SOLUTION OF THE PERSONNEL CLASSIFICATION PROBLEM WITH THE METHOD OF OPTIMAL REGIONS*

PAUL S. DWYER

UNIVERSITY OF MICHIGAN

The personnel classification problem is identified mathematically with other problems in the social and biological sciences. This mathematical problem is shown to be a special case of the general mathematical problem of linear programming. It is proposed here that the personnel classification problem may be solved directly by methods particularly appropriate to it as well as by the simplex method, which is a standard method for solving the general linear programming problem. The method of optimal regions is derived and illustrated in this paper.

1. *Introduction*

The general type of problem considered in this paper is first illustrated with a trivial problem. A business concern needs to fill three jobs which demand different abilities and training. Three applicants who can be hired for identical salaries are available. Because of the different abilities, training, and experiences, however, the value of each applicant to the company depends upon the job in which he is placed. The estimate of the value of each applicant to the company each year if he were to be assigned to any one of the three jobs is given in Table 1.

It is desired to assign the applicants to the jobs in such a way that the total value to the company is as great as possible. In this problem there are $3! = 6$ possible alternative assignments of the three men and the greatest estimated value to the company, \$24,000 per year, is obtained by assigning applicant 1 to job 3, applicant 2 to job 1, and applicant 3 to job 2.

A more general problem is arrived at by denoting by c_{ij} the contribution of individual i to the common effort if he is assigned to job j, in units of the measure of the common effort. Thus in Table 1, $c_{11} = 5$, $c_{12} = 4$, $c_{13} = 7$, $c_{21} = 6$, etc. Sometimes dimensionless c_{ij} may be used to express the relative contributions to the common effort.

A more general problem features N individuals and N jobs in which, as above, the contribution of individual i to the common effort is c_{ij} units if he

is assigned to job j. The assignments are then made so that the sum of the corresponding c_{ij} values is a maximum. We set $x_{ij} = 1$ when individual i is assigned to job j and $x_{ij} = 0$ otherwise. Then

$$\sum_i x_{ij} = 1, \qquad \sum_j x_{ij} = 1, \qquad \sum_{i,j} x_{ij} = N, \qquad (1.1)$$

and the problem is to maximize*

$$T = \sum_{i,j} x_{ij}c_{ij} . \qquad (1.2)$$

There are $N!$ values of T, though some of them may be equal, and for N large it becomes impractical to examine them all to determine the maximum. Improved methods of solution are needed.

Frequently there are many identical jobs or jobs which, though not identical, demand the same basic qualifications indicated by the c_{ij} . Such jobs can be combined into a job category. If there are m categories and if the number of jobs grouped in the jth job category is q_j , we have

$$x_{ij} = 1 \text{ if } i \text{ is assigned to } j \qquad (1.3)$$

$$x_{ij} = 0 \text{ otherwise,}$$

and

$$\left. \begin{array}{l} \sum_{i=1}^{N} x_{ij} = q_j \\[2em] \sum_{j=1}^{m} x_{ij} = 1 \end{array} \right\} \quad \text{with} \quad \sum_{i,j} x_{ij} = \sum_j q_j = \sum_i 1 = N. \qquad (1.4)$$

We wish to maximize the total contribution to the common effort

$$T = \sum_{i,j} x_{ij}c_{ij} . \qquad (1.5)$$

The q_j values are called quotas. This formulation of the problem is illustrated in Table 2.

Sometimes there are different individuals having identical c_{ij} values or at least the values are close enough so that the individuals may be grouped without serious error. In this case we group these individuals into personnel categories. If there are n personal categories and the number of individuals grouped in the ith personnel category is f_i , we have

$$x_{ij} = 1 \text{ if some individual in } i \text{ is assigned to } j \qquad (1.6)$$

$$x_{ij} = 0 \text{ otherwise,}$$

*In certain situations, particularly when the smaller values of c_{ij} indicate the greater contributions to the common effort, the values of T should be minimized.

TABLE 1

Estimated Value Per Year

in Units of $1000

		Job		
		1	2	3
Applicant	1	5	4	7
	2	6	7	3
	3	8	11	2

TABLE 2

Use of Job Categories

$i \backslash j$	1	2	3	...	j	...	m	
1	c_{11}	c_{12}	c_{13}	...	c_{1j}	...	c_{1m}	1
2	c_{21}	c_{22}	c_{23}	...	c_{2j}	...	c_{2m}	1
3	c_{31}	c_{32}	c_{33}	...	c_{3j}	...	c_{3m}	1
...	...	...	...	...	...	...	...	1
i	c_{i1}	c_{i2}	c_{i3}	...	c_{ij}	...	c_{im}	1
...	...	...	...	...	...	...	...	1
N	c_{N1}	c_{N2}	c_{N3}	...	c_{Nj}	...	c_{Nm}	1
q_j	q_1	q_2	q_3	...	q_j	...	q_m	N

