AO, 1 ALGORITHM FOR CITY PLANNING

MYRON B. FIERING*

ABSTRACT

This paper describes a model for the allocation of research funds to a series of urban experiments whose outcomes determine the ultimate disposition and allocation of very much greater amounts of money for urban housing projects. It is shown that the allocation model is a particular application of a technique which can be widely applied in the design of statistical experiments, and the paper describes the algorithm for solving the 0,1 integer programming problem which results from the formulation of the urban model. Central to the working of the model is the derivation of an association matrix which expresses the likelihood that certain experimental procedures will be paired in actual practice.

Key Words: experimental design, housing, city planning, mathematical programming, random sampling, gradient procedure.

Introduction

The operations research literature contains literally dozens of references to solutions of 0,1 programming problems. Quite properly, these many papers focus on the algorithm for extracting the solution and on demonstration of its convergence, uniqueness, and other desirable properties. In most cases the underlying problem — abstracted from the physical, military, management or social sciences — is given short shrift in favor of the more appealing, relevant, and tractable analysis.

This paper reverses the traditional emphasis and concentrates primarily on construction of the objective function for a 0,1 programming problem; passing attention is paid to a new solution algorithm based on a gradient technique. This algorithm does not purport to find the global optimum but rather a series of local optima from which it is possible to decide whether to accept the best available solution and terminate the process, or to run additional trials in the hope of locating a more desirable solution.

---

*Gordon McKay, Professor of Engineering and Applied Mathematics, Harvard University, Cambridge, Massachusetts.

The problem originated from a study of urban housing factors, but is analytically akin to an extension of sampling theory and as such can be modified and rendered generally applicable in that statistical discipline. The original study was undertaken under the auspices of the Department of Housing and Urban Development, and the author acknowledges the assistance provided to him as a consultant to Abt Associates, Inc., which has approved the release and publication of this material; unfortunately, the numerical problem which spawned this solution was not available for publication.

The Problem

Suppose a large sum of money is to be made available for urban communities to spend on housing. The primitive institutional and technological constraints imposed on the housing industry are widely known and seemingly insurmountable; they have led Harvey Brooks to characterize housing as America's "largest cottage industry". The sponsor, whether it be a federal, state or local agency, a foundation, or a private combine, is anxious to overcome as many institutional and technological obstacles as possible, and therefore identifies a number of potential changes in these institutional and technological constraints in the hope that some combination of them might effect major benefits (e.g., cost reduction or quality improvement) in the projected housing development.

The list of such potential changes is very long indeed; a few typical entries are:

1. tax and financing advantages,
2. the use of exotic construction materials,
3. modernization of building codes,
4. modernization of union and restrictive rules, and
5. factory assembly, plumbing, wiring, and drilling of modular components.

The optimal combination of factors is defined as that set of changes which is best for a particular community, and, unhappily, no analysis seems capable of determining this optimum. Experimentation using various combinations in several cities offers some promise, but it is certain to be frightfully expensive and, even worse, doomed to inadequacy because the number of possible combinations is prohibitively large and consequently precludes examination of all but a small fraction of the alternatives. The problem addressed here is whether prior analysis can delineate, for particular cities, certain combinations which are more advantageous than others in that they provide more *information* about the alternatives. Note that this is quite different from attempting to find the optimal combination. The optimal combination maximizes (or minimizes) some housing-oriented benefit function while here we are concerned with *information* and its maximization by selection of experimental modules. To be sure, the two processes (experimentation and construction of the prototype) are intimately and ultimately related, but formally they are quite distinct.

The following paragraphs describe the decision process in formal, analytical terms. The alternatives are called *experiments* and the several places available for experimentation are called *cities* (this notation is made explicit because the locales may in fact not be cities and the technological and institutional changes may not in fact resemble experiments in the scientific sense).

