ELSEVIER

European Journal of Operational Research 111 (1998) 582–588

EUROPEAN
JOURNAL
OF OPERATIONAL
RESEARCH

Theory and Methodology

# Parametric linear programming and cluster analysis

Anito Joseph [a,*], Noel Bryson [b]

[a] Department of Management Science, School of Business Administration, University of Miami, 417 Jenkins Building,
Coral Gables, FL 33124-8237, USA
[b] School of Business, Howard University, Washington, DC 20059, USA

Received 14 April 1997; accepted 29 September 1997

**Abstract**

In the cluster analysis problem one seeks to partition a finite set of objects into disjoint groups (or clusters) such that each group contains relatively similar objects and, relatively dissimilar objects are placed in different groups. For certain classes of the problem or, under certain assumptions, the partitioning exercise can be formulated as a sequence of linear programs (LPs), each with a parametric objective function. Such LPs can be solved using the parametric linear programming procedure developed by Gass and Saaty [(Gass, S., Saaty, T. (1955), Naval Research Logistics Quarterly 2, 39–45)]. In this paper, a parametric linear programming model for solving cluster analysis problems is presented. We show how this model can be used to find optimal solutions for certain variations of the clustering problem or, in other cases, for an approximation of the general clustering problem. © 1998 Elsevier Science B.V. All rights reserved.

*Keywords:* Clustering; Parametric linear programming; Lagrangian relaxation; Integer programming

## 1. Introduction

In general, the cluster analysis problem involves partitioning a set of $n$ objects into $g \leqslant n$ disjoint groups (or clusters). The measure of effectiveness for the clustering exercise may involve a single criterion, e.g., maximizing within cluster similarity, or, there may be additional criteria, e.g., a specific number of groups in the partition. Usually, the properties of a partition are expressed as a sum of its individual group properties, e.g., total cost, total size. The general class of clustering problems is computationally difficult, potentially requiring examination of an exponential number, $\sum_{g=1}^{n} (1/g!)\left[\sum_{i=1}^{g} (-1)^{g-i} [(g!/(i!(g-i)!))](i^n)\right]$, of partitions. Therefore, as $n$ increases, the associated computational burden makes finding an optimal solution impractical and exact solution approaches are abandoned in favor of practical solution approaches.

The linearly ordered clustering problem is a variation of the clustering problem that is more amenable to solution. In this problem, clusters can only be formed from a given sequential ordering of the objects. The number of possible partitions for a sequentially ordered set of $n$ objects is equal to $\sum_{g=1}^{n} (n-1)!/[(g-1)!\,(n-g)!] = 2^{n-1}$. Even

[*] Corresponding author. Fax: +1-305 284 2321.

0377-2217/98/$ 19.00 © 1998 Elsevier Science B.V. All rights reserved.
*PII* S 0 3 7 7 - 2 2 1 7 ( 9 7 ) 0 0 3 7 9 - 2

though this number is exponential, many problem types can be solved in polynomial time. As a result, the linearly ordered clustering problem has been used to approximate the general clustering problem so that practical solutions can be found, Chen and Yu [1], Hwang [2].

The linearly ordered clustering problem has been explored in the literature. Even though its connection to linear programming (LP) has been known, researchers have focused on dynamic programming approaches, tailoring their algorithms to fit the specific problem instance. In this paper, we focus on linear programming (LP) methods and propose a parametric LP model for the linearly ordered clustering problem. Parametric LP methods can be used to find exact solutions in many situations where a second clustering criterion must be satisfied. We use parametric LP to solve a sample of problems taken from the literature. Other variations of the clustering problem where parametric LP can be applied are also identified.

## 2. Clustering linearly ordered objects

Real world applications of the linearly ordered clustering problem occur as a result of one of two broad conditions. Either, (a) objects must be kept in a given order (naturally occurring or as desired by the investigator) Everitt [3], Gordon [4,5], Kernighan [6] or, (b) the mathematical structure of the clustering objective determines that the optimal partition must come from a sequential ordering of objects; Hwang [2], Chakravarty et al. [7], Anily and Federgruen [8].

While dynamic programming has been the popular solution approach used in the literature, the lack of a canonical form and general purpose solution algorithm means that specific algorithms must be devised to handle each application. As the number of objects increase, dynamic programming approaches usually experience inordinate growth in computing resources which results in reduced efficiency. The presence of side conditions can also increase the problem complexity for dynamic programming, and algorithms must be further specialized to handle side conditions. For example, capacity considerations may mean limiting the size of any group, Kernighan [6] or, economic considerations may mean restricting the number of groups in the partition, Chakravarty et al. [7]. Lagrangian relaxation can be used to handle some side conditions, Stanfel [9,10]; however, dynamic programming methods provide no information on the appropriate magnitude or range of values for the Lagrangian multiplier.

