

THE UNIVERSITY OF MICHIGAN
INDUSTRY PROGRAM OF THE COLLEGE OF ENGINEERING

APPLICATION OF MATHEMATICAL PROGRAMMING TO THE SELECTION
OF EXPERIMENTAL FACTOR ARRANGEMENTS WITH RESOURCE CONSTRAINTS

John B. Neuhardt

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in the
University of Michigan
Department of Industrial Engineering
1967

July, 1967

IP-785

PREFACE

I wish to acknowledge the exceptional guidance provided by my chairman, Professor Hugh Bradley, and the continuing efforts of my committee members to provide help whenever it was asked. I further wish to thank Professor Paul Dwyer for initial reading of these ideas, and to acknowledge the helpful suggestions by Professor E. Peterson and Professor Bernard Galler.

I would like to express gratitude to Professor Richard Wilson, Professor Dean Wilson, and Dr. Wilfred Bychinsky for direction and influence in my student graduate career.

I also wish to thank those responsible for my having been chosen recipient of a Rackham Pre-doctoral Fellowship, and a General Motors Student Fellowship.

Finally, I thank my family for confidence, understanding, and infinite patience.

TABLE OF CONTENTS

	Page
PREFACE	ii
LIST OF TABLES	v
LIST OF ILLUSTRATIONS	vi
CHAPTER	
I. INTRODUCTION	1
1.1 The general problem	1
1.2 A simple formulation	2
1.3 Several definitions of the information of a design	3
1.4 A typical problem	6
II. A MINIMUM EIGENVALUE CRITERION	10
2.1 Detailed problem statement	10
2.2 The constraint region and properties of the objective function	12
2.3 The one-factor problem	14
2.4 The gradient search procedure	16
2.5 Obtaining an initial feasible solution	18
2.6 A gradient approximation	20
2.7 Three search illustrations	21
III. A SUM OF SQUARES CRITERION	31
3.1 Sum of squares criterion defined; rationale	31
3.2 A property of the objective function	33
3.3 The noninteger quadratic program	35

CHAPTER	Page
3.4 An algorithm of Kunzi and Oettli	38
3.5 Witzgall's parabolic constraint algorithm	46
3.6 A partial enumeration method of Lawler and Bell; an example	49
3.7 Use as a starting point for the minimum eigenvalue search problem	53
 IV. FURTHER EXAMPLES OF THE SUM OF SQUARES CRITERION	 54
4.1 The equal-cost problem	54
4.2 Three preliminary observations	54
4.3 The one-factor problem	60
4.4 The latin square	62
4.5 The balanced incomplete block	63
4.6 Webb's problem	64
 V. SUMMARY AND AREAS OF FUTURE INVESTIGATION	 67
5.1 Summary	67
5.2 Areas of future investigation	69
 APPENDIXES	 71
A. Eigenvalue Properties	72
B. Obtaining a Feasible Solution	75
C. Computational Considerations	94
 REFERENCES	 102

LIST OF TABLES

TABLE		Page
1.	Integer Three-Factor, Two-Level Search Problem	23
2.	Noninteger Three-Factor, Two-Level Search Problem	24
3.	Integer Search, Minimizing Maximum Eigenvalue	27
4.	Integer Search, Total Number of Observations Large	29
5.	Comparison of Eigenvalue Calculations for Various Matrix Dimensions	96
6.	Experience in Linear Program Calculation Times	97
7.	Optimistic Computation Time per One Step in Gradient Search	97

LIST OF ILLUSTRATIONS

FIGURE		Page
1.	Webb's solution for a $4 \times 4 \times 3$ problem, with twelve observations	26
2.	The quadratic program surface	41
3.	Illustration of the enumeration technique	52
4.	Linear model	77
5.	Additive model	85

CHAPTER I

INTRODUCTION

1.1 The general problem

Stemming from the early writings of R. A. Fisher, developments in the field of experimental design have advanced enormously. Because many authors have shown existence properties for various criteria of optimality, the scientist or engineer can now choose from an abundant storehouse of designs, many of which possess certain optimum properties. However, when the environmental constraints do not readily admit to the selection of any particular known “good” design, the next step is not always apparent. This may be particularly true if the experiment is constrained by cost considerations.

Very often in industrial research, an experimental program must be planned for which . . . existing designs are inadequate. In such cases the consulting statistician may have a tendency to try and alter the thinking of the experimenter so that one of the standard published designs can be used. (51)

Thus, as an expedient measure, the environment may be forced into a theoretical mold, a mold which may fit quite poorly. Even if the environmental constraints are consistent with some known “optimal”

design, the experimenter may not have access to a description of this design that meets his requirements.

One alternative to this problem is to enumerate all possible designs feasible within the constraints. However, such a list could become extremely large, even in the smaller experiments.¹ Another approach would be to develop an iterative procedure that ended in a reasonably good design within the resource constraints. Other than an integer programming formulation for the problem of finding orthogonal latin squares (16, 3), the only work known to this author involving the mathematical programming approach is that of Webb (52), who describes a linear programming formulation for the construction of orthogonal designs. Here, the objective was to minimize the total number of experimental observations.

1.2 A simple formulation

Suppose a factorial design problem has potentially p unique factor combinations that could be observed. Denoting Euclidean p -space as $E^{(p)}$, all vectors $X \in E^{(p)}$ which have integer-valued nonnegative components (each component corresponding to a particular

¹A simple 2^3 factorial design with eight factor combinations that has a constraint of a maximum of five observations per combination could conceivably have 5^8 different possibilities.

factor combination) are possible solutions to the design problem. Defining $I(\mathbf{X})$ to be the information associated with the vector \mathbf{X} , and \mathbf{U} the subspace of $E^{(p)}$ representing the resource constraints, the problem is then stated:

Maximize $I(\mathbf{X})$

Subject to $\mathbf{X} \in \mathbf{U}$

$x_i \geq 0$, integer ($i=1,2,\dots,p$)

Here, x_i is the i^{th} component of \mathbf{X} . In all past work in the literature, $I(\mathbf{X})$ was a nondecreasing function of \mathbf{X} . The objective functions considered in Chapters II and III also satisfy that condition.

1.3 Several definitions of the information of a design

The information of a design has been defined in many ways. To discuss some of these briefly, we consider the factorial model for the expected value of the observation vector $\mathbf{Y} \in E^{(N)}$

$$E(\mathbf{Y}) = \mathbf{Z}\boldsymbol{\alpha}$$

with $\boldsymbol{\alpha}$ the vector of unknown constants and \mathbf{Z} a matrix of known constants related to \mathbf{X} over which the experimenter has control. The exact relation of \mathbf{Z} and \mathbf{X} is shown in Chapter II.