TABLE 3

Use of Personnel Categories

$i \backslash j$	1	2	3	...	j	...	N	f_i
1	c_{11}	c_{12}	c_{13}	...	c_{1j}	...	c_{1N}	f_1
2	c_{21}	c_{22}	c_{23}	...	c_{2j}	...	c_{2N}	f_2
3	c_{31}	c_{32}	c_{33}	...	c_{3j}	...	c_{3N}	f_3
...	...	...	...	...	...	...	...	...
i	c_{i1}	c_{i2}	c_{i3}	...	c_{ij}	...	c_{iN}	f_i
...	...	...	...	...	...	...	...	...
n	c_{n1}	c_{n2}	c_{n3}	...	c_{nj}	...	c_{nN}	f_n
q_j	1	1	1	...	1	...	1	N

TABLE 4

Use of Personnel Categories and Job Categories

$i \backslash j$	1	2	3	...	j	...	m	f_i
1	c_{11}	c_{12}	c_{13}	...	c_{1j}	...	c_{1m}	f_1
2	c_{21}	c_{22}	c_{23}	...	c_{2j}	...	c_{2m}	f_2
3	c_{31}	c_{32}	c_{33}	...	c_{3j}	...	c_{3m}	f_3
...	...	...	...	...	...	...	...	...
i	c_{i1}	c_{i2}	c_{i3}	...	c_{ij}	...	c_{im}	f_i
...	...	...	...	...	...	...	...	...
n	c_{n1}	c_{n2}	c_{n3}	...	c_{nj}	...	c_{nm}	f_n
q_j	q_1	q_2	q_3	...	q_j	...	q_m	N

TABLE 5

Identification of Personnel Classification

with Linear Programming

General	Personnel (Generalization of Table 1)	Personnel (Table 2)	Personnel (Table 3)	Personnel (Table 4)
i	i	i	i	i
j	j	j	j	j
n	N^2	Nm	Nn	mn
λ_i	x_{ij}	x_{ij}	x_{ij}	x_{ij}
c_i	c_{ij}	c_{ij}	c_{ij}	c_{ij}
m	2N	N+m	N+n	m+n
b_j	1	q_j or 1	1 or f_i	f_i or q_j
a_{ij}	1 or 0	1 or 0	1 or 0	X or 0

and

$$\left. \begin{array}{l} \sum_{i=1}^{n} x_{ij} = 1 \\[2em] \sum_{j=1}^{N} x_{ij} = f_i \end{array} \right\} \quad \text{with} \quad \sum_{i,j} x_{ij} = \sum_{i} f_i = \sum_{j} 1 = N. \quad (1.7)$$

The values of f_i are called frequencies. This formulation of the problem is illustrated in Table 3.

A common form of the problem uses both personnel categories and job categories. If, as before, f_i is the number of persons in personnel category i and q_j is the number to be placed in job category j, then

$$x_{ij} = \text{zero or a positive integer indicating the number of persons} \quad (1.8)$$
$$\text{in personnel category } i \text{ assigned to job category } j.$$

Then,

$$\left. \begin{array}{l} \sum_{j}^{m} x_{ij} = f_i \\[2em] \sum_{i}^{n} x_{ij} = q_j \end{array} \right\} \quad \text{with} \quad \sum_{i,j} x_{ij} = \sum_{i} f_i = \sum_{j} q_j = N. \quad (1.9)$$

This formulation of the problem is illustrated in Table 4.

Table 3 is a special case of Table 4 when $q_j = 1$; Table 2 is a special case of Table 4 when $f_i = 1$, and the square N by N matrix results when $f_i = 1$ and $q_j = 1$.

The problem under discussion is closely related to problems previously discussed in *Psychometrika* by Brogden (1), Thorndike (9), Votaw (11) and Lord (5). The formulation of Table 2, used essentially by Brogden (1, 149), Thorndike (9, 233), and Lord (5, 297), appears to be a natural form of the personnel classification problem in which several or many men are to be assigned to a relatively small number of job categories. The formulation in Table 4 has been used by Votaw (11, 257) in identifying the problem as a linear programming problem.

2. Equivalent Mathematical Problems

This problem is essentially the equivalent, mathematically, of problems arising in other applied fields. For example, it is similar to the Hitchcock transportation problem, solutions and theory of which have been supplied by many, including Dantzig (3) and Flood (6). This calls for the shipping of x_{ij} unit parcels from origin i to destination j when a_i parcels are to be shipped from origin i, b_j parcels are to be delivered to destination j, so that the total transportation cost $\sum_{i,j} x_{ij} c_{ij}$ is to be a minimum if c_{ij} is the cost of shipping a unit parcel from origin i to destination j. Equations (1.8), (1.9), and Table

4 are immediately applicable to this problem if f_i is replaced by a_i and q_i by b_i .

Rao (**7**, 322–329) has shown how different problems in biometric classification reduce to the mathematical problem discussed above. In one of these, for example, it is desired to classify an individual into one of m classes on the basis of the measurements of the individual. We desire to minimize the total risk when c_{ij} is the expected risk of erroneous classification resulting from placing individual i in class j when he does not belong there. The equations (1.6), (1.7), and Table 2 are applicable.

Also, von Neumann (**10**) has related the personnel classification problem to a certain zero-sum two-person game.

