

Capturing and Clustering Judges' Policies

ILAN ADLER AND DITSA KAFRY

University of California, Berkeley

Judgment analysis has been used as a method for grouping judges according to the homogeneity of their regression equations. It is demonstrated that (a) contrary to the accepted claim, judgment analysis may fail to generate an optimal grouping, and (b) this problem is, in fact, a special case of a well-known clustering problem for which a variety of cluster analysis techniques are available.

1. INTRODUCTION

The multiple regression model has been used to capture policies of judges with respect to their responses to a set of multiple characteristic stimuli (Slovic & Lichtenstein, 1971). According to this model each stimulus is defined by several characteristics which are quantitatively measured. Every judge is then required to respond to each of the stimuli by providing an overall numerical evaluation. A linear regression equation relating the characteristics (independent variables) to the evaluation (dependent variable) is computed for each judge. The square multiple regression coefficient (R^2) is used as a measure of the efficiency with which the regression equation is capable of predicting the judges' evaluations. The R^2 measures the success in "capturing" the judges' policy by a linear regression equation. Once a policy of a given judge is successfully captured, it can be used to interpret the weights attributed to each of the characteristics and it can be applied to simulate the judges' policy to obtain more consistent evaluations of additional set of stimuli.

This model of policy capturing has been applied to a variety of decision making situations such as: selection and promotion (Christal, 1968), evaluation of beauty of paintings (Holmes & Zedeck, 1973), performance appraisal (Zedeck & Kafry, 1977), and organizational choice (Zedeck, 1977). Obviously, policy capturing can be applied to a variety of other fields such as personality assessment, consumer behavior, perception and information processing, or for studies of need strength and motivation.

In order to obtain a more parsimonious presentation and for reasons of organizational practicality, it is sometimes desired to provide a single regression equation to represent a group of judges who have similar equations. Such clustering is also of interest for understanding the similarities and differences between the policies of groups of judges. Obviously one has to pay the price of such clustering in terms of the predictive efficiency of the group regression equation for the individual judges in that group.

Hence, the problem is to find the clustering (to a given number of groups) which minimizes the loss of predictive efficiency.

The technique of judgment analysis (JAN) has been suggested as an efficient method to solve the grouping problem (Bottenberg & Christal, 1961, 1968). This method has been the focus of several papers discussing its theoretical and practical virtues (e.g., Naylor & Wherry, 1965; Dudycha & Naylor, 1966; Dudycha, 1970) and has been applied in studies of policy capturing (e.g., Christal, 1968; Holmes & Zedeck, 1972; Zedeck & Kafry, 1977). However a close look at JAN and its iterative nature casts doubts on its ability to provide an optimal grouping (with respect to maximizing predictive efficiency). It is the purpose of this paper to demonstrate and explain that JAN does not necessarily generate an optimal grouping. It is also shown that the grouping problem can be transformed to the usual clustering problem and thus can be solved by known cluster analysis techniques. This transformation also provides a useful insight into the nature of the homogeneity criterion applied by JAN.

For the purpose of clarity of the discussion we shall first present the transformation of the grouping problem to a clustering problem; this will be followed by a description of JAN and a demonstration and explanation of its failure to guarantee an optimal grouping. Finally, we conclude with a discussion of the possible applications of cluster analysis techniques to policy capturing problems.

2. THE JUDGES' POLICY CAPTURING CLUSTERING PROBLEM

The policy capturing model is composed of a set of m judges who evaluate n sets of stimuli, where each stimulus consists of p characteristics. Specifically let

X_{ij} = the value of the j th characteristic of the i th stimulus ($i = 1, \dots, n$; $j = 1, \dots, p$),

X^i = the $(1 \times p)$ vector whose j th element is X_{ij} ($i = 1, \dots, n$),

X = the $(n \times p)$ matrix whose i th row is X^i ,

Y_{ki} = the overall numerical evaluation of X^i given by the k th judge ($k = 1, \dots, m$; $i = 1, \dots, n$).

Y^k = the $(n \times 1)$ vector whose i th element is Y_{ki} ($k = 1, \dots, m$).