Least squares estimates, $L_i^t \hat{\alpha}$ of the linear combinations $L_i^t \alpha$ are usually of interest, and most optimality criteria are directed toward the variances of these estimates. It is also assumed that the covariance matrix of Y is $\sigma^2 \mathbf{I}$, where \mathbf{I} is the identity matrix of suitable dimension.

For the moment it will be assumed that $S = Z^t Z$ is nonsingular. In general, however, the matrix S will be singular in this dissertation.¹ When S is nonsingular, there exists a set of $L_i^t \hat{\alpha}$ which have variances $v_i = \sigma^2 S_{ii} / \det(S)$, $i = 1, \dots, p$, S_{ii} is the cofactor of s_{ii} in $S = (s_{ij})$, and "det" denotes the determinant operator.

Under certain restrictions on Z , Plackett and Burnam (42) develop designs which simultaneously minimize the v_i . (Of course, simultaneous minimization of the v_i is generally not possible, but the authors do establish this property for a class of additive designs for which the number of factor-levels are all equal.) If the generalized variance is defined as the product of the v_i , this quantity (except for the multiplicative constant σ^2) is equal to the product of the eigenvalues of the inverse

¹Later, the one-factor case will be an exception, in that the nonsingular matrix will greatly facilitate discussion. Although the α and Z may be transformed in any model to achieve a nonsingular S , interpretations become more difficult and hence this transformation will be avoided.

of S , or S^{-1} , which in turn equals the determinant of S^{-1} . This latter quantity divided by the maximum obtainable was defined by Wald as the design efficiency; latin square arrangements are shown to have efficiency equal to one (50).

Defining average variance to be the sum of the v_i , Kempthorne (35) shows the relationship of this quantity to the harmonic mean of the eigenvalues of S . Kshirsagar (36) extends these results, showing the relationship of the generalized and average variances, and discusses optimal properties of the balanced incomplete block and Youden Square designs.

Turning to the case of a possibly singular S , Ehrenfeld (19) defines the efficiency of a design to be the ratio of the minimum nonzero eigenvalue of S to the maximum such value obtainable over all feasible arrangements. It is this criterion that will be used in Chapter II.

Webb (51) defines an "estimation index" proportional to the generalized variance and a "fitting index" inversely proportional to the average variance. He calculates these indexes for various designs.

1.4 A typical problem

For illustrative purposes, a typical problem will be posed and its characteristics determined throughout the general problem development. Suppose a manufacturer of ceramic parts is interested in maintaining the average level of an observable characteristic γ , a quality measure of the strength of the manufactured material. Suppose that three production variables (controllable) are known to affect the strength, and these are (a) the pulverizing process of the raw materials (process 1 or 2), (b) the pressure (PSI) under which the parts are formed, and (c) the oven temperature at which the parts are "baked." Assuming that only two values (or levels) of each of the three independent variables are of interest, an additive linear model is assumed:

$$E(y_{ijk}) = M + u_i + v_j + w_k \quad (i, j, k = 1, 2)$$

where

E is the expectation operator,

M is a constant strength level

u_i is the differential effect on strength of process one or two ($i = 1, 2$)

v_j is the differential effect on strength of pressure level low or high ($j = 1, 2$)

w_k is the differential effect on strength of oven temperature low or high ($k = 1, 2$)

All the independent differential effects are assumed unknown parameters. It is then desired to estimate linear relations in the parameters of the type $(u_2 - u_1, v_2 - v_1, w_2 - w_1)$, here the differential effects of the independent variables on strength. If several observations of y_{ijk} are to be taken, we label such an experiment in the following tabular manner:

	v_1	v_2
u_1	x_1	x_2
u_2	x_3	x_4
	w_1	

	v_1	v_2
u_1	x_5	x_6
u_2	x_7	x_8
	w_2	

In this table, x_i ($i = 1, 2, \dots, 8$) represents the number of observations that are to be taken with the corresponding settings of the independent variables u, v, w , as noted. For instance, x_5 represents the number of observations with the (u_1, v_1, w_2) factor combination, or variables u, v, w , at levels 1, 1, 2, respectively. Defining the vector $X = (x_1, x_2, \dots, x_8) \in E^{(8)}$, for any information function $I(X)$, nondecreasing in X , we would like to increase X without limit to accomplish a maximum in information, $I(X)$. However, restrictions of some form always exist on X . Classical techniques treat this problem (of the general class called 2^k factorials) with the restriction: $\sum x_i = 8\delta$ for $\delta = 2^{-m}$, (here, $m = 0, 1$). For

$m = 1$, these are fractional factorials in common use (34, 15, 1, 14, 11). In the example, the classical design with known optimal properties is $\delta = 1/2$, an orthogonal one-half replicate. There are two such arrangements, $X_0 = (1, 0, 0, 1, 0, 1, 1, 0)$ and $E_1 - X_0$, where E_1 is a vector with all elements unity.

Suppose, however, that different types of constraints were present in the form $(x_1 + x_2 + x_3 + x_4 \leq 3)$, and $(x_1 + x_2 + x_5 + x_6 \leq 2)$. These types of restrictions will be called resource constraints, the two listed representing an upper limit on the total observations available for variable w at level 1, and variable u at level 1, respectively. Also, if we attach a positive number c_i to the i^{th} factor combination, representing the cost or time to take an observation (increment x_i by 1), the cumulative experimental cost or time may be constrained by the value c_0 . It is these types of problems that will be of interest during the next chapters.

It is noted that the constants (u_i, v_j, w_k) imply a "fixed-effect" model, assumed throughout. Unless specified, it will generally not be assumed that the model is additive (separable in the i, j, k). Also, the central problem is that of selecting a layout, combination, or arrangement of factor levels which will ultimately be observed in a completely randomized design, hence the errors will always be assumed uncorrelated. For example, we consider

properties of latin square “arrangements” rather than latin square
“designs.”

