, \ldots, m \tag{15b}$$

$$\sum_{ij} Q_{ijk} a_i b_j = \lambda c_k, \qquad k = 1, \ldots, p \tag{15c}$$

$$\sum_{ijk} Q_{ijk} a_i b_j c_k = \lambda. \tag{16}$$

As shown in [18], (15) can be reduced to a nonlinear eigenvalue problem

$$\begin{pmatrix} Q(b)Q'(b) & 0 \\ 0 & Q(a)Q'(a) \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \lambda^2 \begin{pmatrix} a \\ b \end{pmatrix} \tag{17}$$

where $Q(b)$ and $Q(a)$ are matrices with the following elements

$$[Q(b)]_{ik} = \sum_j Q_{ijk} b_j, \tag{18a}$$

$$[Q(a)]_{jk} = \sum_i Q_{ijk} a_i. \tag{18b}$$

Problems (15)–(16), or (17)–(18), can be solved numerically by nonlinear programming methods. However, if n, m and p are large integers, or if the matrix considered is of a dimension which is greater than 3, nonlinear programming may not be feasible. Furthermore, there are serious problems concerning existence, convergence and uniqueness of the solution to the nonlinear problem.

In the rest of this paper we consider approximations to the nonlinear LS problem that result in linear problems with simple solutions for the vectors a, b and c.

2. PARALLEL PROPORTIONAL PROFILES

The principle of Parallel Proportional Profiles (PPP) was introduced by Catell [6–8, 21] (see also [16], pp. 8–9, 126–127, 150–151) who suggested that factor scores will change proportionally from one condition (or situation, moment of time, etc.) to another. In our representation of information in Fig. 1 this means that the weights a_i, b_j and c_k must be proportional to some measure of weights for each ith, jth and kth layer of the matrix Q in Fig. 1.

Here we suggest the following mathematical representation. Suppose, we solve the common eigenvalue problem (6) for every plane of the matrix Q. For each given index, k, we have the relations

$$\left.\begin{aligned} &\sum_j Q_{ij(k)} b_j^{(k)} = \lambda^{(k)} a_i^{(k)}, \qquad i = 1,\ldots,n \quad (19a) \\ &\sum_i Q_{ij(k)} a_i^{(k)} = \lambda^{(k)} b_j^{(k)}, \qquad j = 1,\ldots,m \quad (19b) \end{aligned}\right\} k = 1,\ldots,p,$$

where $Q_{ij(k)}$ are matrices parallel to the frontal plane in Fig. 1, and $a^{(k)}$ and $b^{(k)}$ are eigenvectors for the kth layer. Also, we have for each plane in the direction of j

$$\left.\begin{aligned} &\sum_i Q_{i(j)k} a_i^{(j)} = \mu^{(j)} c_k^{(j)}, \qquad k = 1,\ldots,p \quad (20a) \\ &\sum_k Q_{i(j)k} c_k^{(j)} = \mu^{(j)} a_i^{(j)}, \qquad i = 1,\ldots,n \quad (20b) \end{aligned}\right\} j = 1,\ldots,m,$$

where $Q_{i(j)k}$ are the vertical-profile matrices in Fig. 1. The third direction yields, for each i, the following equations

$$\left.\begin{aligned} &\sum_j Q_{(i)jk} b_j^{(i)} = \nu^{(i)} c_k^{(i)}, \qquad k = 1,\ldots,p \quad (21a) \\ &\sum_k Q_{(i)jk} c_k^{(i)} = \nu^{(i)} b_j^{(i)}, \qquad j = 1,\ldots,m \quad (21b) \end{aligned}\right\} i = 1,\ldots,n,$$

where $Q_{(i)jk}$ are the horizontal layers of the matrix in Fig. 1. Multiplying each pair (19)–(21) by the dual vector and summing by this vector's index, taking into consideration (5) or (14), we obtain the eigenvalues

$$\sum_{ij} Q_{ij(k)} a_i^{(k)} b_j^{(k)} = \lambda^{(k)}, \qquad k = 1, \ldots, p \tag{22a}$$

$$\sum_{ik} Q_{i(j)k} a_i^{(j)} c_k^{(j)} = \mu^{(j)}, \qquad j = 1, \ldots, m \tag{22b}$$

$$\sum_{jk} Q_{(i)jk} b_j^{(i)} c_k^{(i)} = \nu^{(i)}, \qquad i = 1, \ldots, n \tag{22c}$$

where all $\lambda^{(k)}$, $\mu^{(j)}$ and $\nu^{(i)}$ are the maximal singular values (square root of the eigenvalues) of the plane matrices in each kth, jth and ith layer, respectively.