To simplify notation we shall assume that the levels of the characteristics are given in terms of deviations from the means, that is:

$$\sum_{i=1}^n X_{ij} = 0 \quad (j = 1, \dots, p).$$

As usual it is assumed that X has a full rank to guarantee the existence of a unique regression coefficients vector.

The objective of the policy capturing model is to try to represent the policy of each judge by a regression equation

$$Y^k = XB^k + e^k,$$

where B^k is a $p \times 1$ vector of the computed least-squares regression coefficients and e^k is the vector of residuals. Throughout the presentation u will be used to denote a column vector of 1's.

Since it is sometimes desired to provide a single regression equation to represent a group of judges, the following clustering problem arises: for a given integer T (the desired number of groups) cluster the judges to T groups G_1, \dots, G_T . The policy of all the judges in Group G_t is then captured and represented by a single regression equation in which $B(G_t)$ is the computed vector of regression coefficients. The objective is to group the judges such that the predictive efficiency of the model is maximized. Given clusters G_1, \dots, G_T the measure of the predictive efficiency $R^2(G_1, \dots, G_T)$ was defined (Bottenberg & Christal, 1968) as one minus the ratio between the sum of the unexplained square residuals of every cluster and the total square deviations of all the evaluations. More specifically

$$R^2(G_1, \dots, G_T) = 1 - \frac{\sum_{t=1}^T SS(G_t)}{\sum_{k=1}^m \sum_{i=1}^n (Y_{ki} - \bar{y})^2}, \quad (1)$$

where

$$SS(G_t) = \sum_{k \in G_t} \left(Y^k - XB(G_t) \right)' \left(Y^k - XB(G_t) \right),$$

and

$$\bar{y} = \frac{1}{m n} \sum_{k=1}^m \sum_{i=1}^n Y_{ki}.$$

Thus mathematically the *policy capturing clustering problem* is: for a given integer T ($1 \leq T \leq m$) find a clustering G_1, \dots, G_T such that $R^2(G_1, \dots, G_T)$ is maximized over all possible clusterings of m judges to T groups.

3. THE TRANSFORMATION OF THE POLICY CAPTURING CLUSTERING PROBLEM TO A GENERAL CLUSTERING PROBLEM

In this section we shall first develop a convenient expression for $SS(G_t)$ which is the key to the transformation of the policy capturing clustering problem to the well-known general clustering problem.

Let B^k be the $p \times 1$ regression coefficient vector for the k th judge, i.e.,

$$B^k = (X'X)^{-1} X'(Y^k - u \bar{y}(k)) \quad (k = 1, \dots, m), \quad (2)$$

where

$$\bar{y}(k) = \frac{1}{n} \sum_{i=1}^n Y_{ki}.$$

Given a group G_t let us assume for simplicity that $G_t = (1, \dots, g)$; that is, the group includes judges 1, \dots , g . We shall first develop a simple explicit expression for $B(G_t)$ the regression coefficient vector obtained by computing a single regression equation for the group G_t . Thus,

$$B(G_t) = \left[\begin{array}{c} (X' \dots X') \\ \left(\begin{array}{c} X \\ \cdot \\ \cdot \\ X \end{array} \right) \end{array} \right]^{-1} \left[\begin{array}{c} (X' \dots X') \\ \left(\begin{array}{c} Y^1 \\ \cdot \\ \cdot \\ Y^g \end{array} \right) - u \bar{y}(G_t) \end{array} \right], \quad (3)$$

where

$$\bar{y}(G_t) = \frac{1}{gn} \sum_{k=1}^g \sum_{i=1}^n Y_{ki},$$

where the dimension of the matrix $(X' \dots X')$ is $p \times (ng)$. Carrying the matrix multiplication in (3) we get

$$\begin{aligned} B(G_t)gX'X)^{-1}X' \sum_{k=1}^g (Y^k - u\bar{y}(k)) \\ = \frac{1}{g} \sum_{k=1}^g (X'X)^{-1}X' (Y^k - u\bar{y}(k)) = \frac{1}{g} \sum_{k=1}^g B^k. \end{aligned} \quad (4)$$