A set of experiments $\{E_1, E_2, .., E_n\}$ is available, and some sub-set of E must be assigned to a group of cities $\{C_1, C_2, .., C_m\}$ so that the total amount of information derived from performing the experiments at the several cities is maximized. Symbolically, we seek an $n \times m$ matrix called Δ such that the element δ_{ij} of Δ is unity if experiment E_i is performed at city C_j and is zero otherwise. The matrix Δ is a decision matrix, as shown in array A-1:

Cities

C_1 C_2 ..... C_m

E_1

E_2

.

Experiments

.

.

E_n

$$\delta_{ij} = \begin{cases} 1: & \text{perform } E_i \text{ at } C_j \\ 0: & \text{do not perform } E_i \text{ at } C_j \end{cases} \quad = \Delta \qquad \text{(A-1)}$$

Decision Matrix

Clearly, certain combinations of (0,1) in Δ are more appropriate than others. For example, the cost of performing a certain sub-set of experiments at city C_j may be quite different from performing the same sub-set of experiments at city C_k, so that if the amount of information obtained from the experiments is equal, then clearly the experiments should be done at that city in which the cost is least. It follows that the optimal solution for Δ must consider a matrix of costs; the cost matrix is also of dimension n $\times$ m and the element c_{ij} is the cost of performing E_i at C_j, as shown in array A-2.

$$\begin{array}{l|c|}
 & C_1 \qquad C_2 \qquad \overset{\text{Cities}}{\cdots\cdots} \qquad C_m \\
\hline
E_1 & \\
E_2 & \\
\vdots & \\
\text{Experiments} & c_{ij} = \text{cost of } E_i \text{ at } C_j \\
\vdots & \\
E_n & \\
\hline
\end{array} \qquad \text{(A-2)}$$

Cost Matrix

We make the fundamental assumption that the information obtained from performing a sub-set of experiments at C_j is precisely the same as that obtained from the same sub-set at C_k. The cities are indistinguishable with respect to results but not with respect to costs. The cost of performing experiment E_i is not independent of the city C_j at which the experiment is performed. Therefore, in all but the most trivial cases, the rows of the cost matrix in array A-2 are not identical so that the cost information cannot be compressed into an n-dimensional vector.

Of course, the political and economic realities encountered in any such experimental enterprise impose a large number of constraints on the specification of the decision matrix Δ. Obedience to geographical distribution, whether mathematically prudent or not, demands that each city C_j get its fair share of the experimental budget. This paper does not purport to judge the worthiness of any particular distributional requirement, but merely presents a technique whereby the cost of geographical distribution can be measured by the difference in information between the optimal experimental procedure and the actual; from the magnitude of this loss it is possible to impute some economic metric to the political luxury of geographical distribution, and to render an informed judgment on the degree of geographical distribution which the experimental procedure ought to accommodate.

Other constraints enter the decision-making process. For example, certain experiments may be uniquely adaptable to certain cities while it may be quite impossible to perform these same experiments elsewhere. Therefore the analysis must have some way of forcing certain elements δ_{ij} to unity and others to zero.

The experiments cannot be scaled or sub-divided; that is, experiment E_i *is* or *is not* performed at city C_j. It is presumed that the level of experimentation (for any experiment) is uniquely determined at any city and that this level is reflected by the cost c_{ij}. At first blush this would seem to make the problem easier because it eliminates the necessity for determining how intensive each potential experiment should be in each of the cities $\{C_i\}$ and replaces it with a bistable, polar decision represented by the pair (0,1). But in fact the converse is true; solving the (0,1) problem is vastly more difficult than solving the corresponding continous problem in which intermediate levels of experimental intensity can be accommodated.

In any event, a set of decisions must be made to define an experimental design on the grid represented by the intersection of experiments $\{E_i\}$ and cities $\{C_j\}$ A set of 0's and 1's are to be located so as to maximize the total amount of information obtained from the experimental design, all subject to appropriate geographical, institutional, and budgetary constraints. If the number of experiments n is of the order of 20, and the number of cities m is of the order of 10, a solution consists of some 200 binary digits. But this unimpressively small number belies the enormous number of feasible combinations and permutations which are somehow inferior to the optimal solution. Sorting through this enormous number of candidates in not a trivial task!

