FACTORS INFLUENCING RATING SCALE RESPONSES: A PROCEDURE FOR PARTITIONING REDUNDANCY

Mohammad H. Sabertehrani, Eastern Michigan University
Robert A. Peterson, University of Texas at Austin

Abstract

This paper attempts to identify those influencing factors other than stimulus objects that have a significant impact on response variability when rating scales are employed. Partitioning redundancy in canonical analysis is introduced as a useful procedure to measure the impact of influencing factors on rating scale responses. Results from an empirical investigation show that non-stimulus object factors such as scale characteristics, type of measurement instrument used, data collection mode, and the environment in which data are collected account for a significant proportion of the explained variance in rating scale responses.

Introduction

The explained variance in responses to a rating scale (EV) can be accounted for only in part by differences in stimulus objects (SO). There are other factors that may influence the variability in responses. Conceptually, these factors can be categorized as scaler characteristics (SC), measurement instrument characteristics (MI), mode of data collection (MD), and the environment in which data collection takes place (EN). Symbolically

EV = f(SO, SC, MI, EN)

Scaler characteristics consist of a wide variety of factors and influences. Many of them are relatively permanent and general in nature, such as demographic, socio-economic, or personality characteristics. Others, like mood, are more transitory, while still others are specific to the research situation (e.g., responses syndromes like haloing or yea saying).

Three categories of factors are methodological in nature. Measurement instrument characteristics include type of question (open or closed end), nature of questionnaire, and so forth. Data collection mode refers to whether data are collected by means of telephone, mail or personal interviews. Environment includes physical, social, and temporal (time of day, etc.) influences.

Canonical Analysis

Since its introduction by Hotelling (1935), canonical analysis has been recommended by many researchers as a multivariate statistical method for relating two or more sets of variables (Anderson 1958, Morrison 1967, Van De Geer 1971, Overall and Klett 1972). Several computational procedures as well as extensions of the technique have been suggested (Bartlett 1941, Horst 1961, Gower 1966, Thordike and Weiss 1973). Also, numerous computer software packages (Cooley and Lohnes 1971, Roskam 1966, Veldman 1978, Dixon and Brown 1979, Nie et al 1975) have been developed that facilitate the application of canonical analysis to research data.

Beyond the traditional test of statistical significance (Bartlett 1941) in canonical analysis, several procedures have been advanced for interpreting canonical relationships. Meredith (1964) recommended the examination of canonical variate-variable correlations (canonical loadings) to investigate the relative importance of individual variables for significant variates. Stewart and Love (1968) suggested the use of a measure of redundancy to explain the proportion of variance in a criterion variate set accounted for by canonical variates in a predictor set. Miller (1975) has derived a significance test--analogous to the F-test in regression analysis--to determine the statistical significance of redundancy.

Although the above procedures have contributed considerably to the interpretation of canonical relationships, there still exist occasions in which researchers need to pinpoint the specific contribution of individual variables and their statistical significance in an overall canonical relationship (i.e., beyond a single variate).

Purpose

The purpose of this paper is twofold. A first purpose is to introduce a procedure for partitioning redundancy in canonical analysis. A second purpose is to demonstrate how the relative importance of various factors that influence scale responses can be evaluated by means of the redundancy partitioning procedure.

Partitioning Redundancy

Redundancy ($R^2_{Y|X}$) is a non-symmetric index of explained variance similar to R^2 in multiple regression. It is frequently useful in interpreting canonical relationships. $R^2_{Y|X}$ summarizes the proportion of total variance in a criterion set of variables (Y) accounted for by canonical variates in a predictor set (X) (Miller and Farr 1971). Stewart and Love applied notions from factor analysis and used the concept of a factor loading to extract variance from the canonical structure matrix (the matrix of correlations of the original variables with the canonical variates). Their index of redundancy estimates the proportion of variance accounted for in the criterion set of variables by a two-step calculation. Specifically, redundancy equals

1. the amount of criterion set variance accounted for by the ith variate of that set $[r^2_{c_i} \mid \min(p,q)]$ multiplied by
2. the proportion of variance that this variate shares with the corresponding canonical variate of the predictor set (λ_i)

summed across all canonical functions.

