Chapter 14: Nonmetric Scaling

Prerequisites: Chapter 7, Section 3.9

14.1 *Additive Conjoint Measurement*

In Chapters 5 through 7 we look at the classical statistical models from which we get our *t*-test, ANOVA, and of course, regression. For example, in a factorial design, we generally assume that our dependent variable is measured at the interval or ratio level, and we test to see if the cell means combine in an additive way, or if instead, we need to include interaction terms. Nonmetric additive conjoint measurement turns this reasoning exactly on its head. In this section we will assume the additivity in order to learn something about the level of measurement of the dependent variable. Or, we can make very weak assumptions about this measurement - i. e. that its merely ordinal - and still capture the main effects of the factorial design.

In this section we will use a linear model, and for the most part we will assume that we have a factorial design. For example, a consumer may be looking at a series of vacation packages. For now, lets just say that each package has five destinations and four prices leading to 20 different packages altogether. Our data might consist of a ranking (or possibly rating if there are not too many ties) of the 20 vacation packages. In any case, we will assume that our data are ordinal. Here are the steps that we will go through.

Step 0. Initialize a second version of the dependent variable,

$$\mathbf{y}^* = \mathbf{y}$$

where $\mathbf{y}$ is the original ordinal-scaled 20 by 1 column vector of the rankings for each of the 20 packages. As we will see, the $\mathbf{y}^*$ vector will be the optimally scaled version of the data.

Step 1. Use least squares to fit an additive model. Our usual notation would have us fit a model looking like

$$\mathbf{y}^* = \mathbf{X}\boldsymbol{\beta} + \mathbf{e} = \hat{\mathbf{y}}^* + \mathbf{e}$$

where $\mathbf{X}$ is a *design matrix* containing, lets say, effect coding (see Section 7.2) for main effects only and the $\boldsymbol{\beta}$ vector contains those effects. We can also use scalar notation that allows us to keep track of data from individual cells in the two way design. For example, looking at destination i and price j we might have

$$y^*_{ij} = \alpha_i + \beta_j + e_{ij} = \hat{y}^*_{ij} + e_{ij}$$

where α_i represents the effect of being in row i (destination i) and β_j captures the effect of being in column j of the design, that is the column with price j. Note that since the $\hat{y}^*_{ij}$ are optimally rescaled anyway we do not have to worry about the grand mean. We can just absorb that in the scaling. Step 1 is a least squares step where we pick the α_i and the β_j (or if you prefer the matrix notation, the elements of the $\boldsymbol{\beta}$ vector) so as to minimize the sum of squared error. But now we are going to go into a second least squares step.

Step 2. Find the monotone transformation that would improve the fit of the above model as much as possible. We will be focusing on this step in this section, but for now, we can say that in Step 1

we modified the parameters to fit the dependent variable, but in this step we are modifying the dependent variable to better fit the additive model. Symbolically we can say that the new values of the y_{ij}^{*} will be

$$y_{ij}^{*} = H_m[y_{ij}, \hat{y}_{ij}^{*}].$$

This says that we will be modifying the optimally scaled dependent values (y^{*}), based on the ordinal data (y) and the linear model of the optimally scaled dependent variable ($\hat{y}^{*}$). The function H_m is monotone, which means that the optimally rescaled dependent variables have to be in the same order as the original ordinal data. Ordinal means that only the order of the numbers is invariant, and the rescaled dependent variable honors that order, meaning that it is equivalent to the original.

Step 3. Refit the additive model.

Step 4. If the model fits OK, stop. Otherwise go back to Step 2 and repeat.

The procedure alternates between two least squares steps with one (Step 1 above) fitting the linear model minimizing the sum of squared error, and the other (Step 2 above) fitting the data minimizing a sum of squares called *STRESS*. We are thus alternating least squares steps, and this technique is part of a family of techniques that are called *Alternating Least Squares* or ALS. We will discuss STRESS in just a little bit. Now let us jump into Step 2 in more detail, a step that we call *optimal scaling*.