CHAPTER II

A MINIMUM EIGENVALUE CRITERION

2.1 Detailed problem statement

The linear estimation problem is examined in more detail in this section. The linear model $E(Y) = Z\alpha$ is postulated with $Y \in \Xi^{(n)}$ a random observation vector with covariance matrix $\sigma^2 I$, α ($m \times 1$) a vector of unknown constants, and Z ($N \times m$) the design matrix over which the experimenter has control. Given Z , an ($m \times m$) information matrix $S = Z^t Z$ is defined, and relative to the set of linearly estimable functions¹ $V = \{L_i^t \alpha; L_i^t L_i = 1\}$, Ehrenfeld (19) shows that the variance of least squares estimators of elements of V , $L_i^t \hat{\alpha}$, attains a maximum value σ^2 / λ , with λ the minimum nonzero eigenvalue of S (provided the rank of S is a maximum $r \leq m$, taken over x_i , for a given row space of Z). The eigenvalues of S are taken to be $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r > 0$. When $r = m$, S is nonsingular, again not generally assumed here.

¹ $L_i^t \alpha$ is linearly estimable if there exists an estimate $\hat{L}_i^t \alpha$, linear in the observations, such that $E[\hat{L}_i^t \alpha] = L_i^t \alpha$

The nature of the factor arrangement problem is now considered. The linear model includes a definition of several factors, each of which can be controlled at one of several "levels." A particular factor-level combination can be thought of as corresponding to an element of the index set $\mathcal{I} = \{1, 2, \dots, p\}$. A set of column vectors $\{\xi_i; \xi_i \in E^{(m)}, i \in \mathcal{I}\}$ is defined whose elements are 0 or 1. Each ξ_i^t becomes a row of Z , specifying which elements of α are included in an observation of Y . Further, the selection by the experimenter of the integers $\{x_i, i \in \mathcal{I}\}$ represents the number of observations that will be taken at each factor combination. Thus, the factor arrangement to be observed is the specification of the vector $\{x_i\} = X \in E^{(p)}$. The matrix Z may be structured as follows:

$$Z = (\xi_1, \xi_1, \dots, \xi_1, \xi_2, \dots, \xi_2, \dots, \xi_p, \dots, \xi_p)^t$$

with each ξ_i repeated x_i times so that

$$S = Z^t Z = \sum_{i=1}^p x_i (\xi_i \xi_i^t).$$

It is seen that the minimum nonzero eigenvalue, λ_r , of S is then a function of X and the information of such a design is then defined as $\lambda_r(X)$. The mathematical problem to be solved is stated:

Maximize $\lambda_r(X)$

Subject to $X \in U$

$$x_i \geq 0, \text{ integer } (i=1, \dots, p)$$

2.2 The constraint region and properties of the objective function

Several pertinent properties of the function $\lambda_r(X)$ are now listed. The eigenvalues of S are defined to be $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r \geq \dots \geq \lambda_p \geq 0$, with $\lambda_r > 0$ if S is of maximal rank r , and $\lambda_j = 0$ ($j=r+1, \dots, m$) in any case. (Note that S is a positive, real symmetric matrix).

As is well known, $\lambda_r(X)$ is a continuous function of X (54). Other properties include:

- (a) $\lambda_r(X)$ is nondecreasing in elements of X ; i.e., if X_1 is element by element greater than X_2 , then

$$\lambda_r(X_1) \geq \lambda_r(X_2)$$

- (b) If s_{ij} is the (i, j) element of S , then

$$\lambda_r(X) \leq \text{Minimum} \{s_{ii}\}$$

- (c) $\lambda_r(kX) = k \lambda_r(X)$, for scalar $k > 0$

- (d) $\sum \lambda_i = \text{Trace}(S) = v \sum x_i$, where v is the (constant) number of 1's in any ξ_i

(e) $\sum \lambda_i^2 = \text{Trace}(S^2) = \sum_{i,j} s_{ij}^2$

- (f) If the linear additive model has g factors, each factor having n_i levels ($i=1, 2, \dots, g$) and if \bar{n} is the maximum

of the n_i , then $\lambda_r(x) \leq N/\bar{n}$, where $N = \sum x_i$

$$(g) \lambda_r(x) \leq [(\sum x_i)^2 - \sum \lambda_i^2]^{1/2} / [r(r-1)]^{1/2} \quad (\text{Reference 12})$$

(Properties b, d, and f are discussed in Appendix A.)

The constraint region U will be assumed to be defined by constraints linear in X . Two types of constraints are considered: resource limitations and an upper bound on total experimental cost or time. The resource limitations are of the form $AX \leq b$, with elements of A zero or one and b integer-valued. To each factor combination is attached a cost or time per observation, c_i , and total experimental cost or time expenditure is assumed to be $\sum x_i c_i$, which has an upper bound constraint c_0 . Consider the example of section 1.4. Suppose that the total observations of factors represented by u , v , and w , at their "low" and "high" levels is bounded by (b_1, b_2) , (b_3, b_4) , and (b_5, b_6) , respectively. Given the $\{c_i\}$ and the constant c_0 , the constraint region U would be defined by the inequalities:

$$\begin{array}{rcl} x_1 + x_2 & + x_5 + x_6 & \leq b_1 \\ & x_3 + x_4 & + x_7 + x_8 \leq b_2 \\ x_1 & + x_3 & + x_5 & + x_7 & \leq b_3 \\ & x_2 & + x_4 & + x_6 & + x_8 & \leq b_4 \\ x_1 + x_2 + x_3 + x_4 & & & & \leq b_5 \\ & & & x_5 + x_6 + x_7 + x_8 & \leq b_6 \\ & & & & \sum_{i=1}^8 c_i x_i \leq c_0 \end{array}$$

Each resource constraint represents the total number of observations allowable for each factor level.

