7 Three-Way and Further Extensions of the Basic Model

There are not three incomprehensibles, nor three uncreated: but one uncreated and one incomprehensible

QUICUNQUE VULT (Creed of St. Athanasius)
Book of Common Prayer, 1662

7.1 Introduction

The remaining programs in the MDS(X) series are either designed for the analysis of three- (and in the case of CANDECOMP, higher-) way data or (as in the case of PREFMAP I and II) are more complex variants of models already encountered in Chapter 6.

The main differentiating characteristic of the programs considered here is the form of the model, or rather models, since both PREFMAP and PINDIS consist of a hierarchy of models of increasing complexity. As in the previous chapter, we shall begin by examining the type of data input to these programs—this provides the best clue to their most fruitful areas of application—and then go on to describe the form of the models employed.

7.1.1 Three- (and higher-) way data

The term three-way data refers to a 'cube' of data (see Figure 7.1, p. 192). Such data occur frequently. The third way usually consists of a set of individuals, occasions, methods, points in time, experimental conditions or geographical locations. Two types of 3-way data are usefully distinguished:

(i) **3-way data which are two-mode,** i.e. consist of a set of ordinary 2-way, one-mode, (dis)similarity matrices, and

(ii) **3-way data which are three-mode,** representing for instance the preferences of a set of subjects (mode 1) for a set of food items (mode 2), where the judgments were made at a number of different occasions (mode 3).

Examples of such 3-way data are:

(a) *Three-way, two-mode data* (A set of pairwise (dis)similarity matrices)

A set of individuals (mode 1) each produce a matrix of pairwise similarity ratings between stimuli (mode 2).

Over a number of weeks (mode 1) the mutual attraction of a set of fraternity members (mode 2) is assessed by averaging the preference scores they give to each other.

A set of individuals rate a set of concepts in terms of a set of semantic differential scales. The ratings between each pair of scales are then correlated. This gives rise to a set of correlation matrices between scales (mode 1), one matrix for each concept (mode 2). (Note that in this case what were originally 3-way, 3-mode data have been reduced by the researcher to 3-way, 2-mode data by aggregating over individuals.)

The frequency with which each pair of plant species (mode 1) co-occurs is tabulated for each of a number of locations (mode 2).

A set of attitude items (mode 1) are rated on a 7-point scale by a set of subjects, and a number of different coefficients of ordinal association (mode 2) are calculated between the items.

A set of five live fish (mode 1) are confronted with different stylised shapes of fish, differing on sexual and other characteristics. Their behaviours are summarised in each case by a matrix of rank correlations representing the similarity of behaviour when presented with stylised fish i as opposed to stylised fish j (mode 2).

(b) *Three-way, three-mode data* (Three distinct sets of entities)

A set of individuals (mode 1) rate a set of automobiles (mode 2) on a set of rating scales (mode 3).

Members of a social group (mode 1) rank each other (mode 2) in terms of emotional closeness. Data are collected on a number of occasions (mode 3).

The input (mode 1)-output (mode 2) matrices between a set of industries is collected for a set of nations (mode 3).

(c) *Higher-way data* (N-way data)

There are examples in the literature of four-way data, e.g. semantic differential experiments on a number of occasions (mode 1), using the same set of individuals (mode 2) to judge the same set of concepts (mode 3) on the same set of rating scales (mode 4). (This is 4-way, 4-mode scaling.)

Each year (mode 1), a (different) set of individuals make pairwise judgments of similarity between a set of names of nations. The investigator wished to distinguish European, North American, Latin American and Third World subjects' judgments, and therefore produced a separate correlation matrix between nations (mode 2) for each sphere of origin (mode 3). This is 4-way, 3-mode data.

In principle, data of any way can be scaled, and the CANDECOMP program accepts up to seven-way data. Users are advised to proceed beyond three-way data with considerable caution. They are in largely uncharted territory.

7.1.2 Organisation of the chapter

The defining characteristics of the MDS(X) programs for analysing three-way and related data are described in Table 7.1. As in previous chapters, characteristics of the *data*, scaling *transformation* and *model* are used to define the programs involved. The models described in this chapter consist mainly of generalised versions of the distance and vector models encountered in Chapter 6. The exact form of the generalisation is specified in Table 7.1 under the headings of dimensional weighting and rotation (in the case of distance models) and vector weighting and translation (in the case of vector models). It will be easier to discuss these increasingly complex transformations in the context of the program(s) where they occur.

Let us first take the programs in the order in which they appear in Table 7.1, an order determined by the type of data they analyse. In the subsequent sections

DATA	MODEL	Specification* T R V W P	TRANSFORM**	MDS (X) PROGRAM	Description
TWO–WAY 2– MODE	DISTANCE	(R W P)	M or L	PREFMAP (I) (PM1)	Dimensional Salience with Idiosyncratic Orientation; Ideal Point.
		(W P)	M or L	PREFMAP (II) (PM2)	Dimensional Salience with Ideal Point.
THREE–WAY 2–MODE (CONFIGURATIONS)	DISTANCE	(W)	L	INDSCAL –S	Dimensional Salience
		((R))	Similarity (S)	PINDIS (PO)	Basic: Procrustes Rotation (General Similarity)
	DISTANCE	(W)	S	PINDIS (P1)	Dimensional Salience
		((R) W)	S	PINDIS (P2)	Dimensional Salience with Idiosyncratic Rotation.
	VECTOR	(V)	S	PINDIS (P3)	Perspective (Fixed Origin)
		(T V)	S	PINDIS (P4)	Perspective with Idiosyncratic Origin
	MIXED	(V W)	S	PINDIS (P5)	Double Weighting
N–WAY N– MODE	VECTOR	(V)	L	CANDECOMP	*N*–way Scalar Products
	Model Specification: Individual: **Translation of origin** **Rotation of axes** **Vector weighting** **Weights (dimensional)** **Point representation of subjects**		***Transform:* **Monotonic** **Linear** **Similarity**		

