Journal of Economic Psychology 8 (1987) 357-384
North-Holland

A NEW STOCHASTIC MULTIDIMENSIONAL UNFOLDING MODEL FOR THE INVESTIGATION OF PAIRED COMPARISON CONSUMER PREFERENCE/CHOICE DATA

Wayne S. DeSARBO *
University of Pennsylvania, Philadelphia, USA

Geert De SOETE **
University of Ghent, Belgium

Jehoshua ELIASHBERG ***
University of Pennsylvania, Philadelphia, USA

Received December 17, 1985; accepted August 4, 1986

This paper presents the development of a new stochastic multidimensional (scaling) unfolding (Coombs 1964) methodology which operates on paired comparison consumer preference or choice data and renders a spatial representation of both consumers and the products or brands they choose. Consumers are represented as ideal points and products as points in a T-dimensional space, where the Euclidean distance between the product points and the consumer ideal points provide information as to the utility of such products to these consumers. The econometric and psychometric literature concerning related models which also operate on such paired comparisons data is reviewed, and a technical description of the new methodology is provided. To illustrate the versatility of the model, a small application measuring consumer preference for several actual brands of over-the-counter analgesics, utilizing one of the optional reparametrized models, is described. Finally, future areas of further research are identified.

* Wayne S. DeSarbo is the Howard and Judy Berkowitz Associate Professor of Marketing, Marketing Department, The Wharton School, University of Pennsylvania, Philadelphia, PA 19104.
** Geert De Soete is Bevoegdverklaard Navorser, Nationaal Fonds voor Wetenschappelijk Onderzoek, University of Ghent, Belgium.
*** Jehoshua Eliashberg is Associate Professor of Marketing, Marketing Department, The Wharton School, University of Pennsylvania, Philadelphia, PA 19104.
Requests for reprints should be sent to G. De Soete, Rijksuniversiteit Gent, Departement Psychologie, Henri Dunantlaan 2, 9000 Gent, Belgium.

Introduction

The method of paired comparisons typically involves the presentation of pairs of stimuli to one or more subjects. The basic experimental unit is the comparison of two stimuli (e.g., products or brands), A and B, by a single subject (e.g., consumer) who, in the simplest case, must choose between them (David 1963). This method was introduced by Fechner (1860) with considerable extensions made popular by Thurstone (1927). Since this paper is concerned with understanding consumer choice behavior, we will be using the terminology of consumers (for subjects) and products/brands (for stimuli). The method of paired comparisons is gainfully utilized in situations where the products to be compared can only be judged subjectively; that is, when it is impossible to make relevant continuous measurement in order to decide which of two products is preferable and by how much. The most frequent applications have been taste testing, consumer tests, color comparisons, personnel ratings, and choice behavior (David 1963). For J products and I consumers, the total number of paired comparisons will be $I\binom{J}{2}$, although a number of incomplete designs are also available (cf. Bock and Jones 1968; Box et al. 1978) for reducing the total number of pairwise judgments under simplifying sets of assumptions. Note that when J is large, the task of making consistent pairwise judgments becomes quite difficult. Oftentimes, intransitivities or circular triads occur in such data where, for example, A may be preferred to B, and B preferred to C, but the same consumer claims to prefer C to A. Therefore, probabilistic models are needed to analyze such paired comparisons data.

The econometric literature on stochastic binary choice models appears to be amenable to an analysis of such paired comparisons data. For a single comparison between product j and k for consumer i, the binary decision by this i-th consumer is represented by a random variable δ_{ijk} that takes on the value 1 if j is selected over k, and zero otherwise. Let p_{ijk} represent the probability that $\delta_{ijk} = 1$. This general class of models assumes that the utility derived from a choice is based on the attributes of that choice (product), the consumer's socio-economic characteristics, and a random error component. Let U_{ij} and U_{ik} denote the respective utilities of products j and k to consumer i, $\boldsymbol{h}_j$ and $\boldsymbol{h}_k$ vectors of characteristics (attributes or features) of the products, and $\boldsymbol{y}_i$ a vector of socio-economic characteristics of the i-th consumer. Then:

$$U_{ij} = \bar{U}_{ij} + e_{ij} = \boldsymbol{h}_j'\boldsymbol{\gamma} + \boldsymbol{y}_i'\boldsymbol{\alpha} + e_{ij}, \tag{1}$$

$$U_{ik} = \bar{U}_{ik} + e_{ik} = \boldsymbol{h}_k'\boldsymbol{\gamma} + \boldsymbol{y}_i'\boldsymbol{\alpha} + e_{ik}, \tag{2}$$

and $\delta_{ijk} = 1$ if the momentary value of $U_{ij} >$ the momentary value of U_{ik}, while $\delta_{ijk} = 0$ if the momentary value of $U_{ik} >$ the momentary value of U_{ij}. Note, $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$ are model parameters which weight respectively the subject and brand characteristics, while e_{ij} and e_{ik} are error terms. Consequently,

$$\begin{aligned} P(\delta_{ijk} = 1) &= P(U_{ij} > U_{ik}) \\ &= P\left[(e_{ij} - e_{ik}) < (\boldsymbol{h}_k - \boldsymbol{h}_j)'\boldsymbol{\gamma}\right] \\ &= F\left((\boldsymbol{h}_k - \boldsymbol{h}_j)'\boldsymbol{\gamma}\right), \end{aligned} \tag{3}$$

where F is the cumulative distribution function of $(e_{ij} - e_{ik})$. According to Judge et al. (1985), the specific type of binary choice model depends on the choice of F. For example, if

$$F\left((\boldsymbol{h}_k - \boldsymbol{h}_j)'\boldsymbol{\gamma}\right) = (\boldsymbol{h}_k - \boldsymbol{h}_j)'\boldsymbol{\gamma},$$

then the linear probability model is defined, and a generalized least-squares procedure suggested by Goldberger (1964) and Zellner and Lee (1965) to correct for problems of heteroscedasticity can be utilized. If one defines

$$F\left((\boldsymbol{h}_k - \boldsymbol{h}_j)'\boldsymbol{\gamma}\right) = \int_{-\infty}^{(\boldsymbol{h}_k - \boldsymbol{h}_j)'\boldsymbol{\gamma}} \frac{1}{\sqrt{2\Pi}} \mathrm{e}^{-t^2/2}\, \mathrm{d}t,$$

then the probit model results. Similarly, if

$$F\left((\boldsymbol{h}_k - \boldsymbol{h}_j)'\boldsymbol{\gamma}\right) = \frac{1}{1 + \mathrm{e}^{-(\boldsymbol{h}_k - \boldsymbol{h}_j)'\boldsymbol{\gamma}}},$$

one has a logit model. Judge et al. (1985) describe the maximum likelihood estimation procedures for obtaining $\boldsymbol{\gamma}$ in the probit and logit

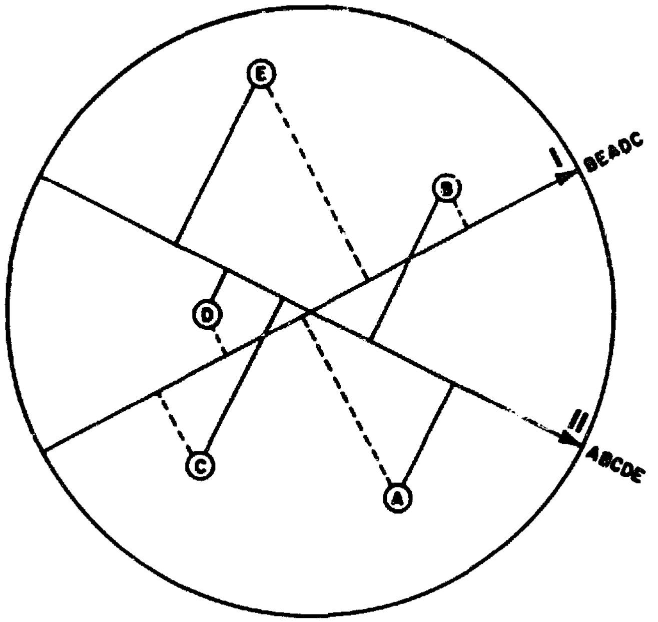

Fig. 1. Two-dimensional illustration of the vector model. (Taken from Carroll and DeSarbo (1985).)

models. Thus, these analyses derive an estimate of γ which denotes the impact of the various attributes or features of the products on overall preference and utility.