2.3 The one-factor problem

This case may be solved by inspection. If the original model $E(Y) = Z\alpha$ has a singular $S = Z^t Z$, this corresponds to a parameter vector $\alpha = (\mu, \beta_1, \beta_2, \beta_3, \dots, \beta_p)^t$. Defining:

$$\beta'_i = \mu + \beta_i, i = 1, 2, \dots, p$$

$$\alpha' = (\beta'_1, \beta'_2, \dots, \beta'_p)^t$$

and Z' to be the matrix Z with the first column removed, we have:

$$E(Y) = Z\alpha = Z'\alpha'$$

Further, $S' = (Z')^t(Z')$ will be a diagonal, nonsingular matrix since now each row of Z' has one unit entry with the remaining components zero.¹

The element s_{ii} corresponds to the number of observations of the i^{th} level (or the number of measurements involving $\beta'_i = \mu + \beta_i$),

¹As mentioned in Chapter I, only in the one-factor problem will the concept of this transformation be utilized. Although textbooks treat such a case for an arbitrary number of factors, the general case of a singular S will be treated for two or more factors.

and is exactly the eigenvalue λ_i of S ($i=1,2,\dots,p$). As mentioned above, the resource constraints are assumed to be limitations on the total observations for each parameter or factor level, so that this problem may be stated:

Maximize w

Subject to $w \leq x_i$ ($i=1,2,\dots,p$)

$$x_i \geq 0, \text{ integer } (i=1,2,\dots,p)$$

$$\sum_i^p c_i x_i \leq c_0$$

$$x_i \leq d_i \quad (i=1,2,\dots,p; d_i \text{ integer})$$

With \underline{d} the minimum of the $\{d_i\}$, κ the largest integer for which $\sum_i c_i \leq c_0/\kappa$, \underline{n} the minimum of (\underline{d}, κ) and X_0 the vector with all components unity, a solution is $\underline{n} X_0$. It is noted that this solution is not necessarily unique since, for a solution $\{x_i^0\} = X^0$ that has $\sum c_i x_i^0$ strictly less than c_0 , some of the $\{x_i^0\}$ could possibly be increased within the constraints even though the objective function, w , remained unchanged. Essentially, all the resources might not be "used up." This shortcoming could be remedied somewhat by considering the problem in stages, the n^{th} stage being the above problem with the additional constraint, $\sum x_i = n$. By solving the stage with resultant λ_r^n , n could be increased until the feasible set was empty, selecting the solution of the m^{th} stage as optimal, where

m is the largest integer such that $\lambda_r^m \geq \lambda_r^q$, for all stages q in the sequence.

2.4 The gradient search procedure

The problem of section 2.1 is not a typical mathematical program. The objective function is not known to be convex, and the feasibility region consists only of vectors with integer coordinates. Standard mathematical programming techniques would appear to fall short in solving this type of problem, with the possible exception of a search technique or enumeration methods of integer programming (38, 7). The latter will be explored in a later chapter. For the present, the gradient search is discussed. Since the feasibility region is integer-valued, standard approaches are not available. Only in the one-dimensional case are such integer or "lattice" search problems considered (53). However, the gradient technique might be applicable if the objective function were relatively "linear" for integer steps. Presuming that this might be the case, the search procedure is reviewed.

The well-known ingredients of a gradient search procedure are the following (see, for instance, ref. 23):

- (a) An initial feasible solution
- (b) For every X , an approximation of the gradient $\nabla \lambda_r(X)$

(c) Calculation of a feasible direction vector dX , where

$$\nabla \lambda_r \cdot dX > 0 \text{ and } X + s dX \in U, \text{ for some scalar } s.$$

(d) Stopping rule(s).

The general premise in the gradient approach is that the step size s , when confined to values for which $s dX$ is integer, is small enough so that $\lambda_r(X) + \nabla \lambda_r \cdot s dX$ is approximately equal to $\lambda_r(X + s dX)$. As will be seen, this is generally not true and the gradient search procedure confined to integer X has limited value.

Relative to the stopping rules, since the objective function is not known to be convex, an X_0 which is a good local optimum need not necessarily be a solution to the global problem. However, if $\lambda_r(X)$ attains a known upper bound, obviously the corresponding X is optimal. For instance, the upper bound of property (b) in section 2.2 might be attained. Also, presuming that we are at a feasible point X , that is where $AX \leq b$, $C^t X \leq c_0$, and given the gradient $\mu = \nabla \lambda_r(X) = [D_1 \lambda_r(X), \dots, D_p \lambda_r(X)]$, Hadley (32) suggests solving the following nonlinear program for a "step" Y given some k :

$$\begin{aligned} &\text{Maximize} && u^t Y \\ &\text{Subject to} && A(X + Y) \leq b \\ & && C^t(X + Y) \leq c_0 \\ & && Y^t Y \leq k \end{aligned}$$

If the solution $u^t Y_0$ is zero, search could be halted with this local maximum.

2.5 Obtaining an initial feasible solution

For convenience in presentation in the remainder of this chapter, it will be assumed that the additional constraint, $\lambda_r(X) > 0$, is placed on the solution X .¹ Because it is desired to keep the time-consuming eigenvalue calculations (Appendix C) to a minimum, methods for obtaining a "good" initial solution are now discussed. The problem of obtaining a feasible solution is to find an X^0 so that the four constraints are satisfied:

$$(a) \quad AX^0 \leq b \quad a_{ij} = 0, 1 \quad (i=1, \dots, r; j=1, \dots, p), \quad X^0 \geq 0$$

$$(b) \quad C^t X^0 \leq c_0 \quad c_j \geq 0 \quad (j=1, \dots, p), \quad X^0 \geq 0$$

$$(c) \quad x_i^0 \text{ integer} \quad (i=1, \dots, p)$$

$$(d) \quad \lambda_r(X^0) > 0$$

Should either constraint (a) or (d) be eliminated, advantage can be taken of the resulting problem structure. For instance, consider

¹This constraint is superfluous in the maximization problem, but is necessary in the gradient search method since the gradient approximation that will be used is valid only when $\lambda_r > 0$.

the problem of obtaining an X^0 satisfying constraints b, c, and d. Note that requirement d is satisfied only if the space generated by the set $\{\xi_i^0; x_i^0 > 0\}$ is of dimension r , i.e., there must be r linearly independent vectors corresponding to the nonzero x_i . Therefore, if we can choose a minimum cost collection, \mathcal{Y} , of r vectors from the ξ_i and set $x_i^0 = 1$ if $\xi_i \in \mathcal{Y}$, with $x_j^0 = 0$ otherwise, the resultant $X^0 = (x_1^0, \dots, x_p^0)$ is a feasible solution (if $C^t X^0 \leq c_0$) or no feasible solution exists ($C^t X^0 > c_0$) since X^0 is the minimal cost vector subject to constraints b, c, and d. A method for finding the set \mathcal{Y} may be found in Appendix B.

Suppose now that constraint (d) is ignored and again we seek a minimum cost solution and formulate the problem (inserting equalities in the resource constraint for simplicity):

$$\begin{aligned}
 (2.5.1) \quad & \text{Minimize} && C^t X \\
 & \text{Subject to} && AX = b \\
 & && x_i \geq 0, \text{ integer } (i = 1, 2, \dots, p)
 \end{aligned}$$

If, for the solution X^0 to this integer program, $C^t X^0 > c_0$, obviously no feasible solution exists. On the other hand, if $C^t X^0 \leq c_0$ and $\lambda_r(X^0) > 0$ one has obtained a feasible solution.