Table 7.1 *Analysis of 3-way and related data by MDS(X) programs*

of the chapter, by contrast, the order will proceed from the simplest to the more complex models.

We have already encountered PREFMAP as a program for mapping two-way, two-mode data into a user-provided configuration according to the vector and simple distance models (see 6.2.1 and 6.2.4). In this chapter these models are extended to include a weighted distance model (PREFMAP-II) and a rotated and weighted distance model (PREFMAP-I). As before, these are chiefly used to analyse sets of preference or, in general, similarity rankings or ratings when the user wishes to represent both the stimuli and subjects in the same solution.

The most common form of three-way data is two-mode, and the most popular form of analysis is the INDSCAL model. This model interprets differences between the subjects (third-way) as arising from differences in the weights (interpreted as importance or salience) ascribed to the dimensions of a common configuration. Because of its conceptual simplicity it makes a natural starting point for discussing more complex models, and is explained in 7.2.1.

An alternative approach to studying individual differences is to scale each matrix separately as an initial stage and then compare the configurations obtained. The PINDIS hierarchy of models provides a successively complex set of models for comparing configurations. There is no reason why the configurations should be obtained in this way; any set of configurations referring to the same set of objects, however obtained and of whatever dimensionality, may legitimately be input. The PINDIS models are discussed last, in section 7.4.1, due to their greater complexity.

This chapter also deals with three-way, three-mode and higher-way data, which may be analysed using a generalisation of the scalar products or factor models already encountered in the last chapter (e.g. MDPREF). The basic ideas of canonical decomposition, used to implement these models, are discussed in 7.2.2, following the exposition of the INDSCAL model which turns out to be a special case.

7.2 Individual Differences and Dimensional Salience

Three-way, two-mode data appear very frequently in the form of a set of (dis)similarity matrices. A typical example occurs when psychologists have subjects make pair-comparison estimates of the similarity between stimuli (such as colour chips) and wish to examine how individuals differ among themselves in the way they perceive colour. (This is the origin of the acronym: INdividual Differences SCALing.) Sociologists often have correlation matrices between a given set of variables for a number of different survey subgroups, and wish to see how the subgroup matrices differ (see 7.2.1.3). Plant ecologists may have co-occurrence matrices for a number of species, one for each of a number of sites chosen to differ on given criteria, and wish to inspect the differences between the sites.

Each example poses a similar methodological problem of aggregation. If the data for each element of the third-way differ to a substantial degree then there is little communality and it is hard to see how they are to be compared at all. If, by contrast, subjects differ in no systematic way but simply represent minor random fluctuations, then there is no point in making anything of the differences. However, if the data are simply pooled together as a single matrix at the outset then all information about differences—whether systematic or random—is lost.

Drawing on ideas developed by Horan (1969), Carroll and Chang (1970) propose the following way of thinking about such individual differences. Suppose each individual (group, or element in the third-way) makes use of a variety of attributes or dimensions in judging the stimuli (the exposition is easiest in psychological terms, but the model generalises easily to encompass any other sort of third-way element). Then define a master or *group space*, which consists of all the dimensions which the subjects happen to use. Each individual subject's space can now be thought of as a special case of the group space—as a reduction of the group space, since she is using some subset of the total available dimensions. This is termed the subject's 'private space'.

In Horan's original formulation, every individual was simply thought of as either using, or not using, each group space dimension, so each 'private space' could be represented by a sequence of 1s and 0s, indicating whether the subject used (1) or did not use (0) the dimensions of the group space. This 'all or nothing' approach was modified by Carroll and Chang by postulating that each subject attaches a *varying* (positive) weight to each dimension which represents the *degree* of salience (or importance, or attention or relevance or centrality) of that dimension to her judgments. So each individual i can be thought of as having an idiosyncratic set of weights, symbolised by $w_a^{(i)}$: the weight given to dimension a of the group space by individual i. These weights hence represent the way in which the subjects differ in the importance attached to each of the dimensions. An individual who attaches equal importance to each of the dimensions will have a set of weights of the same value, and it is such a subject whom the group space actually represents. Others by contrast will attach different weights to different dimensions of the group space and thus systematically distort the group space into the 'subjective metric' of their own private space.

The INDSCAL model presents a way of interrelating these 'private spaces' and provides one-way of comparing how subjects (or elements of the third-way) differ among themselves, but only, be it noted, by accounting for the individual differences in terms of differing weights being associated with the *same* dimensions: INDSCAL is explicitly a *dimensional* model.