Hence the regression coefficient vector of the group is the simple average of the regression coefficient vectors of the judges in the group. Similarly if we define $\hat{y}(G_t)$ as the predicted evaluation vector of the group then,

$$\hat{Y}(G_t) = XB(G_t) + \bar{y}(G_t) = \frac{1}{g} \left[\sum_{k=1}^g XB^k + \bar{y}(k) \right] = \frac{1}{g} \sum_{k=1}^g \hat{Y}^k, \quad (5)$$

where \hat{Y}^k is the predicted evaluation vector of the k th judge based on his/her own regression equation. Using these results we now construct useful expression for $SS(G_t)$. By definition

$$\begin{aligned}
SS(G_t) &= \sum_{k=1}^g (Y^k - \hat{Y}(G_t))' (Y^k - \hat{Y}(G_t)) \\
&= \sum_{k=1}^g (Y^k - \hat{Y}^k + \hat{Y}^k - \hat{Y}(G_t))' (Y^k - \hat{Y}^k + \hat{Y}^k - \hat{Y}(G_t)) \\
&= \sum_{k=1}^g (Y^k - \hat{Y}^k)' (Y^k - \hat{Y}^k) + \sum_{k=1}^g (\hat{Y}^k - \hat{Y}(G_t))' (\hat{Y}^k - \hat{Y}(G_t)) \\
&\quad + 2 \sum_{k=1}^g (Y^k - \hat{Y}^k)' (\hat{Y}^k - \hat{Y}(G_t)).
\end{aligned}$$

However, the last term is the covariance between the residual least-squares estimates, $Y - \hat{Y}$, and a linear function of X . Hence, this last term equals to zero (Castellan, 1973), thus we have

$$SS(G_t) = \sum_{k=1}^g (Y^k - \hat{Y}^k)' (Y^k - \hat{Y}^k) + \sum_{k=1}^g (\hat{Y}^k - \hat{Y}(G_t))' (\hat{Y}^k - \hat{Y}(G_t)). \quad (6)$$

Therefore, the sum of squared residuals within the group is equal to the sum of square residuals associated with the regression equations of each of the judges plus the sum of square deviations of the predicted evaluation of each judge from the predicted evaluation of the group. It is important to note that the second term of $SS(G_t)$ is equal to $(1/g) \sum_{k=1}^g \sum_{j=k+1}^g (\hat{Y}^k - \hat{Y}^j)' (\hat{Y}^k - \hat{Y}^j)$, which expresses the mean Euclidian distance between all the predicted vectors of group G_t .

Using $SS(G_t)$ in the expression of $R^2(G_1, \dots, G_T)$ (see (1)) we thus get that for every clustering G_1, \dots, G_T with g_t as the number of groups in G_t we have

$$\begin{aligned}
R^2(G_1, \dots, G_T) &= 1 - \frac{\sum_{k=1}^m (Y^k - \hat{Y}^k)' (Y^k - \hat{Y}^k) + \sum_{t=1}^T \frac{1}{g_t} \sum_{k,j \in G_t} (\hat{Y}^k - \hat{Y}^j)' (\hat{Y}^k - \hat{Y}^j)}{\sum_{k=1}^m \sum_{i=1}^n (Y_{ki} - \bar{y})^2}.
\end{aligned}$$

It is obvious that since $\sum_{k=1}^m (Y^k - \hat{Y}^k)' (Y^k - \hat{Y}^k)$ and $\sum_{k=1}^m \sum_{i=1}^n (Y_{ki} - \bar{y})^2$ are constant for all possible groupings G_1, \dots, G_T the clustering problem is equivalent to the following problem: Given m vectors $\hat{Y}^1, \dots,$

\hat{Y}^m and an integer T ($1 \leq T \leq m$), cluster the \hat{Y}^k vectors to T groups G_1, \dots, G_T as to minimize

$$D^2(G_1, \dots, G_T) = \sum_{t=1}^T \frac{1}{g_t} \sum_{k,j \in G_t} (\hat{Y}^k - \hat{Y}^j)' (\hat{Y}^k - \hat{Y}^j).$$

This problem is the well-known clustering problem with the predicted evaluation vectors $\hat{Y}^1, \dots, \hat{Y}^m$ as the points to be clustered and the Euclidian distance as the measure of homogeneity. Therefore known cluster analysis techniques can be employed for its solution. A discussion of the availability of such methods is presented in the concluding section.