The similarity of (15) and (22) is obvious. The summation of the matrix Q_{ijk} by the indices i and j in (22a) is made with weights $a_i^{(k)}$ and $b_j^{(k)}$, and in (15c) with weights a_i and b_j (which are constant over all layers k). As a result, we obtain in (22a) the vector $\lambda^{(k)}$ with p components; in (15c) we obtain λc_k, also a vector (multiplied by a normalizing constant) of size p. We have an analogous similarity in (22b) and (15b), and in the pair (22c) and (15a). That is, the LS approximation for a whole matrix by one set of vectors (12) yields relations (15) for determining the vectors a, b, and c, which we can interpret as mean vectors of the sets $a^{(k)}$ and $a^{(j)}$, $b^{(k)}$ and $b^{(i)}$, $c^{(i)}$ and $c^{(j)}$ respectively. Comparing the right-hand sides of equations (22) and (15) we can obtain expressions for these mean vectors a, b and c:

$$\lambda a_i = \nu^{(i)}, \qquad i = 1, \ldots, n \tag{23a}$$

$$\lambda b_j = \mu^{(j)}, \qquad j = 1, \ldots, m \tag{23b}$$

$$\lambda c_k = \lambda^{(k)}, \qquad k = 1, \ldots, p \tag{23c}$$

Thus, we obtain a result that corresponds to the PPP principle: elements a_i of the vector a for the 3-dimensional matrix Q are proportional (23a) to the maximal singular numbers $\nu^{(i)}$ (square roots of the maximal eigenvalues) of the 2-dimensional matrices (21) identified by the fixed index i as the layers of the plane matrices $Q_{(i)jk}$. The same is true for the other vectors: the components b_j of b for the matrix Q are proportional to the singular numbers $\mu^{(j)}$ (23b) of the 2-dimensional matrices $Q_{i(j)k}$ in (20), which can be obtained as the jth layers of the matrix Q_{ijk}. Finally, using (23c), the components c_k of the vector c for the 3-dimensional matrix Q are proportional to the singular values in (19) for each kth layer $Q_{ij(k)}$.

Thus, we have a very simple procedure for obtaining the vectors a, b, c for the matrix Q; i.e. for $k = 1$ we solve eigenvalue problem (19) for the first frontal layer $Q_{ij(1)}$ of the matrix Q_{ijk}. Problem (19) can be reduced to the two usual eigenvalue problems in the same way that (6) was reduced to (8). Hence, it is sufficient to solve only one of the problems (8a) or (8b). From this solution we need only the maximal eigenvalue $(\lambda^{(1)})^2$. Following this scheme, and taking one frontal layer (with $k = 1, 2, \ldots, p$) after another, we can get all maximal numbers $(\lambda^{(k)})^2$ for each k. Taking their square roots and dividing these singular values by the sum of the eigenvalues, we obtain the normalized (see (14)) eigenvector c for the matrix Q:

$$c_k = \frac{\lambda^{(k)}}{\sum_{\tilde{k}=1}^{p} (\lambda^{(\tilde{k})})^2}, \qquad k = 1, \ldots, p \tag{24a}$$

Using the same procedure in the directions i and j of the matrix Q we obtain the other normalized vectors

$$a_i = \frac{\nu^{(i)}}{\sum_{\tilde{i}=1}^{n} (\nu^{(\tilde{i})})^2}, \qquad i = 1, \ldots, n \tag{24b}$$

$$b_j = \frac{\mu^{(j)}}{\sum_{\tilde{j}=1}^{m} (\mu^{(\tilde{j})})^2}, \qquad j = 1, \ldots, m \tag{24c}$$

With these estimates of the eigenvectors we can calculate the singular value λ for the 3-dimensional matrix using equation (16).

The PPP principle can be easily generalized to 4, or more, dimensional matrices. For example, in the case of a 4-dimensional matrix we define the following LS approximation

$$S = \sum_{ijkt} (Q_{ijkt} - \gamma a_i b_j c_k d_t)^2 \to \min, \tag{25}$$

and similarly to (13–16) we obtain the system

$$\sum_{jkt} Q_{ijkt} b_j c_k d_t = \gamma a_i, \qquad i = 1, \dots, n, \tag{26a}$$

$$\sum_{ikt} Q_{ijkt} a_i c_k d_t = \gamma b_j, \qquad j = 1, \dots, m, \tag{26b}$$

$$\sum_{ijt} Q_{ijkt} a_i b_j d_t = \gamma c_k, \qquad k = 1, \dots, p, \tag{26c}$$

$$\sum_{ijk} Q_{ijkt} a_i b_j c_k = \gamma d_t, \qquad t = 1, \dots, \tau, \tag{26d}$$

$$\sum_{ijkt} Q_{ijkt} a_i b_j c_k d_t = \gamma. \tag{27}$$

The 4-way matrix Q_{ijkt} can be subsequently reduced to 3-way matrices by fixing each one of its indices. In this case we first obtain the eigenvalues of the 2-way submatrices, use them for constructing eigenvectors of 3-way submatrices (24), and obtain the singular values of these 3-way submatrices. These singular values are then used as components of the eigenvectors (26) of the 4-way matrix, which in turn yields the singular value (27) for the initial matrix Q_{ijkt} in (25).

3. VARIANTS THAT USE CYCLIC MATRICES

3a. Plane's approximation

In this section we consider a number of variants of the approximation for the 3-way matrix Q by planes constructed from the vectors a, b and c.