7.2.1 The INDSCAL model*

The Carroll-Chang model is described in full in their definitive 1970 article. A lucid and extended exposition, relating INDSCAL to other forms of three-way scaling is given in Carroll and Wish (1973) and a wide range of applications is discussed in Wish and Carroll (1974). Elementary treatments are given in chapter 4 of Kruskal and Wish (1978), in Spence (1978) and in the MDS(X) documentation. In this section we shall concentrate chiefly on the basic characteristics of the model and upon the interpretation of an INDSCAL-S solution. Further details of the estimation procedure in INDSCAL-S are contained in Appendix A7.2.

Before using INDSCAL-S or embarking upon interpretation of an INDSCAL solution, it is essential to understand clearly the characteristics of the *group space*, the *subject space*, the *private spaces* and their interrelationships.

*Hereafter, INDSCAL refers to the model and INDSCAL-S to the version of the program in MDS(X).

(i) The *group space* (denoted **X**) consists of a configuration of p stimulus points in a user-chosen number of dimensions r. The orientation of the axes of this space are *uniquely determined* in the sense that any change in their orientation destroys the optimality of the INDSCAL solution. The INDSCAL axes are often found to be readily interpretable.

This group space acts as the 'reference configuration' to which all the subjects' private spaces may be referred and from which they may be all derived. The group space need not in fact describe any actual subject, and the configuration should not itself be interpreted if it turns out that it is simply a compromise between the configurations of groups of subjects with very different patterns of individual weights.

(ii) The *private space of each subject* i (denoted $\mathbf{Y}^{(i)}$) is a configuration of the p points in r dimensions. Within each private space, distances between stimuli are straightforwardly Euclidean.*

(iii) The *subject space* (denoted **W**) is simply a useful graphical way of comparing subjects in terms of their sets of dimensional weights. It has the same dimensions as the group space and each subject is represented by a vector located by the value of the weights on each of the dimensions.

These basic ideas are illustrated in Figure 7.1 by reference to a simple artificial 2-dimensional example (see also Carroll 1972, p. 105 et seq., Carroll and Wish 1973, p. 57 et seq., and Kruskal and Wish 1978, p. 61 et seq. for similar examples). In this expository example, there are 3 objects and 16 subjects, so the data would consist of 16 lower triangular matrices between the 3 objects. The overall 2-dimensional group space configuration, **X**, consists of 3 points which make an equilateral triangle (representing equal distance between the objects). The private spaces, $\mathbf{Y}^{(i)}$, for subjects 1 and 2 are also presented. Note that in the private spaces the configuration of points no longer forms an equilateral triangle but rather an isosceles triangle (two sides remain the same length but the third is foreshortened). Clearly, *the distances between stimulus points are different within each private space.* The two private spaces are nonetheless related: they may be derived from the reference group space by a simple process of differentially stretching or shrinking the axes of the group space by the square root of the subject's 'importance weights'. In other words, the co-ordinates in the private space (say, for subject 1) are simply a weighted version of the group space co-ordinates. To obtain subject i's private space, we take the co-ordinates of the p stimulus points on the 1st dimension of the group space (x_{ji}) and rescale (stretch or shrink) them by the square root of subject i's weight for this dimension ($\sqrt{w_a^{(i)}}$): that is,

$$y_{ja}^{(i)} = \sqrt{w_a^{(i)}}\,x_{ja}$$

Then the distance between the stimuli j and k in subject i's private space will be:

$$d_{jk}^{(i)} = \sqrt{\sum_a \left(\sqrt{w_a^{(i)}}\,x_{ja} - \sqrt{w_a^{(i)}}\,x_{ka}\right)^2}$$

*In INDSCAL the private space of each subject is *estimated* as a distortion of the group space directly from the data. In PINDIS, by contrast, each subject's 2-way data are *first scaled* and then input in the form of configuration co-ordinates into the program.

or, in simplified form (taking the weight outside the squared term):