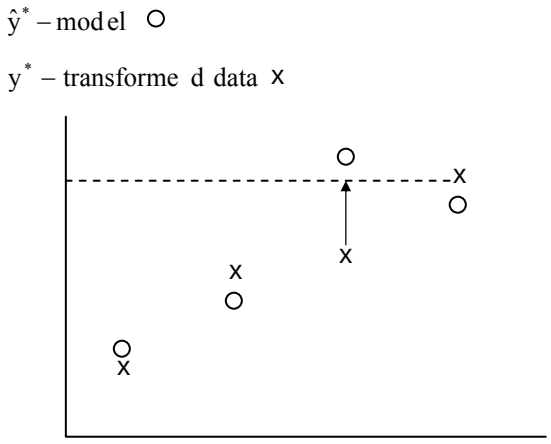
For the time being, to make our life easier, we will assume that there are no ties in the data. If we sort the data, and revert back to one subscript that refers to each datum's sort sequence, the original ordinal data would look like

$$y_1 < y_2 < \cdots < y_n . \tag{14.1}$$

In that case, to honor or maintain this order, we impose on the optimally scaled values the following constraints:

$$y_1^{*} \le y_2^{*} \le \cdots \le y_n^{*} . \tag{14.2}$$

We can picture the situation by looking at what is known as a *Shepard Diagram*, named after the Stanford psychologist Roger Shepard,



The x's represent the optimally rescaled data, and the 0's are the predictions of that data from the additive model. In particular, focus on the third data point. Due to the monotone restrictions, it cannot pass the fourth data point. It can come right up to its value and tie it however, since Equation (14.2) allows for equality. The program will move the y^* value as close to its predicted value as it can without violating the monotone constraints. In analytic terms, we minimize the following quantity subject to the inequalities above:

$$\text{STRESS} = \sqrt{\frac{\sum_i (y_i^* - \hat{y}_i^*)^2}{\sum_i (\hat{y}_i^* - \bar{\hat{y}}^*)^2}} \tag{14.3}$$

where $\bar{\hat{y}}^*$ is the average of the $\hat{y}_i^*$. In effect, the denominator normalizes the value of STRESS. It is the numerator that is where the action is. But remember, the formula is subject to the series of inequalities given in Equation (14.2).

There are two ways to handle ties in the data. The primary approach occurs when $y_i = y_j$ implies $y_i^* \leq y_j^*$. Here we are treating the data as fundamentally continuous with thresholds. The secondary approach is when when $y_i = y_j$ implies $y_i^* = y_j^*$. Here we treat the data as truly discrete, and we fully honor equalities.

The output from this procedure consists of the y_{ij}^*, called the *utilities*, and the α_i and the β_j, called *part-worths*. These can all be used to simulate various market conditions.

14.2 *Multidimensional Scaling*

Many of the models in this and other sections represent choice situations. Which brand do you like more and which do you like less? *Multidimensional Scaling*, often abbreviated MDS, is designed to get at the consumer's perception of the brands rather than their preferences for them. Later on we will bring preference back into the model, but for now we will focus on the way that the consumer sees the brands. We will also focus on nonmetric MDS, meaning that we will assume that the data are ordinal. Nonetheless, we will be able to fit a model with interval scaled parameters, just as we did in the section on conjoint measurement.

The MDS data collection procedure is one of the least obtrusive methods that exist in the world of marketing research. The respondent's job is to rank (or rate) pairs of brands as to how similar they are. We might imagine a simplified experimental design with three brands A, B and C. The respondent will tell us which of the three possible pairs, AB, AC, BC, are the most similar. Then which pair is the next most similar, and so forth until all of the pairs are ranked as to their similarity.

MDS uses a geometric model for *similarity* or *proximity judgments*. Brands judged highly similar, according to the model, are represented near each other in a *perceptual space*. Conversely, brands judged dissimilar find themselves distant in this perceptual space of r dimensions. Later we will get back to the dimensionality of the space. Now, lets think about the similarity judgment between brand i and brand j, and call it d_{ij}. The optimally rescaled data will be called d_{ij}^* while the predicted rescaled data, that is predicted from the distance model, will be called $\hat{d}_{ij}^*$. As before we will be using Alternating Least Squares. The perceptual space reveals the aspects of the brands that the consumer considers salient when looking at those brands. The steps in the algorithm are

Step 0. Initialize $d_{ij}^* = d_{ij}$.

Step 1. Fit the distance model $\hat{d}_{ij}^*$ to the d_{ij}^*.