The psychometric literature attempts to utilize spatial models to display the structure in such paired comparisons data in representing consumers and products. There have been a number of *unidimensional* scaling procedures proposed to obtain scale values for products from such (aggregated) paired comparisons data (Torgerson 1958; Bock and Jones 1968; Thurstone 1927). More recently, multidimensional scaling models have been devised to account for the multidimensional nature of the products. Here, two general classes of models have been typically utilized to represent such preference/choice data: vector and unfolding models. A vector or scalar products multidimensional scaling model (Tucker 1960; Slater 1960) represents consumers as vectors and products as points in a T-dimensional space. Fig. 1 represents a hypothetical two-dimensional portrayal of such a representation where there are two consumers (represented by two vectors I and II) and five products (represented by the letters A–E). Here, preference order or utility for a given consumer is assumed to be given by the projection of products onto the vector representing that consumer. For example, for consumer I, product B has the highest utility, then E, then A, then D, and finally

C. For consumer II, the order of utility (from highest to lowest) is A, B, C, D, and E. The goal of the analysis here is to estimate the 'optimal' vector directions and product coordinates in a prescribed dimensionality. An intuitively unattractive property of the vector model is that it assumes preference or utility to change monotonically with all dimensions. That is, it assumes that if a certain amount of a thing is good, even more must be even better. (The iso-utility contours therefore are parallel straight lines perpendicular to a consumer's vector.) According to Carroll (1980), this is not an accurate representation for most quantities or attributes in the real world (perhaps with the exception of money, happiness, and health).

There has been some work concerning analyzing such paired comparisons via such vector or scalar products models. Bechtel et al. (1971) have developed a scalar products model for examining *graded* paired comparisons responses (i.e., where customers indicate which of two products are preferred and to what extent). They impose restrictions on sums and variances, and constraints on various parameters to insure uniqueness of the solution (e.g., orthogonality). Zinnes and Wolff (1977) have developed a probabilistic multidimensional Thurstonian model for spatially representing the structure in different–same judgments using a threshold perspective. Cooper and Nakanishi (1983) have developed two logit models (vector and ideal point) for the *external* analysis of paired comparisons data. Recently, Carroll (1980) suggested the wandering vector model for the analysis of paired comparisons data. According to this vector model, it is assumed that each consumer can be represented by a vector and that individual consumers will prefer that brand from a pair having the largest projection on that vector. The direction cosines of this vector specify the relative weights the consumer attaches to the underlying dimensions. The wandering vector model assumes that a consumer's vector wanders or fluctuates from a central vector in such a way that the distribution of the vector termini is multivariate normal. De Soete and Carroll (1983) develop a maximum likelihood method for fitting this model and propose various extensions of the original Carroll model to accommodate additional sources of error and graded paired comparisons. DeSarbo et al. (1986) have extended the work of Carroll (1980) in developing a probabilistic vector model not requiring replications in estimating a vector for each consumer, and having user options for including linear restrictions on configuration coordinates.

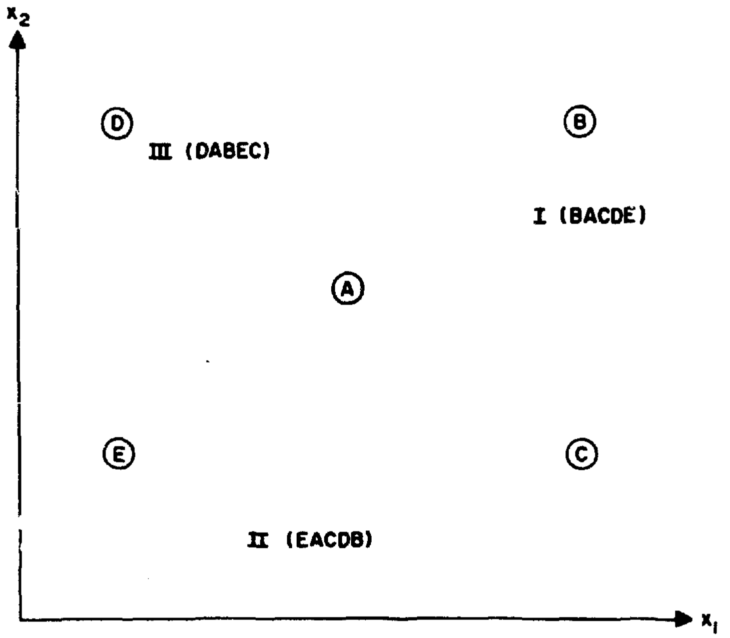

Fig. 2. Two-dimensional illustration of the simple ideal point model. (Taken from Carroll and DeSarbo (1985).)

The other major type of psychometric model to represent such preference/choice data is the unfolding model (Coombs 1964). We will discuss only the simple unfolding model of Coombs (1964) (cf. Carroll (1980) and Carroll and DeSarbo (1985) for a discussion of the simple, weighted, and general unfolding model, and the work by DeSarbo and Rao (1984) on GENFOLD2 – an unfolding methodology which accommodates the estimation of all three types of unfolding models). In the simple unfolding model, both consumers and products are represented as points in a T-dimensional space. The points for consumers represent ideal products, or optimal sets of dimension values. As generalized to the multidimensional case by Bennett and Hays (1960), the farther a given product point is from a consumer's ideal point, the less utility that consumer has for that product. This notion of relative distance implies a Euclidean metric on the space which implies that, in $T = 2$ dimensions, iso-utility contours are families of concentric circles centered at a consumer's ideal point. Carroll (1980) demonstrates that the vector model is a special case of this unfolding model where the ideal points go off to infinity. Fig. 2 illustrates a hypothetical two-dimensional space from an unfolding perspective. Here there are three consumers represented by ideal points labelled I, II, and III, and five products labelled A–E. The figure specifies the preference/utility order for each consumer as a function of distance away from the respective ideal point. The objective in unfolding analyses is to estimate the 'optimal' set of ideal points and product coordinates in a prescribed dimensionality.