## 3. A parametric linear programming model

It is well known that the cluster analysis problem can be formulated as an integer programming problem (IP), Vinod [11], Rao [12], called a set partitioning problem (SPP). The SPP belongs to a class of computationally difficult problems and this approach has not been effective for solving the general class of cluster analysis problems. The linear ordering of objects for clustering means that the matrix representation of the clusters has special structure which can be exploited to solve the problem efficiently.

The SPP formulation of the single-criterion clustering problem for $n$ objects may be stated as:

$$\mathbf{P}_1\text{: } \{\min \mathbf{CX} \mid \mathbf{AX} = \mathbf{1};\ \mathbf{0} \leqslant \mathbf{X} \leqslant \mathbf{1} \text{ and binary}\},$$

where $\mathbf{X} = \{X_j\text{: } j = 1, \ldots, N\}$; $X_j = 1$ if group $j$ is in the partition, 0 otherwise; $N$ is the total number of possible groups; $\mathbf{C} = \{C_j\}$, and $C_j$ is the criterion measure for group $j$; $\mathbf{A} = \{A_j\}$, and $A_j$ is an $n$-dimensional vector associated with group $j$ such that $a_{ij} = 1$ indicates that object $i$ is a member of group $j$ and $a_{ij} = 0$ otherwise; $\mathbf{0}$ and $\mathbf{1}$ are $n$-dimensional vectors of 0's and 1's, respectively.

Restricting the clusters to be formed from a linear ordering of objects means that each column of $\mathbf{A}$ has its 1's in consecutive rows. Such a matrix is described as an interval matrix, Nemhauser and Wolsey [13], and interval matrices are known to be totally unimodular. The LP relaxation of $\mathbf{P}_1$ for the linearly ordered set of objects would thus possess the integrality property and its solution is a valid partition. Note that, for the linearly constrained problem, the number of columns $N$ in $\mathbf{P}_1$ is equal to $n(n + 1)/2$. This is a polynomial function of $n$, in comparison to the total number of possible partitions which is an exponential function of $n$. This is because the LP model represents the possible partitions as extreme points of the LP polytope. Hence, LP provides an efficient data structure for storing the problem information.

In some cases, a second clustering criterion can be handled as an additional constraint in the LP model. Problem $\mathbf{P}_2$ is formulated: $\{\min \sum_j C_j X_j \mid \mathbf{AX}= \mathbf{1};\ \sum_j D_j X_j = K;\ X_j \text{ binary}\}$, where $D_j$ represents the coefficient for group $X_j$ in the second criterion, and $K$ is a specified value that must be achieved for the second criterion. An additional constraint means that the special structure may not be preserved and the solution to the LP relaxation of $\mathbf{P}_2$ may be non-integer. This constraint can be moved to the objective function and a new problem formulated:

$$\mathbf{P}_3\text{:}\ \left\{ \min \sum_j C_j X_j + \theta \left( \sum_j D_j X_j \right) \mid \mathbf{AX} = \mathbf{1};\ X_j \text{ binary} \right\}.$$

Problem $\mathbf{P}_3$ is totally unimodular and given a value $\theta$, LP techniques will find an integer solution for the relaxed problem. Further, problem $\mathbf{P}_3$ is a linear programming problem with a parametric objective function. Hence, the parametric programming approach proposed by Gass and Saaty [14] can be applied to solve problem $\mathbf{P}_3$ exactly.

## 4. Parametric programming approach

The parametric programming procedure of Gass and Saaty [14] solves the following linear programming problem for every value of $w$:

$$\mathbf{P}_4\text{:}\ \{\min Z_0(\mathbf{X}) + w Z_1(\mathbf{X}) \mid \mathbf{AX} = \mathbf{b};\ \mathbf{X} \geqslant 0\}$$

where $Z_0(\mathbf{X})$ and $Z_1(\mathbf{X})$ are linear functions of $\mathbf{X}$, and $w$ is a scalar multiplier. Using the simplex method, the parametric programming procedure first solves problem $\mathbf{P}_4$ for $w = w_0$, where $w_0$ is some arbitrary small number. The result is that either there is some efficient extreme-point solution $X'$, or that there is no finite minimum for $w = w_0$.