$$d_{jk}^{(i)} = \sqrt{\sum_a w_a^{(i)} \left(x_{ja} - x_{ka} \right)^2}$$

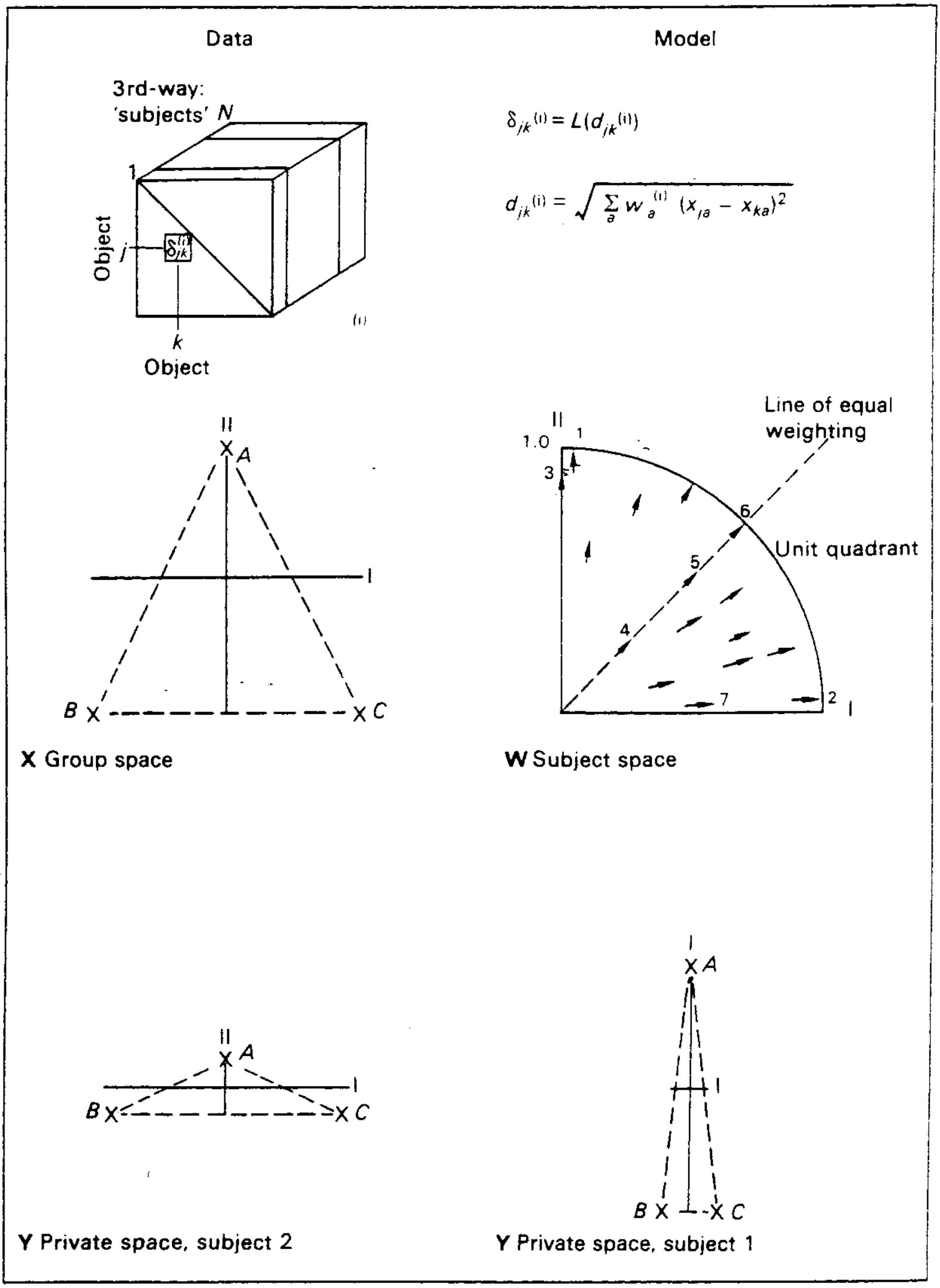


Figure 7.1 *Basic INDSCAL model*

This last equation gives the general form of the INDSCAL *and other weighted distance models:* A 'subject' *i*'s judgment of the (dis)similarity between objects *j* and *k* is taken to be a (linear) function of the overall distance between stimuli *j* and *k* in the group space, after that space has been differentially rescaled (stretched and shrunk) by the subject's set of weights into the 'subjective metric' of the subject concerned.

7.2.1.1 The group stimulus space: its properties and interpretation

The group stimulus space functions as the basic reference configuration from which the private configurations of individual subjects can be derived by differentially shrinking or stretching the dimensions by the (square root) of the corresponding weights.

The INDSCAL dimensions actually represent the (orthogonal) directions where the variation among subjects is the greatest: it is for this reason that they are normally easy to interpret. These dimensions are uniquely identified, in the sense that if the original dimensions are rotated and new subject dimension weights calculated, the resulting solution will explain the subjects' data less well than the original solution.† If it turns out that the extent of individual differences is not great, then such a reduction in explained variance is likely to be small, but in the normal way the reduction is usually fairly substantial. Unless there are compelling reasons of interpretability or little subject variation, the INDSCAL axes should be regarded as fixed.

In most MDS solutions encountered so far, the final configuration is rotated to principal axes—that is, dimensions are chosen which have the statistically convenient property that co-ordinates of the points are not correlated across dimensions. *This is not (generally) true of an* INDSCAL *group space:* the dimensions of greatest subject variation will usually give rise to a configuration where the co-ordinates of the stimulus points are to a greater or lesser extent correlated.‡ The information about the extent of this correlation between pairs of axes is contained in the output from INDSCAL-S in the matrix of scalar products between dimensions ('sums of products') for matrix 2.

The INDSCAL group stimulus space configuration should therefore be interpreted with caution: strictly speaking it represents a subject who weights the dimensions equally, and if a significant number of subjects' weights depart markedly from equality then there is a danger of trying to interpret a configuration which is in no sense representative. That said, methods for external interpretation of INDSCAL *dimensions*—and especially linear property-fitting (see 4.4.1)—are particularly appropriate, since the dimensions are *not* arbitrary and it is important to try to tie down their meaning as accurately as possible. Good examples of the use of

†The unique orientation of axes in the INDSCAL model means that the solution is unique up to permutation of axes, which is equivalent to saying that the only permissible rotation of the dimensions which preserves all significant information is through multiples of 90°. However, the actual size of the configuration is arbitrary, and is therefore normalised so that the variance of the projections on each of the co-ordinate axes is unity and the centroid of the configuration provides the origin (Carroll and Wish 1973, p. 30).

‡An option SOLUTIONS (1) exists in the MDS(X) version to obtain a solution where the axes are as close as possible to being uncorrelated. Such a solution will normally be sup-optimal compared to the ordinary solution.

property fitting to validate or confirm the interpretation of INDSCAL dimensions occur in the classic Carroll and Chang (1970) paper and elsewhere.