Using the direct product of all pairs of a, b and c for approximation of all layers of Q corresponds to the following LS approach

$$S = \sum_k \left[\sum_{i,j} (Q_{ijk} - \lambda a_i b_j)^2 \right] + \sum_j \left[\sum_{i,k} (Q_{ijk} - \lambda a_i c_k)^2 \right] + \sum_i \left[\sum_{jk} (Q_{ijk} - \lambda b_j c_k)^2 \right]. \tag{28}$$

From (13) and normalization (14) we have

$$0 \cdot a + pxb + mz'c = \lambda(p + m)a, \tag{29a}$$

$$px'a + 0 \cdot b + nyc = \lambda(p + n)b, \tag{29b}$$

$$mza + ny'b + 0 \cdot c = \lambda(n + m)c, \tag{29c}$$

where

$$x \equiv \bar{Q}_{ij} \equiv \frac{1}{p} \sum_{k=1}^{p} Q_{ijk}, \tag{30a}$$

$$y \equiv \bar{Q}_{jk} \equiv \frac{1}{n} \sum_{i=1}^{n} Q_{ijk}, \tag{30b}$$

$$z \equiv \bar{Q}_{ki} \equiv \frac{1}{m} \sum_{j=1}^{m} Q_{ijk}. \tag{30c}$$

(29) represents a general eigenvalue problem with a symmetric matrix on the left-hand side and a diagonal matrix on the right-hand side:

$$\begin{pmatrix} 0 & px & mz' \\ px' & 0 & ny \\ mz & ny' & 0 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \lambda \begin{pmatrix} (p+m)I_n & 0 & 0 \\ 0 & (p+n)I_m & 0 \\ 0 & 0 & (n+m)I_p \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix}, \tag{31}$$

where I_n, I_m, I_p are identity matrices of orders n, m and p, respectively.

It is possible to redefine (28) and use the average values of the sums to obtain

$$S = \frac{1}{p}\sum_k^p \left[\sum_{ij}(Q_{ijk} - \lambda a_i b_j)^2\right] + \frac{1}{m}\sum_j^m \left[\sum_{i,k}(Q_{ijk} - \lambda a_i c_k)^2\right] + \frac{1}{n}\sum_i^n \left[\sum_{j,k}(Q_{ijk} - \lambda b_j c_k)^2\right]. \tag{32}$$

As in (28–31) we can obtain a simple eigenvalue problem from (32)

$$\begin{pmatrix} 0 & x & z' \\ x' & 0 & y \\ z & y' & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \mu \begin{pmatrix} a \\ b \\ c \end{pmatrix} \tag{33}$$

with the submatrices x, y and z defined in (30), and eigenvalue $\mu = 2\lambda$. It is interesting to note that if we approximate the averaged layers by plane matrices constructed from direct vector products, that is, using the following LS approximation

$$S = \sum_{ij}\left(\frac{1}{p}\sum_k^p Q_{ijk} - \lambda a_i b_j\right)^2 + \sum_{ik}\left(\frac{1}{m}\sum_j^m Q_{ijk} - \lambda a_i c_k\right)^2 + \sum_{jk}\left(\frac{1}{n}\sum_i^n Q_{ijk} - \lambda b_j c_k\right)^2, \tag{34}$$

we obtain precisely the same solution for a, b and c, as in (33). Eigenvalue problem (33) can be solved by standard methods for symmetric matrices. System (33) may also be reduced to a system with only two vectors. For example, from (33) we can express one vector as follows

$$c = \frac{1}{\mu}(za + y'b), \tag{35}$$

and insert it into the first two subsystems in (33). This operation results in a quadratic eigenvalue system for the vectors a and b

$$\left\{\begin{pmatrix} I_n & 0 \\ 0 & I_m \end{pmatrix} \cdot \mu^2 - \begin{pmatrix} 0 & x \\ x' & 0 \end{pmatrix} \cdot \mu - \begin{pmatrix} z'z & z'y' \\ yz & yy' \end{pmatrix}\right\} \begin{pmatrix} a \\ b \end{pmatrix} = 0. \tag{36}$$

It is possible to transform (36) to a problem that is linear in μ (see [30], Chap. 9, Pt 61). Denote

$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \mu \begin{pmatrix} a \\ b \end{pmatrix}, \tag{37}$$

whereby system (36) becomes:

$$\begin{pmatrix} 0 & 0 & I_n & 0 \\ 0 & 0 & 0 & I_m \\ z'z & z'y' & 0 & x \\ yz & yy' & x' & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ \alpha \\ \beta \end{pmatrix} = \mu \begin{pmatrix} a \\ b \\ \alpha \\ \beta \end{pmatrix}. \tag{38}$$

Eigenvector problem (38) is of order $2(n + m)$, while (33) is of order $n + m + p$. Hence, when $p \gg n + m$ it is preferable to solve problem (38) which is non-symmetric, but may be of a much smaller dimension than problem (33). For example, if 10 attributes were estimated for 5 objects by 100 individuals (see Fig. 1) it is possible to eliminate the last direction and solve problem (38), which is of order 30, instead of problem (33), which is of order 115.