Let NB($X'$) be the index set of non-basic variables associated with $X'$ and, let the reduced costs have the form $(h_j + wg_j)$ where $h_j$ and $g_j$ are the reduced costs in terms of $Z_0(\mathbf{X})$ and $Z_1(\mathbf{X})$, respectively. For optimality we require that $(h_j + wg_j) \leqslant 0$. Therefore, $X'$ is the optimal solution of problem $\mathbf{P}_4$ for any $w \in [w'_L, w'_U]$ where $w'_L$, $w'_U$ are determined as follows:

$$\begin{aligned} w'_L = &-h_{qL}/g_{qL} = \max\{-h_j/g_j\text{: } g_j < 0;\\ &j \in \mathrm{NB}(X')\} \text{ or}\\ &-\infty \quad \text{if } g_j \geqslant 0 \ \text{ for all } j \in \mathrm{NB}(X') \end{aligned}$$

$$\begin{aligned} w'_U = &-h_{qU}/g_{qU} = \min\{-h_j/g_j\text{: } g_j > 0;\\ &j \in \mathrm{NB}(X')\} \text{ or}\\ &+\infty \quad \text{if } g_j \leqslant 0 \ \text{ for all } j \in \mathrm{NB}(X'). \end{aligned}$$

The procedure is terminated if either of the following conditions occurs: (a) $w'_U = +\infty$, (b) $w'_U$ is finite but all the corresponding $a_{i,qU} \leqslant 0$. Otherwise, the simplex method introduces $x_{qU}$ into the basis and eliminates the basic variable in the usual manner. Gass and Saaty [14] established that the resulting basis yields a minimum for at least one value of $w$, and that if $[w''_L, w''_U]$ is the interval for which the resulting basis yields a minimum then $w'_U = w''_L$.

Thus, the parametric programming procedure generates the set of efficient, extreme points that solve the single parameter LP problem for all $w$ such that $-\infty \leqslant w \leqslant +\infty$ and will identify the parametric interval that is associated with each such extreme point. It can be seen that problems $\mathbf{P}_3$ and $\mathbf{P}_4$ are equivalent therefore, the parametric programming procedure may be used to solve problem $\mathbf{P}_3$ for all relevant values of $\theta$, and:

(a) the precise range of values for $\theta$ for each solution vector is determined;

(b) only a single LP problem will have to be solved.

## 5. Solving an approximation of a general clustering problem

We apply the parametric linear programming model to a clustering problem studied in Stanfel [9,10]. The grouping criterion to be minimized is the average dissimilarity within groups relative to the average dissimilarity between groups. This form of objective function is common in the cluster analysis literature and assesses a partition on its performance both within the groups and between the groups. A distance measure is used to represent dissimilarity/similarity between objects. The clustering problem can be expressed as

$$\mathbf{P}_1\text{: } \min F = \frac{\sum_i \sum_{r>i} d_{ir} Y_{ir}}{\sum_i \sum_{r>i} Y_{ir}} - \frac{\sum_i \sum_{r>i} (1 - Y_{ir}) d_{ir}}{\sum_i \sum_{r>i} (1 - Y_{ir})},$$

where $d_{ir}$ is the distance between objects $i$ and $r$, and $Y_{ir}$ is a binary variable that is used to indicate whether objects $i$ and $r$ are placed in the same group ($Y_{ir} = 1$) or in different groups ($Y_{ir} = 0$), $i$, $r \in \{1, \dots, n\}$.

Using the development of Stanfel [9], the nonlinear objective function is transformed into a linear function by letting $M = n(n - 1)/2$, $\sum_i \sum_{r>i} d_{ir} = C$, and restricting $\sum_i \sum_{r>i} Y_{ir}$ to equal some constant $K$. The objective function can be expressed as

$$G(K) = \left[ M \sum_i \sum_{r>i} d_{ir} Y_{ir} / (K(M - K)) \right] - [C/(M - K)].$$

Minimizing the objective function, $F$, becomes equivalent to minimizing $\sum_i \sum_{r>i} d_{ir} Y_{ir}$ given a value for $K$. Note that, when all objects are in separate groups then, $K = 0$, and $F = -C/M$; when all objects belong to one group, then $K = M$, and $F = \sum_i \sum_{r>i} d_{ir} Y_{ir} / \sum_i \sum_{r>i} Y_{ir} = C/M$. Also, it may not be possible for $K$ to attain all values between 0 and $M$, for example, when $n = 5$, then, $K \in \{0,1,2,3,4,6,10\}$, and there are no partitions such that $K = 5$, 7, 8, or, 9.