One final observation is made relative to this last approach. If the model $E(Y) = Z\alpha$ is assumed additive, each vector of the set

$\{\xi_i\}$ (potentially a row of Z), corresponds precisely to a column of the matrix A above. This is due to the way the resource constraints were defined. As a result, the constraint $\lambda_r(x) > 0$ can be adjoined in (2.5.1) by requiring instead that the solution be non-degenerate. It is possible that an algorithm could be devised to enforce such nondegeneracy, although no such treatment is known to this author. Algorithms for the other extreme, enforcing a high degree of degeneracy, have been devised for the fixed-charge linear programming transportation problem.

2.6 A gradient approximation

For a gradient calculation, Goertzel and Tralli (26) consider the matrix S with an eigenvalue $\lambda > 0$ and corresponding eigenvector P of unit length, perturbed by ΔS , resulting in a corresponding eigenvalue $\lambda' = \lambda + \Delta\lambda$ and eigenvector $P' = P + \Delta P$, so that:

$$(2.6.1) \quad SP = \lambda P$$

$$(S + \Delta S)(P + \Delta P) = (\lambda + \Delta\lambda)(P + \Delta P)$$

Expanding the latter expression, simplifying due to equation (2.6.1), and multiplying by P^t results in:

$$P^t S(\Delta P) + P^t (\Delta S)P + P^t (\Delta S)(\Delta P) = P^t \lambda (\Delta P) + P^t (\Delta\lambda)P + P^t (\Delta\lambda)(\Delta P)$$

The leading terms on each side cancel (since, from 2.6.1, $P^t S = P^t \lambda$); utilizing the fact that $P^t P = I$ results in:

$$\Delta \lambda = \frac{P^t(\Delta S)P + P^t(\Delta S)\Delta P}{1 + P^t(\Delta P)}$$

Since ΔP involves information we do not have until $\lambda' = \lambda + \Delta \lambda$ is calculated, we approximate the above by:

$$\Delta \lambda \doteq P^t(\Delta S)P$$

Defining $\Delta S = \Delta x_i (\xi_i \xi_i^t)$ then:

$$\frac{\Delta \lambda}{\Delta x_i} \doteq P^t(\xi_i \xi_i^t)P = v_i^2$$

with $v_i = \xi_i^t P$. The change in λ is approximated by the square of a scalar, the inner product of ξ_i^t and the eigenvector associated with λ . This inner product is simply the sum of those elements of P which correspond to the nonzero elements of ξ_i (zero or one).

2.7 Three search illustrations

A computer program was written to generate these gradients for a given point X , and the "linearity" approximation investigated for various design configurations. Only for designs which allowed the total $\sum_i^p x_i$ to be greater than p did the optimization seem to merit the gradient approach over, say, pure random search. This

is obviously a serious deficiency, if true in general, since much of the interest in experimental design lies in problems where the total number of observations must be far less than the total factor-level combinations, ρ . After generating several series of these computations and investigating several heuristic rules unsuccessfully, work was halted along this line of investigation.

Three search histories are next illustrated, one of which resulted in a local optimum solution. In each of the following, the problem to be solved is:

$$\begin{aligned} &\text{Maximize} && \lambda_r(X) \\ &\text{Subject to} && \sum x_i \leq N \\ &&& x_i \geq 0, \text{integer } (i=1, \dots, p) \end{aligned}$$

The resource constraint region was simplified to more easily investigate the properties of $\lambda_r(X)$. Since $\lambda_r(X)$ is nondecreasing in X , the inequality $\sum x_i \leq N$ may be replaced by the equality $\sum x_i = N$.

Illustration one.--A simple three-factor, two-level experiment is considered, the type described in (1.4), with the maximum number of observations N , equal to four. The solution to this problem, noted in (1.4), is either $X_0 = (1, 0, 0, 1, 0, 1, 1, 0)$ or $E_1 - X_0$, where E_1 is a vector with all elements unity. Here S has maximal rank four,

so $r = 4$. The following rules are defined to increase (and decrease if necessary to insure feasibility) one component at a time, with the minimum step, 1 or -1, for a maximum change in λ_4 :

- (a) Given a feasible solution, $X = \{x_i\}$, calculate the values of (i, j) such that $d_{ij} = \left[\frac{\Delta \lambda_4}{\Delta x_i} - \frac{\Delta \lambda_4}{\Delta x_j} \delta_j \right]$ is maximized. In this example, $\delta_j = 1$ if $\sum x_i = 4$, otherwise $\delta_j = 0$. Let the new X have elements $\{x_k + \sigma_k\}$, where

$$\sigma_k = \begin{cases} 1, & k = i \\ -1, & k = j \text{ and } \sum x_i = 4 \\ 0, & \text{otherwise} \end{cases}$$

- (b) Stopping Rule: If $\lambda_4 = N / \bar{n} = 4/2$, stop.

Table 1 illustrates the history for one initial feasible solution. D represents the maximum d_{ij} or predicted change in λ_4

TABLE 1
 INTEGER THREE-FACTOR, TWO-LEVEL SEARCH PROBLEM

X	λ_4	$D(i, j)$	(i, j)
(1,1,0,0,1,0,1,0)	0.59	1.2	(6,1)
(0,1,0,0,1,1,1,0)	0.30	1.2	(4,5)
(0,1,0,1,0,1,1,0)	0.30	1.8	(1,2)
(1,0,0,1,0,1,1,0)	2.00 ^a	0.0	

^aComputation stopped with optimal λ_4 .

Although the optimum was found in the last row, the gradient technique was not responsible since the increase expected in λ_4 at each step (D) is seen to be far in excess of what was obtained. Relaxing the integer requirement, and introducing steps of size 0.5 resulted in the history shown in Table 2.

