14. PROFIT (PROperty FITting)

Page No.

2.1.2.1 Input of properties

Each property consists of a set of values, one for each stimulus in the configuration. All properties must be in the same format and this is specified by means of the INPUT FORMAT card which precedes the READ MATRIX card which reads the properties. Each property is preceded however, by a label card and this label is printed in the output.

2.1.3 Example

To illustrate the use of the PROFIT program we take the configuration reported by Wish (Wish et al, 1972). In their study individuals (subjects) gave ratings on a scale of the degree of similarity between pairs of nations (stimuli). The averaged ratings were used to obtain a four-dimensional MDS solution where a larger distance between a pair of points in this space indicates a greater dissimilarity between the nations concerned. After visual inspection of the plots the authors interpreted the dimensions as shown in figure 1a and 1b.

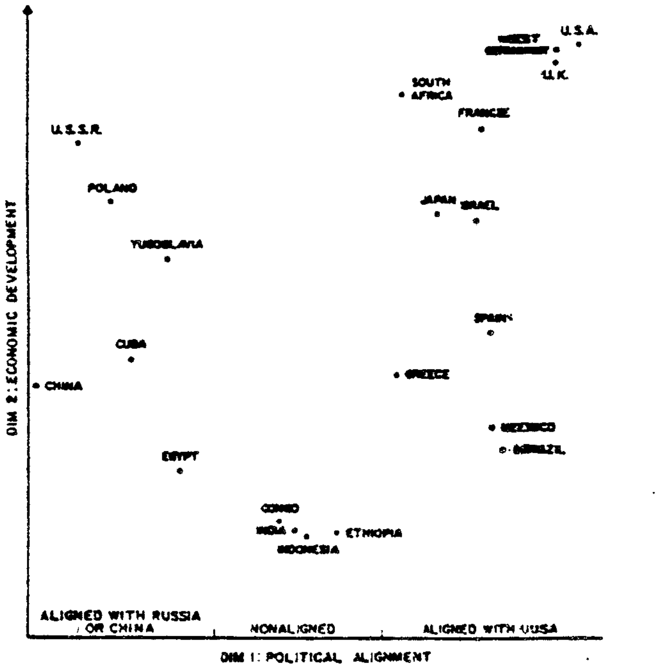

figure 1a

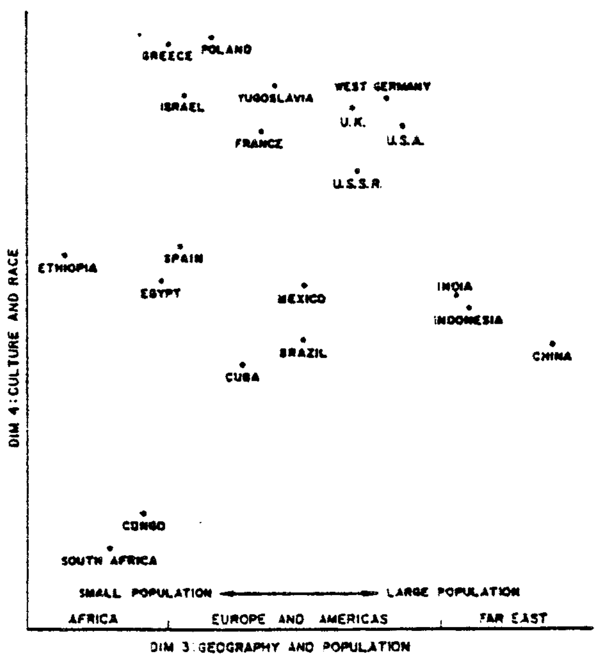

figure 1b

1. OVERVIEW

Concisely: PROFIT (PROperty FITting) provides external analysis of a configuration by a set of property ratings or rankings in row-conditional format by a scalar products (vector) model using either a linear or continuity transformation of the data.

According to the categories developed by Carroll and Arabie (1979) PROFIT may be described as:

Data:		Model:	
	Two-mode		Scalar-product
	Two-way		Two set of points
	Asymmetric		One space
	Dyadic		External
	Ordinal or Interval/Ratio		
	Row-conditional		
	Complete		

1.1 ORIGINS, VERSIONS AND ACRONYMS

PROFIT was developed by J.D. Carroll and J.J. Chang at Bell Laboratories and originally documented in Chang and Carroll (1968).