Failure of Standard Techniques

Consider an experimental design from which it is desired to evaluate two effects, and let these effects be measured by experiments E_I and E_2. Traditional experimental design calls for four experiments: (i) both E_I and E_2

absent, (ii) E_I absent and E_2 present, (iii) E_I present and E_2 absent, and (iv) both E_I and E_2 present. From this arrangement it is possible to evaluate the results of each effect alone and in combination, with conclusions usually cushioned by the limits of statistical significance. The number of combinations which must be considered in a complete factorial experiment is 2^n, which is 1024 for the number of experiments $n = 10$. In this study, each city is associated with a unique combination of experiments; that is, one of the 1024 possible arrangements can be tried at each city. With 20 cities, the total number of different assignments is the total number of combinations of 1024 items taken 20 at a time, or

$$ {}_{1024}C_{20} = \binom{1024}{20} = \frac{1024!}{1004!\ 20!} > 10^{60} \quad , \tag{1} $$

a truly staggering quantity. One of these assignments is best in the sense that it gives more information than any other, and it is our task to find that one.

We are stymied on several accounts. First, with only 20 opportunities (i.e., cities) for experimentation rather than 1024, it is patently impossible to perform a factorial experiment which would (i) uniquely isolate the effect of any factor and (ii) specify interactions between that factor and all other combinations of factors. Second, accepting the constraint of 20 cities, it is clearly impossible to consider the systematic extraction of potential experiments because their number is so formidable. As a corollary, the variable cost structure which represents the fact that c_{ij} might differ substantially from c_{ik} makes impractical a randomized block experimental design. Third, we have not yet come to grips with the essential problem of what it is that constitutes a "good" experiment, having devoted ourselves mainly to the vague notion that good experiments provide lots of information while poor ones do not.

Modifications of the factorial design include such schemes as Latin Square, Graeco-Latin Squares, randomized blocks, and other techniques which can be studied in any one of many standard references. But these techniques specifically ignore the cost of experimentation at the several alternative locations, the value of information at the several locations, and the difficulties associated with establishing a criterion of performance for the experimental design. Consequently the standard techniques are rejected in this analysis, and it is necessary to consider techniques of mathematical programming.

Formulation of the Experimental Design as a Programming Problem

If the constraints on cities, experiments and combinations can be written as inequalities, mathematical programming offers the preferred solution. The decision variables are the δ_{ij}, the elements of array A-1. There is a constraint on the total budget for the experimental design; this is expressed by the inequality

$$\sum_{i=1}^{n} \sum_{j=1}^{m} \delta_{ij}\ c_j \leq B\ , \tag{2}$$

where B is the total budget for the program. Moreover, there are two constraints on the budgetary allowance for each city C_j, expressed by

$$\sum_{i=1}^{n} \delta_{ij}\ c_{ij} \leq B_j\ ,\ \forall_j\ , \tag{3}$$

$$\sum_{i=1}^{n} \delta_{ij}\ c_{ij} \geq B^*_j\ ,\ \forall j\ , \tag{4}$$

where B_j is the maximum budget allocated to city C_j and B^*_j is the minimum. Judicious manipulation of the B_j and B^*_j is tantamount to imposition of geographic distribution, and if all the B^*_j are zero then geographical distribution is not a consideration in the optimal assignment of experiments.

It is also clear that there must be some control exercised over the number of locations at which any experiment is performed. For example, if E_i is performed at every C_j, there is no basis on which to determine its effect because there is no "untreated" city. Conversely, it must be performed someplace, in at least one C_j. These constraints are expressed by

$$\sum_{j=1}^{m} \delta_{ij} \leq N_i,\quad \forall_i\ , \tag{5}$$

$$\sum_{j=1}^{m} \delta_{ij} \geq N^*_i\ ,\ \forall_i\ , \tag{6}$$

where N_i and N^*_i are the upper and lower bounds, respectively, on the frequency of experiment E_i.