It should be noted that this transformation of the judges' grouping problem to the general clustering problem provides a useful insight to the measuring of the criterion of maximizing the predictive efficiency (R^2).

As is evident from the previous discussion the application of R^2 as the criterion results in clusters which are homogeneous with respect to the Euclidian distances among the *predicted* evaluation vectors. Hence, such clustering should be undertaken only if this is a desirable measure of homogeneity.

4. THE JAN METHOD

It was demonstrated in the previous section that grouping judges' policies is a special case of the general clustering problem for which a variety of methods have been developed. However, one can find in the literature only one technique that has been applied to the judges' grouping problem: the Judgment Analysis (JAN) method (Bottenberg & Christal, 1968).

In this section we shall present this method and then illustrate by a numerical example that it may fail to generate an optimal solution to the grouping problem.

Starting with a maximal possible number of groups (which is obviously equal to the number of judges) the JAN method gradually decreases the number of clusters by combining two previously defined clusters that minimize the reduction in the overall predictive efficiency (R^2). This process stops when all the judges are combined to form a single cluster. Alternatively, the process could be terminated either when a given number of groups is obtained, or when a predetermined lower bound of R^2 is reached.

Though the JAN method in its original form (Bottenberg & Christal, 1968) was presented in terms of maximizing the predictive efficiency (R^2), we prefer to present the method in terms of the equivalent problem of minimizing the mean-square Euclidian distances among the predicted

evaluation vectors within the groups. We believe that such a presentation is simple, computationally efficient and consistent with our approach to the grouping problem. Let us now present the JAN method in detail.

Preparation Step

The regression equations of the judges are computed and then used to compute the predicted evaluation vectors $\hat{Y}^1, \dots, \hat{Y}^m$ of the m judges.

First Step

The two vectors \hat{Y}^i, \hat{Y}^j with the minimal Euclidian distance are combined into one cluster.

General Step

Given previously defined clusters G_1, \dots, G_T (where G_t is the set of indices corresponding to the judges in the t th cluster), the two clusters G_t, G_r whose grouping resulted in minimal increase of the mean Euclidian distances among the predicted evaluation vectors in the combined group are grouped together. That is, groups G_t and G_r for which the sum of the mean-square Euclidian distances within clusters, $D(G_1, \dots, G_t \cup G_r, \dots, G_T)$, is minimal are combined.

Termination Rule

The process terminates when all the judges form one cluster (or according to the alternative termination rules which were mentioned previously).

It should be noted that the predictive efficiency $R^2(G_1, \dots, G_T)$ of the clusters formed at step $m - T$ is easily computed by using the expression that was developed in the previous section

$$R^2(G_1, \dots, G_T) = 1 - \frac{\sum_{k=1}^m (Y^k - \hat{Y}^k)' (Y^k - \hat{Y}^k) + D^2(G_1, \dots, G_T)}{\sum_{k=1}^m \sum_{i=1}^n (Y_{ki} - \bar{y})^2} \quad (9)$$

(Note that $\sum_{k=1}^m (Y^k - \hat{Y}^k)' (Y^k - \hat{Y}^k)$ and $\sum_{k=1}^m \sum_{i=1}^n (Y_{ki} - \bar{y})^2$ can be computed in the preparation step based on the computed judges' regression equation while $D^2(G_1, \dots, G_T)$ is provided at the end of every step.)

Obviously this hierarchical method is very simple to program and highly efficient. However, a close look at the method casts doubt on its ability to produce an optimal grouping. It is clear that once a group is formed it can never be separated in subsequent steps. But it seems quite possible that two (or more) judges who form a group in one of the early stages will belong to different clusters in the optimal grouping of later stages.

In the following example we present such a case, thus illustrating that the JAN method does not necessarily generate an optimal clustering.