INDSCAL-S can also be used in an external mode if the user provides the program with a group stimulus space configuration (which remains fixed in orientation) and the INDSCAL analysis then concentrates entirely upon estimating from the subjects' data the subject weights for this configuration. (External use is achieved in INDSCAL-S using the FIX POINTS (1) option). External analysis of this sort has two main uses: (i) to scale a large number of subjects' data and (ii) to compare a number of different data sets by referring them to a common reference configuration. Thus if the user has, say, 500 matrices for analysis, it is sensible to choose a manageable sample of those matrices and scale them. The resulting group stimulus space can then be fixed, and the subject weights can then be estimated for as many batches of subject matrices as desired.* An example of the second use occurs where a replication has been made of a previous study and the researcher wishes to investigate the extent to which her subjects' data compare to the weights obtained in the earlier study. The original group space configuration is fixed under this option, and the subjects' weights may then be estimated and compared to those of the original study.

7.2.1.2 The subject space: its properties and interpretation

When subjects' data are input to INDSCAL-S they are normalised to have equal weight, which has the effect of giving each subject's data equal influence on the solution. This fact, in conjunction with the normalisation of the group stimulus space described above, gives rise to several nice properties of the subject space which are useful to bear in mind when interpreting an INDSCAL solution:

(i) The subject's weight on a dimension is (approximately) equal to the correlation of the intervals between stimulus co-ordinates on that dimension and the corresponding pairwise dissimilarity values in the subject's data

(ii) Consequently, the *squared* subject's weight on a dimension is (approximately) equal to the proportion of variance in the subject's data that can be accounted for by that dimension (Wish and Carroll 1974, p. 452).

(iii) Therefore, the squared distance from the origin of the subject space to a subject's point in that space is (approximately) equal to the proportion of variance in the subject's data accounted for by the full INDSCAL solution.

If the dimensions of the INDSCAL solution are uncorrelated, then the word 'exactly' replaces the word 'approximately' in the above three sections. Thus in the subject space portrayed in Figure 7.1, subjects 4, 5 and 6 provide an example of subjects who weight the dimensions equally: they differ only in the fraction of their data explained by the model, with the data of subject 6 perfectly accounted for. Similarly, subjects 7 and 2 have the same pattern, giving virtually exclusive salience to dimension I, whilst subject 3 uses only dimension II. Looking at the pattern in terms of goodness of fit, the data of subjects 1, 6 and 2 are totally accounted for, whilst those of subject 4 are very poorly explained.

*See Coxon and Jones 1979, pp. 54–9, and especially T3.17, for an example using a balanced set of 68 matrices to obtain the group space configuration by reference to which 286 subjects' subject weights were estimated.

Note that only *positive* weights are allowed by the INDSCAL model. If, as occasionally happens, a very small negative weight occurs in may be considered as approximation to a zero weight; if it is substantial it can only be interpreted as indicating that the basic model does not hold for the data of the subject concerned.

The significant information in the subject space is contained, then, (1) in the *direction* in which a point is located from the origin, since any points lying on line from the origin have weights in the same ratio, and (2) in the *distance* (of a subject vector) from the origin, representing how well the subject's data are explained by the model.

Before embarking on any systematic analysis of INDSCAL subject weights, it is important to know something of the stability of INDSCAL solutions (see Jones and Waddington 1973; MacCallum 1977).

(i) Simulation studies show that, even in circumstances of high error in the data, recovery of the group space configuration and its dimensional orientation is excellent, but that

(ii) the stability of the subject space is far less stable and much more subject to fluctuation in the presence of error.

The temptation to use cluster analysis on subject weights should be strongly resisted: the separations of subject points are in no sense ordinary distances and their location is far from stable. The question of whether any linear procedures such as ANOVA and its multivariate variants should be used on INDSCAL weights remains contentious. MacCallum (1977) and others often strongly counsel against their use: Carroll and others think that a more lenient approach is called for.

Usually the user will want to compare subjects in terms of the patterns of relative salience given to dimensions. This is best done by concentrating on the angular separation between subject vectors: the smaller the angle of separation, the more similar is the pattern of weights. In the two-dimensional case, it is usually a simple matter to see closely collinear 'sheaves' of subject vectors in the subject space, and such bunching can also be detected visually in three dimensions. Beyond that, statistical analyses of different subject vectors should be used (see Mardia 1972; Coxon and Jones 1979, pp. 128–36 for use in an MDS context). A simple alternative for two-dimensional data is simply to take the ratio of the weights for each subject and, since the distribution of such ratios is usually markedly positively skew, it often makes sense to correct this by taking a logarithmic transformation of the weight ratios.

An alternative to Carroll and Chang's representation of subject weights has been suggested recently by Young (1978). The Young Plot allows the amount of variation explained to be represented independently of the relative salience of the subject weights, and is illustrated in Figure 7.3b (p. 199).

The Young Plot

The Young Plot charts each subject in terms of two things—the relative salience ascribed to one dimension over another (on the horizontal axis) and how well the subject's data are fit by the model (the vertical axis). The first is measured by the ratio of the two-dimensional weights—which can be interpreted trigonometrically as the tangent of the angular separation between a subject's vector and the line of equal weighting in the conventional representation of INDSCAL subject space.* The goodness of fit is given simply by the squared correlation r^2 between the subject's original data and the values predicted by the INDSCAL model. This information is provided separately in an INDSCAL run.