Finally, to constrain the decision variables δ_{ij} to the values 0,1, the constraint

$$\delta_{ij} = \delta^2_{ij} \tag{7}$$

is imposed; this is satisfied by the two values $\delta_{ij} = 0$ and $\delta_{ij} = 1$, and further ensures that the decision variables are non-negative and obey the constraint

$$\delta_{ij} \geq 0. \tag{8}$$

The programming problem is to maximize the total information, heretofore undefined but written functionally as

$$I = I(\Delta), \tag{9}$$

subject to the constraints represented by equations (2) through (8). Because of equation (7), it is clear that the problem cannot be cast as a linear programming problem and that one of the more sophisticated relatives of this blessedly simple technique must be employed.

The following sections develop a suitable non-linear objective function corresponding to equation (9), and because of the inherent difficulty of non-linear programming problems, provide an algorithm for approximating the analytical solution.

The Objective Function

Suppose there are three cities and three factors or experiments to be investigated. There are 2^3 combinations in a complete factorial experiment, and all the combinations represented by the complete factorial arrangement cannot be accommodated in the available cities because $2^3 > 3$. Array A-3 shows the eight possible combinations, no more than three of which may be utilized in the experiment:

$$\begin{array}{cc|cccccccc|}
 & & \multicolumn{8}{c}{\text{Combination}} \\
 & & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\
\hline
 & E_1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\
\text{Factor} & E_2 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\
 & E_3 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\
\hline
\end{array} \tag{A-3}$$

It is necessary to maintain some diversity in the experimental design, so that there would be little benefit in repeating any column or combination in more than one city. In the more general case, for which different cities exert unique effects, this statement would not be evident *a priori*; however, under the assumption that cities are indistinguishable with respect to effects (but not with respect to costs), replication is not indicated.

Which columns, then, are more appropriate for the limited experimental effort to be undertaken in the several cities? Consider a square matrix A, of dimension n $\times$ n, whose elements represent the degree of association which exists between each pair of experiments or factors. For example, it might happen that in housing practice, when the several factors are incorporated into prototype construction projects, certain experiments tend to occur together while others tend to preclude each other. If experiment E_1 is some institutional change which, if implemented, strongly implies that E_2 would be incorporated while E_3 would generally be bypassed, the matrix A has the following general form

$$A \approx \begin{array}{c} \\ E_1 \\ E_2 \\ E_3 \end{array} \begin{array}{ccc} E_1 & E_2 & E_3 \\ \hline \multicolumn{1}{|c}{0} & + & \multicolumn{1}{c|}{-} \\ \multicolumn{1}{|c}{+} & 0 & \multicolumn{1}{c|}{} \\ \multicolumn{1}{|c}{-} & & \multicolumn{1}{c|}{0} \\ \hline \end{array} \qquad (A\text{-}4)$$

It is inappropriate to continue to label the rows and columns as *experiments* E_i because A represents the degree of association encountered in practice, not under the controlled conditions which constitute an experiment. However, for the sake of notational consistency, the symbol E will be used throughout and the context will make its significance abundantly clear. By definition, the elements along the main diagonal of A are equal to zero. Because $E1$ and $E2$ tend to occur together, the elements $a12$ and $a21$ are positive; similarly, a_{13} and a_{31} are negative. The matrix A should not be thought of as a correlation matrix because there is no implication that E_1 and E_2 force the output of the experiment (whatever that may be) in the same direction, nor conversely for the negatively associated pair $E_1 - E_3$. The elements of A do not specify reinforcement or antagonism in the usual statistical sense, but merely the fact that political and social reality dictate which pairs of experiments are likely to be run together, which are likely to be run individually, and which are independent.

Numerical values are assigned to the elements a_{ij}; the array A-4 merely specifies the signs for several of the constituent pairs. Elements whose absolute values are large reflect strong association or dissociation, and conversely for small values; the proposed solution is independent of an arbitrary scale factor, so that while it might be convenient to adjust the elements of A to lie within the range $-1 \leq a_{ij} \leq 1$, it is unnecessary to do so.

Suppose a study of the three available experiments suggests

E_1: adoption of a performance-based building code,

E_2: availability of an attractive financing scheme for housing,

E_3: acceptance by trade unions of liberalized restrictive practices.