The restricted problem may be formulated as a set partitioning problem:

$$\mathbf{P}_2\text{: } \left\{ \min \sum_j C_j X_j \mid \mathbf{AX} = \mathbf{1}; \sum_j M_j X_j = K;\ X_j \text{ binary} \right\}$$

where $C_j = \sum_i a_{ij} a_{rj} d_{ir}$, $i < r$, is the total of the within group distances in group $j$; $n_j$ = the number of objects in group $j$, and $M_j = n_j(n_j - 1)/2$. Stanfel [9] proposed to solve an approximation of problem $\mathbf{P}_2$ by (i) assuming that the objects must be kept in a given linear order, and (ii) using Lagrangian relaxation to move the side constraint to the objective function. The new problem $\mathbf{P}_3$ does possess the integrality property:

$$\mathbf{P}_3\text{: } \left\{ \min \sum_j (C_j + vM_j) X_j \mid \mathbf{AX} = \mathbf{1};\ X_j \in (0, 1) \right\}.$$

Stanfel's [9] solution approach involved (a) establishing a range and step length for the multiplier; (b) generating problems for a range of multiplier values; (c) applying the simplex method to each of the problem generated; (d) using dynamic programming techniques to price and select the next column for entry into the basis. In a later approach, Stanfel [10], the simplex method is eliminated, and the problem is solved using only dynamic programming methods.

In a parametric programming approach, if we let $Z_0(\mathbf{X}) = \sum_j C_j X_j$, $Z_1\ (\mathbf{X}) = \sum_j -M_j X_j$, $w = -v$ {and $\mathbf{b} = \mathbf{1}$, then problem $\mathbf{P}_3$ is just a special case of problem $\mathbf{P}_4$ with $w \in (0, +\infty)$ and so the parametric programming procedure may be used to find optimal solutions for all relevant values of $v$. Since the parametric procedure obtains the exact values and ranges for the optimal multipliers, there is no chance of missing an optimal solution.

## 6. Computational results

We demonstrate the parametric linear programming model using eight examples of problem type $\mathbf{P}_3$. Examples E1–E4 are taken from Stanfel [9] and E5–E8 are taken from Stanfel [10]. The object coordinates are given in Table 1.

Example E1 has 20 objects; examples E2–E4 all have 50 objects, example E5 has 91 objects, and examples E6–E8 all have 100 objects. The examples E5 and E6 show obvious structure; for example E6 the structure is well defined both within and between groups. We used the parametric programming feature of IBM's Optimization Subroutine Library (OSL) to solve the four problems. Table 2 contains descriptions and corresponding test results.

Table 1
Object coordinates

| | |
|---|---|
| E1 | 0, 1, 5, 7, 10, 15, 17–20, 25, 27, 30, 32, 34, 35, 40, 45, 47, 55 |
| E2 | 0, 1, 3, 4, 7, 9–14, 16–22, 26–32, 40, 41, 43, 45, 46, 49, 51, 53–56, 58, 60, 70, 73, 76, 78, 79, 80, 86–88, 95, 96, 99 |
| E3 | 0, 1, 5, 7, 10, 15, 17–20, 25, 27, 30, 32, 34, 35, 40, 45, 47, 55–57, 60, 62, 64–67, 70–81, 85–89, 92–96 |
| E4 | 0–3, 7, 9–14, 17–20, 23, 26–36, 40, 42–45, 50, 51, 53, 54, 56–58, 60, 65, 70, 75, 80, 82–84, 90, 91, 98 |
| E5 | 0–9, 30–59, 120–149, 210–220, 270–279 |
| E6 | 0–9, 30–39, 60–69, 90–99, 120–129, 150–159, 180–189, 210–219, 240–249, 270–279 |
| E7 | 0, 2, 3, 5, 6, 8, 10, 12, 15, 20, 45–54, 70–89, 130–134, 160–189, 200, 202, 205, 228, 229, 255, 260, 265, 270, 275, 300, 302, 304, 308, 316, 325, 329, 333, 339, 346, 400, 430, 460, 500, 505 |
| E8 | 0–10, 12, 15–17, 20, 22, 24, 26, 30–33, 60, 63, 64, 66, 68, 71, 73, 75, 77, 80, 82, 83, 85, 86, 89, 92, 94, 95, 97, 100, 102, 104, 105, 108, 112, 114, 115, 119, 120, 122, 170, 174, 175, 179, 182, 183, 185, 190, 191, 195, 198, 202, 205, 210, 212, 214, 216, 219, 223, 225, 404, 410, 415, 420, 425–427, 429, 431, 433, 435, 439, 442, 446, 449, 451, 453, 455, 460, 463, 610, 612, 615, 620, 623, 625, 631 |