These problems and other similar ones are all special cases of the general problem of linear programming. Dantzig (**2**) has presented a mathematical model which is applicable to many problems, after reduction, in linear programming as follows:

Find the values of λ_1 , λ_2 , $\cdots$, λ_n which maximize the linear form

$$\lambda_1 c_1 + \lambda_2 c_2 + \cdots + \lambda_n c_n$$

subject to the conditions that $\lambda_j \geq 0, j = 1, 2, \cdots, n$, and

$$\lambda_1 a_{11} + \lambda_2 a_{12} + \cdots + \lambda_n a_{1n} = b_1$$
$$\lambda_1 a_{21} + \lambda_2 a_{22} + \cdots + \lambda_n a_{2n} = b_2$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\lambda_1 a_{m1} + \lambda_2 a_{m2} + \cdots + \lambda_n a_{mn} = b_m ,$$

where a_{ij} , b_i , c_j are constants ($i = 1, 2, \cdots, m; j = 1, 2, \cdots, n$).

The personnel classification problem is a special case of this problem. The correspondence between the general problem and (a) the problem expressed in the individual form of which Table 1 is a special case when $N = 3$, (b) the forms of Tables 2, 3, and 4 is indicated in Table 5.

Since the classification problem, which is hereafter used to designate the mathematical problem (including minimization) encompassing the alternative problems described above, is a special case of linear programming, it follows that methods of solution of the general linear programming problem are available. In particular the simplex method is available. This has been pointed out by Votaw (**11**) to be a solution of the classification problem and furthermore a solution that in some cases can be carried out with a high speed computing machine. Even though modifications of the simplex method, as applied to the classification problem, are possible, this is not the only useful method available. It is the purpose of this paper to introduce the method of optimal regions and to show how it can be used.

It does not appear at first that the personnel classification problem is considerably simpler, mathematically, than the general linear programming problem. However, examination of the last line of Table 5 shows that the a_{ij} are integral. This fact seems to indicate a much simpler solution than is possible with the less restricted a_{ij}, for example, those which usually appear in game theory. The implication is that we should solve the classification problem directly from the mathematical model rather than to use a method of solution which is designed to handle the more general problem.

3. *Different Forms of the Problem*

Tables 1, 2, 3, and 4 illustrate four different forms in which the problem may appear. We may call the form of Table 1, which features an N by N matrix, the *individual form* since neither individuals nor jobs are grouped; the form of Table 2, the *quota form* since the quotas for each job category are specified; the form of Table 3, the *frequency form* since the frequencies of the personnel categories are specified; and the form of Table 4, the *frequency-quota form*. Actually the frequency form appears seldom in classification problems so we can drop it from further consideration and use the term "frequency form" to described the situation in Table 4. In fact, this frequency form is usually not the natural form of the classification problem but results from the grouping of individuals which is demanded when the problem is to be worked by the previously available, simplex method.

Each of these forms describes a situation in which the number of men is identical with the number of jobs. Alternative situations can be reduced formally to the forms above with the introduction of dummy men, or dummy jobs, having $c_{ij} = 0$. This is illustrated in Table 6, where the c_{ij} values for ten men and two job categories are given. In problem (a) the quotas are $q_1 = 4$ and $q_2 = 4$. It is sufficient to have a dummy job category with $q_3 = 2$ to complete the assignment. In problem (b) the quotas are $q_1 = 6$ and $q_2 = 6$. Two dummy men are introduced.

Occasionally a problem arises in which no quotas are specified. The solution in this case is extremely simple. Assign each person according to his job profile, i.e., to the job category for which his c_{ij} value is largest. This solution gives the greatest possible T, irrespective of quotas, and is sometimes included with that of the quota solution in order to show the effect of the quota restrictions.

4. *The Case of Two Job Categories*

There is no assignment problem when there is only one job category. For the case of complete quotas all individuals are assigned to this job category and, for the case of partial quotas, the individuals are assigned in the decreasing order of the c_{i1} values.

The problem really begins with two job categories. As was pointed out

TABLE 6

(a). Personnel Surplus

q_j	4	4	2
i	c_{11}	c_{12}	c_{13}
1	23	13	0
2	23	17	0
3	30	20	0
4	26	22	0
5	28	25	0
6	29	27	0
7	29	30	0
8	25	33	0
9	25	37	0
10	30	49	0

(b). Personnel Deficiency

q_j	6	6
i	c_{11}	c_{12}
1	23	13
2	23	17
3	30	20
4	26	22
5	28	25
6	29	27
7	29	30
8	25	33
9	25	37
10	30	49
11	0	0
12	0	0

TABLE 8

Subtraction of k from c_{12}

	4	6				
	4-1	5-1				
	5-1	4-1				
	6-0	4-0				
q_j	4	6				
i	c_{11}	c_{12}	$J^{(1)}$	$J^{(2)}$	$J^{(3)}$	J
1	23	13	1	1	1	1
2	23	17	1	1	1	1
3	30	20	1	1	1	1
4	26	22	1	1	1	1
5	28	25	1	1	1,2	2
6	29	27	1	1,2	2	2
7	29	30	2	2	2	2
8	25	33	2	2	2	2
9	25	37	2	2	2	2
10	30	49	2	2	2	2
c_1	0	0	308			
c_1	0	-2		308		
c_1	0	-3				303