Although several unidimensional stochastic unfolding models have been proposed in the literature (Bechtel 1968, 1976; Coombs et al. 1961; Sixtl 1973; Zinnes and Griggs 1974), only three *multidimensional* probabilistic unfolding models have been developed to accommodate paired comparisons data. (One option in GENFOLD2 (DeSarbo and Rao 1984, 1986), which typically operates on two-way, two-mode dominance or profile data, is to collapse the paired comparison matrix for each consumer into a vector of marginals, and then analyze the resulting two-way, two-mode matrix of (dominance) integer counts. Unfortunately, there is typically a substantial loss of information involved in collapsing such individual paired comparison matrices to perform this analysis.) The first one by Schönemann and Wang (1972; Wang et al. 1975) is based on the well-known Bradley-Terry-Luce model and consequently assumes strong stochastic transitivity. In the multidimensional unfolding model proposed by Zinnes and Griggs (1974), it is assumed that the coordinates of both the consumer and the product points are independently normally distributed with a common variance. (Note that recently this assumption has been relaxed in Zinnes and MacKay's (1983) approach to probabilistic multidimensional scaling.) Zinnes and Griggs assume that for each element of the products pair a consumer independently samples a point from his or her ideal point distribution. The Zinnes–Griggs model is defined by

$$p_{ijk} = \text{Prob}\left\{F''\left(T, T, d_{ij}^2, d_{ik}^2\right) \leq 1\right\}, \tag{4}$$

where $F''(\nu_1, \nu_2, \lambda_1, \lambda_2)$ denotes the doubly noncentral F distribution with degrees of freedom ν_1 and ν_2 and noncentrality parameter λ_1 and λ_2, and d_{ij}^2 (respectively d_{ik}^2) the Euclidean distance between the mean point of consumer i and the mean point of product j (respectively k). More recently, De Soete et al. (1986) have proposed the wandering ideal point model for the analysis of such paired comparisons data as an extension of the wandering vector model. According to this model, it is assumed that each consumer can be represented by an ideal point and that they will prefer that product from a pair which has the smallest Euclidean distance from that ideal point (which can be arbitrarily set at the origin of the space). This model assumes that a consumer's ideal point wanders or fluctuates from a central ideal point in such a way that the distribution of the ideal point coordinates is multivariate normal. De Soete et al. (1986) develop a maximum likelihood method for fitting this model and show that it is the only existing probabilistic multidimensional unfolding model requiring only moderate stochastic transitivity.

Unfortunately, the De Soete et al. (1986) model requires replications of paired comparison matrices per consumer to estimate more than one ideal point. This turns out to be a rather difficult task in terms of data collection. Without such replications, only one centroid ideal point can be estimated for a sample of I consumers. Assuming considerable heterogeneity in the sample, the single centroid ideal point may be estimated with considerable high variances. In addition, no provision is currently available to explore individual differences (with replications) as a function of specified consumer differences (such as demographic characteristics), or have similar reparametrizations on products (vis-à-vis attributes or features).

We propose an alternative probabilistic MDS unfolding model which also operates on paired comparisons. Our model can estimate separate consumer ideal points without requiring within-consumer replications. A variety of possible model specifications are provided where ideal points and/or product coordinates can be reparametrized as a function of specified background variables which aids in the understanding of consumer choice behavior. We describe the model structure and the various program options. Next, a small illustration is provided whereby consumer preferences for several brands of over-the-counter analgesics are investigated utilizing one of the reparametrized models. Finally, further research avenues are discussed.

Methodology

Research objectives

As stated, the objective of this paper is to develop a new MDS methodology which operates on paired comparison judgments so that consumers and products can be displayed spatially, thus permitting inferences concerning the nature of the consumer behavior under investigation. In doing so, two sub-objectives will be addressed. The first concerns the ability to investigate the nature of individual (consumer) differences on preference/choice and its measurement, while the second involves modeling the effect of specific product features on the measurement of preference/choice. The technical aspects of the models are described in the next section where consumer preference is the latent construct of interest, since this will be the nature of the application to be presented later. The discussion section will suggest further potential applications to the investigation of still other latent constructs.

The model

Let:

i = $1 \ldots I$ consumers,
j, k = $1 \ldots J$ brands,
t = $1 \ldots T$ dimensions,
l = $1 \ldots L$ brand features
n = $1 \ldots N$ consumer variables,
X_{ij} = the j-th brand presented to the i-th consumer,

$$\delta_{ijk} = \begin{cases} 1 & \text{if consumer } i \text{ finds } X_{ij} \text{ more satisfying than } X_{ik}, \\ 0 & \text{else,} \end{cases}$$

H_{jl} = the l-th feature/attribute value for the j-th brand,
Y_{in} = the nth variable value for the i-th consumer,
a_{it} = the t-th coordinate for consumer i,
b_{jt} = the t-th coordinate for brand j,
α_{nt} = the impact coefficient of the n-th consumer variable on the t-th dimension,
γ_{lt} = the impact coefficient of the l-th brand variable on the t-th dimension.

Now, define a latent consumer 'dispreference' (inversely related to preference) or disutility construct:

$$V_{ij} = U_{ij}^* + e_{ij}, \tag{5}$$

where:
V_{ij} = the (latent) dispreference/disutility of brand j to consumer i,
$U_{ij}^* = \sum_{t=1}^{T}(a_{it} - b_{jt})^2$,
e_{ij} = error.

We assume that:

$$e_{ij} \sim N(0, \sigma_i^2),$$

(where σ_i^2 is the variance parameter for the i-th consumer),

$$\text{Cov}(e_{ij}, e_{ik}) = 0, \quad \forall i,\ j \neq k,$$

$$\text{Cov}(e_{ij}, e_{i'k}) = 0, \quad \forall i \neq i',\ j,\ k. \tag{6}$$

Suppose that the consumer i is presented two objects (e.g., products or brands) j and k and is asked to select the one that is 'more preferred'. Then:

$$P(\delta_{ijk} = 1) = P(V_{ij} < V_{ik})$$

$$= P\left[e_{ij} - e_{ik} < 2\sum_{t=1}^{T} a_{it}(b_{jt} - b_{kt}) + \sum_{t=1}^{T}(b_{kt}^2 - b_{jt}^2)\right]$$

$$= \Phi\left(\frac{2\sum_{t=1}^{T} a_{it}(b_{jt} - b_{kt}) + \sum_{t=1}^{T}(b_{kt}^2 - b_{jt}^2)}{\sqrt{2\sigma_i^2}}\right), \tag{7}$$

where Φ denotes the standard normal distribution function. Similarly,

$$P(\delta_{ijk} = 0) = 1 - \Phi\left(\frac{2\sum_{t=1}^{T} a_{it}(b_{jt} - b_{kt}) + \sum_{t=1}^{T}(b_{kt}^2 - b_{jt}^2)}{\sqrt{2\sigma_i^2}}\right). \tag{8}$$

The general form of the likelihood function, assuming independence over subscripts i, j, and k, is:

$$L = \prod_{i=1}^{I} \left[\prod_{j<k}^{J} \prod \Phi(\cdot)^{\delta_{ijk}} \prod_{j<k}^{J} \prod (1 - \Phi(\cdot))^{1-\delta_{ijk}} \right], \tag{9}$$

where:

$$\Phi(\cdot) = \Phi \left(\frac{2 \sum_{t=1}^{T} a_{it}(b_{jt} - b_{kt}) + \sum_{t=1}^{T} (b_{kt}^2 - b_{jt}^2)}{\sqrt{2\sigma_i^2}} \right). \tag{10}$$

Taking logs, we obtain the log likelihood function:

$$\ln L = \sum_{i=1}^{I} \sum_{j<k}^{J} \sum \delta_{ijk} \ln \Phi(\cdot) + \sum_{i=1}^{I} \sum_{j<k}^{J} \sum \left((1 - \delta_{ijk}) \ln(1 - \Phi(\cdot)) \right). \tag{11}$$

We use a maximum likelihood procedure to estimate $\boldsymbol{A} = ((a_{it}))$, $\boldsymbol{B} = ((b_{jt}))$, and $\boldsymbol{\sigma} = (\sigma_i)$, given $\boldsymbol{\Delta} = ((\delta_{ijk}))$ and T. The conjugate gradient algorithm used for estimation is described in the appendix.