Thus, summing redundancy on corresponding criterion-predictor variates--min (p,q)--gives

$$R^2_{Y|X} = \sum_{i=1}^{\min(p,q)} [r^2_{c_i} | \min(p,q)]\lambda_i$$

$$i = 1,2,\ldots,\min(p,q)$$

where p = the number of variables in predictor set

q = the number of variables in criterion set

r^2_{ci} = the eigenvalue of the ith variate extracted from criterion set Y

As the above computation formula indicates, the redundancy procedure permits determination of each variate's partial contribution to overall accounted-for variance:

$$R^2_{Y|X} = R^2_{1\ Y|X} + R^2_{2\ Y|X} + \ldots R^2_{i\ Y|X} \ldots + + R^2_{\min(p,q)\ Y|X}$$

or

$$R^2_{Y|X} = \sum_{i=1}^{\min(p,q)} R^2_{i\ Y|X}$$

By partitioning redundancy, the amount of variance in the criterion variables accounted for by each predictor variable can be calculated. The notion of partial redundancy associated with a particular variate can be extended to include the original variables which comprise that variate. Specifically, partitioning partial redundancies across the individual predictor variables can be done by a two-step procedure:

1. the proportion of the squared correlations between the ith predictor variate and the jth predictor variable ($a_{ji}/\sum_{j=1}^{p} a_{ji}$) (computed from the canonical structure matrix) is multiplied by

2. the amount of redundancy in criterion set accounted for by the ith predictor variate ($R^2_{i\ Y|X}$).

Thus

$$V_{ji} = \frac{a_{ji}}{\sum_{j=1}^{p} a_{ji}} R^2_{i\ Y|X}$$

or

$$R^2_{i\ Y|X} \sum_{j=1}^{p} V_{ji}$$

and

$$R^2_{Y|X} = \sum_{i=1}^{\min(p,q)} \sum_{j=1}^{p} V_{ji}$$

where

V_{ji} = the amount of variance in the criterion set explained by the jth variable for the ith predictor variate;

a_{ji} = the amount of variance in the ith predictor variate explained by the jth predictor variable; and

$R^2_{i\ Y|X}$ = the amount of variance in the criterion set explained by the ith predictor variate.

Test of Significance

Redundancy was tested for statistical significance through a method developed by Miller (1975). By means of a Monte Carlo study, Miller developed and validated a sampling distribution and a siginificance test for the bimultivariate redundancy test statistic similar to the F-ratio in multiple regression. His findings suggest that a function of canonical redundancy is distributed as F just as $[R^2/(1 - R^2)][df_2/df_1]$ in multiple regression follows the F distribution under the null hypothesis. The F-ratio for redundancy is

$$F = \left[\frac{R^2_{Y|X}}{1 - R^2_{Y|X}}\right] \left[\frac{(N - p - 1)q}{pq}\right]$$

where $df_1 = pq$ and $df_2 = (N - p - 1)q$

The significance test described above was generalized to include so-called part redundancies. An interpretation of part redundancy $R^2_{Y|X'(j=1,2,\ldots,p)}$ is that it is a simple squared correlation between criterion set Y and the residual predictor subset X' from which the effects of predictor variable x_j are taken out. Indeed, the part redundancy is a squared correlation of the residuals of set Y and the subset X', after the effects of x_j have been taken out. The F-ratio for part redundancy is formulated as

$$F = \frac{R^2_{Y|X'(j=1,2,\ldots,p)}}{(1 - R^2_{Y|X})/(N - p - 1)q}$$

Illustration

To demonstrate the redundancy partitioning procedure in the context of a canonical correlation problem, an example study drawn from Peterson (1981) is used. The sample consisted of 347 students from two undergraduate marketing classes at the University of Texas at Austin. Students were randomly assigned to one of six cells in a 2 x 2 x 2 (stimulus object by scale polarity by measurement setting) experimental design. Each student was requested to complete a three-page questionnaire containing questions about retail store image, shopping behavior, self-perception, mood, and demographics. The dependent variables were 10 7-point rating scale questions relating to retail store characteristics such as prices, merchandise quality, sales personnel, and physical layout of store.

Three experimental treatments were designed to induce variations in stimulus object, measurement instrument, and measurement setting characteristics. The respective treatments were store (VAR1), polarity (VAR2), and the environment in which a questionnaire was completed (VAR3). Two department stores in Austin were presented as stimulus objects--Foley's and J. C. Penney. Approximately half of the study individuals answered the 10 rating-scale image questions about Foley's and the other half evaluated Penney's. Polarity of the image scales was manipulated to introduce the second experimental treatment. About half of the students were administered the image scales in a unipolar format while the other half were presented bipolar scales.

The environment in which students answered the questions was also varied. Two-thirds of the respondents completed the scales in a classroom setting while the remaining third completed their questionnaires outside the classroom and returned them at their convenience.

Results

As can be seen in Table 1, approximately 17 percent of the variance in rating scale responses was accounted for by the 7 predictor variables. The contribution of non-stimulus object variables was relatively large; nearly one third of the relative variance was explained by non-stimulus object factors. The major influencing factors in this group were subject characteristics. Shopping frequency was the most important such factor and accounted for approximately 8 percent of the relative variance explained. General interest in department store and sex also accounted for statistically significant proportions of the explained variance. Environment, polarity and mood had no significant impact on response variation.