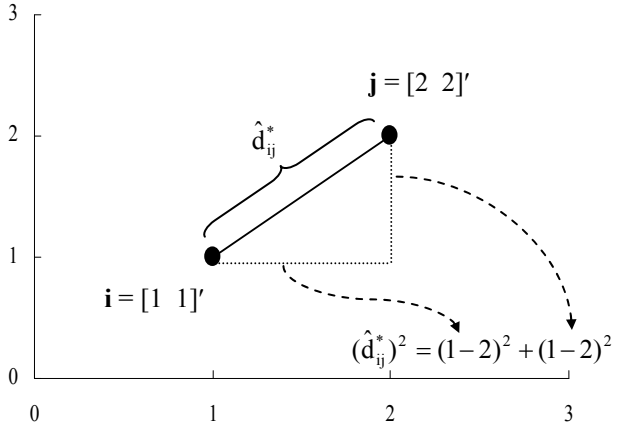
Step 2. Optimally rescale the d_{ij}^* to the $\hat{d}_{ij}^*$ honoring the order of the d_{ij}.

Step 3. Quit if done or go back to Step 1.

As before, in Step 2 we will be minimizing STRESS. However, at this point it would be wise to look at the distance model used in Step 1. We will be modeling the proximities as distances,

$$\hat{d}_{ij}^* = \sqrt{\sum_{m}^{r} (x_{im} - x_{jm})^2}. \tag{14.4}$$

Some of you may remember this equation from a high school geometry course. It is the *Euclidean distance* between points i and j in a space of r dimensions. The parameter x_{im} represents the coordinate for brand i on the m^{th} dimension. Assuming that r = 2, we might look at a graph of the situation:



In the case pictured, the distance between i and j is $\sqrt{2}$.

The flexibility of MDS can hardly be overstated. There are at least three categories of methods that allow us to capture similarity or proximity:

Direct

Ask for pairwise ratings or rankings
Have respondents sort objects into categories
Pick the pair of pairs most similar (Method of tetrads)
For each member of a trio, indicate which other brand it is most similar to (Method of triads)

Attribute Based

Calculate correlations over measures
Calculate distances over measures

Behavioral

Traffic volume, phone calls, trade or migration between two cities, regions, etc.
Switching proportions between brands
Confusability
Cross elasticities
Percent agreement, Chi Square, other measures of association

14.3 *Other Distance Models*

In addition to the classic Euclidean formula, other formulae qualify as distances, which are also called *metrics*. More accurately, we might use the word metric. Four axioms must be satisfied for a set of numbers to qualify as a metric:

Identity $\hat{d}^*_{ii} = 0,$ (14.5)

Non-negativity $\hat{d}^*_{ij} \geq 0,$ (14.6)

Symmetry $\hat{d}^*_{ij} = \hat{d}^*_{ji}$, and (14.7)

Triangle inequality $\hat{d}^*_{ij} + \hat{d}^*_{jk} \geq \hat{d}^*_{ik}$. (14.8)

Of course the already noted classic Euclidean Distance equation,

$$\hat{d}^*_{ij} = \sqrt{\sum_m^r (x_{im} - x_{jm})^2},$$

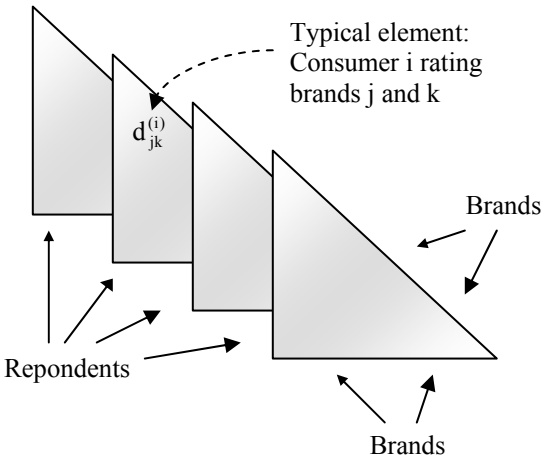
satisfies all four requirements. But other distance models are possible. For example, there is a very flexible formula called the *Generalized Minkowski Metric* in which

$$\hat{d}^*_{ij} = \left[\sum_m^r | x_{im} - x_{jm} |^a \right]^{1/a}. \quad (14.9)$$

When a = 2 you get the classic Euclidean formula. When a = 1 you get a metric known as the *City Block Metric*. This is the distance between two places where all angles have to be 90° and the triangle inequality holds as an equality. This metric is often used when the objects being scaled are perceptually decomposable. As $a \to \infty$ you get the *supremum metric* in which respondents only notice the biggest difference.