3b. Canonical correlation and principal component analysis

In this section we consider the application of Canonical Correlation Analysis (CCA) to the determination of the vectors a, b, c. For each pair of the three projections (30) we can construct vectors of common size; for example,

$$xb = \zeta, \qquad z'c = \eta, \tag{39}$$

where ζ and η have n components, and may be interpreted as estimates of the vector a from the side of the matrix x and from the side of the matrix z. ζ and η can be constructed by maximizing the following canonical correlation problem

$$\rho(b, c) = c'zxb - \frac{\lambda}{2}(b'x'xb - 1) - \frac{\mu}{2}(c'zz'c - 1). \tag{40}$$

Maximization of (40) results in $\lambda = \mu$ and the eigenvector problem (see [17])

$$\begin{pmatrix} 0 & x'z' \\ zx & 0 \end{pmatrix}\begin{pmatrix} b \\ c \end{pmatrix} = \lambda\begin{pmatrix} x'x & 0 \\ 0 & zz' \end{pmatrix}\begin{pmatrix} b \\ c \end{pmatrix}. \tag{41}$$

Generalizing (39)–(41) to account for all three pairs of correlations we have

$$\rho = \rho(b, c) + \rho(a, b) + \rho(a, c) = c'zxb + b'yza + a'xyc, \tag{42}$$

with one normalization

$$\varphi \equiv a'(xx' + z'z)a + b'(x'x + yy')b + c'(zz' + y'y)c - 3 = 0. \tag{43}$$

Therefore, the function for maximization is

$$S = \rho - \frac{\lambda}{2}\varphi, \tag{44}$$

with ρ and φ defined by (42) and (43), respectively. Maximizing (44) we obtain a generalization of problem (41) for three vectors

$$\begin{pmatrix} 0 & z'y' & xy \\ yz & 0 & x'z' \\ y'x' & zx & 0 \end{pmatrix}\cdot\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \lambda\begin{pmatrix} (xx' + z'z) & 0 & 0 \\ 0 & (yy' + x'x) & 0 \\ 0 & 0 & (zz' + y'y) \end{pmatrix}\cdot\begin{pmatrix} a \\ b \\ c \end{pmatrix}. \tag{45}$$

If we add the right-hand side of (45) (without λ) to both sides of (45), we obtain the system

$$\begin{pmatrix} (xx' + z'z) & z'y' & xy \\ yz & (yy' + x'x) & x'z' \\ y'x' & zx & (zz' + y'y) \end{pmatrix}\cdot\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \nu\begin{pmatrix} (xx' + z'z) & 0 & 0 \\ 0 & (yy' + x'x) & 0 \\ 0 & 0 & (zz' + y'y) \end{pmatrix}\cdot\begin{pmatrix} a \\ b \\ c \end{pmatrix}, \tag{46}$$

with eigenvalue $\nu \equiv \lambda + 1$.

The matrix on the left-hand side of (46), denoted A, is equal to the matrix of the second moments of the symmetric matrix of initial information which is used in (33), i.e.

$$A \equiv \begin{pmatrix} (xx' + z'z) & z'y' & xy \\ yz & (yy' + x'x) & x'z' \\ y'x' & zx & (zz' + y'y) \end{pmatrix} = \begin{pmatrix} 0 & x & z' \\ x' & 0 & y \\ z & y' & 0 \end{pmatrix}^2. \tag{47}$$

Taking into account (47) we introduce the following vector of order $(n + m + p)$

$$\gamma = \begin{pmatrix} 0 & x & z' \\ x' & 0 & y \\ z & y' & 0 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix}. \tag{48}$$

In addition, define

$$D = \gamma'\gamma = \begin{pmatrix} a \\ b \\ c \end{pmatrix}' A \begin{pmatrix} a \\ b \\ c \end{pmatrix} \tag{49}$$

and let

$$S = D - v \cdot \varphi, \tag{50}$$

where φ is defined in (43). Optimization of (50) yields the CAA solution (46).

If we use in (50) the normalization

$$\varphi = a'a + b'b + c'c - 3 = 0, \tag{51}$$

instead of (43), we obtain the eigenvector problem

$$A \begin{pmatrix} a \\ b \\ c \end{pmatrix} = v \begin{pmatrix} a \\ b \\ c \end{pmatrix} \tag{52}$$

which is the Principal Component (PC) solution for the maximization of the variance in (49).

3c. Cyclic matrices: a generalization

Consider the initial matrix Q in Fig. 1 and its three-facet matrices x_{ij}, y_{jk} and z_{ki} in (30). Using the interpretations of these matrices we can write several equations that project each one of the vectors a, b and c on the other two. Namely,

$$xb = \lambda_1 a, \tag{53a}$$

$$za = \lambda_2 c, \tag{53b}$$

$$yc = \lambda_3 b. \tag{53c}$$

The dual equations of (53) are

$$x'a = \lambda_1 b, \tag{54a}$$

$$z'c = \lambda_2 a, \tag{54b}$$

$$y'b = \lambda_3 c. \tag{54c}$$

The first set of equations in (53a) and (54a) are identical to equations (7) for the matrix x, and the other equations express analogous relations for y and z. Substituting in succession one of equations (53) into another we obtain

$$xyza = \lambda a, \tag{55a}$$

$$zxyc = \lambda c, \tag{55b}$$

$$yzxb = \lambda b, \tag{55c}$$

where the constant λ is equal to a product of the constants in (53). By similar substitutions we obtain from (54)

$$z'y'x'a = \lambda a, \tag{56a}$$

$$y'x'z'c = \lambda c, \tag{56b}$$

$$x'z'y'b = \lambda b. \tag{56c}$$

Combining equations (55) and (56) we obtain three separated eigenvector problems for the vectors a, b and c,

$$(xyz + z'y'x')a = \mu a, \tag{57a}$$

$$(yzx + x'z'y')b = \mu b, \tag{57b}$$

$$(zxy + y'x'z')c = \mu c. \tag{57c}$$

Each of problems (57) is defined for the sum of a matrix and its transpose, i.e. for a symmetric matrix.