E_1 and E_2 are strongly associated; E_1 and E_3, for the particular cities involved, are thought to be strongly exclusive; E_2 and E_3 are very nearly independent in that the realization of one does not imply much about the other. The matrix A, or association matrix, is

$$A = \begin{array}{c} \\ E_1 \\ E_2 \\ E_3 \end{array} \begin{array}{ccc} E_1 & E_2 & E_3 \\ \hline 0 & 0.8 & -0.6 \\ 0.8 & 0 & -0.1 \\ -0.6 & -0.1 & 0 \end{array} \quad \text{(A-5)}$$

(These values are abstracted from the Abt study cited earlier.) Because E_1 and E_2 are strongly associated, combinations 1, 2, 7, and 8 of array A-3 appear to be most promising because in each of these E_1 and E_2 are performed or bypassed jointly. By the same reasoning, combinations 2 and 7 appear to be more suitable than 1 or 8 because E_1 and E_3 are strongly opposed (that is, a_{13}, being negative, suggests that both are unlikely to occur simultaneously). It therefore follows that however the several combinations are ranked, numbers 2 and 7 should fare better than their competitors.

The several combinations are ranked by a simple algorithm. Each score is the weighted sum of elements a_{ij} of the matrix A, with $i \neq j$, and with the weighting factors being positive or negative depending on whether E_i and E_j are run together or not. For combination 1, (0,0,0), all experiments are bypassed so that the score is the sum of elements in the matrix A; for convenience we use only the elements above (or below) the main diagonal, thereby taking advantage of the symmetry of A. Thus for combination 1 the score is $S_1 = 0.8 - 0.6 - 0.1 = 0.1$, as shown in Table 1. Similarly, combination 2, (0,0,1), ignores E_1 and E_2; E_3 is performed, so that the score includes $-a_{13}$; or $+ 0.6$. Finally, because E_2 and E_3 are not jointly performed or bypassed, the sign on a_{13} is negative and the total score is $S_2 = 0.8 + 0.6 + 0.1 = 1.5$, as shown below. The scores do not represent a physical parameter but rather the degree of "independent return" from the combination k.

Combination, k	1	2	3	4	5	6	7	8
Score, S_k	0.1	1.5	−1.3	−0.3	−0.3	−1.3	1.5	0.1

Table 1 — Scoring System for Experiments

Due to the symmetry of the factorial experiments and the matrix A, the scoring system is indifferent between complementary conbinations. However, if A is not symmetric because of peculiar institutional constraints, the scoring system would necessarily have to consider all elements of A rather than the triangular portion alone.

Table 1 shows the score S_k corresponding to all combinations k, where k runs from 1 to 2^n. If combination 2 were performed at all 3 sites, no information would be derived from the experimental design because there would be no standard against which to compare the information or effects derived from different experimental arrangements.

Suppose city C_i is assigned experimental design or combination number k_i and city C_j is assigned experimental design k_j. The sum of scores for both cities is

$$S_{ij} = S_{k_i} + S_{k_j} , \tag{10}$$

where S_{ij} depends solely on the experimental arrays at each city and not on any measure of replication between them. The contribution to the total score which is due to the pair of cities C_i and C_j is the sum over all experiments

$$\sum_{k=1}^{n} (\delta_{ki} - \delta_{kj})^2 \; S_{ij} \tag{11}$$

which, in effect, assigns a weighting factor of unity to those elements of the experimental design which are different in the two cities and a weighting factor of zero to those elements which are identical. Calculation of the total score is then simply a matter of summing up over all possible pairs of cities in the decision matrix, so that the total amount of information derived from the experimental design may be written

$$I = \sum_{i=1}^{m} \sum_{j=i+1}^{m} \left[\sum_{k=1}^{n} (\delta_{ki} - \delta_{kj})^2 \; S_{ij} \right] , \tag{12}$$

which is to maximized. It is a trivial matter to include an arbitrary weighting factor in each element of the matrix S_{ij}, this factor to represent some *a priori* evaluation of the importance of particular experiments at particular places. If such a factor, say λ_{ijk}, is added, the total information to be maximized is

$$I = \sum_{i=1}^{m} \sum_{j=i+1}^{m} \left[\sum_{k=1}^{n} (\delta_{ki} - \delta_{kj})^2 \lambda_{ijk} S_{ij} \right] . \tag{13}$$