TABLE 7

Use of Differences

q_j	4	6						
i	c_{11}	c_{12}	$c_{11}-c_{12}$	J	c_{1J}	$c_{11}-c_{12}-3$	J	$c_{12}+3$
1	23	13	10	1	23	7	1	16
2	23	17	6	1	23	3	1	20
3	30	20	10	1	30	7	1	23
4	26	22	4	1	26	1	1	25
5	28	25	3	2	25	0	2	28
6	29	27	2	2	27	-1	2	30
7	29	30	-1	2	30	-4	2	33
8	25	33	-8	2	33	-11	2	36
9	25	37	-12	2	37	-15	2	40
10	30	49	-19	2	49	-22	2	52
								T=303

TABLE 9

Individual Form with Optimal Regions
Units of $1000 Per Year

$J \rightarrow$	1	2	3		
$q_j \rightarrow$	1	1	1		
i	c_{11}	c_{12}	c_{13}		J
1	5	4	7	3	3
2	6	7	3	1,2	1
3	8	11	2	1,2	2
$v_j \rightarrow$	8	11	7		24

TABLE 10

Quota Form with Optimal Regions

q	4	1	4	1						
$q_j^{(4)}$	4-1	1-0	3-2	0-1						
$q_j^{(3)}$	4-1	1-0	4-1	0-0						
$q_j^{(2)}$	5-0	2-0	4-0	0-0						
$q_j^{(1)}$	6-0	0-0	4-0	0-0						
$q_j^{(0)}$	4-0	2-0	0-0	4-0						
q_j	4	1	4	1	$J^{(0)}$	$J^{(1)}$	$J^{(2)}$	$J^{(3)}$	$J^{(4)}$	J
10	23	13	16	14	1	1	1	1	1	1
20	23	17	27	37	4	3	3	3	3	3
30	30	20	20	16	1	1	1	1	1	1
40	26	22	22	15	1	1	1	1,3	1,3	3
50	28	25	30	41	4	3	3	3	3,4	4
60	29	27	20	13	1	1	1	1	1	1
70	29	30	36	38	4	3	3	3	3	3
80	25	33	20	22	2	1	1	1	1	1
90	25	37	33	38	4	3	3	3	3	3
100	30	49	25	26	2	1	2	2	2	2
$v_j^{(0)}$	0	0	0	0	344					
$v_j^{(1)}$	29	49	27	41		289				
$v_j^{(2)}$	30	48	27	40			308			
$v_j^{(3)}$	31	48	27	39				308		
$v_j^{(4)}$	31	48	27	38						315

by Thorndike (9) and Lord (5), however, this problem, for complete quotas, is easily solved without extensive theory by the use of subtraction.

Consider the c_{ij} values of Table 7 with $q_1 = 4$ and $q_2 = 6$. It is only necessary to form the difference $c_{i1} - c_{i2}$ for all i and to assign the individuals to job category 1 in the order of the largest value of this difference and to job category 2 in the order of the smallest value of this difference. The method is illustrated in Table 7.

The symbol J is used to indicate the assignment so that $x_{iJ} = 1$ and any other $x_{ij} = 0$. The value of c_{iJ} is the value of c_{ij} corresponding to x_{iJ} and we have

$$T = \sum_{i,j} x_{ij} c_{ij} = \sum_{i} x_{iJ} c_{iJ} . \qquad (4.1)$$

In Table 7 this is 303 units where $q_1 = 4$ and $q_2 = 6$.

The sum which results when assignment is made to job category 1 when $c_{i1} - c_{i2} > 0$, to job category 2 when $c_{i2} - c_{i1} > 0$, and to either job category 1 or job category 2 when $c_{i1} - c_{i2} = 0$, is certainly an optimal one though the desired quotas are not necessarily met. Now since the assignments depend on the differences $c_{i1} - c_{i2}$, we can add any constant to $c_{i1} - c_{i2}$ without changing the resulting assignments. The value, $- 3$, is added to each of the values of $c_{i1} - c_{i2}$ to form $c_{i1} - c_{i2} - 3$ in Table 7 so that the positive values correspond to assignments to job category 1, the negative values (including 0) to job category 2, and the quotas are met by assigning the men in accordance with the sign of $c_{i1} - c_{i2} - 3$.

The solutions for different quotas can be obtained by forming, for different values of k,

$$c_{i1} - c_{i2} + k = (c_{i1} - 0) - (c_{i2} - k). \qquad (4.2)$$

Assignment is then made

to the first job category when $c_{i1} - 0 > c_{i2} - k,$

to the second job category when $c_{i1} - 0 < c_{i2} - k,$ $\qquad (4.3)$

to either job category when $c_{i1} - 0 = c_{i2} - k.$

The problem of Table 7 is worked in Table 8 by subtracting different values of k. The values so subtracted are placed at the bottom of the c_{i2} column. Assignments are made using (4.3). The resulting assignments are summarized at the top of the table. The first of each pair of entries indicates the number of unique assignments and the second the number of ties. On the basis of these results a new k is selected until the quotas are met. The value of T is placed in the lower right-hand corner at each step.