Problems (57) point to the simplest way to determine each of the vectors a, b and c independently, through a matrix of order n, m or p (57). Combining (57a)–(57c) yields an eigenvector problem with block-diagonal matrix; that is,

$$\begin{pmatrix} xyz + z'y'x' & 0 & 0 \\ 0 & yzx + x'z'y' & 0 \\ 0 & 0 & zxy + y'x'z' \end{pmatrix} \cdot \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \mu \begin{pmatrix} a \\ b \\ c \end{pmatrix}. \tag{58}$$

It is possible to describe the methods in Subsections 3a and 3b above on the basis of equations (53), which can be written as an eigenvector problem as follows,

$$P \cdot \begin{pmatrix} a \\ b \\ c \end{pmatrix} = g \begin{pmatrix} a \\ b \\ c \end{pmatrix} \tag{59}$$

with

$$P \equiv \begin{pmatrix} 0 & x & 0 \\ 0 & 0 & y \\ z & 0 & 0 \end{pmatrix}. \tag{60}$$

P is a cyclic (block-permutation) matrix of order $(n + m + p)$. System (54) is given by

$$P' \cdot \begin{pmatrix} a \\ b \\ c \end{pmatrix} = g \begin{pmatrix} a \\ b \\ c \end{pmatrix}, \tag{61}$$

i.e. it is an eigenvector problem with the transposed matrix (60). The sum of (59) and (61) also yields an eigenvector problem with the matrix

$$P + P' = \begin{pmatrix} 0 & x & z' \\ x' & 0 & y \\ z & y' & 0 \end{pmatrix}, \tag{62}$$

that is, the LS solution (33) to problem (32).

The constructions

$$P^2 + (P')^2 = \begin{pmatrix} 0 & 0 & xy \\ yz & 0 & 0 \\ 0 & zx & 0 \end{pmatrix} + \begin{pmatrix} 0 & z'y' & 0 \\ 0 & 0 & x'z' \\ y'x' & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & z'y' & xy \\ yz & 0 & x'z' \\ y'x' & zx & 0 \end{pmatrix} \tag{63}$$

and

$$PP' + P'P = \begin{pmatrix} xx' & 0 & 0 \\ 0 & yy' & 0 \\ 0 & 0 & zz' \end{pmatrix} + \begin{pmatrix} z'z & 0 & 0 \\ 0 & x'x & 0 \\ 0 & 0 & y'y \end{pmatrix} = \begin{pmatrix} xx' + z'z & 0 & 0 \\ 0 & yy' + x'x & 0 \\ 0 & 0 & zz' + y'y \end{pmatrix} \tag{64}$$

yield the matrices on the left and right-hand side of the CCA solution in (45).

From the relation

$$(P + P')^2 = [P^2 + (P')^2] + [PP' + P'P] \tag{65}$$

which is obtained by adding the left-hand side matrices in (63) and (64), we get A in (47), which is the relevant matrix for the PC analysis (see (52)). Finally, problems (55) are eigenvector problems with the matrix

$$P^3 = \begin{pmatrix} xyz & 0 & 0 \\ 0 & yzx & 0 \\ 0 & 0 & zxy \end{pmatrix} \tag{66}$$

Problem (56) corresponds to the transposed matrix $(P')^3$, and problem (57) is connected with the matrix $P^3 + (P')^3$, that is, the matrix in (58).

Note that the non-negative matrix (60) is an irreducible cyclic matrix of imprimitivity index $k = 3$; which means that for every eigenvalue λ (including $\lambda_{\max} = r(P)$—the so-called spectral radius of a non-negative matrix) there exist eigenvalues $\lambda_l = \lambda \cdot \exp(i2\pi l/3)$, $l = 0, 1, 2$. It is well known that the matrix P^3 with an imprimitivity index $k = 3$ is block-diagonal with equal eigenvalues for each block (for example, their spectral radius $r^k(P) = \lambda^3_{\max}$). Clearly, $(P')^3$ and P^3 have identical eigenvalues.

Thus, the specific features of cyclic matrices (see, for example, [22], pp. 209–210, [1, 20, 23]) suggest a systematic construction and simplification of different variants of solutions for many-facet matrices such as principal components, canonical correlations, plane approximation and their generalizations.

4. THE LOGARITHMIC METHOD

If the elements of the data matrix Q_{ijk} are positive numbers, we can apply the logarithmic method (LM) to determine the vectors a, b, c. This type of estimation is in [16], pp. 39, 513–514, [22, 2].