14.4 *Individual Differences in Perception*

Up to this point in this chapter, we have been looking at two way single mode data. What this means is that we have a data matrix with rows and columns and thus it is said to be *two way data*. There is just a single mode, however, since both ways of the matrix are indexed by brands. Now we will look at what happens when we have *three way data* representing *two modes*. The second mode will be individual subjects. A diagram for this sort of data is given below:



A typical element in the dataset would be $d^{(i)}_{jk}$, the similarity judgment for brands j and k made by subject i. If the numbers from each matrix are not comparable, as they would be if each subject was engaging in rank-ordering, the data are called *matrix conditional*.

Data such as these can be analyzed using the *Weighted Euclidean Model*, also known as the *INDSCAL model* (INDividual Differences SCALing). That model is given now:

$$\hat{d}_{jk}^{(i)*} = \left[\sum_{m}^{r} w_{im} (x_{jm} - x_{km})^2 \right]^{1/2}, \tag{14.10}$$

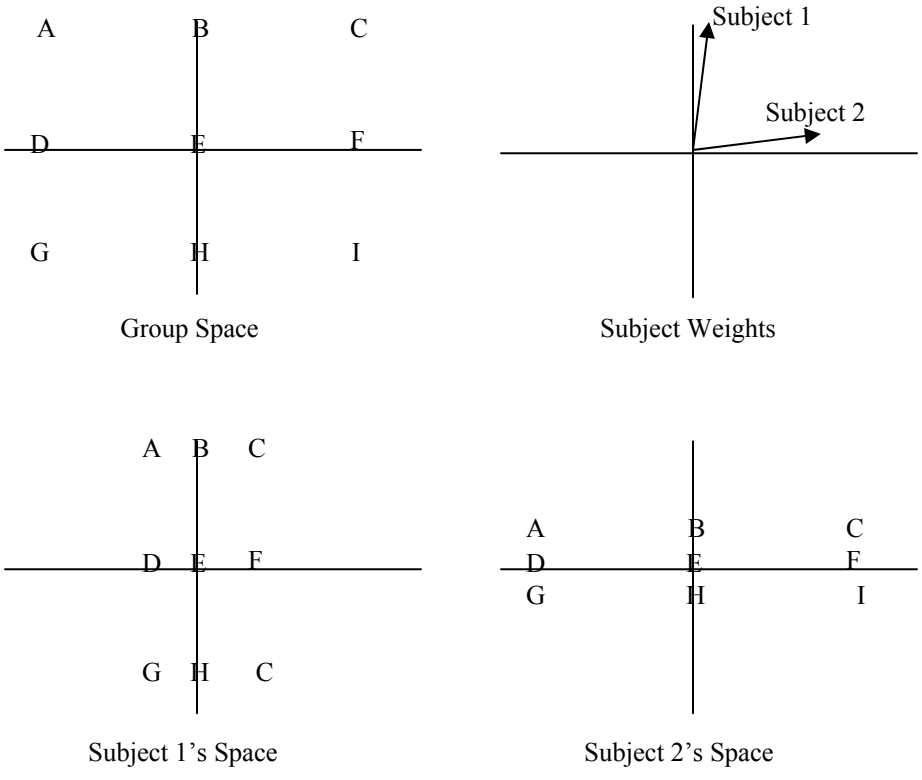
with the added element of the weights, the w_{im}, which represent the importance placed on dimension m by individual i. The coordinates, the x_{jk}, are still coordinates in this model but they are the coordinates of the brand in the *group space*. Each individual has their own coordinates which create a Euclidean space in which the axes have been stretched or shrunk. The coordinate for individual i would be

$$y_{jm}^{(i)} = x_{jm} \cdot w_{im}^{1/2} \tag{14.11}$$

so that

$$\hat{d}_{jk}^{(i)*} = \left[\sum_{m}^{r} (y_{jm}^{(i)} - y_{km}^{(i)})^2 \right]^{1/2} \tag{14.12}$$

using these "customized" coordinates. A diagram of how all this looks appears below:



Subject 1, who has a high weight for dimension 2 and a very small weight for dimension 1 has a space where brands that are separated on the second dimension are very dissimilar, but brands

whose only difference lies along dimension 1 (for example, brands A, D and G), seem quite similar to this person. Subject 2 has the opposite pattern.