TABLE 2

NONINTEGER THREE-FACTOR, TWO-LEVEL SEARCH PROBLEM

X	λ_4	$D(i,j)$	(i,j)
(1,1,0,0,1,0,1,0)	0.59	0.6	(6,1)
(0.5,1,0,0,1,0.5,1,0)	0.82	0.98	(4,2)
(0.5,0.5,0,0.5,1,0.5,1,0)	1.0	0.5	(3,5)
(0.5,0.5,0.5,0.5,0.5,0.5,1,0)	1.2	0.8	(8,7)
(0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5)	2.0	0.0	

Again, the optimal solution $\lambda_4=2$ is obtained, but with noninteger components in X . The estimate D is still much higher than the actual change in λ_4 . It was found that "steps" in $\{x_i\}$ had to be reduced to 0.1 or 0.2 in order that $\Delta\lambda_4$ be approximately linear in ΔX and hence a proposed step direction to maximize $\Delta\lambda_4$ be successful. However, as seen in the second case of this example, with no provision to "force" an integer solution, the search converges toward the point ($x_i = \text{constant for all } i$), in this case.

Illustration two.--Two fundamental problems associated with the search technique are illustrated by the previous example. These problems also exist in the example in this section, so a different approach is shown. This problem has three factors at 4, 4, and 3 levels. We suppose the number of observations is limited to twelve. A design for this problem was proposed by Webb (51), although no comments as to its properties were given. The proposed solution was $x_i = 1$, for $i = 1, 9, 11, 16, 20, 24, 27, 29, 31, 38, 42, 46$ and $x_i = 0$, otherwise (Figure 1). In this problem $r = 9$, and the value for λ_q for the above values of the $\{x_i\}$ is $\lambda_q = 2.0$. Since $\lambda_q \leq 12/4 = 3$, it might be expected that a better design exists for $\sum x_i = 12$. It is the opinion of this author after many computer runs that no better design does exist and 2.0 is the maximum value attainable for λ_q . Integer search similar to that in the previous example resulted in little success. A different approach was tried as follows: Rather than begin with a point X at the boundary of the feasible region $\sum x_i = 12$, begin with an interior point. Since $\lambda_q = 0^1$ for many of these points it was decided to take a unit step (when in the interior) which minimized λ_1 , the maximum eigenvalue. It

¹The gradient approximation for $\lambda_r(X)$ as developed earlier is not valid for $\lambda_q = 0$.

x_1	x_{19}	x_{25}	x_{47}
x_4	x_{16}	x_{28}	x_{36}
x_7	x_{19}	x_{31}	x_{43}
x_{10}	x_{22}	x_{34}	x_{46}

x_2	x_{14}	x_{26}	x_{38}
x_5	x_{17}	x_{29}	x_{41}
x_8	x_{20}	x_{32}	x_{44}
x_{11}	x_{23}	x_{35}	x_{47}

x_3	x_{15}	x_{27}	x_{39}
x_6	x_{18}	x_{30}	x_{42}
x_9	x_{21}	x_{33}	x_{45}
x_{12}	x_{24}	x_{36}	x_{48}

Fig. 1.--Webb's solution (51) for a $4 \times 4 \times 3$ problem, with twelve observations (solution encircled).

was hoped that by suppressing λ_1 , the increase in X_i would be more apportioned to the lower-valued eigenvalues. Table 3 is a history of this approach. The initial value of X was $X_{35} = 1$, and all other $X_i = 0$. The column labeled dX indicates at each step which X_i was increased from zero to one.

TABLE 3
INTEGER SEARCH, MINIMIZING MAXIMUM EIGENVALUE

dX	λ_1	λ_9
(initial $X_{35} = 1$)	4.0	0
$X_1 = 1$	5.0	0
$X_{18} = 1$	6.0	0
$X_{44} = 1$	7.6	0
$X_{42} = 1$	9.8	0
$X_{25} = 1$	11.4	0
$X_{20} = 1$	13.2	0
$X_{10} = 1$	14.8	0
$X_3 = 1$	16.7	0
$X_{16} = 1$	18.7	0
$X_{32} = 1$	20.4	0.14
$X_{30} = 1^a$	22.5	1.32

^aComputation halted at boundary.

Another initial value, $x_1 = 1$, was tried, and the value of $\lambda_q = 1.32$ was (with the same approach) also obtained. Since 1.32 was quite a bit lower than 2.0, this approach was for the most part abandoned for others mentioned in section 2.5, and Chapter III. It is noted that gradient calculations for this problem were found valid for a range in step size of 0.2 to 0.25.

Illustration three.--The search approach has one area of application, that in which the total number of observations allowable is greater than the number of factor combinations. This is illustrated again merely with several gradient calculations and resultant values of λ_r for the 4x4x3 factor-level problem of the last illustration. However, it is assumed that the constraint region is given by $\sum x_i \leq 61$. Searching on the boundary gave results shown in Table 4. In this example, the initial solution was as follows:

$$x_i = 5, \quad i = 1, 9, 11, 16, 20, 24, 27, 29, 31, 38, 42, 46, \text{ and}$$

$$x_i = 0 \text{ for all other}$$

Because of the severe nonlinear character of the minimum eigenvalue function in areas of many experimental arrangement problems, no obvious extension of the gradient search method is evident. Several extensions were attempted, however.

TABLE 4
 INTEGER SEARCH, TOTAL NUMBER OF
 OBSERVATIONS LARGE

Step	λ_9	Suggested Change	Predicted Change in λ_9
1	4.4	$X_{46}, 0 \rightarrow 1$	1.8
2	6.2	$X_{34}, 0 \rightarrow 1$ $X_{42}, 5 \rightarrow 4$	1.4
3	7.5	$X_{38}, 5 \rightarrow 4$ $X_{47}, 0 \rightarrow 1$	1.0
4	8.5	$X_{24}, 5 \rightarrow 4$ $X_{48}, 0 \rightarrow 1$	0.6
5	9.0		

A heuristic rule which allowed noninteger steps (steps which preserved approximate linearity of the objective function), but attempted to weight heavily directions toward integer points, met with little success. In illustration two, section 2.7, one result of the method of beginning the search with an interior point of the resource constraint region was given. This led to a specification of an alternate objective function (the maximum eigenvalue, which was minimized) during the initial search. Although this approach did not yield direct results of any merit, the idea of an alternate but related objective function pointed to a host of possibilities. The result of following one such direction is next discussed in Chapter III.