1.2 PROFIT IN BRIEF

PROFIT takes as input both a configuration of stimulus points and a set of rankings or ratings of the same set of stimuli. These rankings and ratings are usually estimates of different properties of the stimuli. The program locates each property as a vector through the configuration of points, so that it indicates the direction over the space in which the property is increasing. The fitting is accomplished by maximising the correlation between the original property values and the projection of the stimuli onto the vector. This correlation may be either linear or non-linear (continuity).

1.3 RELATION OF PROFIT TO OTHER PROGRAMS IN THE MDS(X) SERIES

1. PROFIT using the linear option is formally identical to Phase 4 (vector model) of the preference mapping program PREFMAP, also using the linear option. (Note that PREFMAP phase IV may also be used with a quasi-non-metric option, providing a form of ordinal property fitting).

2. An internal form of the point-vector model (i.e. where the input configuration is not fixed but is generated from the data) is available in MDPREF.

3. An option within PARAMAP allows a rectangular or row-conditional (two-way, two mode) array of data to be input for internal analysis using a continuity (kappa) transformation between the data and the solution. But only the stimuli are represented in the solution.

2. DESCRIPTION OF THE PROGRAM

2.1 DATA

There are two parts to the input data for PROFIT.

2.1.1 The configuration

The configuration consists of the coordinates for a set of objects (stimuli) on a number of dimensions. This may be an *a priori* configuration (Coxon, 1974) or one resulting from another multidimensional scaling analysis, or, indeed, from a factor analysis. The configuration is input to the program by means of the READ CONFIG card under its associated INPUT FORMAT card and may be presented either stimuli (rows) by dimensions (columns) or dimensions (rows) by stimuli (columns). In this latter case the parameter MATFORM should be given the value 1. Since the configuration is not substantially altered by the PROFIT algorithm, analysis can only take place in a given dimensionality and attempts to specify more than one value on the DIMENSIONS card will cause an error.

2.1.2 The properties

Each of the "properties" which PROFIT will seek to represent as vectors in the configuration, is a set of values which distinguish the stimuli on a particular criterion. These may be physical values (as in the following example) or subjective evaluations of the stimuli on criteria other than that or those used to generate the original configuration. For instance, a simple use of the program might be to map into a MINISSA representation of the perceived similarities between a set of stimuli, information about the subjects' preferences of the same stimuli.

We may wish to concentrate on the following properties of the nations concerned:

1) Gross National Product per Capita, 1965

2) Total Population, 1965

3) Population Growth Rate, Total Time Span (1950-1965)

4) Ethno-linguistic Fractionalization

5) Soviet Aid per Capita, 1954/5 - 1965

6) Total U.S. Economic and Military Aid per Capita (1958-1965)

These aggregate data were obtained under the direction of Taylor (Taylor et al, 1973) and the list could be expanded to contain as many of the 300 and more variables which they report for each country. The deck set up for two properties of this example is given in section 4.

2.2 THE MODEL

PROFIT seeks to represent the properties as vectors over the configuration of points. The analysis is external in as much as the configuration is regarded as being fixed: the stimulus points cannot be moved to make the fit of the vectors better.

The fitted vector is regarded as indicating the direction in which the given property is increasing. As a theory this implies that preference increases continually, never reaching a maximum (corresponding to the economic concept of insatiability).

The property values are then correlated with the projections of the stimuli onto the vector in the following way. The vector is drawn through the origin of the space.* The perpendicular projections

*This is for convenience only. In actual fact any vector parallel to this will give an identical result, since it is only the projections which are significant.

from the origin to the bases of the projections calculated. It is this final set of measurements (the distances from the origin to the projections) which is correlated with the original property values and it is this correlation which is the index of goodness-of-fit between data and solution. Two options are available to the user in calculating this correlation. The program will either calculate and maximise the (linear) product-moment correlation between data and solution or a (non-linear) "smoothness" or "continuity" measure (or, indeed, both). These are chosen by means of the REGRESSION parameter.