The INDSCAL model can be conveniently represented in matrix notation. Place the coordinates for brand j in the vector

$$\mathbf{x}'_{j\cdot} = [\,x_{j1} \quad x_{j2} \quad \cdots \quad x_{jm} \quad \cdots \quad x_{jr}\,] .$$

Here, the dot subscript reduction operator on the symbol $\mathbf{x}'_{j\cdot}$ comes from Equation (1.2), and basically is used to hold the place of the second subscript, the one for the dimensions. We also put the subject weights on the diagonal of the matrix $\mathbf{W}^{(i)}$ as

$$\mathbf{W} = \begin{bmatrix} w_{i1} & 0 & \cdots & 0 \\ 0 & w_{i2} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & w_{ir} \end{bmatrix} .$$

Then the INDSCAL model is

$$\hat{d}_{jk}^{(i)*} = \left[\mathbf{x}'_{j\cdot}\mathbf{W}^{(i)}\mathbf{x}_{k\cdot}\right]^{1/2}$$

and the original, unweighted Euclidean model is a special case where $\mathbf{W}^{(i)} = \mathbf{I}$ for all i, or in other words, where

$$\hat{d}_{jk}^{*} = \left[\mathbf{x}'_{j\cdot}\mathbf{x}_{k\cdot}\right]^{1/2} .$$

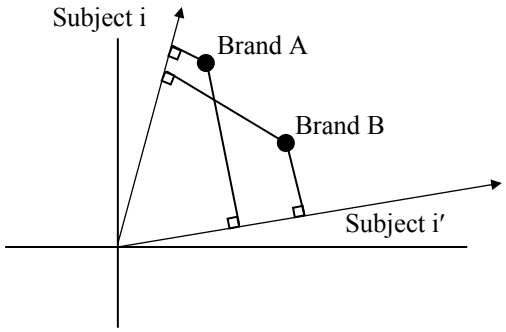
14.5 *Preference Models: The Vector Model*

In this model, we will be representing not just which brands are similar to which others, but which brands the consumers like the best. The vector model can be estimated from a variety of data types, but here we will assume that we have similarity judgments and preferences rankings or ratings.

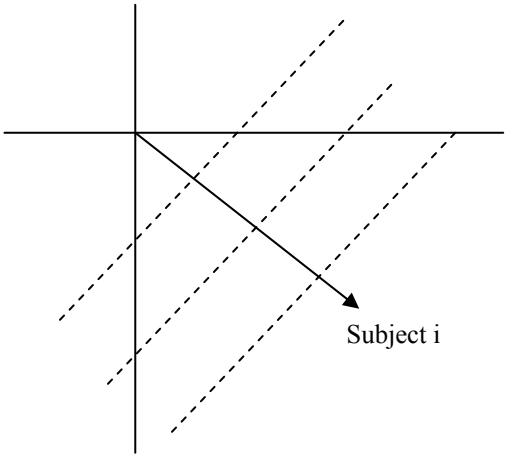
In the Vector model, the brands are represented as points in the perceptual space, as before. Each brand has a set of coordinates on the r dimensions in the perceptual space,

$$\mathbf{x}'_{j\cdot} = [\,x_{i1} \quad x_{i2} \quad \cdots \quad x_{im} \quad \cdots \quad x_{ir}\,] .$$

But now, each subject is also represented in the space, by a vector. The projection of the brand on that vector determines the preference for it. The situation is illustrated below:



Subject i prefers Brand A to Brand B, as the projection for A exceeds the value for B on that subjects *preference vector*. Subject i′, on the other hand, prefers B to A as that person's projections line up in the opposite order. Note that the projections of the brands onto each of the subject vectors occur at right angles to those subject vectors. It is instructive to look at *isopreference contours* for a particular subject. A subject will be indifferent between any two brands sitting on the same isopreference contour since both have equal appeal. These contours are graphed below:



According to the model, our Subject i would be completely indifferent between any brands that would appear on the same dashed line. However, the subject would prefer a brand on a line farther from the origin to one on a line closer in.