Consider the following data: Given four judges, one characteristic, and five stimuli (hence $m = 4$, $p = 1$, and $n = 5$), where

$$X = \begin{pmatrix} -2 \\ -1 \\ 0 \\ 1 \\ 2 \end{pmatrix}; Y^1 = \begin{pmatrix} 6 \\ 9 \\ 15 \\ 19 \\ 26 \end{pmatrix}; Y^2 = \begin{pmatrix} 15 \\ 22 \\ 26 \\ 32 \\ 35 \end{pmatrix}; Y^3 = \begin{pmatrix} 28 \\ 29 \\ 36 \\ 39 \\ 48 \end{pmatrix}; Y^4 = \begin{pmatrix} 35 \\ 44 \\ 47 \\ 54 \\ 55 \end{pmatrix}.$$

Preliminary Step

Computing the regression equation for each of the judges we obtain:

$$B^1 = 5; \quad B^2 = 5; \quad B^3 = 5; \quad B^4 = 5,$$

and the predicted evaluation vectors are

$$\hat{Y}^1 = \begin{pmatrix} 5 \\ 10 \\ 15 \\ 20 \\ 25 \end{pmatrix}; \hat{Y}^2 = \begin{pmatrix} 16 \\ 21 \\ 26 \\ 31 \\ 36 \end{pmatrix}; \hat{Y}^3 = \begin{pmatrix} 26 \\ 31 \\ 36 \\ 41 \\ 46 \end{pmatrix}; \hat{Y}^4 = \begin{pmatrix} 37 \\ 42 \\ 47 \\ 52 \\ 57 \end{pmatrix}.$$

Since we are interested in the value of R^2 in each iteration we compute

$$\sum_{k=1}^4 (Y^k - \hat{Y}^k)'(Y^k - \hat{Y}^k) = 40,$$

and

$$\sum_{k=1}^m \sum_{i=1}^n (Y_{ki} - \bar{y})^2 = 3850.$$

To facilitate the computation, we shall present in every step a matrix A whose (t, r) -th element is equal to the increase in the mean-square Euclidian distances resulting by combining groups G_t with G_r . Thus at each step the two groups corresponding to the minimal value in A are combined.

Step 1

$$G_1 = \{1\} \quad G_2 = \{2\} \quad G_3 = \{3\} \quad G_4 = \{4\}.$$

$$D^2(G_1, G_2, G_3, G_4) = 0,$$

so

$$R^2(G_1, G_2, G_3, G_4) = 1 - \frac{40 + 0}{3850} = .990.$$

Computing the A matrix we get

	G_2	G_3	G_4
G_1	302.5	1102.5	2560
G_2		250	1102.5
G_3			302.5

Thus we combine $\{2\}$ and $\{3\}$, so $G_1 = \{1\}$; $G_2 = \{2,3\}$; $G_3 = \{4\}$, and $D^2\{G_1, G_2, G_3\} = 302.5$,

$$R^2(G_1, G_2, G_3) = 1 - \frac{40 + 302.5}{3850} = .925.$$

Step 2

The new A matrix is

	G_2	G_3
G_1	800.8	2757.5
G_2		800.8

Thus we combine either G_1 with G_2 , or G_2 with G_3 , and let us arbitrarily combine G_1 with G_2 so $G_1 = \{1,2,3\}$, $G_2 = \{4\}$,

$$D^2(G_1, G_2) = 302.5 + 800.8 = 1103.3,$$

$$R^2(G_1, G_2) = 1 - \frac{40 + 1103.3}{3850} = .703.$$

Step 3

The new A matrix is:

	G_2
G_1	1706.7

Obviously we combine the last two groups G_1, G_2 thus obtaining $G_1 = \{1,2,3,4\}$,

$$D^2(G_1) = 1103.3 + 1706.7 = 2810$$

$$R^2(G_1) = 1 - \frac{40 + 2810}{3850} = .260.$$

In summary, the JAN method generated the following groups:

Number of groups	Groups	R^2
4	$\{1\}; \{2\}; \{3\}; \{4\}$.990
3	$\{1\}; \{2,3\}; \{4\}$.925
2	$\{1,2,3\}; \{4\}$.703
1	$\{1,2,3,4\}$.260

However, it can be easily shown that $\{1,2\}; \{3,4\}$ is a better grouping for two groups than the one produced by the JAN since $D^2(\{1,2\}, \{3,4\}) = 605$ so $R^2(\{1,2\}; \{3,4\}) = 1 - (40 + 605)/3850 = .832$ which is greater than .703.