Despite its name, the non-linear procedure does not fit curves rather than straight lines into the space. Rather, the function which links data (property values) to solution (point projections) is not constrained to being linear and may instead be drawn from the wider class of non-linear functions. In PROFIT, the particular index of non-linear badness-of-fit is KAPPA, which ensures local monotonicity. This means that in the Shepard diagram the function plot might be upwardly monotone in the lower range and downwardly monotone in the upper range since it is the variations between data values adjacent (or close) to each other which are crucial in calculating the index: Kappa maintains only the smoothness or continuity of the function between adjacent values (hence "local" monotonicity). In the algorithm this is done by giving adjacent (or close) data values a heavy weight. The user is given the option of varying this weight to give varying importance to different aspects of the data (v.i).

2.2.1 The Algorithm

Since the linear and non-linear procedures differ from each other quite considerably, we discuss them here separately.

2.2.1.1 The linear procedure

1. The columns of the configuration are normalised.

2. The XMAT matrix is computed.

For each property in turn:

3. The direction cosines of the vectors are computed.

4. The projections of the points onto the vectors are computed.

5. The correlation between the projections and the property values is computed.

6. The cosines corresponding to the angles between each pair of vectors are computed.

7. The configuration and vector-ends are plotted using both normalised and original coordinates.

2.2.1.2 The non-linear procedure

1. The configuration is normalised.

For each property:

2. KAPPA and ZSQ measures of alienation and correlation respectively are computed.

3. The cosines of the angles between the vectors and the original axes are calculated.

4. The projections of the points onto the vectors are calculated.

When all properties have been thus treated:

5. The cosine of the angle between each pair of vectors is calculated.

6. The configuration of points and vectors is plotted in original and normalised co-ordinates.

2.3 FURTHER OPTIONS

2.3.1 Linear vs. non-linear regression

Because the results of non-linear analysis are more difficult to evaluate, it is often tempting to start with the more familiar linear regression. The linear procedure is however merely a special case of the non-linear and, since usually we do not possess prior information on the form of the relation expected between property values and stimulus projections, the more general non-linear analysis may be preferred as an exploratory technique.

The PROFIT program always reports the product-moment correlation coefficient. It is quite possible that a relatively low value for the non-linear continuity measure, KAPPA and a high value for the (linear) correlation coefficient will be found. This would indicate that the relation is indeed linear and PROFIT should then be run with the linear option in order to test this assumption and provide the information on the (linearly) best fitting property vector.

2.3.2 Non-linear measures of goodness-of-fit

In the case of linear property fitting, the product moment correlation is a suitable measure of goodness-of-fit between the data and the solution. In the non-linear case no such familiar index is available. Rather an index KAPPA (κ) which is a badness-of-fit measure is minimized. Intuitively this measure is minimized whenever the form of the function relating the data to the solution becomes smoother or more continuous locally, whatever its actual overall shape may be. Thus it may be considered as an index of 'local' monotonicity. The definition of KAPPA is given in Appendix 2.

2.3.2.1 The use of the weight parameter

The weighting function plays a crucial role in the definition of KAPPA. This function can take on three different values and each value defines a different "flavour" of KAPPA. The choice of flavour depends crucially on the characteristics of the property values.

2.3.2.1.1 When WEIGHT (Ø)

This is the general definition of non-linear correlation and no restrictions are placed on the data. Therefore, this index can always be applied to examine the extent to which the property values (data) and the projections of the stimulus points (solution) are related by a smooth or continuous function.

2.3.2.1.2 When WEIGHT (1)

In this case, it is assumed that the property values are equally spaced. So the level of measurement of the properties is in effect taken to be ordinal *if* the order is specified with equal intervals. To do this any equally spaced values may be chosen, such as
1, 2, 3,...N or 5, 10, 15,...5N.

There is no restriction on the characteristics of the stimulus configuration when using this option. This option limits the calculation of KAPPA to adjacent points. In this case, κ becomes equivalent to Von Neumann's η (the ratio of the mean square successive difference) as defined in Von Neumann (1941). See below (2.3.2.2.2) for the use of BCO in conjunction with this option.