In order to express these ideas mathematically, we need to be able to identify each consumer's vector. It will be convenient to pick a point on the vector at a distance of 1 unit from the origin. Doing so, we then have

$$\mathbf{y}'_{i\cdot} = \begin{bmatrix} y_{i1} & y_{i2} & \cdots & y_{im} & \cdots & y_{ir} \end{bmatrix}$$

and since the distance from this point to the origin is 1, we have

$$\sqrt{\mathbf{y}'_{i\cdot}\mathbf{y}_{i\cdot}} = \sqrt{\sum_{m}^{r} y_{im}^2} = 1.$$

The preference of person i for brand j, which is the projection of the brand's point onto the subject's preference vector so as to create a 90° angle with that vector, is given by

$$\hat{s}_{ij} = \frac{\sum_{m}^{r} y_{im} x_{jm}}{\left[\sum_{m}^{r} y_{im}^2\right]^{1/2}}$$

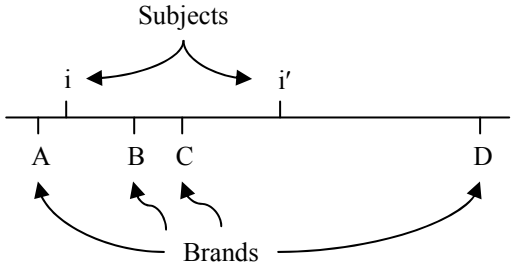
but since the denominator is 1, we have simply

$$\hat{s}_{ij} = \sum_{m}^{r} y_{im} x_{jm} = \mathbf{y}'_{i\cdot} \mathbf{x}_{j\cdot} . \tag{14.13}$$

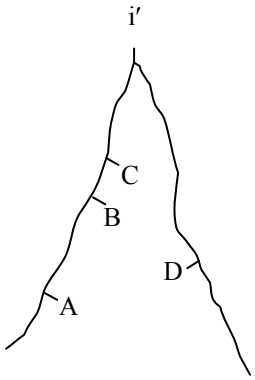
The vector model holds for many different product attributes. For example, for price, less is always better. For miles per gallon, more is preferred to less. But there are some product attributes for which the vector model makes ridiculous predictions. For example, it is quite possible that I would like a car that is larger than a sub-compact. But does this mean I would always want a larger and larger car? The vector model predicts that I would prefer a 2 mile long car to a sub-compact. This brings to mind the story of the porridge, which might be too hot, it might be too cold, or it might be perfect. To model perceptual attributes that act like this requires that we turn to the notion of an ideal point.

14.6 *Preference Models: The Ideal Point Model*

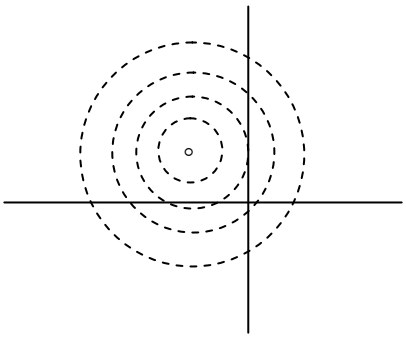
In the vector model, we represent individuals as a direction in the perceptual space. In the ideal point model, individuals, as well as brands, become points. In that sense we have a *joint space*. The situation is illustrated below with a one dimensional joint space.



We have picked two hypothetical respondents: i and i′. In the ideal point model, the closer a brand is to you, the more you like it. Thus i has the preference sequence: A-B-C-D while i′ prefers the brands in the following order: C-B-D-A. If the dimension were like a string, the preference of subject i′ could be determined by picking up that string at his or her ideal point:

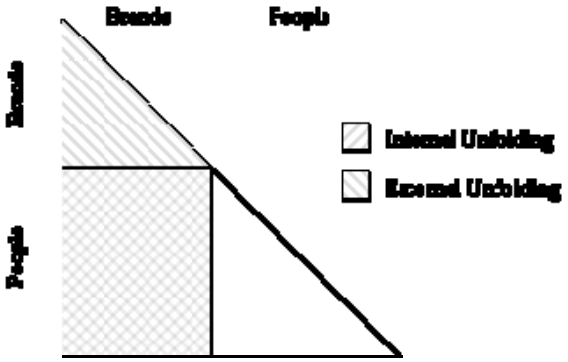


In fact, using this technique we actually perform the opposite mathematical operation. Given a set of consumers' preference rankings, we unfold the string, or more generally the r-dimensional space, to deduce the underlying position of the brands and the individuals' ideal points. In fact, this technique is sometimes known as *unfolding*. Looking at a two dimensional space, we can see the isopreference contours for the unfolding model:



Again, the subject whose ideal point is located at the center of those concentric circles, will be indifferent between any pair of brands appearing on the same circle. A brand on an inner circle will be preferred to a brand on a more outer circle.