If the values of c_{i1} are plotted as x's and the values of c_{i2} plotted as y's, the line $x - y + k = 0$ divides the space into "optimal regions" with the

number of points in each indicating the resulting quotas. We move the line parallel to its original position till the desired quotas are met. These optimal regions for specified quotas correspond in the m-dimensional case to the regions bounded by hyperplanes discussed by Brogden (1) and Lord (5).

5. *The Conditions of Solution**

For the mathematical problem of classification we need conditions of solution which are generalizations of (4.3). To expedite the analytic solution, y_{ij}^2 is used in place of $x_{ij} \geq 0$ in (1.8) so that the expressions (1.9) become

$$\left. \begin{aligned} \sum_j^m y_{ij}^2 = f_i \\[2mm] \sum_i^n y_{ij}^2 = q_j \end{aligned} \right\} \quad \text{with} \quad \sum_{ij} y_{ij}^2 = \sum_i f_i = \sum_j q_j = N. \quad (5.1)$$

We desire to maximize

$$T = \sum_{ij} y_{ij}^2 c_{ij} , \quad (5.2)$$

where y_{ij}^2 is 0 or a positive integer.

An analytic derivation of the condition of solution using differential calculus can be obtained from an extension of the problem in which y_{ij}^2 is any real number defined by

$$0 \leq y_{ij}^2 \leq N . \quad (5.3)$$

Since c_{ij} is also finite, T is bounded and there is at least one set of y_{ij}^2 resulting in the largest T. If these y_{ij}^2 are integers they constitute a solution of the desired problem. We use Lagrange multipliers for the side conditions (5.1) and have

$$\phi = \sum_{ij} c_{ij} y_{ij}^2 + \sum_i u_i(f_i - \sum_j y_{ij}^2) + \sum_j v_j(q_j - \sum_i y_{ij}^2). \quad (5.4)$$

Taking partial derivatives with respect to each y_{ij} we find

$$\frac{\partial \phi}{\partial y_{ij}} = 2(c_{ij} - u_i - v_j)y_{ij}$$

$$\frac{\partial^2 \phi}{\partial y_{ij}^2} = 2(c_{ij} - u_i - v_j) \quad (5.5)$$

$$\frac{\partial^2 \phi}{\partial y_{ij} \, \partial y_{i'j'}} = 0, \quad i \neq i', j \neq j'.$$

*The initial steps in the development of this section follow the outline of a proof presented by Professor D. F. Votaw, Jr., at a conference on personnel classification problems held at Personnel Research Branch, Department of the Army, Washington, D. C. in August, 1952.

Hence the necessary conditions for a maximum sum, where the maximum sum need not be greater than any other sum but must be at least as large, are

$$\left.\begin{array}{c} (c_{ij} - u_i - v_j)y_{ij} = 0 \\ (c_{ij} - u_i - v_j) \leq 0. \end{array}\right\} \tag{5.6}$$

For assigned values, $x_{ij} > 0$, $y_{ij} > 0$ and the necessary conditions are

$$c_{ij} - u_i - v_j = 0, \tag{5.7}$$

and in all other cases, $x_{ij} = 0$, $y_{ij} = 0$ and the necessary conditions are

$$c_{ij} - u_i - v_j \leq 0. \tag{5.8}$$

Since there is always a maximum sum in the sense defined above (for the complete assignment problem) there must be at least one set of u_i and v_j satisfying (5.7) and (5.8). Hence when the sum is maximal we have

$$\begin{cases} c_{ij} = u_i + v_j & \text{for assigned values} \\ c_{ij} \leq u_i + v_j & \text{for unassigned values.} \end{cases} \tag{5.9}$$

These conditions are used as the conditions of solution of the simplex method (3, 369) (11, 259). The conditions themselves are independent of the x_{ij} which are positive integers (including zero) indicating the number of assignments.

These conditions of solution are illustrated by application to the trivial problem of Table 1 in which $f_i = 1$, $q_i = 1$ so $x_{ij} = 0$ or 1. The values $v_1 = 0$, $v_2 = 2$, $v_3 = 1$, $u_1 = 6$, $u_2 = 6$, $u_3 = 9$ satisfy the first equation of (5.9) where the assignments are the 13, the 21, and the 32 terms. The other values of $c_{ij} - u_i - v_j$ are all negative so the solution is a maximum.

If we denote the assigned values by J, and the others by j, $x_{iJ} > 0$, $x_{ij} = 0$ and

$$\begin{cases} c_{iJ} - u_i - v_J = 0 \\ c_{ij} - u_i - v_j \leq 0 \end{cases} \tag{5.10}$$

are the conditions of solution.