Let us apply the LS minimization to the relative deviations, δ_{ijk}, as follows

$$\frac{Q_{ijk}}{\lambda a_i b_j c_k} = 1 + \delta_{ijk}. \tag{67}$$

Thus, we set the loss function

$$F = \sum_{ijk} [\ln(1 + \delta_{ijk})]^2 = \sum_{ijk} (q_{ijk} - \mu - \alpha_i - \beta_j - \gamma_k)^2 \to \min, \tag{68}$$

with the notation

$$q_{ijk} = \ln Q_{ijk}, \qquad \mu = \ln \lambda, \qquad \alpha_i = \ln a_i, \qquad \beta_j = \ln b_j, \qquad \gamma_k = \ln c_k. \tag{69}$$

The first-order conditions of (68), subject to the normalizations

$$\sum_{i=1}^{n} \alpha_i = \sum_{j=1}^{m} \beta_j = \sum_{k=1}^{p} \gamma_k = 0$$

yield

$$\alpha_i = \bar{q}_i - \bar{q} \equiv \frac{1}{mp} \sum_{jk} q_{ijk} - \bar{q}, \tag{70a}$$

$$\beta_j = \bar{q}_j - \bar{q} \equiv \frac{1}{np} \sum_{ik} q_{ijk} - \bar{q}, \tag{70b}$$

$$\gamma_k = \bar{q}_k - \bar{q} \equiv \frac{1}{\mathrm{nm}} \sum_{ij} q_{ijk} - \bar{q}, \tag{70c}$$

$$\mu = \bar{q} \equiv \frac{1}{nmp} \sum_{ijk} q_{ijk}, \tag{70d}$$

where $\bar{q}_t$ represents the average over other indexes (i.e., the mean value for each given t).

Similar solution can be obtained for all many-way matrices. For example, for a 4-way matrix we have

$$\alpha_i = \bar{q}_i - \bar{q} \equiv \frac{1}{mpl} \sum_{j=1}^{m} \sum_{k=1}^{p} \sum_{t=1}^{l} q_{ijkt} - \bar{q}, \tag{71a}$$

and

$$\mu = \bar{q} \equiv \frac{1}{nmpl} \sum_{ijkt} q_{ijkt}. \tag{71b}$$

The expressions for the other vectors are self-evident.

From (69) we obtain the formulas for the initial vectors,

$$a_i = \frac{1}{\lambda}\left(\prod_{j=1}^{m} \prod_{k=1}^{p} Q_{ijk}\right)^{1/(mp)}, \qquad i = 1, \ldots, n, \tag{72a}$$

$$b_j = \frac{1}{\lambda}\left(\prod_{i=1}^{n} \prod_{k=1}^{p} Q_{ijk}\right)^{1/(np)}. \qquad j = 1, \ldots, m, \tag{72b}$$

$$c_k = \frac{1}{\lambda}\left(\prod_{i=1}^{n} \prod_{j=1}^{m} Q_{ijk}\right)^{1/(nm)}, \qquad k = 1, \ldots, p, \tag{72c}$$

with the normalizing term

$$\lambda = \left(\prod_{i=1}^{n} \prod_{j=1}^{m} \sum_{k=1}^{l} Q_{ijk}\right)^{1/(nmp)}. \tag{72d}$$

Hence, each component of the vectors a, b and c is a geometric mean of the elements in the corresponding layer of the matrix Q, and λ is the geometric mean of all the matrix elements.

5. AN APPLICATION: THE QUALITY OF A UNIVERSITY

We tested the methods presented in Sections 2–4 by the following example on the measurement of the quality of a university. Quality of a university is a major factor in the demand for higher education. Most studies use the average Scholastic Aptitude Test (SAT) score of students as a measure of the university's quality (see [19]). Here we suggest that the demand for a university is dependent on the perceived quality of the university by the potential applicants (see [26]). The basic assumption in this paper is that the perceived level of overall quality of a university is a weighted average of n attributes. A group of $p = 101$ individuals (university applicants) was asked to evaluate $n = 5$ attributes of $m = 5$ universities in Israel. In addition, they were asked to evaluate the "overall quality" of each university, and the weight (in shares) that each of the n attributes should get in order to obtain their measure of overall quality of a university. The evaluations were made on a scale of 1 (bad, very low) to 7 (very good, very high). Let Q_{ijk} (an integer ranging from 1 to 7) be the evaluation of attribute i for university j by the kth individual.

The universities are Tel Aviv (UT), Ben Gurion (UB), Haifa (UH), College of Management (UM) and Bar Ilan (UI). The attributes are:

- AG—the distance from the university to the individual's projected location of residence during the forthcoming academic year.
- AA—academic level.
- AT—attitude of the faculty and administration of the university towards the students.
- AP—the possibility to study towards a higher degree in the area chosen for the B.A. studies.
- AS—the availability of courses and programs in the university.

The data for individual k ($k = 1, \ldots, p$), can be represented by the matrix shown in Fig. 2, where Q_{jk} stands for the (quoted) overall measure of the quality of university j by individual k, and W_{ik} is the (quoted) weight that individual k gives to attribute i in forming his measure of overall quality for a university.