This completes the formulation of the problem as a non-linear (0,1) integer programming problem subject to linear constraints, but unhappily there is little promise of a solution! The essential feature in formulating the problem is an elaborate structure involving the matrix A; this is necessitated by the fact that there is no specific measure for the benefit accruing to any combination of experiments at a particular city, so the usual notion of economic benefits is replaced by a formulation which measures the absence of replication and, simultaneously, is strongly influenced by the closeness with which experimental arrangements agree with the format and political constraints within which actual construction projects are presumed to operate. Again, it is assumed that the information obtained from any one city is as useful as that obtained from any other; the extent to which this is untrue can be accommodated by assigning a range of values to the parameters λ_{ijk}. For the example cited here, and for the larger problem described above, all values of λ are set at unity so that the program does not distinguish between information obtained at the several cities.

To summarize, the decision variables are the values of δ which appear in equation (13); it is desired to find that set of δ's which maximizes the information defined in equation (13), subject to the several constraints in equations (2) through (8). The next section is devoted to obtaining a numerical approximation to the exact solution of the programming problem.

Numerical Approximation to the Solution

The algorithm developed for this problem is a steepest ascent or gradient technique which starts from a random feasible solution as defined by a decision matrix Δ and proceeds therefrom to a new matrix Δ_I which is a local optimum in the sense that interchanging any adjacent (0,1) pair produces either an infeasible solution or a lessening of the information $I(\Delta)$. Another random feasible start is then made, and a new local optimum Δ_2 is reached; after several random starts, the most advantageous (or locally) optimal value of $I(\Delta)$ is chosen to approximate the (globally) optimal $I(\Delta)$ and the corresponding decision matrix, Δ, is specified as the experimental design to be implemented.

The following steps are executed in the algorithm:

1. Read all control data; the association matrix A; the cost matrix C; the several budgets B, B^*, N, N^*; and any predetermined combinations which specify that E_I is, or is not, to be performed at C_j.
2. Calculate the score S_k for all experimental combinations, $k = 1,2,\ldots,2^n$. Clearly, for n large, the number of combinations is formidable.
3. Select a pair of rectangularly distributed random sampling integers in the range $1 \leq i \leq n$, $1 \leq j \leq m$, and assign $\delta_{ij} = 1$ unless one or more of the following conditions prevails:
   (a) the intersection δ_{ij} has been precluded by the input data,
   (b) the intersection δ_{ij} has already been established to be unity,
   (c) the budget at C_j is exceeded,
   (d) the limitation on N_i is exceeded, or
   (e) the total budget B is exceeded.

   Continue to put values $\delta_{ij} = 1$ until condition (e) is violated, whereupon a quasi-feasible solution is defined. This solution obeys any constraints on maxima, but not necessarily those on minima. These latter constraints are, for the moment, neglected.
4. Calculate the information $I(\Delta)$.
5. Isolate that pair of experiments E_p, E_q for which a_{pq} is a maximum, and adjust Δ as follows:
   (a) if $a_{pq} > 0$, try to make $\delta_{pk} \neq \delta_{qk}$, starting with that city C_k which minimizes the cost of making the exchange;
   (b) if $a_{pq} < 0$, try to make $\delta_{pk} \neq \delta_{qk}$, again starting with that C_k which minimizes the cost;
   (c) all adjustments are made subject to the budgetary and frequency constraints.
6. Move to a new pair of experiments E_r, E_s for which $|a_{rs}|$ is second-largest, and perform step 5 again. Continue iterating in this way, each time using the largest remaining $|a_{ij}|$. Finally, when all distinct pairs are exhausted (no more than $n(n-1)/2$ pairs are possible, and many of these may have $a_{ij} = 0$), a local optimum is reached.
7. Store the value $I(\Delta)$ corresponding to the decision vector Δ_1.