6. A Generalization of Brogden's Condition

It seems preferable to eliminate the values of the u_i from (5.10) by subtraction and to write the conditions of solution as

$$c_{iJ} - c_{ij} \geq v_J - v_j \tag{6.1}$$

or the equivalent

$$c_{iJ} - v_J \geq c_{ij} - v_j . \tag{6.2}$$

These inequalities are the conditions of optimal solution on which the method of optimal regions is based. For the case of complete quotas there are always

maximal sums in the sense defined above and in this case the condition (6.2) holds. Hence there must be values of v_i (and v_J) for expressing this condition in any numerical problem. This is consistent with the formulation of the problem as given by Lord (**5**, 300), which calls for thed etermination of $a_i = v_i$. These values of v_i are expressible in terms of the number system in which the c_{ij} are expressed.

Now (6.1), and hence (6.2), is in a sense a generalization of a condition used by Brogden in connection with his method of differences of critical rejection scores. He suggested that (**1**, 151) "Assignment is made to the first of two assignments when the value of the difference variable exceeds that of the difference score." In the present notation we might write Brogden's condition as

$$c_{iJ} - c_{ii} \geq r_J - r_i \, , \tag{6.3}$$

where r_J and r_i are rejection scores with no assignment being made when $c_{ii} - r_i$ are all negative for all j and fixed i. The condition (6.1), or its equivalent (6.2), might well be called Brogden's generalized condition.

Rao (**7**, 344) has proved a lemma which is very similar to (6.2). He has worked out an ingenious graphical-mechanical solution of the classification problem when $m = 3$ (**7**, 327–329).

7. *The Method of Optimal Regions*

We are now in a position to apply the method of optimal regions, which consists in taking values of v_i , finding the corresponding quotas, and changing the v_i until the quotas are met. The method is illustrated first in Table 9 with the trivial problem of Table 1.

Since one man is to be selected from each column we take the initial v_i to be the highest c_{ii} value in that column. Assignments using (6.2) show that, when the tie situation resulting from its application to the c_{ii} values of man No. 3 is resolved in favor of assignment to the second job, the quotas are satisfied and the greatest sum is $24,000 per year.

For some purposes we wish to know what is the maximum sum if there are no quota restrictions. In Table 9, for example, this amount is $25,000 per year, which is obtained by adding the largest value in each row. The quota restrictions do not result here in a large percentage loss.

Assignments made with (6.2) with $v_i = 0$ give this completely maximal assignment. In the following illustrations a preliminary step to the method of optimal regions features the use of $v_i = 0$. The first step then features the values of v_i which are estimates of the v_i leading to the desired quotas.

The method is applied in Table 10 to a problem in the quota form. This is an abbreviation of a problem used by Brogden (**1**, 149). Every tenth man in his list was taken and the values of c_{ii} were obtained from his values by multiplying by 10 and adding 25. The quotas adopted for this problem are

$q_1 = 4$, $q_2 = 1$, $q_3 = 4$, $q_4 = 1$. The greatest possible T, without regard to quotas, is 344 units.

The initial values of v_i, the $v_i^{(1)}$, are determined by selecting the fourth largest c_{i1}, the largest c_{i2}, the fourth largest c_{i3}, and the largest c_{i4}. The tentative assignments are made using (6.2) and the resulting quotas are indicated in rows above the table. The first entry is the number of unique assignments and the second is the number of ties.

The resulting $q_i^{(1)}$ are compared with the desired q_i. The initial assignment gives an excess of 2 in the first job category, a deficiency of 1 in the second, and a deficiency of 1 in the fourth. An excess indicates the value of v_i is too low and a deficiency indicates the value of v_i is too high. The value of $v_1^{(1)}$ is raised by one unit to form $v_1^{(2)}$ and the values of $v_2^{(1)}$ and $v_4^{(1)}$ are lowered by one unit to form $v_2^{(2)}$ and $v_4^{(2)}$. The values of $J^{(2)}$ and $q_i^{(2)}$ are computed. Values of $v_i^{(3)}$ are then assigned. After four steps with the assignment of ties, the quotas are met.

The method is called the method of optimal regions, since the changing of the values of v_i may be interpreted as the moving of hyperplanes parallel to original positions in such a way that the optimal solutions involving the number of points within the resulting regions (or on the boundaries) eventually satisfy the desired quotas.

The method of optimal regions is based on the fact that bounded values of v_i exist which identify the regions associated with the desired quotas. A primitive form of the method calls for the calculation of the quotas of each different set of v_i values (or perhaps the calculation of those sets close to an approximate set such as $v_i^{(1)}$). This form of the method is not practical except in trivial problems. A less formal variation of the method, suitable for problems of hand calculation but not of machine calculation, calls for estimates of the values of the increments to the $v_i^{(k)}$ in accordance with the values of $q_i^{(k)} - q_i$. There is no guarantee, of course, unless an appropriate estimation scheme is specified, that such estimates will lead to an improvement at any specific step, but in many hand computation problems the alert computer can arrive at the result in relatively few steps. Larger numbers of job categories, larger numbers of rows, and large positive correlations between the columns of c_{ij} values tend to increase the number of steps.