There are two ways of collecting data for this model, and in fact for the vector model described above. You can collect internal data, which means that each individual rates or ranks their preference towards each brand, or you can combine that with perceptual ratings or rankings of the similarity of the brands. The situation is represented by the data matrix below.



Here we placed people and brands in the same lower triangular matrix. *Internal unfolding* utilizes just the people × brands rectangular part of this matrix, while *external unfolding* adds the brand × brand information.

The ideal point model is a distance model, so the formula for the preference of subject i for brand j is just the distance between subject i's ideal point and brand j's position in the joint space,

$$\begin{aligned}\hat{s}_{ij} &= \left[\sum_{m}^{r}(y_{im} - x_{jm})^2\right]^{1/2} \\ &= \left[(\mathbf{y}_{i\cdot} - \mathbf{x}_{j\cdot})'(\mathbf{y}_{i\cdot} - \mathbf{x}_{j\cdot})\right]^{1/2}.\end{aligned} \tag{14.14}$$

We can use metric or alternating least squares versions of this technique, and it is possible to have a version where there are individual differences in perception of the joint space, as we had with the INDSCAL model.

References

Conjoint Measurement

Green, Paul E. and V. Srinivasan (1978) "Conjoint Analysis in Consumer Research: Issues and Outlook," *Journal of Consumer Research*, 5 (September), 103-21.

Green, Paul E. and V. Srinivasan (1990) "Conjoint Analysis in Marketing: New Developments with Implication for Research and Practice," *Journal of Marketing,* 54 (October), 3-19.

Jaccard, James, David Brinberg and Lee J. Ackerman (1986) "Assessing Attribute Importance: A Comparison of Six Methods," *Journal of Consumer Research,* 12 (March), 463-8.

Validity and Conjoint Measurement

Green, Paul E., Wayne S. DeSarbe and Pradeep K. Kedia (1980) "On the Insensitivity of Brand-Choice Simulations to Attribute Importance Weights," *Decision Sciences,* 11, 439-50.

Green, Paul E., Kristiaan Helsen and Bruce Shandler (1988) "Conjoint Internal Validity under Alternative Profile Presentations," *Journal of Consumer Research,* 15 (December), 392-7.

Green, Paul E., Abba M. Kreiger, and Pradep Bansal (1988) "Complete Unacceptable Levels in Conjoint Analysis: A Cautionary Note," *Journal of Marketing Research*, 25 (August), 293-300.

Hagerty, Michael R. (1986) "The Cost of Simplifying Preference Models," M*arketing Science,* 5 (Fall), 298-319.

Moore, William L. and Morris B. Holbrook (1990) "Conjoint Analysis on Object with Environmentally Correlated Attributes: The Questionable Importance of Representative Design," *Journal of Consumer Research* 16 (March), 490-7.

Reibstein, David, John E.G. Bateson and William Boulding (1988) "Conjoint Analysis Reliability: Empirical Findings," *Marketing Science*, 7 (Summer), 271-86.

Safizadeh, M. Hossein (1989) "The Internal Validity of the Trade-off Method of Conjoint Analysis," *Decision Sciences* 20 (Summer), 451-61.

Teas R. Kenneth (1985) "An Analysis of the Temporal Stability and Structural Reliability of Metric Conjoint Analysis Procedures," *Journal of the Academy of Marketing Science,* 13 (Winter/Spring), 122-42.

Wittink, Dick R. and Phillipe Cattin (1981) "Alternative Estimation Methods for Conjoint Analysis: A Monte Carlo Study," *Journal of Marketing Research,* 18 (February), 101-6

Conjoint: Choice Simulation and Aggregation

Green, Paul E. and Abba M. Krieger (1988) "Choice Rules and Sensitivity Analysis in Conjoint Simulators," *Journal of the Academy of Marketing Science*, 16 (Spring), 114-27.