Unlike De Soete et al.'s (1986) original formulation of the wandering ideal point model, the model proposed here posits no *explicit* distribution on the consumer's ideal points. Rather, it is a type of random utility model (Thurstone 1927) where the latent construct being modeled (e.g., choice, preference, satisfaction, etc.) is specified following McFadden (1976), with the introduction of an error term. Both wandering vector and ideal point models, again, require replications of paired comparisons data for each consumer in order to estimate more than a single vector/ideal point, since the consumer vector/ideal point is modeled as being explicitly normally distributed. However, both the wandering vector and ideal point models have the advantage of implying only moderate stochastic transitivity, whereas the present model, like Thurstone's (1927) Law of Comparative Judgment Case V and the three econometric models previously discussed, implies strong stochastic transitivity.

Program options

The probabilistic ideal point model developed here can accommodate a number of different model specifications and options. One can perform either an internal analysis (where the user estimates both consumer ideal points and brand coordinates) or an external analysis (where the user can fix one or more sets of coordinates throughout the analysis). The user can also select among a number of methods of generating starting estimates, including a user-defined option. Also, since the unfolding model is invariant to orthogonal transformations, options are provided to rotate either A or B to principal axes for possible enhancement of subsequent interpretation. One can also estimate individual σ_i's, or constrain all σ_i to be identical to each other. It can be shown that this common variance parameter can, without loss of generality, be set to a positive constant.

Perhaps the most valuable program option concerns the possibility of reparametrizing consumer ideal point coordinates and/or brand coordinates as functions of prespecified background features or attributes. That is, one may reparametrize consumer ideal points via:

$$a_{it} = \sum_{n=1}^{N} Y_{in}\alpha_{nt}, \tag{12}$$

and/or brand coordinates via:

$$b_{jt} = \sum_{l=1}^{L} H_{jl}\gamma_{lt}. \tag{13}$$

As in *CANDELINC* (Carroll et al. 1979), Three-Way Multivariate Conjoint Analysis (DeSarbo et al. 1982), and *GENFOLD2* (DeSarbo and Rao 1984), one can use these reparametrization options to examine what impact such features/attributes have on the derived solution. This can often aid in interpreting the resulting solution, and render insight into the motivations of consumers in the choice process.

As mentioned, these reparametrizations can aid in the interpretation of the derived dimensions (cf. Bentler and Weeks 1978; Bloxom 1978; Noma and Johnson 1977; De Leeuw and Heiser 1980; Heiser 1981; Heiser and Meulman 1983), and can replace the post-analysis property-fitting method often used to attempt to interpret results. In addition, as shall be discussed, the imposition of these sets of reparametrizations can provide an effective tool for understanding the nature of the particular latent construct under investigation (e.g., preference or choice).

It should be noted that when a linear function replaces a brand or consumer coordinate, the number of background variables in the linear function cannot exceed the number of entities that exist for those variables. For example, if J brands have L attributes, $J \geq L$ since one can only identify at most JT coordinates (excluding rotational indeterminacy). Similarly, if I consumers have N background variables, $I \geq N$ since one can only identify at most IT coordinates (excluding rotational indeterminacy). Thus, in most applications, such reparametrizations actually improve the degrees of freedom of the model by reducing the number of parameters to be estimated.

Degrees of freedom

One typically collects $I(J(J-1)/2)$ independent paired comparison responses in an application. Defining the degrees of freedom of the model as the effective number of free model parameters, one can calculate the degrees of freedom for the various models accommodated by this methodology. These are shown in table 1, where it is assumed one is interested in estimating σ_i, $\forall i$. Note that in all models an

Table 1
Model degrees of freedom calculations for the various models.

Model	No. of free model parameters
$U_{ij}^* = \sum_{t=1}^{T} (a_{it} - b_{jt})^2$	$T(I+J)+(I-1)-\frac{T(T-1)}{2}-T$
$U_{ij}^* = \sum_{t=1}^{T} \left[a_{it} - \left(\sum_{l=1}^{L} H_{jl}\gamma_{lt} \right) \right]^2$	$T(I+L)+(I-1)-\frac{T(T-1)}{2}$
$U_{ij}^* = \sum_{t=1}^{T} \left[\left(\sum_{n=1}^{N} Y_{in}\alpha_{nt} \right) - b_{jt} \right]^2$	$T(N+J)+(I-1)-\frac{T(T-1)}{2}$
$U_{ij}^* = \sum_{t=1}^{T} \left[\left(\sum_{n=1}^{N} Y_{in}\alpha_{nt} \right) - \left(\sum_{l=1}^{L} H_{jl}\gamma_{lt} \right) \right]^2$	$T(N+L)+(I-1)-\frac{T(T-1)}{2}$

adjustment of $T(T-1)/2$ is required due to the well-known invariance of such unfolding models to orthogonal transformations (cf. Carroll (1980) for a complete discussion of this and associated problems of interpretation). Also, in models where b_{jt} is not reparametrized, one can add a constant vector $\boldsymbol{c}$ to all brand vectors $\boldsymbol{b}_j$ and not affect the choice probabilities in eq. (7). This, necessitates a subtraction of T from the degrees of freedom. Finally, in estimating σ_i, one can set the overall scale by fixing one $\sigma_i = 1$.

An illustration

Study design

A convenience sample of $I = 7$ members of the administration staff of the University of Pennsylvania were asked to take part in a small study designed to measure preferences for various brands of existing over-the-counter (OTC) analgesic pain relievers. This was to be a pretest for a much larger study to be conducted later. We purposely restrict our attention to small sample so that we can report the specific details of the results. These respondents were initially questioned as to the brand(s) they currently use (as well as frequency of use) and their personal motivations for why they chose such a brand(s) (e.g., ingredients, price, availability, etc.). They were then presented eight existing OTC analgesic brands:

Brands	Plotting code	Brands	Plotting code
1. Anacin	A	5. Datril	E
2. Bayer	B	6. Excedrin	F
3. Bufferin	C	7. Tylenol	G
4. Cope	D	8. Vanquish	H

Initially, they were presented colored photographs of each brand and its packaging, together with price per 100 tablets, ingredients, package claims, and manufacturer (cf. DeSarbo and Carroll 1985). Table 2 presents summaries of selected portions of the descriptions for each of the brands. Each subject/consumer was requested to read this information and return to it at any time during the experiment if he/she so

Table 2
Brand descriptions. [a]

Brand	Price	mg. of Aspirin	mg. of Acetaminophen	mg. of buffered ingredients	mg. of caffeine
Anacin	3.47	400	0	0	32
Bayer	3.41	325	0	0	0
Bufferin	4.70	324	0	16	0
Cope	6.33	421	0	75	32
Datril	2.65	0	325	0	0
Excedrin	4.92	250	250	0	65
Tylenol	4.53	0	325	0	0
Vanquish	7.67	227	194	75	0
Plot codes	V	W	X	Y	Z

[a] Based on a store audit conducted in New Jersey in 1983.

wished. After a period of time, they were asked to make paired comparison preference judgments for all possible 28 pairs of brands. They were told that they had to choose one from each pair (i.e., no ties were allowed). The presented pairs were randomized for each subject. Table 3 presents the raw paired comparisons data collected for each of the seven consumers (Δ). A simple way to examine such paired comparisons data in a condensed fashion is to compute 'dominance counts' for each brand by consumer. These counts are merely integers which measure the net amount of times a brand is chosen over another brand. Positive counts indicate that a consumer typically has selected this brand more times over other brands than vice versa. Negative counts, then, indicate that a consumer has typically selected other brands over this one than vice versa. Table 4 presents the two-way, two-mode matrix of dominance counts by consumer and brand. As Table 4 shows, consumer 1 prefers Datril and Tylenol; consumers 2, 3 and 5 prefer Bayer and Bufferin; consumer 4 prefers Datril; consumer 6 prefers Bufferin and Excedrin: and consumer 7 prefers Bayer, Bufferin, and Tylenol. These judgments were quite consistent with their responses to actual usage and motivations for use asked previous to the collection of Δ. Here, consumer 1 uses Datril and Tylenol exclusively since he is allergic to aspirin; consumer 2 prefers plain aspirin (e.g., Bayer) and typically buys the brand on scale; consumer 3 prefers Bufferin and Ibuprophen brands (not tested here) since she tends to experience stomach discomfort with plain aspirin, and thus prefers

Table 3
Raw paired comparisons data for $I = 7$ subjects.