8. Make a new random start, as described in step 3, and continue until an appropriate number of Δ_i are investigated. The stopping point is defined, in part, by the relative smoothness of the function $I(\Delta)$ and by the execution time required to locate a local optimum.
9. Finally, identify the approximate global optimum and either terminate the solution or start again with new input data, as described in step 1, to determine the sensitivity of the solution to a range of input parameters.

Example

Continuing our numerical example, we assume the following parameters and constraints:

$$B = 2$$

$$B_j = 2, \forall j$$

$$B_j^* = 0, \forall j$$

$$c_{ij} = 1, \forall i,j$$

$$N_i = 2, \forall i$$

$$N_i^* = 0, \forall i$$

$$\lambda_{ijr} = 1, \forall i,j,k$$

and the trial design or decision matrix:

		City 1	2	3
	1	1	0	0
Experiment	2	0	1	0
	3	0	0	0

Thus, from Array A-3, we have $k_1 = 5$, $k_2 = 3$, and $k_3 = 1$. For these 2 cities we have $S_{12} = S_5 + S_3 = -1.6$. If now we sum the products

$(1-0)^2(-1.6) + (0-1)^2(-1.6) + (0-0)(-1.6) = -3.2$ taken down columns 1 and 2 of the decision matrix, equation 11 is evaluated. It is a simple matter to sum down every pair of columns:

1 and 2		-3.2
1 and 3	$-0.2\,[(1-0)^2 + (0-0)^2 + (0-0)^2] =$	-0.2
2 and 3	$-1.2\,[(0-0)^2 + (1-0)^2 + (0-0)^2] =$	-1.2
		$I = -4.6$

where I is the information, equation 12.

Each possible (and feasible) decision matrix is associated with a value of I; for small problems, we could draw an exhaustive list. However, even for this simple problem with $m = n = 3$, it is too demanding to do so by hand. Evaluation by computer shows a total of 28 feasible matrices, with I maximized when the decision matrix is

		City 1	City 2	City 3
Experiment	1	0	1	0
	2	0	1	0
	3	0	0	0

Inferential Analysis

It is prudent to assume that the response surface representing the information function $I(\Delta)$ has many local optima so that most randomly selected experimental designs will lead to globally non-optimal solutions. The extent of this shortcoming does not depend on the number of local peaks; rather it is a function of the difference between the global solution and the best of the local optima, a quantity which cannot be estimated with any degree of certainty but which does lend itself to certain statistical theorems turning on sampling reliability. For example, the probability that the best of p random and independent trials lies in the upper 100ϕ percent of all possible solutions is $1-(1-\phi)^n$; for example, if $n = 30$ and $\phi = 0.1$, the probability that the best of 30 trials lies in the upper 10 percent of all possible solutions is $1-(0.9)^{30} = 0.957$. This says nothing about the difference between the best of the sample and the global optimum, but it does provide a lower limit to the reliability of random sampling techniques because by use of gradient methods, the reliability of the result is improved and consequently is better than that which can be ascribed to the unimproved random sampling. This result is independent of the number of decision variables required to characterize any trial design or decision matrix.

The suitability of the best local solution can be estimated by a study of the range of the other local maxima. If the surface appears to be regular and fairly smooth, small values of n are tolerable. If, however, the surface shows abrupt changes of elevation and slope, a more extensive sampling investigation is warranted, provided, of course, that the cost of so doing does not appear to exceed the potential improvement which might be gained in the response.

Conclusion

A computer program to implement the solution algorithm was written in FORTRAN IV for the IBM 7094. It can accommodate up to ten cities and twenty experiments, and has been run successfully on matrices of this size within four to six minutes of computation time. The results have been encouraging, showing major improvements (that is, better response) over the best manual solutions for a wide range of budgetary and geographic constraints. These solutions are being implemented, and it is hoped that a second paper can report on the results of field testing. However, more significant than this numerical achievment is the formalism for casting an urban problem in precise operational terms. It is here, at the interface between mathematics and the social sciences, that the real excitement in modern operational analysis is to be found.