Kamakura, Wagner A. (1988) "A Least Squares Procedure for Benefit Segmentation with Conjoint Experiments," *Journal of Marketing Research,* 25 (May), 157-67.

Wiley, James B. and James T. Low (1983) "A Monte Carlo Simulation of Two Approaches for Aggregation Conjoint Data,"*Journal of Marketing Research,* 20 (November), 405-16.

Conjoint: Extensions

Green Paul E. (1984) "Hybrid Models for Conjoint Analysis: An Expository Review," *Journal of Marketing Research,* 21 (May), 155-69.

Tantiwong, Duang Tip and Peter C. Wilton (1985) "Understanding Food Store Preferences among the Elderly Using Hybrid Conjoint Measurement Models," *Journal of Retailing,* 61 (Winter), 35-64.

Green Paul E., J. Douglas Carroll and Stephen M. Goldberg (1981) "A General Approach to Product Design Optimization Via Conjoint Analysis," *Journal of Marketing,* 45 (Summer), 17-37.

Green, Paul E. and Abba M. Krieger (1985) "Models and Heuristics for Product Line Selection," *Marketing Science* 4 (Winter), 1-19

Teas R. Kenneth (1987) "Magnitude Scaling of the Dependent Variable in Decompositional Multiattribute Preference Models," *Journal of the Academy of Marketing Science,* 15 (Fall), 64-73.

MDS in Marketing Research

Cooper, Lee G. (1983) "A Review of Multidimensional Scaling in Marketing Research," *Applied Psychological Measurement*, 7 (Fall), 427-50

Green, Paul E. and Frank J. Carmone (1969) "Multidimensional Scaling: An Introduction and comparison of Nonmetric Unfolding Techniques," *Journal of Marketing Research*, 6 (August), 330-41.

Lehmann, Donald R. (1972) "Judged Similarity and Brand-Switching Data as Similarity Measures," *Journal of Marketing Research,* 9 (August), 331-4.

Perreault, William D. and Forrest W. Young (1980) "Alternating Lease Squares Optimal Scaling: Analysis of Nonmetric Data in Marketing Research," *Journal of Marketing Research,* 17 (February), 1-13.

Moore, William L. and Morris B. Holbrook (1982) "On the Predictive Validity of Joint Space Models in Consumer Evaluations of New Concepts," *Journal of Consumer Research*, 9 (September), 206-10

MacKay, David B. and Joseph L. Zinnes (1986) "A Probabilistic Model for the Multidimensional Scaling of Proximity and Preference Data," *Marketing Science,* 5 (Fall), 324-44.

Schiffman, Susan S., M. Lance Reynolds and Forrest W. Young (1981) *Introduction to Multidimensional Scaling*. New York: Academic.

Shocker, Allen D. and V. Srinivasan (1974) "A Consumer Based Methodology for the Identification of New Product Ideas," *Management Science,* Series B, 20 (February), 921-37.

MDS Models: Applications and Extensions

DeSarbo Wayne S. and Vithala R. Rao (1986) "A Constrained Unfolding Methodology for Product Positioning," *Marketing Science,* 5 (Winter), 1-19.

Gavish, Bezale, Dan Horsky and Kizhanatham Srikanth (1983) "An Approach to the Optimal Positioning of a New Product," *Management Science*, 29 (November), 1277-97.

Green, Paul E. and Abba M. Krieger (1989) "Recent Contributions to Optimal Product Positioning and Buyer Segmentation," *European Journal of Operational Research*, 41 (July), 127-41.

Harshman, Richard A., Paul E. Green, Yorman Wind and Margaret E. Lundy (1982) "A Model for the Analysis of Asymmetric Data in Marketing Research," *Marketing Science*, 1 (Spring), 205-42.

Holbrook, Morris B. and Duglas V. Holloway (1984) "Marketing Strategy and the Structure of Aggregate, Segment-specific, and Differential Preferences," *Journal of Marketing,* 48 (Winter), 62-7.

Sudharshan, D., Jerrod H. May and Allan D. Shocker (1987) "A Simulation Comparison of Methods for New Product Location," *Marketing Science* 6 (Spring), 182-201.