		A	B	C	D	E	F	G	H
Subject 1	A		0	0	1	0	1	0	0
	B			0	1	0	1	0	0
	C				1	0	1	0	1
	D					0	1	0	0
	E						1	1	1
	F							0	0
	G								1
	H								
Subject 2	A		0	0	1	1	0	1	1
	B			1	1	1	1	1	1
	C				1	1	1	1	1
	D					0	0	0	0
	E						0	1	1
	F							1	1
	G								1
	H								
Subject 3	A		0	0	1	0	1	1	1
	B			0	1	1	1	1	1
	C				1	1	1	1	1
	D					0	0	0	0
	E						1	1	1
	F							0	1
	G								1
	H								
Subject 4	A		0	0	1	0	0	0	1
	B			1	1	0	0	0	1
	C				1	0	0	0	1
	D					0	0	0	1
	E						1	1	1
	F							0	1
	G								0
	H								
Subject 5	A		0	0	1	1	0	1	1
	B			0	1	1	1	1	1
	C				1	1	1	1	1
	D					0	0	0	0
	E						0	1	1
	F							1	1
	G								1
	H								

Table 3 (continued)

		A	B	C	D	E	F	G	H
Subject 6	A		0	0	0	0	0	0	0
	B			0	0	0	0	0	0
	C				1	1	0	1	1
	D					0	0	1	0
	E						0	0	0
	F							0	1
	G								1
	H								
Subject 7	A		0	0	1	0	0	0	1
	B			1	1	1	1	0	1
	C				1	0	1	1	1
	D					1	0	0	1
	E						1	1	1
	F							0	1
	G								1
	H								

brands which safeguard against this; she also strongly dislikes brands with caffeine; consumer 4 uses Datril and Tylenol since they have no aspirin which causes her stomach upset; consumer 5 is another aspirin user (Bufferin), but can be persuaded to switch among brands of aspirin by coupons and promotions; consumer 6 uses Bufferin, Tylenol, or Vanquish since she also worries about aspirin's side effect causing stomach upset; and, consumer 7 claimed he uses any *major* brand of OTC analgesic – any that are advertised heavily since he believes these are more effective.

Table 4
Dominance counts.

	Brand							
	A	B	C	D	E	F	G	H
Subject								
1	−3	−1	3	−5	7	−7	5	1
2	1	7	5	−7	−1	3	−3	−5
3	1	5	7	−7	3	−3	−1	−5
4	−3	1	−1	−5	7	3	3	−5
5	1	5	7	−7	−1	3	−3	−5
6	−7	−5	5	−1	−1	5	3	1
7	−3	5	3	−3	3	−1	3	−7

Analysis

We conducted an analysis of Δ (in table 3) in $T = 1$, 2 and 3 dimensions with $\sigma_i = 1$, $\forall i$, and the reparametrization option via $\boldsymbol{B} = \boldsymbol{H}\gamma$, where $\boldsymbol{H}$ is defined via the five feature variables (standardized to zero mean and unit variance) presented in table 2. This specification was preferred since $\boldsymbol{H}$ contains features that these consumers said were important in their choice of a specific OTC analgesic brand. All consumers were encouraged to read this information contained on the color photographs of the brand and packaging prior to the paired comparison task. Table 5 presents the statistical results of these three analyses. Assuming the asymptotic test is appropriate, the $T = 2$ dimensional solution is identified as the 'best' solution. Note, even if this statistic were inappropriate, the values of the proportion of correct predictions by dimension also confirms this decision, given only a 2% increase in moving from the $T = 2$ to $T = 3$ dimensional solution.

Fig. 3 presents the joint space representation for the $T = 2$ dimensional solution. Consumer ideal points are represented by the integers 1–7, brand points by the letters A–H, and brand features as vectors, given the nature of eq. (13), by letters V–Z. The figure clearly depicts the structure in the data. Consumers 1 and 4 prefer Datril and Tylenol, consumers 2, 3 and 5 prefer Bayer; consumer 6 is located near Tylenol, and Vanquish; consumer 7 is near the centroid of major brands including Anacin, Bayer, Tylenol, Datril, Bufferin, and Excedrin. This is consistent with the information (presented earlier) collected in the initial part of the study concerning brand usage and motivations.

The vectors of $\boldsymbol{H}$ in fig. 3 also render insight into dimensional interpretation. Under restrictive assumptions concerning normalization and orthogonality, one can show that the cosine of the angle a vector

Table 5
Statistical results for analgesics data.

T	Model degrees of freedom	ln L	Deviance	Proportion of correct predictions	Difference in deviance
1	12	−101.47	202.94	0.73	–
2	23	−83.81	167.62	0.85	35.32 [a]
3	33	−78.98	157.96	0.87	9.66 (NS)

[a] $p \leq 0.001$.

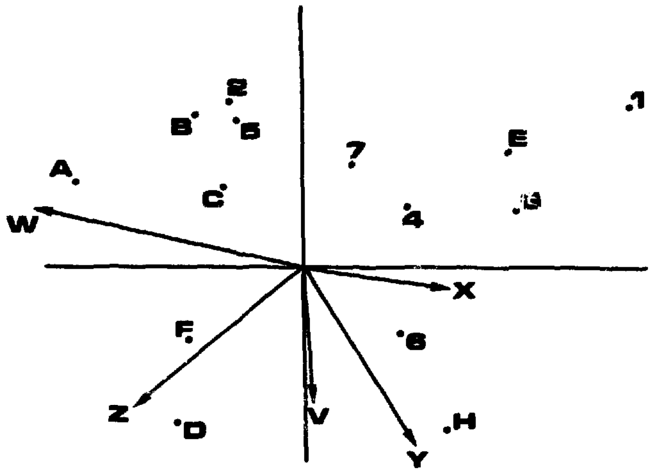

Fig. 3. Two-dimensional solution for consumer preference data.

makes with a dimension is related to the correlation of that vector with that dimension. Table 6 presents these correlations between the two dimensions of ***B*** and the five columns of ***H***. The first dimension is clearly an Aspirin vs. Acetaminophen (Aspirin substitute) dimension. Here Datril and Tylenol load positively (high Acetaminophen), while Cope, Anacin, and Bayer load negatively (high Aspirin). The second dimension is dominated by price and to some extent buffered ingredients, where brands like Vanquish are high on both, while brands like Bayer are low on both. These interpretations make sense given the scatter of brands and ideal points. The interrelationships between the seven consumers and eight brands is very consistent with the preliminary information collected before the paired comparisons data.

Another useful set of statistics to investigate concerns the correlations between dimensions for ***A***, ***B*** and γ. These will, of course, vary

Table 6
Correlations between ***B*** and ***H***.