A more formal variation of the method, and one which is better adapted to machine computation, calls for the calculation, and ordering for fixed j and k, of the values of $c_{ij} - v_i^{(k)} - c_{iJ} + v_J^{(k)}$. These values are all non-positive as indicated by (6.2). This detailed method of optimal regions calls for the determination of the increments to the $v_i^{(k)}$ by selecting $q_i - q_i^{(k)}$ largest values of $c_{ij} - v_i^{(k)} - c_{iJ} + v_J^{(k)}$ in the same way that the values of $v_i^{(1)}$ are selected using the largest q_i values of c_{ij}. There is no absolute guarantee that the process will converge to a solution, though the method of exhaustion does guarantee a solution after a finite number of steps. The method is simply

an iterative process in which, at each step, we see where we are and where we want to go. We take steps, which for the most part can be mechanized, toward the objective.

In practice there seems to be little difficulty in obtaining a solution in a relatively small number of steps except when the objective is not clear because of large numbers of ties. These make it difficult to discover if the q_i can be obtained from the $q_i^{(k)}$. This is precisely the problem faced by Smith (**8**, 45) with the method of bounding sets. A solution has been given by Votaw and Dailey (**12**, 16) which is practical, for hand computation, when the number of job categories is less than eight, and for larger values with machine computation. Variations in the method of interchange (**4**) may also be used in solving this problem, which has been called "the quota problem" by Votaw and Dailey (**12**, 15).

Because of the multiplicity of ties in rows having high frequencies, the quota problem is commonly more serious in connection with problems naturally expressed in the frequency form (transportation problems) than those naturally expressed in the quota form (personnel classification problems). As an illustration involving multiple ties the method of optimal regions is applied to a problem naturally expressed in the frequency form in Table 11. This is essentially a problem used by Votaw (**3**, 259) to illustrate the simplex method. The form of Table 11 presents both the statement of the problem and its solution. The values of $v_j^{(1)}$ are determined by finding the 35th of the c_{i1} values, the 35th of the c_{i2} values, and the 30th of the c_{i3} values, considering the frequencies and measuring from the largest values of c_{ij}.

Ties, many of them, result in the first and fourth man categories. Can these ties be resolved so as to satisfy the quotas? A necessary condition is

TABLE 11

Frequency Form with Optimal Regions

q_j	35	35	30				
$q'_j{}^{(1)}$	20–60	20–20	0–40				
$q_j{}^{(0)}$	40–40	0–20	0–60				
q_j	35	35	30	f_i	$J^{(0)}$	$J^{(1)}$	x_{ij}
1	9	2	9	40	1,3	1,3	10 30
2	1	8	8	20	2,3	2	20
3	7	2	1	20	1	1	20
4	9	8	0	20	1	1,2	5 15
$v_j{}^{(0)}$	0	0	0	840			(35) (35) (30)
$v_j{}^{(1)}$	9	8	9				825

TABLE 12

Quota Form with Rejection Scores

q_j	2	1	2	1				
$q_j{}^{(2)}$	3–0	0–0	2–0	1–0				
$q_j{}^{(1)}$	1–1	0–2	2–0	1–0				
$q_j{}^{(0)}$	4–0	2–0	0–0	4–0				
q_j	2	1	2	1				
10	23	13	16	14	1	–	–	–
20	23	17	27	37	4	–	–	–
30	30	20	20	16	1	1	1	1
40	26	22	22	15	1	–	–	–
50	28	25	30	41	4	4	4	4
60	29	27	20	13	1	–	1	1
70	29	30	36	38	4	3	3	3
80	25	33	20	22	2	–	–	–
90	25	37	33	38	4	3	3	3
100	30	49	25	26	2	1,2	1	2
$r_j{}^{(0)}$	0	0	0	0	236			
$r_j{}^{(1)}$	30	49	33	41		189		
$r_j{}^{(2)}$	29	49	33	41			199	
$r_j{}^{(3)}$	29	47	33	41				218

that the available values of $q_j^{(1)}$, including ties, must be greater than q_i for each value of j. We see at once from the upper rows that

$$g_1 = 20 + 60 \geq q_1 = 35; \qquad g_2 = 20 + 20 \geq q_2 = 35;$$

$$g_3 = 0 + 40 \geq q_3 = 30.$$

If this necessary condition were not met, we would take values of $v_j^{(2)}$ in the direction of meeting it. Here we apply the additional necessary conditions involving combinations of job categories:

$$g_{12} = 80 \geq q_{12} = 70; \qquad g_{13} = 80 \geq q_{13} = 65; \qquad g_{23} = 80 \geq q_{23} = 65;$$

$$g_{123} = 100 \geq q_{123} = 100.$$

The satisfaction of all these necessary conditions constitutes a sufficient condition and the quotas can be met. The resulting values of $x_{i,j}$ are shown at the right of Table 11. The value of T is 825 units and the quota restrictions cause a loss of but 15 units.

Frequently, as in this case, the method of optimal regions gives a very satisfactory solution of problems naturally expressed in the frequency form, even though the matter of ties sometimes demands close examination. The method is recommended, however, primarily for problems which are naturally expressed in a quota form featuring a relatively small m.