The Young Plot and its construction from a set of subject weights is illustrated in Figure 7.3b and is very simple to read. Subjects located in the centre of the horizontal axis (such as the Labour group of voters in this example) weight the dimensions equally: the more that dimension I dominates over II the further left the subject point is, so the non-voters group has the most dominant weight for dimension I and the Conservative voters group has the most dominant weight for dimension II. The goodness of fit is simply read up the vertical axis. In this example the greatest differentiation is between the 'other parties' group whose data are not well explained (being largely Scottish and Welsh Nationalist party supporters they are presumably dancing to a different piper) and the others.

The most important advantages of the Young Plot are that it gives accurate representation of patterns of dimensional weighting and of goodness of fit independently of dimensional correlation, and concentrates the user's attention onto the angular separation (relative salience) of patterns of subject weights rather than on the proximity of points portrayed somewhat misleadingly in a conventional subject space. The Young Plot can also be modified in various ways—to portray patterns of three-dimensional weights, or to compare relative salience of weights with any other variable of interest (see Coxon and Jones 1980, p. 59 et seq.).

7.2.1.3 An example: political party imagery

Alt et al. (1976) carried out a survey of 2,462 British voters after the 1974 British election. The questionnaire included 20 attitudinal items—political party features (items 1–7), the parties' handling of contentious issues (items 8–10), blame (11–12), taxes and pensions (13–14) and policy positions (15–17). These are reproduced in Table 7.2. Each pair of items was cross-tabulated and the association between them measured by Goodman and Kruskal's gamma, which preserves weak monotonicity of the item categories (see 2.2.2 above). The respondents were divided into five subgroups, viz

A Conservative voters
B Labour voters
C Liberal voters
D Other voters (principally Scottish and Welsh National Parties)
E Non-voters

Each of these subgroups were then treated as a 'pseudo-subject', and gamma coefficients were calculated for each subgroup, hence providing a (5 × 20 × 20) array for input to INDSCAL. The group space configuration is given in Figure 7.2 and

*The tangent of the angle which the subject vector makes with the first dimension (tan θ_1) is defined as the ratio of the weight on dimension II to the weight on dimension I. Tan (θ_1 − 45°) measures this predominance of dimension II over dimension I as a deflection (angular departure) from the line of equal weighting.

Item No.	*Symbol*	*Title*
1	K	Keeps/breaks promises
2	D	Divides/unites country
3	B	Bloody-minded/reasonable
4	G	Good for one/all classes
5	E	Extreme/moderate
6	Ca	Capable/not capable
7	SF	Stands firm/gives way
8	P	Prices
9	M	Miners' strike
10	S	Strikes
11	PB	Blame for prices
12	MB	Blame for miners' strike
13	T	Taxation
1	Pe	Pensions
15	CM	Common Market
16	N	Nationalisation
17	SS	Social services
18	W	Wage controls
19	C	Communists
20	R	Reliability

Table 7.2 *Items in political party imagery study* (Alt et al. 1976) (Reproduced by permission of the journal *Quality and Quantity*)

the subject weight plots are given in Figure 7.3. Alt et al. identify dimension I as 'image consciousness' (by which they mean an ideologically-based concern with both political style and performance) and dimension II as 'policy consciousness' (concerned primarily with welfare and related policy issues). Note from the shape of the group stimulus space configuration that the two dimensions are clearly positively correlated. The authors do not provide this information, but our estimate is $r_{I,II} = 0.23$.

Further interpretation of the group space should wait upon inspection of the subject weights (Figure 7.3). Even a cursory examination of the subject space (a) and more obviously of the Young diagram (b), shows very considerable differences in the goodness of fit and in the relative salience of two-dimensional weights between the subgroups.

But just how significant are these relative differences in weights, given what we know of the relative instability of INDSCAL weights? Alt et al. use an unusual form of internal validation. They divide their subjects into a number of pseudo-groups based upon 'irrelevant' factors (such as male/female) and random criteria (exclusive but randomly constructed subgroups and overlapping random subsamples of subjects) and proceed to estimate weights for each group, keeping the reference configuration fixed. Only if the voter subgroup differences exceed the