5. CONCLUSIONS

Capturing judges' evaluation policies by regression equations and grouping them while retaining maximum predictive efficiency (R^2) has been used in a variety of decision making situations. In these studies the hierarchical Judgment Analysis (JAN) technique has been applied. It was demonstrated here that this technique does not necessarily generate an optimal clustering. However, it has been shown that the judges' grouping problem can be transformed to an equivalent general clustering problem. In the latter case the objective is to minimize the mean-square Euclidian distances among the judges' predicted evaluation vectors within each cluster. Hence, the judges' grouping problem can be solved by applying a variety of known cluster analysis techniques.

To the best of our knowledge the only clustering method which guarantees an optimal solution is that suggested by Jensen (1969). Unfortunately, its application is limited to about 25 judges. Therefore, whenever a larger problem is presented, it is recommended to use several out of the multitude of available efficient heuristic techniques choosing the one which generates the best solution. An extensive list of such techniques (together with computer programs) can be found in Anderberg's book (1973). It is believed that following such a procedure could lead to better clustering.

REFERENCES

- Andenberg, M. R. *Cluster analysis for applications*. New York: Academic Press, 1973.
- Bottenberg, R. A., & Christal, R. E. Grouping criteria—A method which retains maximum predictive efficiency. *The Journal of Experimental Education*, 1968, 36, 28–34.
- Bottenberg, R. A., & Christal, R. E. *Iterative technique for clustering criteria which retains optimum predictive efficiency*. Lackland Air Force Base, Texas. Personnel Laboratory, Wright Air Development Division, March 1961.
- Castellan, N. J., Jr. Comments on the "Lens Model" equation and the analysis of multiple-cue judgment tasks. *Psychometrika*, 1973, 38, 87–100.
- Christal, R. E. JAN: A technique for analyzing group judgment. *The Journal of Experimental Education*, 1968, 36, 24–27.
- Dudycha, A. L. A Monte-Carlo evaluation of JAN: A technique for capturing and clustering raters' policies. *Organizational Behavior and Human Performance*, 1970, 5, 501–516.
- Dudycha, A. L., & Naylor, J. C. The effect of variations in the cue R matrix upon the obtained policy equation of judges. *Educational and Psychological Measurement*, 1966, 26, 583–603.
- Holmes, G. P., & Zedeck, S. Judgement analysis for assessing paintings. *The Journal of Experimental Education*, 1973, 41, 26–30.

- Jensen, R. E. A dynamic programming algorithm for cluster analysis. *Operations Research*, 1969, 17, 1034–1055.
- Naylor, J. C., & Wherry, R. J., Sr. The use of simulated stimuli and the 'JAN' technique to capture and cluster the policies of raters. *Educational and Psychological Measurement*, 1965, 4, 969–986.
- Slovic, P., & Lichtenstein, S. Comparison of bayesian and regression approaches to the study of information processing in judgement. *Organizational Behavior and Human Performance*, 1971, 6, 649–744.
- Ward, J. H., Jr. Hierarchical grouping to optimize an objective function. *Journal of American Statistical Association*, 1963, 58, 236–244.
- Ward, J. H., Jr., & Hook, M. E. Application of an hierarchical grouping procedure to a problem of grouping profiles. *Educational and Psychological Measurement*, 1963, 23, 69–82.
- Zedeck, S. An information processing model and approach to the study of motivation. *Organizational Behavior and Human Performance*, 1977.
- Zedeck, S., & Kafry, D. Capturing raters' policies for processing evaluation data. *Organizational Behavior and Human Performance*, 1977, 18, 269–294.

RECEIVED: July 16, 1979