	Dimension	
	I	II
Price	0.073	−0.909
Aspirin	−0.908	−0.248
Acetaminophen	0.811	−0.052
Buffered ing.	0.042	−0.836
Caffeine	−0.590	−0.403

Table 7
Proportion of correct predictions by subject.

Subject	Proportion
1	0.93
2	0.93
3	0.93
4	0.68
5	0.93
6	0.79
7	0.79
Total	0.85

according to the particular orthogonal rotation utilized to interpret the results. They are also indicators of dimensionality since extracting dimensions which are highly intercorrelated is not deemed as parsimonious in rendering independent views of the structure in $\boldsymbol{\Delta}$. Here, $\text{Cor}(\boldsymbol{A}_1, \boldsymbol{A}_2) = -0.192$, $\text{Cor}(\boldsymbol{B}_1, \boldsymbol{B}_2) = 0.005$, and $\text{Cor}(\gamma_1, \gamma_2) = 0.425$

Table 8
Additional output.

	Brand							
	A	B	C	D	E	F	G	H
Latent disutility scores								
Subject								
1	17.307	10.350	9.632	19.239	0.993	15.001	1.576	9.947
2	1.764	0.070	0.557	8.082	4.442	0.388	5.384	10.685
3	5.472	1.704	3.187	14.998	4.194	9.804	6.041	15.516
4	5.960	3.024	1.832	6.561	0.811	3.959	0.676	3.890
5	1.710	0.092	0.366	7.316	4.107	3.832	4.890	9.681
6	7.481	5.904	3.309	3.460	3.123	2.498	1.862	0.827
7	4.184	1.520	0.947	6.879	1.351	3.818	1.630	5.806
Selection probabilities by subject and brand								
Subject								
1	0.027	0.046	0.049	0.025	0.475	0.031	0.299	0.047
2	0.032	0.817	0.102	0.007	0.013	0.013	0.011	0.005
3	0.106	0.341	0.182	0.039	0.139	0.059	0.096	0.037
4	0.038	0.075	0.124	0.035	0.279	0.057	0.335	0.058
5	0.039	0.718	0.181	0.009	0.016	0.017	0.014	0.007
6	0.040	0.050	0.090	0.086	0.095	0.119	0.160	0.360
7	0.062	0.169	0.272	0.037	0.191	0.067	0.158	0.044

indicating that the two dimensions appear to be explaining quite different aspects of the structure in Δ.

Table 7 presents the proportion of correct predictions (a matching coefficient) by consumer. Here, all consumers, with the possible exception of consumer 4, are fit well by the model.

Note that if one were to assume that the reciprocals of U_{ij}^* were measures of 'utility', then one could convert these $s_{ij} = 1/U_{ij}^*$ scores into choice probabilities of selection via the Luce and Suppes (1965) formula:

$$q(i, j) = \frac{s_{ij}}{\sum_{j=1}^{J} s_{ij}}, \tag{14}$$

where $q(i, j)$ is the probability consumer i selects product j. Table 8 presents the latent disutility scores and selection probabilities computed for this small sample of $I = 7$ and $J = 8$. Again, these have face validity given the preliminary information collected for each of these seven consumers. According to the selection probabilities, consumer 1 is most likely to buy Datril; consumers 2, 3 and 5, Bayer; consumer 4, Tylenol; consumer 6, Vanquish; and, consumer 7, Bufferin.

GENFOLD2 results

In order to investigate the validity of the solution, the two-way, two-mode matrix of integer dominance counts in table 4 was run with GENFOLD2 (DeSarbo and Rao 1984, 1986), a metric, two-way unfolding methodology which also allows for the reparametrization option in eq. (13). A two-dimensional solution was obtained which accounted for 93.87% of the variance. Fig. 4 presents the resulting two-dimensional solution utilizing the same plot codes. In order to compare figs. 4 and 3, a configuration matching procedure was utilized to compare the two solutions. Again, given the rotational indeterminacies associated with such unfolding solutions, canonical correlation was utilized as an 'approximate' configuration matching methodology. Table 9 presents the canonical correlation measures for these various configurations and the joint space

$$R = \begin{bmatrix} A \\ B \end{bmatrix}.$$

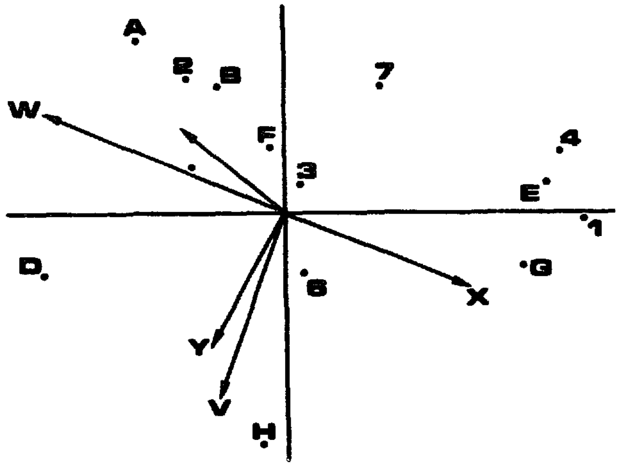

Fig. 4. Two-dimensional GENFOLD2 solution.

Table 9
Configuration matching coefficients.

A	B	γ	$R=\begin{bmatrix} A \\ B \end{bmatrix}$
$\lambda_1=0.908$ $\lambda_2=0.263$	$\lambda_1=0.996$ $\lambda_2=0.857$	$\lambda_1=0.957$ $\lambda_2=0.435$	$\lambda_1=0.953$ $\lambda_2=0.631$

As shown, there is clearly at least one dimension in common between the two sets of solutions. The canonical correlations are lowest for the ideal points. This makes sense given the fact that collapsing the paired comparison judgments into integer dominance counts loses substantial information by consumer including which specific brands dominate which other brands. The canonical correlations for the brands shows two congruent dimensions that match. This is not as strong as for $\boldsymbol{R}$ and γ, yet there is some congruence. Thus, another unfolding methodology applied to a set of two-way data derived from Δ produced somewhat similar results.

Discussion

Summary and implications for latent construct measurement

A new MDS methodology for the spatial analysis of paired comparisons data has been presented and contrasted to existing econometric and psychometric methodologies in terms of model structure, stochastic assumptions, input requirements, and model specification options. The model, its assumptions, and the variety of different reparametrization options available for various analyses have been described. A small application of the methodology to a measurement problem in consumer preference was described in some detail, where five hypothesized determinants were combined via the $\boldsymbol{H}$ matrix. An analysis was performed were brand coordinates (actual OTC analgesic brands) were directly reparametrized in terms of these five features. The procedure produced two dimensions dominated by aspirin/acetaminophen and price/buffered ingredients respectively, as well as approximations of each consumer's latent preference scale and probability of selection for each of the eight brands.

This methodology should prove equally viable for various other applications where paired comparisons are collected. It can aid in similar measurement problems concerning latent, unobservable constructs such as utility, satisfaction, choice, similarity, risk, intention/attitude, etc. With the various reparametrization options for a_{it} and b_{jt}, additional flexibility is provided for investigating determinants of both individual differences (e.g., demographic information) and stimulus differences.

Such reparametrization options would also be valuable in utilizing this methodology for an external type of preference MDS analysis generally referred to as conjoint analysis. Here, a design matrix is presented defining hypothetical object brands profiles, and a dominance judgment such as preference or intention to buy is asked in paired comparison form. The methodology then derives the contribution of each object design variable to the resulting derived dimensions. This has proven to be of substantial interest to the marketing profession for product design applications (see DeSarbo et al. 1982).