The method is applicable to non-trivial problems. A classification problem with 1152 man and 7 job categories, and highly correlated columns of $c_{i,}$, was worked with it in eight iterations.

8. *The Case of Incomplete Quotas*

A problem involving incomplete quotas can be worked with the method of optimal regions after the problem has been transformed to a problem with complete quotas by the use of the device illustrated in Table 6. However, the method of critical rejection scores may be used instead. This method is especially effective when the quotas are (relatively) small numbers. The method of rejection scores is a slight modification of the method of differences of critical rejection scores introduced by Brogden (1). This modification uses a method of solution very similar in form to that of the method of optimal regions described above.

Brogden used rejection scores in indicating a preliminary assignment. If r_j is a rejection score, assignments are tentatively made to job category j only if

$$c_{ij} - r_j \geq 0. \tag{8.1}$$

The use of (8.1) may lead to the tentative assignment to more than one job category. Conflicting assignments are then resolved by assigning the individual

to job category J in accordance with Brogden's condition (6.3). This may be written in the alternative form

$$c_{iJ} - r_J \geq c_{ii} - r_i . \tag{8.2}$$

It is not necessary to use (8.1) and then Brogden's condition. The same result can be obtained with the use of (8.2) and the additional specification $c_{iJ} - r_J \geq 0$. If this is not satisfied for every value $j = J$, no assignment is made.

The technique of the method follows that of the method of optimal regions with r_J playing the role of v_J , r_i of v_i , except that no assignment is made to individual i when every $c_{ii} - r_i < 0$.

The method is illustrated in Table 12 where the problem of Table 10 with $q_1 = 2$, $q_2 = 1$, $q_3 = 2$, $q_4 = 1$ is solved with the method of rejection scores. The values $r_i = 0$ are first used. The six largest c_{iJ} resulting are added to obtain the maximum sum disregarding quotas.

The method of rejection scores can be used to solve the complete assignment problem. In this case one takes successively smaller values of the r_i until the quotas are met.

9. *Concluding Remarks*

The personnel classification problem, which is mathematically the equivalent of several other problems in the social and biological sciences, is really a special problem in linear programming. Though the techniques, such as the simplex method, of the general linear programming problem are available. it seems preferable to direct the mathematical solution toward the simpler mathematical model of the classification problem. When this is done a condition for assignment, which is a generalization of Brogden's condition, is available and serves as the foundation of the method of optimal regions. This method, which is based on concepts related to those introduced by personnel workers such as Brogden, Thorndike, and Lord, seems particularly effective when applied to the personnel classification problem.

REFERENCES

1. Brogden, Hubert E. An approach to the problem of differential prediction. *Psychometrika*, 1946, 11, 139–154.
2. Dantzig, George B. Maximization of a linear function of variables subject to linear inequalities. Chapter XXI. Activity analysis of production and allocation (T. C. Koopmans, Editor). Cowles Commission Monograph No. 13, Wiley, New York, 1951.
3. Dantzig, George B. Application of the simplex method to a transportation problem. Chapter XXIII of the monograph of (2).
4. Dwyer, P. S. Basic theory and methods in optimal classification of personnel. Personnel Research Branch, Department of the Army, 1953.
5. Lord, Frederick M. Notes on a problem of multiple classification. *Psychometrika*, 1952, 17, 297–304.
6. Flood, Merrill M. On the Hitchcock distribution problem. Symposium on linear inequalities and programming. SCOOP. U. S. Air Force and National Bureau of

Standards. Washington, D. C. April, 1952, 74–99. See also *Pac. Jour.* of *Math.*, 1953, **3**, 369–386.

7. Rao, C. R. Advanced statistical methods in biometric research. New York: Wiley, 1952.

8. Smith, Robert B. Hand computational methods for the classification problem. Air Training Command, Human Resources Research Center, September 1951.

9. Thorndike, Robert L. The problem of classification of personnel. *Psychometrika*, 1950, **15**, 215–235.

10. von Neumann, John. A certain zero-sum two-person game equivalent to the optimal assignment problem. Contributions to the theory of games, Vol. II, Annals of Mathematics Study No. 28, edited by H. W. Kuhn and A. W. Tucker. Princeton: Princeton Univ. Press, 1953.

11. Votaw, D. F., Jr. Methods of solving some personnel-classification problems. *Psychometrika*, 1952, **17**, 255–266.

12. Votaw, D. F., Jr. and Dailey, J. T. Assignment of personnel to jobs. Research Bulletin 52-24, Air Training Command, Human Resources Research Center, Lackland Air Force Base, August, 1952.

Addition at proof reading: The detailed method of optimal regions has been further improved with the introduction of devices for automatic determination of the successive sets of u_j which converge to a solution. Preliminary transformations lead to initial values of v_j which, in all the problems worked thus far, become solutions with two or three iterations. For example, the 1152 $\times$ 7 problem mentioned above has been solved by determining sets of v_j from a preliminary transformation followed by two steps of the detailed method of optimal regions.