CHAPTER III

A SUM OF SQUARES CRITERION

3.1 Sum of squares criterion defined; rationale

The minimum eigenvalue criterion gave rise to solution methods that were far from adequate. An alternate criterion is now proposed which appears similar to the minimum eigenvalue formulation and is much more attractive relative to methods of solution. Let $\{x_i^0\} = X^0$ be a solution to the following minimum eigenvalue problem:

$$\begin{array}{ll} \text{Maximize} & \lambda_r(X) \\ \text{Subject to} & X \in U \\ & X \geq 0, \text{ integer-valued} \end{array}$$

Problem (I)

It is assumed that U is bounded. Let $\mathcal{V}_m = \{X; \sum x_i = m, x_i \geq 0, i=1, \dots, p\}$

Consider the following problem:

$$\begin{array}{ll} \text{Minimize} & \sum_i \lambda_i^2(X) \\ \text{Subject to} & X \in U \\ & X \geq 0, \text{ integer-valued} \\ & X \in \mathcal{V}_m \end{array}$$

Problem P(m)

From property (g), section 2.2, we have:

$$(3.1.1) \quad k_1 \lambda_r^2(X) \leq k_2 - \sum_i \lambda_i^2(X), \quad X \in \Psi_m$$

where k_1, k_2 are constant provided $\sum x_i$ is fixed (in this case, equal to m). Now define, from the optimal solution $X^0 = \{x_i^0\}$ of Problem (I), the quantity $M_0 = \sum x_i^0$. Provided M_0 were known, Problem $P(M_0)$ would then seek to maximize an upper bound on $\lambda_r(X)$ through the inequality (3.1.1) by stipulating that $X \in \Psi_{M_0}$. We are attempting to replace Problem (I) entirely, however, so that M_0 will generally not be known (unless of course it is specified in the constraint region). This obstacle is circumvented somewhat by considering a (finite, if U bounded) sequence of problems $\{P(m); m=1, \dots, M\}$. Let the solution to the problem $P(m)$ be X^m , so that a finite sequence $\{X^m\}$ is generated. The solution to the problem is then defined to be X^q , for which it is true that $\lambda_r(X^q) \geq \lambda_r(X^j)$ for all X^j in the sequence $\{X^j\}$. The resultant solution will be called the sum of squares solution, satisfying the sum of squares criterion. That is, if X^m solves problem $P(m)$ above, then the sum of squares solution, X^q , satisfies:

$$\lambda_r(X^q) = \underset{m}{\text{Maximum}} [\lambda_r(X^m)]$$

It has been impossible to establish that the minimum eigenvalue and sum of squares criterion yield the same solution for an arbitrary feasibility region U , defined by linear inequalities. Such is the case for specific instances, however, and these are discussed in Chapter IV. In addition to inequality (3.1.1), it is pointed out that for $\sum x_i = m$, the minimization of $\sum_i \lambda_i^2$ attempts to "equalize" all the $\{\lambda_i\}$ and hence in some sense maximizes λ_r . The most appealing characteristic of the sum of squares criterion is that algorithms are available for its solution. These will be the topic of the remainder of this chapter.

3.2 A property of the objective function

From section 2.2, property (e), we have: $\sum_i \lambda_i^2 = \sum_{i,j} s_{ij}^2$ where $\{s_{ij}\} = S = \sum_K x_K (\xi_K \xi_K^t)$. The quantity v (property d of that section) was defined to be the constant number of ones in any of the vectors ξ_i . Since the elements of each ξ_i are zero or one, elements of $\xi_i \xi_i^t$ will also be zero or one. We then have:

$$\begin{aligned}
 SS &= \left(\sum_K x_K \xi_K \xi_K^t \right) \left(\sum_i x_i \xi_i \xi_i^t \right), \text{ and} \\
 \sum \lambda_i^2 &= \text{trace } S^2 = \text{trace} \left\{ \sum_{i,k} x_i x_k (\xi_i \xi_i^t) (\xi_k \xi_k^t) \right\} \\
 &= \text{trace} \left\{ \sum_{i,k} x_i x_k (\xi_k^t \xi_i) (\xi_k^t \xi_i) \right\} \\
 &= \sum_{i,k} x_i x_k (\xi_k^t \xi_i)^2 \\
 &= X^t Q X
 \end{aligned}$$

$$\text{where } Q = [(\xi_K^t \xi_i)^2]$$

The basic fact used in this development is trace $AB = \text{trace } BA$ and is applied to: $\text{trace} (\xi_K^t \xi_i) = \text{trace} (\xi_i^t \xi_K) = \xi_i^t \xi_K = \xi_K^t \xi_i$.

A balanced model is defined to be one in which all row (or column) sums of Q are equal. The additive model is always balanced. In case h -factor interaction terms (nonadditive model, parameters involving h factors) are included, the model is balanced if all h factor interaction terms possible (all factor-level effects taken h at a time) are included in the vector α . For instance, in a two-factor, two-level factorial with all two-factor interaction terms present,

$$\alpha = (\mu, \alpha_1, \alpha_2, \beta_1, \beta_2, \alpha\beta_{11}, \alpha\beta_{12}, \alpha\beta_{21}, \alpha\beta_{22}),$$

so that the four vectors $\{\xi_i\}$ are:

$$\xi_1 = (1, 1, 0, 1, 0, 1, 0, 0, 0)$$

$$\xi_2 = (1, 0, 1, 1, 0, 0, 1, 0, 0)$$

$$\xi_3 = (1, 1, 0, 0, 1, 0, 0, 1, 0)$$

$$\xi_4 = (1, 0, 1, 0, 1, 0, 0, 0, 1)$$

and the model is balanced. In this case,

$$Q = \begin{bmatrix} 16 & 4 & 4 & 1 \\ 4 & 16 & 1 & 4 \\ 4 & 1 & 16 & 4 \\ 1 & 4 & 4 & 16 \end{bmatrix}$$

For an example of an unbalanced model, suppose $\alpha = (\alpha_1, \alpha_2, \beta_1, \beta_2, \alpha\beta_{11})$.