Future research

There are several clear avenues for future investigation concerning this methodology. First, a rigorous Monte Carlo study should be performed whereby the performance of this new methodology is investigated while a number of data, model, and algorithmic factors are experimentally varied. Second, the small properties of the estimators and the χ^2 test should be examined in order to justify the asymptotic

χ^2 test for nested models with incidental parameters and no within-consumer replications. Finally, further work should proceed in order to test this methodology with additional different data sets.

Appendix

The algorithm

Maximum likelihood methods are utilized to estimate the desired set of parameters to maximize ln L (or minimize $-\ln L$) in expression (11). The method of conjugate gradients (Fletcher and Reeves 1964) with automatic restarts is utilized to solve this nonlinear, unconstrained optimization problem. The partial derivatives of ln L in expression (11) with respect to the various parameters are:

$$\frac{\partial \ln L}{\partial a_{it}} = 2\sum\sum_{j<k}^{J} \frac{\delta_{ijk} - \Phi(\cdot)}{\Phi(\cdot)[1-\Phi(\cdot)]} \frac{\phi(\cdot)(b_{jt} - b_{kt})}{\sqrt{2\sigma_i^2}}, \tag{A.1}$$

$$\frac{\partial \ln L}{\partial b_{jt}} = 2\sum_{i=1}^{I} \sum_{k>j}^{J} \frac{\delta_{ijk} - \Phi(\cdot)}{\Phi(\cdot)[1-\Phi(\cdot)]} \frac{\phi(\cdot)(a_{it} - b_{jt})}{\sqrt{2\sigma_j^2}}, \tag{A.2}$$

$$\frac{\partial \ln L}{\partial \gamma_{lt}} = 2\sum_{i=1}^{I} \sum\sum_{j<k}^{J} \frac{\delta_{ijk} - \Phi(\cdot)}{\Phi(\cdot)[1-\Phi(\cdot)]} \frac{\phi(\cdot)(H_{jl} - H_{kl})(a_{it} - b_{kt})}{\sqrt{2\sigma_i^2}}, \tag{A.3}$$

$$\frac{\partial \ln L}{\partial \alpha_{nt}} = \sum_{i=1}^{I} \sum\sum_{j<k}^{J} \frac{\delta_{ijk} - \Phi(\cdot)}{\Phi(\cdot)[1-\Phi(\cdot)]} \frac{\phi(\cdot)Y_{in}(b_{jt} - b_{kt})}{\sqrt{2\sigma_i^2}}, \tag{A.4}$$

$$\frac{\partial \ln L}{\partial \sigma_i} = \sum\sum_{j<k}^{J} \frac{\delta_{ijk} - \Phi(\cdot)}{\Phi(\cdot)[1-\Phi(\cdot)]}\phi(\cdot) \times \frac{\left[2\sum_{t=1}^{T} a_{it}(b_{jt} - b_{kt}) + \sum_{t=1}^{T}\left(b_{kt}^2 - b_{jt}^2\right)\right]}{\sqrt{2\sigma_i^4}}, \tag{A.5}$$

where $\phi(\cdot)$ represents the evaluation of the standard normal density at $(\cdot)$. For sake of convenience, assume that the relevant parameters to be estimated are contained in the vector $\boldsymbol{\Theta}$, and that $\nabla\boldsymbol{\Theta}$ is the vector of partial derivatives for this set of parameters.

The complete conjugate gradient procedure can be summarized as follows:

(i) Start with initial parameter estimates $\boldsymbol{\Theta}^{(1)}$; set the iteration counter $MIT = 1$.

(ii) Set the first search direction $S^{(1)} = \nabla\Theta^{(1)}$.
(iii) Find $\Theta^{(2)}$ according to the relation:

$$\Theta^{(2)} = \Theta^{(1)} + u^{(1)}S^{(1)}, \tag{A.6}$$

where $u^{(1)}$ is the optimal step length in the direction $S^{(1)}$. The optimal step size is found by quadratic interpolation methods. Set $MIT = 2$.
(iv) Calculate $\nabla\Theta^{(MIT)}$ and set

$$S^{(MIT)} = \nabla\Theta^{(MIT)} + \frac{(\nabla\Theta^{(MIT)})'(\nabla\Theta^{(MIT)})}{(\nabla\Theta^{(MIT-1)})'(\nabla\Theta^{(MIT-1)})}S^{(MIT)}. \tag{A.7}$$

(v) Compute the optimal step length $u^{(MIT)}$ in the direction $S^{(MIT)}$, and find

$$\Theta^{(MIT+1)} = \Theta^{(MIT)} + u^{(MIT)}S^{(MIT)}. \tag{A.8}$$

(vi) If $\Theta^{(MIT+1)}$ is optimal, stop. Otherwise set $MIT = MIT + 1$ and go to step (iv) above (i.e., undertake another iteration).

It has been demonstrated empirically that conjugate gradient procedures can avoid the typical 'cycling' often encountered with steepest descent algorithms (cf. Rao 1979). In addition, they demonstrate valuable quadratic termination properties (Himmelblau 1972) – i.e., conjugate gradient procedures will typically find the globally optimum solution for a quadratic function in n steps, where n is the number of parameters to be solved. Note that since $\ln L$ in (11) has an upper bound of zero, and since each estimating stage (or iteration) of the likelihood maximization can be shown to increase $\ln L$, one can use a limiting sums argument (Courant 1965) to prove convergence to at least a locally optimum solution. Several Monte Carlo runs on small synthetic data sets revealed that the procedure does recover (known) configurations.

This conjugate gradient method is particularly useful for optimizing functions of several parameters since it does not require the storage of any matrices (as is necessary in Quasi-Newton and second derivative methods). However, as noted by Powell (1977), the rate of convergence of the algorithm is linear only if the iterative procedure is 'restarted' occasionally. Restarts have been implemented in the algorithm automatically depending upon successive improvement in the objective function.

A number of goodness-of-fit measures are computed for this model:

(1) The ln likelihood function: $\ln L$,
(2) A deviance measure (Nelder and Wedderburn 1972; McCullagh and Nelder 1983):

$$D = -2\left[\sum_{i=1}^{I}\sum_{j<k}^{J}\sum \delta_{ijk}\ln(\hat{p}_{ijk}) + (1-\delta_{ijk})\ln(1-\hat{p}_{ijk})\right]$$

$$= -2\ln L, \tag{A.9}$$

where $\hat{p}_{ijk}$ is the estimated probability that consumer i finds brand j more preferable than brand k as expressed in eq. (7). Note that one can test nested models as the difference between respective deviance measures. This difference is (asymptotically) χ^2 distributed with the difference in model degrees of freedom providing the appropriate χ^2 test degrees of freedom. This test is 'theoretically' appropriate in testing dimensionality as well as the various models described in table 1 because of the nested terms. Recall, this is an asymptotic test, however. One obvious problem with this approach concerns incidental parameters in the likelihood function (i.e., parameters whose order vary according to the order of Δ, such as the a_{it}'s). According to Anderson (1980), maximum likelihood estimators in such cases may not be consistent. (Takane (1983) also noticed this problem with an item response model he created and utilized a marginal likelihood loss function to integrate out such (subject) incidental parameters.) This is particularly relevant in the present case since there are no replications. This is clearly an avenue of needed research for the proposed procedure if the asymptotic test is to be used for hypothesis testing.