2.3.2.1.3 When WEIGHT (2)

If the property values tend to be highly clustered into two or more groups of values, then the PROFIT program can be used to determine whether this is also the case for the projections of the stimuli on the fitted vector. To do this we must choose the property values in such a way that it becomes possible to discriminate the clusters. Ordinal level of measurement is sufficient, provided the property values are equally spaced. By defining the maximum distance between two points which are to be taken as falling in the same grouping, the program then selects the clusters. This maximum distance is set using the BCO parameter (see 2.3.2.2.3 below).

The weight factor will now have the effect of restricting attention to property distances which are close to each other (in effect, in the same grouping) and ignoring values outside the BCO value. In this case, κ can be shown to be the equivalent of the "correlation ratio" (Carroll 1964, see also Nie et al, 1975).

2.3.2.2 The use of the BCO parameter

This parameter has a different use and meaning when used in conjunction with different WEIGHT options.

2.3.2.2.1 When WEIGHT Ø

In the general case a value of Ø for BCO (the default) will make the weighting function be undefined for equal property values. If there are equal property values and BCO (Ø) the program will terminate. Thus this option in effect assumes that there are no ties between the property values. If ties do occur among your property values then a small value of BCO (say .001) should be used. This will allow calculation of the weight factor even when the property values are equal. A large value for BCO has the effect of allowing Kappa to decrease indefinitely and is not recommended.

2.3.2.2.2 When WEIGHT (1)

When Van Neumann's η is approximated, then the value of the BCO parameter has a more simple explanation than in the previous case. Now BCO simply gives the size of the equal intervals. Note that if WEIGHT (1), which is the default value, then BCO (Ø) has no meaning and some other value must be specified.

2.3.2.2.3 When WEIGHT (2)

In this case the BCO parameter gives the maximum distance allowed between points in the hypothetical clusters described above in 2.3.2.1.3. Again in this case, the default value BCO (Ø) has no meaning, and must be over-ridden by some other value.

3. INPUT PARAMETERS

3.1 LIST OF PARAMETERS

Keyword	Default Value	Function
REGRESSION	1	1: Linear regression only will be performed. 2: Non-linear regression. 3: Both regressions will be performed (independently).
MATFORM	Ø	Ø: The input configuration is punched stimuli (rows) by dimensions (columns). 1: The input configuration is punched dimensions (rows) by stimuli (columns).
WEIGHT	Ø	(See Section 2.3.2). Ø: Carroll's index of continuity. 1: Van Neumann's ratio of the mean square successive difference. 2: the "correlation ratio".
BCO	Ø	(See Section 2.3.2).

3.2 NOTES

1. The # OF PROPERTIES card may be used in PROFIT in place of the # OF SUBJECTS card.

2. The READ CONFIG card is obligatory.

3. Since the non-linear option involves calculation of large powers of the data values, exponent overflow may occur. In this case the data values should be made smaller. This might be done by changing the format statement so as to divide the values by, say, 100.

3.3 PROGRAM LIMITS

Maximum dimensionality: 10

Maximum number of points: 60

Maximum number of properties: 20

3.4 PRINT, PLOT AND PUNCH OPTIONS

The general format for printing, plotting and punching output is described in the Overview. In the case of PROFIT, the available options are as follows:

3.4.1 PRINT options

The PRINT DATA command will echo both the input stimulus configuration and the property values.

Keyword	Form	Description
INITIAL	p × r	The matrix of stimulus points as normalised by the program. This will differ in linear and non-linear approaches.
CORRELATIONS (Default)	1 × N	The following are printed: 1(a) the correlations for each property (linear regression). (b) the eigenroots associated with each vector (non-linear regression).
PROPERTIES (Default)		The following are printed:
	N × r	1. The direction cosines between each of the fitted vectors and each dimensions in the normalised space.
	N × r	2. The direction cosines between each vector and each dimension of the original space.
	N × N	3. The cosines of the angles between the vectors.
RESIDUALS		A table of residuals is printed i.e. obtained distances - original distances.