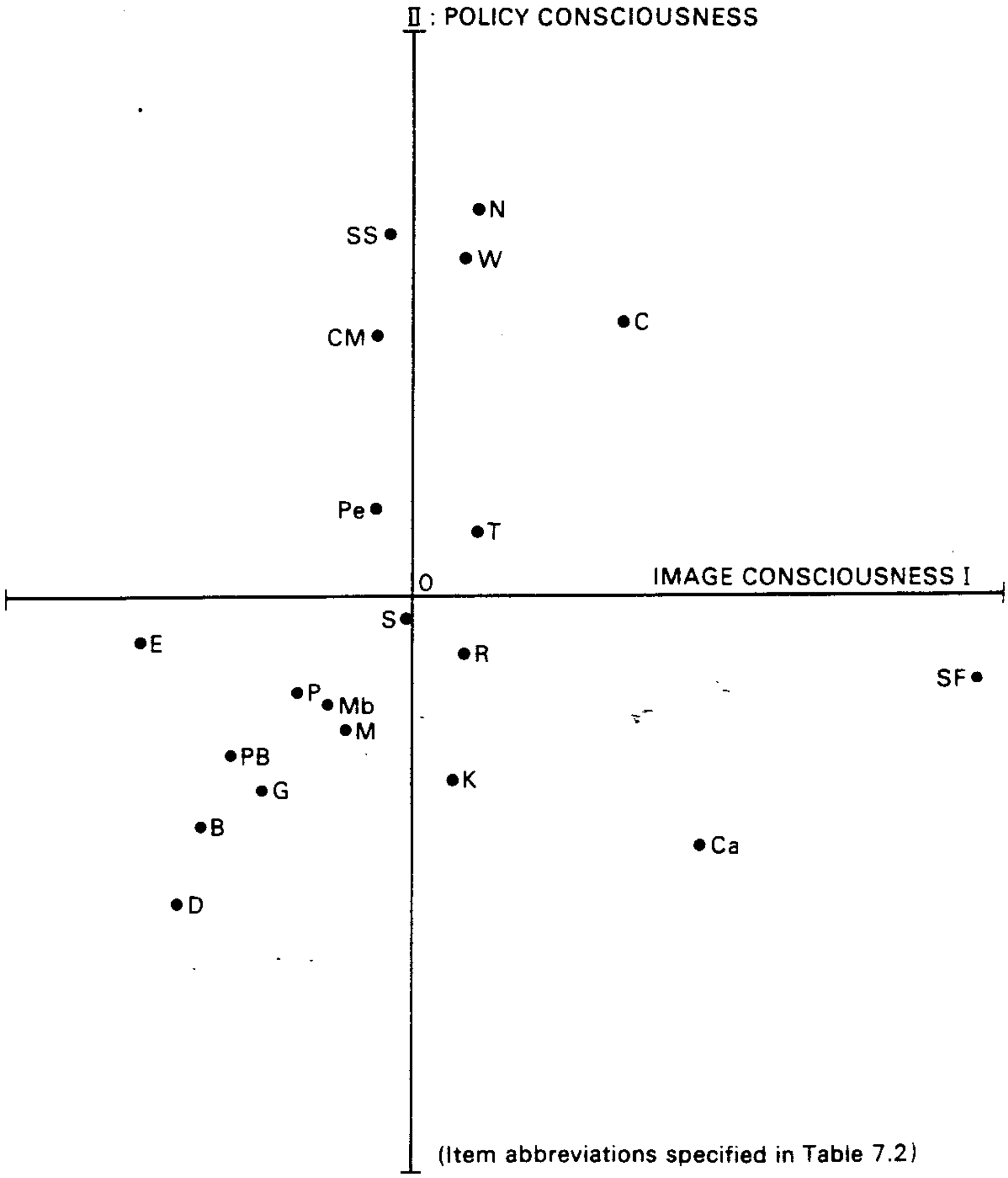
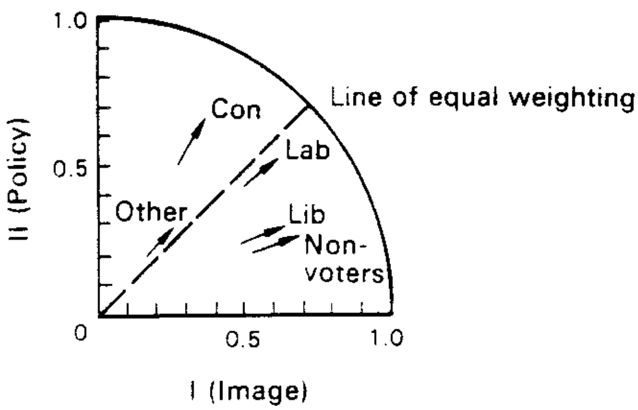


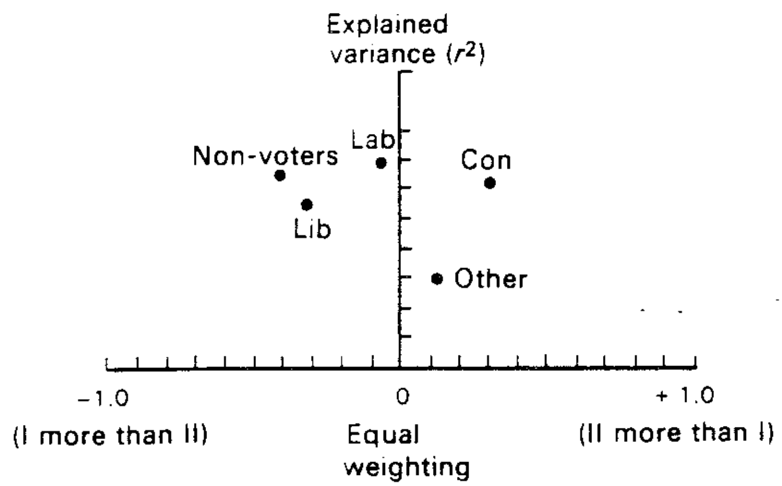
Figure 7.2 *INDSCAL group space: political imagery study*

random differences do they consider them to be sufficient to merit separate treatment. It turns out that the differences among voter subgroups greatly exceed those found for the random groups, especially on the first dimension. The authors then construct the private space for each subgroup (pp. 308–9), and comment:

> The relative unidimensionality of the items for Liberals and Non-voters is apparent. For them, big differences between items only occur between those most clearly reflecting 'style' and 'performance'. In contrast, voters for the two major parties use both dimensions in differentiating items, and the previously mentioned differences between these groups are also evident. Particularly striking is how small the group space looks—how undifferentiated all the items appear—to 'voters for other parties'. These results are substantively not necessarily surprising: the items were, after all, re-scaled as inter-party



(a) INDSCAL Subject Space



(b) Young diagram of (a)

Group	(i) Dimension weights I	II	(ii) Angle w.r.t. dim. I	(iii) Deflection from line of equal weighting*	(iv) Goodness-of-fit (r^2)
A. Conservative	0.32	0.60	61.93°	0.30	0.62
B. Labour	0.57	0.51	41.82°	-0.06	0.69
C. Liberal	0.56	0.28	26.57°	-0.33	0.55
D. Other	0.21	0.26	51.07°	0.11	0.31
E. Non-voter	0.64	0.26	22.11°	-0.42	0.66

*Tan (θ_1 - 45°)

Figure 7.3 *Subject weights plots: political imagery study*

comparisons. The clarity and parsimony with which INDSCAL recovers this property of the group space is, nevertheless, impressive.

(Alt et al. 1976, p. 310)

This example of the use of INDSCAL shows well how, with a little initiative and imagination a model developed within an individually-based psychological tradition can be adapted with considerable success to analyse survey data referring to several thousand respondents (see also Coxon and Jones 1977).

APPENDIX 7.2 NOTES ON THE ESTIMATION PROCEDURE IN INDSCAL

Full details of the alternating least squares procedure for estimating the parameters of the INDSCAL model are contained in Carroll and Chang (1970), Carroll and Wish (1973) and in the MDS(X) documentation of INDSCAL-S.

What follows here is a brief introduction to the basic method of analysis.

(i) The basic model assumes that the subject's data dissimilarities are a linear function of the distances of the solution

$$\delta_{jk}^{(i)} = L(d_{jk}^{(i)})$$

where the distances refer to the ith individual's (private) space, i.e.

$$d_{jk}^{(i)} = \sqrt{\sum_a (y_{ja} - y_{ka})^2}$$

The first step, as in other metric models, is to *convert the subject's data ('relative distances') into estimates of distances* by calculating an additive constant, as in classic scaling (see 5.2.3.2), which will make the data satisfy the triangle inequality.

(ii) *The data 'distances' are then converted into estimated scalar products*, as in classic metric scaling (see Appendix A5.2), with their origin at the centroid of the points. At this stage, each subject's data are normalised to have equal influence. The relationship between the estimated scalar products (b_{jk}) and the private space co-ordinates is simply:

$$b_{jk}^{(i)} = \sum_a y_{ja}^{(i)} y_{ka}^{(i)} \qquad (1)$$

(iii) The INDSCAL model, stated in its distance form is:

$$\delta_{jk}^{(i)} = \sqrt{\sum_a w_a^{(i)}\left(x_{ja} - x_{ka}\right)^2}$$

and the relationship between the group space co-ordinates (x_{ja}) and the private space co-ordinates (y_{ja}) is:

$$y_{ja}^{(i)} = \sqrt{w_a^{(i)}}x_{ja} \qquad (2)$$

Substituting (2) into (1) gives

$$b_{jk}^{(i)} = \sum_a w_a^{(i)} x_{ja} x_{ka} \qquad (3)$$

This is the three-way scalar products formulation of the INDSCAL *model.* For notational simplicity, it helps to rewrite (3), putting subject references (i) as subscripts:

$$b_{ijk} = \sum_a w_{ia} x_{ja} x_{ka} \qquad (3a)$$

(iv) The estimation of the subject weights (w_{ia}) and group space co-ordinates (x_{ja}) in INDSCAL is performed by a variant of the three-way canonical decomposition model (see 7.2.2), which ensures that the second and third ways (x_{ja} and x_{ka}) are in fact identical.

(v) The INDSCAL model has been shown by Schönemann (1972) to have an exact algebraic solution—for perfect data. In the case of errorful data, an iterative process (which may use Schönemann's method to provide an initial configuration) is employed, using an alternating procedure. It consists of finding a preliminary estimate for the two stimulus weights (x_{ja} and x_{ka}), fixing them, and then estimating (by least squares) the subject weights w_{ia}. Then the x_{ja} are estimated, with the w_{ia} and x_{ka} fixed, and so on.

When a satisfactory approximation to the data is obtained, the process terminates, ways 2 and 3 are set equal, a final estimate of way 1 (subject weights) is made and the weights are then appropriately normalised before being output.

A caution

All variants of alternating least squares estimation procedures are susceptible to a greater or lesser extent to local minimum solutions. In any event, users should be prepared for this eventuality: often ten runs with different starting configurations are necessary before one can be virtually certain that one has an optimal solution. In any event, it would be foolhardy to rely on less than three. In the repeated runs, the group space configurations will probably be very similar *except for slight differences in orientation.* Since subject weights refer directly to a particular orientation and will often change considerably under relatively small rotations of the dimensions, particular attention should therefore be paid to how the group space dimensions change and to the individual and overall goodness of fit measures.