Conclusions

Although the influence of non-stimulus object factors on rating scale responses was not large in absolute magnitude, they were sufficiently strong to provide clear evidence that non-stimulus object factors may confound or even distort research findings derived from rating scale responses. To isolate the impact of potentially confounding influences in multivariate analyses, redundancy partitioning is a techinque that can highlight these influences.

Future research is needed to come up with a more comprehensive conceptual framework that takes into account other potentially confounding influences on scale responses and on other aspects of self-report data. Ultimately some typology of influencing factors needs to be constructed so that rating scale data can be better evaluated with regard to the manner in which they are collected.

References

Anderson, T. W. (1958), Introduction to Multivariate Statistical Analysis. New York: Wiley.

Bartlett, M. S. (1941), "The Statistical Significance of Canonical Correlations," Biometrika, 32 (January), 29-38.

Cooley, William A. and Paul R. Lohnes (1962), Multivariate Procedures for the Behavioral Sciences. New York: Wiley.

Dixon, W. J. and M. B. Brown (1979), Biomedical Computer Programs. Berkeley, CA: University of California Press.

Gower, J. C. (1966), "A Q-technique for the Calculations of Canonical Variates," Biometrika, 53 (December), 588-90.

Horst, Paul (1961), "Relations Among M Sets of Measures," Psychometrika, 26 (June), 129-49.

Hotelling, Harold (1935), "The Most Predictable Criterion," Journal of Educational Psychology, 26 (February), 139-42.

Meredith, William (1964), "Canonical Correlations with Fallible Data," Psychometrika, 29 (March), 55-65.

Miller, John K. (1975), "The Sampling Distribution and a Test for the Significance of the

Table 1

REDUNDANCY ACROSS PREDICTOR VARIABLES

	Canonical Variates							Absolute Variance (%)	Relative Variance (%)	F	Probability
	1	2	3	4	5	6	7				
Environment	.00	.02	.09	.02	.03	.03	.00	.19	1.13	.68	n.s.
Store	11.11	.13	.00	.00	.00	.00	.00	11.24	67.03	40.51	<.01
Polarity	.02	.02	.04	.14	.00	.00	.00	.22	1.31	.79	n.s.
Frequency	1.46	.77	.07	.00	.01	.00	.00	2.31	13.78	8.33	<.01
Interest	.85	.34	.12	.00	.01	.00	.01	1.33	7.93	4.79	<.01
Sex	1.07	.03	.12	.02	.00	.02	.00	1.26	7.51	4.54	<.01
Mood	.00	.01	.05	.01	.14	.01	.00	.22	1.31	.79	n.s.
All Variables	14.51%	1.32%	.49%	.19%	.19%	.06%	.01%	16.77%	100.00%	8.64	<.01

Bimultivariate Redundancy Statistic: A Monte Carlo Study," Multivariate Behavioral Research, 19 (April), 233-44.

Miller, John K. and S. David Farr (1971), "Bimultivariate Redundancy: A Comprehensive Measure of Interbattery Relationship," Multivariate Behavioral Research, 6 (July), 313-24.

Morrison, Donald F. (1967), Multivariate Statistical Methods. New York: McGraw-Hill.

Nie, Norman H., D. H. Hull, J. G. Jenkins, K. Steinbrenner, and D. H. Bent (1975), Statistical Package for the Social Sciences, Second Edition. New York: McGraw-Hill.

Overall, J. E. and D. J. Klett (1972), Applied Multivariate Analysis. New York: McGraw-Hill.

Peterson, Robert A. (1981), "An Exploratory Investigation of Mediating Factors in Retail Store Image Responses," Advances in Consumer Research, Vol. 8, Kent Monroe (ed.), Ann Arbor, MI: Association for Consumer Research, 662-64.

Roskam, E. (1966), "A Program for Computing Canonical Correlations on IBM 1620," Educational and Psychological Measurement, 26 (Spring), 193-98.

Stewart, Douglas and William Love (1968), "A General Canonical Correlation Index," Psychological Bulletin, 70, (September), 160-63.

Thordike, R. M. and Weiss, D. J. (1973), "A Study of the Stability of Canonical Correlations and Canonical Components," Educational and Psychological Measurement, 33 (Spring), 123-134.

Van De Geer, J. P. (1971), Introduction to Multivariate Analysis for the Social Sciences. San Francisco, CA: Freeman.

Veldman, Donald J. (1978), The Prime System Computer Programs for Statistical Analysis. Research and Development Center for Teacher Education, University of Texas at Austin.