Given the high degree of primal degeneracy in this problem class, multiple pivots are sometimes required to obtain a new distinct solution. The ratio of No. of Pivots to No. of Distinct Solutions is 5.74 for E1 with 20 objects, 7.75 for E2, 7.92 for E3 and 8.98 for E4 with 50 objects. For E5 with 91 objects, this ratio is 10.52, while for E6, E7 and E8 with 100 objects each, the ratio values are 9, 9.30, and 8.74, respectively. These ratios are small and do not indicate a serious problem with degeneracy. Note that the optimal solution obtained for example E4 is superior to that obtained by Stanfel [9]; this is because the range of values explored for the Lagrangian multiplier did not include the optimal value.

The results show that the total number of pivots required by the parametric procedure represents only a minute fraction of the total number of possible partitions (or extreme points of the problem), e.g. for $n = 20$, the total number of pivots required represents 155/524, 288 or 0.0295% of the total number of extreme points. This indicates that LP is highly efficient for solving the linearly ordered clustering problem.

In example E1, the optimal solution was found after 19 pivots. For examples E2–E4, the optimal solution was found after 301, 462, and 530 pivots, respectively. For examples E5–E8, the optimal solution was found after 868, 503, 1144, and 1147 pivots. The relatively small number of pivots to obtain the optimal solution for E6 is in keeping with its well-defined structure, both within and between groups. Given the number of objects, the number of pivots to find the optimal solution increases as the presence of structure decreases. The less defined the problem structure, the smaller the

Table 2
Computational results

| Problem | No. of objects | Optimal $F$ | No. of pivots | No. of distinct solutions | Solution time (s) |
|---|---|---|---|---|---|
| E1 | 20 | −17.93 | 155 | 27 | 1.59 |
| E2 | 50 | −35.58 | 496 | 64 | 24.56 |
| E3 | 50 | −38.86 | 539 | 68 | 27.07 |
| E4 | 50 | −32.43 | 557 | 62 | 29.18 |
| E5 | 91 | −114.80 | 905 | 86 | 17.64 |
| E6 | 100 | −106.33 | 846 | 94 | 19.05 |
| E7 | 100 | −234.86 | 1144 | 123 | 90.64 |
| E8 | 100 | −304.44 | 1162 | 133 | 81.96 |

number of clusters in the optimal partition. Therefore, for problems E4, E7, and E8, the number of pivots to arrive at the optimal solution is large compared to the total number of pivots to solve the respective problem.

The advantages of the parametric programming procedure for solving the class of clustering problem studied in this paper are most obvious when one considers how the search parameters for the dynamic programming model were determined. In order to specify appropriate values and step sizes for the Lagrangian multiplier, Stanfel [10] first conducted sample runs and studied problem behavior to deduce appropriate search parameters. These sample runs suggested the following characteristics: (a) nearly unimodal behavior of $F$ as $v$ decreases; (b) global optimal solutions persist for relatively wide ranges of multiplier values $v$, this also occurs for other optimal solutions; (c) integer values of $v$ exist which will find the global optimal solution.

Based on these observations a heuristic procedure for setting and altering the step size $v$ was then developed. Stanfel [9,10] noted the disadvantages of using a heuristic procedure for multiplier generation. Namely, (1) different multipliers may result in the same solution, and (2) the optimal solution may be omitted because the appropriate value of the multiplier was missed. The characteristics (a)–(c) also reflect a relatively well-defined behavior for the given objective function and present favorable conditions for the use of heuristic methods for multiplier generation. This well-defined behavior does not necessarily persist for other clustering criteria, and additional study would be required to develop search parameter values.

## 7. Extensions

If a second clustering criterion can be expressed as a linear constraint on the set of possible clusters, then a parametric LP solution approach should be explored. For example, a typical clustering problem minimizes the within group sum of squared distances about the group centroid and requires a specific number of groups $g$ in the optimal partition, Gordon [5], Everitt [3]. Thus $F = \sum_j C_j X_j$, where $C_j$ is the sum of squared distances for group $j$, and an additional constraint $\sum_j X_j = g$ is added to the LP. A firm value for $g$ is usually not known at the outset and a number of values for g are investigated. This is a straightforward application of the parametric LP procedure and a search will provide solutions for the instances where a multiplier exists for $\sum_j X_j = g,\ 1 \leqslant g \leqslant n$. An inventory management application is discussed in Chakravarty et al. [7]; Chen and Yu [1] discuss an application in concurrency control in data sharing environments.