3.4.2 PLOT options

Option	Description
INITIAL	The stimulus configuration plotted in pairs of dimensions with both original and normalised co-ordinates marked (up to r(r-1) 2 plots).
FINAL	Both stimulus points and property vectors plotted together original and normalised co-ordinates (up to r(r-1)2 plots).
SHEPARD	N plots of original property values against projections on fitted vectors giving the shape of the linking function.
RESIDUALS	Histogram of residual values.

By default only the first two dimensions of the joint space are plotted.

3.4.3 PUNCH options

Option	Description
SPSS	This command produces a file containing the following variables: I property J stimulus DATA original value on property *i* of stimulus *j* FITTED projection on fitted vector RESID difference between original and fitted values.
SOLUTION	Two matrices are punched: i) the matrix of stimulus points as normalised, and ii) the matrix of direction cosines for the fitted vectors.

4. EXAMPLES

4.1 TEST RUN

```
col 1                 col 16

RUN NAME              TEST DATA AS IN SECTION 2
N OF STIMULI          21
N OF PROPERTIES       2
DIMENSIONS            4
PARAMETERS            REGRESSION(3), BCO(.001)
COMMENT               * * * *
                      NOTICE THAT BOTH LINEAR AND NON-LINEAR OPTIONS
                      ARE TO BE USED AND THAT THE SMALL VALUE IS
                      GIVEN TO BCO BECAUSE THERE ARE TIES IN THE DATA
                      (SEE SECTION 2.3.2.2.1)
                      * * * *
INPUT FORMAT          (4F4.3)
COMMENT               * * * *
                      THIS FORMAT STATEMENT REFERS TO THE CONFIGURATION
                      TO FOLLOW ...
                      * * * *
READ CONFIG
  <here follows the configuration in four dimensions>
INPUT FORMAT          (11F5.0)
COMMENT               * * * *
                      ... WHILE THIS REFERS TO THE PROPERTIES
                      * * * *
READ MATRIX
POPULATION GROWTH RATE 1950-1965
1.60 0.50 1.10 1.10 4.70 1.10 2.40 0.80 0.80 3.10 3.40
1.70 2.00 2.10 1.40 2.50 1.50 2.20 1.20 1.60 1.60
ETHNO-LINGUISTIC FRACTIONALISATION
505 325 026 261 199 015 877 099 436 071 305
694 886 764 657 044 118 038 754 028 666
PLOT                  SHEPARD
COMPUTE
FINISH
```

BIBLIOGRAPHY

Carroll, J.D. and P. Arabie (1979) Multidimensional scaling, in (1980) Annual Review of Psychology, Palo Alto, Ca,: Annual Reviews.

Carroll, J.D. and J.J. Chang (1964) A general index of non-linear correlation and its application to the problem of relating physical and psychological dimensions, unpublished paper, Bell Telephone Laboratories.

Chang, J.J. and J.D. Carroll (1968) How to use PROFIT, a computer program for property fitting by optimizing non-linear or linear correlation, unpublished paper, Bell Telephone Laboratories.

Coxon, A.P.M. (1974) The mapping of family composition preferences: a scaling analysis, Soc.Sci.Res., 3, pp 191-210.

Miller, J.E., R.N. Shepard and J.J. Chang (1964) An analytic approach to the interpretation of multidimensional scaling solution. Paper presented at A.P.A. 1964. Abstract in Am.Psych., 19, pp 579-80.

Neumann, J. von, et al (1941) The mean square successive difference, Am.Math.Stat., 12, pp 153-62.

Taylor, C.L. et al (1973) World handbook of political and social indicators, (2nd edition), Ann Arbor, Michigan.

Wish, M. (1972) Differences in the perceived similarity on nations, in A.K. Romney, R. Shepard and S.B. Nerlove (eds.) Multidimensional Scaling: Theory and Applications, New York, Seminar.

APPENDIX 1: RELATION OF PROFIT TO OTHER PROGRAMS OUTSIDE THE MDS(X) SERIES

No programs outside the MDS(X) series (and the corresponding Bell Laboratories versions) implement a continuity or "smoothness" scaling transformation, and therefore no parallel programs exist for the non-linear version of PROFIT.