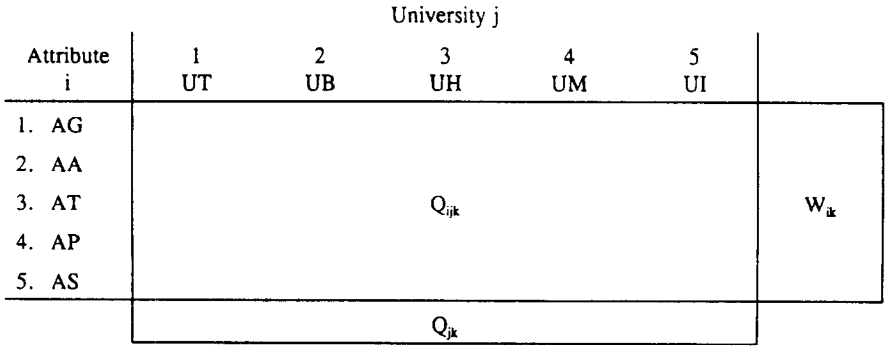

Fig. 2.

Table 1. Weights of attributes

	ATTRIBUTES				
Method of estimation	AG	AA	AT	AP	AS
$\bar{W}$	0.14(5)	0.32(1)	0.19(3)	0.16(4)	0.19(2)
LM	0.17(5)	0.22(1)	0.22(2)	0.18(4)	0.21(3)
PC	0.18(5)	0.22(1)	0.20(3)	0.19(4)	0.21(2)
PPP	0.19(5)	0.21(1)	0.21(3)	0.19(4)	0.21(2)
NLS	0.19(5)	0.21(1)	0.20(3)	0.19(4)	0.21(2)

Note: Ranks across all attributes are given in parentheses.

Table 2. Qualities of universities

	UNIVERSITIES				
Method of estimation	UT	UB	UH	UM	UI
	Tel Aviv	Ben Gurion	Haifa	College	Bar-Ilan
$\bar{Q}$	5.94(1)	4.81(3)	4.56(4)	3.95(5)	5.31(2)
LM	5.89(1)	4.32(4)	4.45(3)	4.30(5)	5.60(2)
PC	5.87(1)	4.47(3)	4.42(4)	4.27(5)	5.23(2)
PPP	5.70(1)	4.59(3)	4.47(4)	4.47(5)	5.33(2)
NLS	5.72(1)	4.57(3)	4.49(4)	4.40(5)	5.38(2)

Note: Ranks across all universities are given in parentheses.

It is apparent that average (across individuals) measures of quality for each university (the vector a) can be obtained in several ways; that is, by using Q_{jk}, by weighting the observations Q_{ijk} with the quoted weights W_{ik}, or by using the Q_{ijk} alone. Similarly, one can estimate the weights of the attributes (the vector b) in forming the overall quality measure of the jth university by using W_{ik}, regressing Q_{ik} on Q_{ijk}, or by using Q_{ijk} alone (which is the most common case since Q_{jk} and W_{ik} are, usually, unavailable in actual applications).

Tables 1 and 2 present the estimates of a and b for the methods described in the preceding sections together with the weighted averages of the weights ($\bar{W}$) and qualities ($\bar{Q}$) that were reported by the individuals, where,

$$\bar{W}_i = \sum_j \sum_k Q_{ijk} \cdot Q_{jk}$$
$$\bar{Q}_j = \sum_i \sum_k Q_{ijk} \cdot W_{ik} \tag{73}$$

and

$$\bar{W} \equiv (\bar{W}_1, \ldots, \bar{W}_n)', \qquad \bar{Q} \equiv (\bar{Q}_1, \ldots, \bar{Q}_m)'.$$

That is, the W_{ik}'s are the weights used to construct $\bar{Q}$, and the Q_{jk}'s are the weights used to construct $\bar{W}$. The values of $\bar{Q}$ and $\bar{W}$ can be viewed as first step approximations to the vectors b and a, respectively (see equations (6) and (15)). NLS stands for the general nonlinear programming method, see equation (12). The numbers in parentheses give the relative rank of the attribute or university, respectively. The vector c (which includes 101 elements), which may be interpreted as the classification of students according to their attitude to university quality was also computed, but is not reported here.

The similarity of the results within the eigenvector methods, as well as the LM, is striking. With one exception, these methods yield results that are similar to the weights of the attributes and university qualities as quoted by the individuals in the sample. The exception is the high value that academic achievement received by the individuals as compared to the implied weights/qualities that are calculated from individual data (Q_{ijk}'s only).

6. SUMMARY

This paper presents several methods that approximate a many-way matrix by its outer vector product. The nonlinear least squares (NLS) approximation of this matrix can be easily used for a small matrix. However, for a complex system, which is described by many directions (facets of the matrix) and/or many elements in each direction, the NLS method becomes infeasible even when evaluated by very fast computers. Thus, some approximation to the NLS, as those in Sections 2–4, must be used.

Here we show the relations among various approximation methods and apply them to a medium-size problem in which we estimate 111 parameters. The outcome of the real example in this paper is encouraging in the sense that the results of the approximation methods are very close to the more general NLS method.