Another popular clustering criterion minimizes the difference between the average within-group sum of squared distances and the average between-group sum of squared distances, i.e., $F = \left(\sum_i \sum_{r>i} d_{ir}^2 Y_{ir} / \sum_i \sum_{r>i} Y_{ir}\right) - \left(\sum_i \sum_{r>i} d_{ir}^2 (1 - Y_{ir}) / \sum_i \sum_{r>i} (1 - Y_{ir})\right)$. This criterion can be handled in a similar manner as was done in Section 5. The resulting problem can then be solved using parametric LP.

The parametric LP approach can handle many side conditions by manipulating the columns of the constraint matrix. For example, columns for clusters that do not satisfy capacity limits or membership conditions are eliminated from the problem matrix. Limits on group cost can also be handled by eliminating the relevant columns. Eliminating columns reduces the problem size and the computational effort required for solution. In addition, the special structure of the linearly constrained clustering problem, can be exploited in achieving efficient input and storage of problem data.

## 8. Conclusion

In this paper a technique for locating exact solutions for an important class of grouping problems has been presented. It offers an improvement over other techniques in that the user does not have to settle for near optimal solutions, and it does not require that the user specify a procedure for generating the set of Lagrangian multipliers. The approach is practical, examining a minute fraction of the total possible partitions and requiring relatively small computational effort to do so. The storage requirements are a polynomial function of the number of objects rather than the total number of possible partitions.

The technique presented here has implications for solution strategies for classes of partitioning problems beyond that considered in this paper. The problem class considered here is also an approximation for other classes of partitioning problems for which there are no known strategies for their exact solution. Thus it might be useful to solve such problems by obtaining the exact solutions for a set of associated approximate problems rather than to attempt to obtain an approximate solution for the original problem. There are no existing data on the effectiveness of linearly ordered approximations of the general clustering problem. This is so probably because it was not feasible to conduct such studies efficiently. The technique presented in this paper can be used to study approximation schemes for different clustering problems and find exact solutions efficiently. Therefore, future research projects will use the parametric LP approach to investigate the performance of linearly ordered approximation schemes for various clustering problems.

## References

[1] M.-S. Chen, P.S. Yu, Optimal design of multiple hash tables for concurrency control, IEEE Transactions on Knowledge and Data Engineering 9 (1997) 384–390.

[2] F. Hwang, Optimal partitions, Journal of Optimization Theory and Application 34 (1981) 1–10.

[3] B.S. Everitt, Cluster Analysis, Halsted Press, Wiley, New York, 1993.

[4] A. Gordon, Classification in the presence of constraints, Biometrics 29 (1973) 821–827.

[5] A. Gordon, Classification, Chapman and Hall, London, 1980.

[6] B. Kernighan, Optimal sequential partitions of graphs, Journal of the Association for Computing Machinery 18 (1) (1971) 34–40.

[7] A. Chakravarty, J. Orlin, V. Rothblum, A partitioning problem with additive objective with an application to optimal inventory groupings for joint replenishment, Operations Research 30 (1982) 1018–1020.

[8] S. Anily, A. Federgruen, Structured partioning problems, Operations Research 39 (1991) 130–149.

[9] L. Stanfel, An algorithm using Lagrangian relaxation and column generation for one-dimensional clustering problems, in: J. Rustagi, S. Zanakis (Eds.), Optimization in Statistics, The Institute of Management Science, Providence, RI, 1982, pp. 165–185.

[10] L. Stanfel, Recursive Lagrangian method for clustering problems, European Journal of Operational Research 27 (3) (1986) 332–342.

[11] H. Vinod, Integer programming and the theory of grouping, Journal of The American Statistical Association 64 (1969) 506–519.

[12] M. Rao, Cluster analysis and mathematical programming, Journal of the American Statistical Association 66 (1971) 622–626.

[13] G.L. Nemhauser, L.A. Wolsey, Integer and Combinatorial Optimization, Wiley, New York, 1988.

[14] S. Gass, T. Saaty, The computational algorithm for the parametric objective function, Naval Research Logistics Quarterly 2 (1955) 39–45.