The linear version of PROFIT can be thought of as a linear multiple regression program: predicting property values from a linear combination of dimensional co-ordinates of the stimuli involved. Strictly speaking, any multiple regression program can therefore be used to implement linear PROFIT.

A number of MDS programs outside the MDS(X) series have the capability of external scaling with linear (metric) or ordinal (non-metric) transformation functions. (Guttman-Lingoes SSA-1; KYST; ALSCAL4) - but only for an ideal point (distance) model. However, none of these allow the possibility of using a vector (scalar products) model. Currently the only accessible equivalent of linear PROFIT occurs in the PRINCIPALS model in the Young - de Leeuw - Takane ALSCAL series.

APPENDIX 2: TECHNICAL DESCRIPTION OF PROFIT*

A2.1 The linear method

Let

$$\underset{\sim}{X} \equiv \{x_{ja}\} \qquad j = 1,\ldots,p \quad a = 1,\ldots,r$$

be the matrix of coordinates of p points in an r-space

and where

$$\underset{\sim}{X}'\underset{\sim}{X} = \underset{\sim}{D}$$

where $\underset{\sim}{D}$ is a diagonal matrix.

Also, define

$$\underset{\sim}{p} = \underset{\sim}{p}^{i}_{j} \qquad i = 1,\ldots,N \quad j = 1,\ldots,p$$

as the row vector of property values for subject i

$$\underset{\sim}{t} = \underset{\sim}{t}^{i}_{a} \qquad i = 1,\ldots,N \quad a = 1,\ldots,r$$

as the column vector of direction cosines of the fitted vector

$$\underset{\sim}{h} = \underset{\sim}{h}^{i}_{j} \qquad i = 1,\ldots,N \quad j = 1,\ldots,p$$

as the row vector of projections of the p points onto the fitted vector.

The problem is to find $\underset{\sim}{t}$ and $\underset{\sim}{h}$ such that

$$| \underset{\sim}{p} - \underset{\sim}{h} |^2 = \min | \underset{\sim}{p} - \underset{\sim}{h} |^2$$

---

*This appendix is based on Chang & Carroll 19   which is used with permission.

when

$$\underset{\sim}{h} = \underset{\sim}{X}\,\underset{\sim}{t}$$

i.e. a least-squares solution for $\underset{\sim}{t}$

This is given by

$$\underset{\sim}{t} = (\underset{\sim}{X}'\underset{\sim}{X})^{-1}\,\underset{\sim}{X}'\underset{\sim}{p}$$

while

$$\underset{\sim}{h} = \underset{\sim}{X}\,\underset{\sim}{t} = \underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{X})^{-1}\,\underset{\sim}{X}'\underset{\sim}{p}$$

A2.2 <u>The non-linear method</u>

The general index of non-linear correlation (KAPPA: κ) between an independent variable p and a dependent x was given by Carroll (1964) as:

$$\kappa = \frac{1}{S^2} \sum_{i \neq j} w_{ij}\,(x_i - x_j)^2$$

where

$$w_{ij} = f(|p_i - p_j|)$$

and f a monotone decreasing function

and

$$S^2 = \frac{1}{N} \sum_{i} (x_i - \bar{x})^2$$

In PROFIT the independent p corresponds to one property and the dependent x to the projections of the points onto the vector. The program seeks to <u>minimize</u> κ.

The matrix

$$\underset{\sim}{X} \equiv \{x_{ja}\} \qquad j = 1,\ldots,p \quad a = 1,\ldots,r$$

is transformed so that

$$\underset{\sim}{X}'\underset{\sim}{X} = \underset{\sim}{I}$$

and the matrix

$$\underset{\sim}{X}'\underset{\sim}{A}\,\underset{\sim}{X}$$

is computed, where

$$A \equiv \begin{cases} a_{ij} = -w_{ij} & (i \neq j) \\ a_{ii} = \sum_{j \neq i} w_{ij} \end{cases}$$

The smallest eigenvalue of $\underset{\sim}{X}'\underset{\sim}{A}\,\underset{\sim}{X}$ is the minimum value of κ and the corresponding vector gives the direction cosines of the new vector.