(3) The proportion of correct predictions in Δ. Here, the proportion of times the solution correctly predicts δ_{ijk} is calculated for the total sample as well as for each consumer.

References

Anderson, E.B., 1980. Discrete statistical models with social science applications. New York: Elsevier.

Bechtel, G.G., 1968. Folded and unfolded scaling from preferential paired comparisons. Journal of Mathematical Psychology 5, 333–357.

Bechtel, G.G., 1976. Multidimensional preference scaling. The Hague: Mouton.

Bechtel, G.G., L.R. Tucker and W. Chang, 1971. A scalar product model for the multidimensional scaling of choice. Psychometrika 36, 369–388.

Bennett, J.F. and W.L. Hays, 1960. Multidimensional unfolding: determining the dimensionality of ranked preference data. Psychometrika 25, 27–43.

Bentler, P.M. and D.G. Weeks, 1978. Restricted multidimensional scaling models. Journal of Mathematical Psychology 17, 138–151.

Bloxom, B., 1978. Constrained multidimensional scaling in N spaces. Psychometrika 43, 397–408.

Bock, R.D. and L.V. Jones, 1968. The measurement and prediction of judgment and choice. San Francisco, CA: Holden-Day.

Box, B.E.P., W.G. Hunter and J.S. Hunter, 1978. Statistics for experimenters. New York: Wiley.

Carroll, J.D., 1980. 'Models and methods for multidimensional analysis of preferential choice (or other dominance data)'. In: E.D. Langermann and H. Feger (eds.), Similarity and choice. Bern: Huber. pp. 234–289.

Carroll, J.D. and W.S. DeSarbo, 1985. 'Two-way spatial models for modeling individual differences in preference'. In: E.C. Hirschman and M.B. Holbrook (eds.), Advances in consumer research, Vol. 12. Ann Arbor, MI: Association for Consumer Research. pp. 571–576.

Carroll, J.D., S. Pruzansky and J.B. Kruskal, 1979. CANDELINC: a general approach to multidimensional analysis of many-way arrays with linear constraints on parameters. Psychometrika 45, 3–24.

Coombs, C.H., 1964. A theory of data. New York: Wiley.

Coombs, C.H., M. Greenberg and J. Zinnes, 1961. A double law of comparative judgment for the analysis of preferential choice and similarities data. Psychometrika 26, 165–171.

Cooper, L.G. and M. Nakanishi, 1983. Two logit models for external analysis of preference. Psychometrika 48, 607–620.

Courant, R., 1965. Differential and integral calculus, Vol. I (2nd ed.). New York: Wiley.

David, H.A., 1963. The method of paired comparisons. New York: Hafner.

De Leeuw, J. and W. Heiser, 1980. 'Multidimensional scaling with restrictions on the configuration'. In: P.R. Krishnaiah (ed.), Multivariate analysis. New York: Elsevier. pp. 501–522.

DeSarbo, W.S. and J.D. Carroll, 1985. Three-way metric unfolding via alternating weighted least squares. Psychometrika 50, 275–300.

DeSarbo, W.S. and V.R. Rao, 1984. GENFOLD2: a set of models and algorithms for the GENeral UnFOLDing analysis of preference/dominance data. Journal of Classification 1, 147–186.

DeSarbo, W.S. and V.R. Rao, 1986. A constrained unfolding model for product positioning. Marketing Science 5, 1–19.

DeSarbo, W.S., R.L. Oliver and G. De Soete, 1986. A probabilistic multidimensional scaling vector model. Applied Psychological Measurement 10, 79–98.

DeSarbo, W.S., J.D. Carroll, D.R. Lehmann and J. O'Shaughnessy, 1982. Three-way multivariate conjoint analysis. Marketing Science 1, 323–350.

De Soete, G. and J.D. Carroll, 1983. A maximum likelihood method for fitting the wandering vector model. Psychometrika 48, 553–556.

De Soete, G., J.D. Carroll and W.S. DeSarbo, 1986. The wandering ideal point model: a probabilistic multidimensional unfolding model for paired comparisons data. Journal of Mathematical Psychology 30, 28–41.

Fechner, G.T., 1860. Elemente der Psychophysik. Leipzig: Breitkopf und Härtel.

Fletcher, R. and C.M. Reeves, 1964. Function minimization by conjugate gradients. Computer Journal 7, 149–154.

Goldberger, A.S., 1964. Econometric theory. New York: Wiley.

Heiser, W., 1981. Unfolding analysis of proximity data. Leiden: Reprodienst Psychologie RUL.

Heiser, W. and J. Meulman, 1983. Constrained multidimensional scaling including confirmation. Applied Psychological Measurement 7, 381–404.

Himmelblau, D.M., 1972. Applied nonlinear programming. New York: McGraw-Hill.

Judge, G.G., W.E. Griffiths, R.C. Hill, H. Lütkepohl and T. Lee, 1985. The theory and practice of econometrics. New York: Wiley.

Luce, R.D. and P. Suppes, 1965. 'Preference, utility, and subjective probability'. In: R.D. Luce, R.R. Bush and R. Galanter (eds.), Handbook of mathematical psychology, Vol. 3. New York: Wiley. pp. 249–410.

McCullagh, P. and J.A. Nelder, 1983. Generalized linear models. New York: Chapman and Hall.

McFadden, D., 1976. Quantal choice analysis: a survey. Annals of Economic and Social Measurement 5, 363–390.

Nelder, J.A. and R.W.M. Wedderburn, 1972. Generalized linear models. Journal of the Royal Statistical Society A 135, 370–384.

Noma, E. and J. Johnson, 1977. Constraining nonmetric multidimensional scaling configurations (Tech. Rep. 60). Ann Arbor, MI: The University of Michigan, Human Performance Center.

Powell, M.J.D., 1977. Restart procedures for the conjugate gradient method. Mathematical Programming 12, 241–254.

Rao, S.S., 1979. Optimization: theory and applications. New York: Wiley.

Schönemann, P.H. and M.M. Wang, 1972. An individual difference model for the multidimensional analysis of preference data. Psychometrika 37, 275–305.

Sixtl, F., 1973. Probabilistic unfolding. Psychometrika 38, 235–248.

Slater, P., 1960. The analysis of personal preferences. British Journal of Statistical Psychology 13, 119–135.

Takane, Y., 1983. An item response model for the multidimensional analysis of multiple-choice data. Handout for a talk given at the 1983 Psychometric Society Meeting at Los Angeles, CA.

Thurstone, L.L., 1927. A law of comparative judgment. Psychological Review 34, 273–286.

Torgerson, W.S., 1958. Theory and methods of scaling. New York: Wiley.

Tucker, L.R., 1960. 'Intra-individual and inter-individual multidimensionality'. In: H. Gulliksen and S. Messick (eds.), Psychological scaling: theory and applications. New York: Wiley. pp. 155–167.

Wang, M.M., P.H. Schönemann and J.G. Rusk, 1975. A conjugate gradient algorithm for the multidimensional analysis of preference data. Multivariate Behavioral Research 10, 45–99.

Zellner, A. and T.H. Lee, 1965. Joint estimation of relationships involving discrete random variables. Econometrika 33, 382–394.

Zinnes, J.L. and R.A. Griggs, 1974. Probabilistic, multidimensional unfolding analysis. Psychometrika 39, 327–350.

Zinnes, J.L. and D.B. MacKay, 1983. Probabilistic multidimensional scaling: complete and incomplete data. Psychometrika 48, 27–48.

Zinnes, J.L. and R.P. Wolff, 1977. Single and multidimensional same–different judgments. Journal of Mathematical Psychology 16, 30–50.