Despite the proliferation of multi-mode problems in economics, political science, management, and marketing research (see [4, 5, 10, 24]), multi-mode methods are used only infrequently in these areas because they are complicated to calculate and interpret. We hope that the simple methods described in this paper will be helpful in solving those problems that require evaluation of multi-mode data, and will be useful in the promotion of multivariate statistical analysis for practical purposes.

REFERENCES

1. A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences.* Academic Press, New York (1979).
2. Y. M. Bishop, S. E. Feinberg and P. W. Holland, *Discrete Multivariate Analysis: Theory and Practice.* MIT Press, Cambridge, Mass. (1975).
3. J. D. Carroll and J. J. Chang, Analysis of individual differences in multi-dimensional scaling via an N-way generalization of "Eckart–Young" decomposition. *Psychometrika* **35** (No. 3), 283–319 (1970).
4. J. D. Carroll and S. Pruzansky, The CANDECOMP-CANDELINC family of models and methods for multidimensional data analysis. In: H. G. Law *et al.* (Eds.), *Research Methods for Multimode Data Analysis.* Praeger, New York (1984).
5. E. C. Carterette (Ed.), *Handbook of Perception,* Vol. 2. Academic Press, New York (1974).
6. R. B. Cattell, "Parallel proportional profiles" and other principles for determining the choice of factors by rotation. *Psychometrika* **9**, 267–283 (1944).
7. R. B. Cattell, The scientific use of factor analysis in behavioral and life sciences. Plenum Press (1978).
8. R. B. Cattell and A. K. S. Cattell, Factor rotation for proportional profiles: analytical solution and an example. *Br. J. statist. Psych.* **8**, 83–92 (1955).
9. C. Eckart and G. Young, The approximation of one matrix by another of lower rank. *Psychometrika* **1**, 211–218 (1936).
10. P. E. Green, F. J. Carmone and D. P. Wachspress, Consumer segmentation via latent class analysis. *J. consumer Res.* **3**, 170–174 (1976).
11. H. H. Harman, *Modern Factor Analysis.* University of Chicago Press (1967).
12. D. H. Krantz, R. C. Atkinson, R. D. Luce and P. Suppes (Eds), *Contemporary Developments in Mathematical Psychology,* Vol. 2. W. H. Freeman, San Francisco (1974).
13. M. G. Kendall and A. Stuart, *The Advanced Theory of Statistics,* Vol. 2: *Statistical Inference and Statistical Relationship.* Hafner (1974).
14. M. G. Kendall and A. Stuart, *The Advanced Theory of Statistics,* Vol. 3: *Designated Analysis, and Time Series.* Hafner (1973).
15. J. Lastovicka, The extension of component analysis to four-mode matrices. *Psychometrika* **46** (No. 1), 47–57 (1981).
16. H. G. Law, C. W. Snyder, J. A. Hattie and R. P. McDonald (Eds), *Research Methods for Multimode Data Analysis.* Praeger, New York (1984).

17. S. Lipovetsky, *Canonical Correlations, Redundancy Analysis, Multiple Regression, and Principal Components of the Intercorrelations for Two Groups of Features. Industrial Laboratory* (Trans. from Russian). Plenum Press, New York, pp. 1005–1013 (1985).
18. S. Lipovetsky, *Component Analysis of Tables with Many Inputs. Industrial Laboratory* (Trans. from Russian). Plenum Press, New York, pp. 474–481 (1982).
19. C. F. Manski and D. A. Wise, *College Choice in America.* Harvard University Press, Cambridge, Mass. (1983).
20. M. Marcus and H. Minc, *A Survey of Matrix Theory and Matrix Inequalities.* Prindle, Weber & Schmidt, Boston, Mass. (1964).
21. J. R. Nesselroade and R. B. Cattell (Eds), *Handbook of Multivariate Experimental Psychology.* Plenum Press, New York (1988).
22. T. L. Saaty, *The Analytic Hierarchy Process.* McGraw–Hill, New York (1980).
23. E. Seneta, *Non-Negative Matrices.* George Allen & Unwin, London (1973).
24. J. N. Sheth (Ed.), *Multivariate Methods for Market and Survey Research.* American Marketing Association, Chicago (1977).
25. P. Slater (Ed.), *The Measurement of Intrapersonal Space by Grid Technique: Dimensions of Intrapersonal Space*, Vol. 2. Wiley, London (1977).
26. A. Tishler, The demand for colleges in Israel: preliminary results. Faculty of Management, Tel Aviv University (1992).
27. L. R. Tucker, *Implications of Factor Analysis of Three-way Matrices for Measurements of Change.* In: C. W. Harris (Ed.) *Problems in Measuring Change.* University of Wisconsin Press, Madison, Wis. (1963).
28. L. R. Tucker, The extension of factor analysis to three-dimensional matrices. In: N. Frederiksen and H. Gulliksen (Eds), *Contributions to Mathematical Psychology.* Holt, Rinehart & Winston, New York (1964).
29. L. R. Tucker, Some mathematical notes on three-mode factor analysis. *Psychometrika* **31** (No. 3), 279–311 (1966).
30. J. H. Wilkinson, *The Algebraic Eigenvalue Problem.* Clarendon Press, Oxford (1965).