Then

$$\xi_1 = (1, 1, 0, 1, 0, 1)$$

$$\xi_2 = (1, 0, 1, 1, 0, 0)$$

$$\xi_3 = (1, 1, 0, 0, 1, 0)$$

$$\xi_4 = (1, 0, 1, 0, 1, 0)$$

and in this instance,

$$Q = \begin{bmatrix} 16 & 4 & 4 & 1 \\ 4 & 9 & 1 & 4 \\ 4 & 1 & 9 & 4 \\ 1 & 4 & 4 & 9 \end{bmatrix}$$

3.3 The noninteger quadratic program

As mentioned previously, the advantage of the quadratic objective function is that problem $P(m)$ of section 3.1 is a quadratic program for which solution methods are available. One approach would be to ignore the integer requirement, with $P(m)$ then being

stated as (for simplicity, we define the vector X so that the constraint region is specified by linear equalities):

$$\begin{aligned} \text{Minimize} \quad & X^t Q X \\ \text{Subject to} \quad & AX = b \\ & x_i \geq 0, \quad i = 1, \dots, p \end{aligned}$$

Here, Q is positive semidefinite or positive definite so that $X^t Q X$ is convex and sufficient conditions for an optimal solution may be derived through the Kuhn-Tucker Conditions (32): X^* is optimal if $AX^* = b$, $x_i^* \geq 0, i = 1, \dots, p$, and there exists a row vector v^* with $v^* \geq 0$ and Δ^* of appropriate dimension so that:

$$\begin{aligned} 2QX^* + A^t \Delta^* - v^* &= 0 \\ v_j^* x_j^* &= 0, \quad j = 1, 2, \dots, p \end{aligned}$$

Algorithms for the definite and semidefinite cases are known when the integer requirement on the $\{x_i\}$ is ignored (32, 17, 8, among others). An optimal solution is now illustrated for the three-factor, two-level problem of Chapter I (page 6), in which the constraints are defined to be:

$$AX=b \begin{cases} x_1+x_2 & & +x_5+x_6 & & =2 \\ & x_3+x_4 & & +x_7+x_8 & =2 \\ x_1+x_2+x_3+x_4 & & & & =2 \\ & & x_5+x_6+x_7+x_8 & & =2 \\ x_1 & +x_3 & +x_5 & +x_7 & =2 \\ & x_2 & +x_4 & +x_6 & +x_8 & =2 \end{cases}$$

$$x_i \geq 0, i=1, \dots, 8$$

Assuming the additive model as stated in section 1.4, the matrix Q in the quadratic objective function is:

$$Q = \begin{bmatrix} 16 & 9 & 9 & 4 & 9 & 4 & 4 & 1 \\ 9 & 16 & 4 & 9 & 4 & 9 & 1 & 4 \\ 9 & 4 & 16 & 9 & 4 & 1 & 9 & 4 \\ 4 & 9 & 9 & 16 & 1 & 4 & 4 & 9 \\ 9 & 4 & 4 & 1 & 16 & 9 & 9 & 4 \\ 4 & 9 & 1 & 4 & 9 & 16 & 4 & 9 \\ 4 & 1 & 9 & 4 & 9 & 4 & 16 & 9 \\ 1 & 4 & 4 & 9 & 4 & 9 & 9 & 16 \end{bmatrix}$$

The initial Tableau of Dantzig's Method (17) would appear:

b	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	Δ_1	Δ_2	Δ_3	Δ_4	Δ_5	Δ_6	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8
0	2Q								$-A^t$						I_8							
b	A								0						0							

It may be seen that the following solves the problem:

$$X = (1,0,0,1,0,1,1,0)$$

$$\Delta = (20, 20, 20, 20, 16, 16)$$

$$v = (0,0,0,0,0,0,0,0)$$

Since the vectors v and X are orthogonal, and satisfy the conditions in the tableau, they are optimal. Fortunately in this case, the solution was integer-valued, which generally may not be true. This solution is also not unique.

Since the above quadratic algorithm does not guarantee the integer property of the solution, we will investigate several methods which do, two algorithms which take advantage of the quadratic objective function and an enumeration method which takes advantage of its nondecreasing characteristic.

3.4 An algorithm of Kunzi and Oettli

In a recent paper (37), Kunzi and Oettli developed an algorithm for the integer quadratic program for the case when Q is positive definite:

$$\begin{array}{ll}
 \text{Minimize} & X^t Q X \\
 \text{(I) Subject to} & A X \leq b \\
 & x_i \geq 0, \text{ integer } (i = 1, \dots, p)
 \end{array}$$

It is assumed that the feasibility region is bounded away from zero, so that the trivial solution $x_i = 0$ ($i = 1, \dots, p$) is not feasible. The form of the proof of this algorithm is changed slightly for present purposes, with the addition of the case when Q may be positive semidefinite.

First consider a solution to problem (I) without the integer constraint, not necessarily unique if Q is positive semidefinite (uniqueness is not needed as the authors point out). Let the objective function of the solution have the value $f(y) = y^t Q y = d$. An equivalent problem (to be shown) to that of (I) is the following:

$$\begin{array}{ll}
 \text{Minimize} & \pi \\
 \text{(II)} & \text{Subject to: } AX \leq b \\
 & x_i \geq 0, \text{ integer } (i = 1, \dots, p) \\
 & \pi \geq 2 \bar{x}^t Q X / \bar{x}^t Q \bar{x}
 \end{array}$$

where \bar{X} is any point for which $\bar{x}^t Q \bar{x} = d$, i.e., any point on the curve $x^t Q x = d$. Later, a constructive procedure which generates solutions to (II) is defined, but first it is shown that a solution to (II) also solves (I). This is done in four steps following the next definition.

Define, for any point X_a , the point $\bar{X}_a = r_a X_a$, where $r_a = \sqrt{d / X_a^t Q X_a}$. Note that $\bar{x}_a^t Q \bar{x}_a = d$, so that \bar{X}_a lies on the

intersection of $X^t Q X = d$ and a line through the origin and X_a (Figure 2).

Lemma 1: Given X_a , the point \bar{X}_a solves:

Maximize $2X_a^t Q X$ (Q positive definite or positive semidefinite)

Subject to $X^t Q X = d$
 $X \geq 0$

Proof:

(i) Let $L(X, \Delta) = 2X_a^t Q X - \Delta(X^t Q X - d)$ be defined over the convex set $(X \geq 0, \Delta \geq 0)$.

(ii) $L(X, \Delta)$ is the sum of three concave functions, $2X_a^t Q X$, $X^t(-\Delta Q)X$, and Δd , hence $L(X, \Delta)$ is concave. Therefore to illustrate a global maximum, it is sufficient to show that

$$\nabla_X L(X^*, \Delta^*) = 0 \quad (32), \quad (\text{for feasible } X^*).$$

(iii) Let $X^* = \bar{X}_a, \Delta^* = 1/r_a$. Then, $\nabla_X L(\bar{X}_a, 1/r_a) = 2QX_a - (2/r_a)(Q\bar{X}_a) = 0$. The Lemma is true for positive definite or positive semidefinite Q .

Theorem 1: If X_a solves (I), then (X_a, π_a) solves (II) where

$$\pi_a = 2\bar{X}_a^t Q X_a / d$$

Proof:

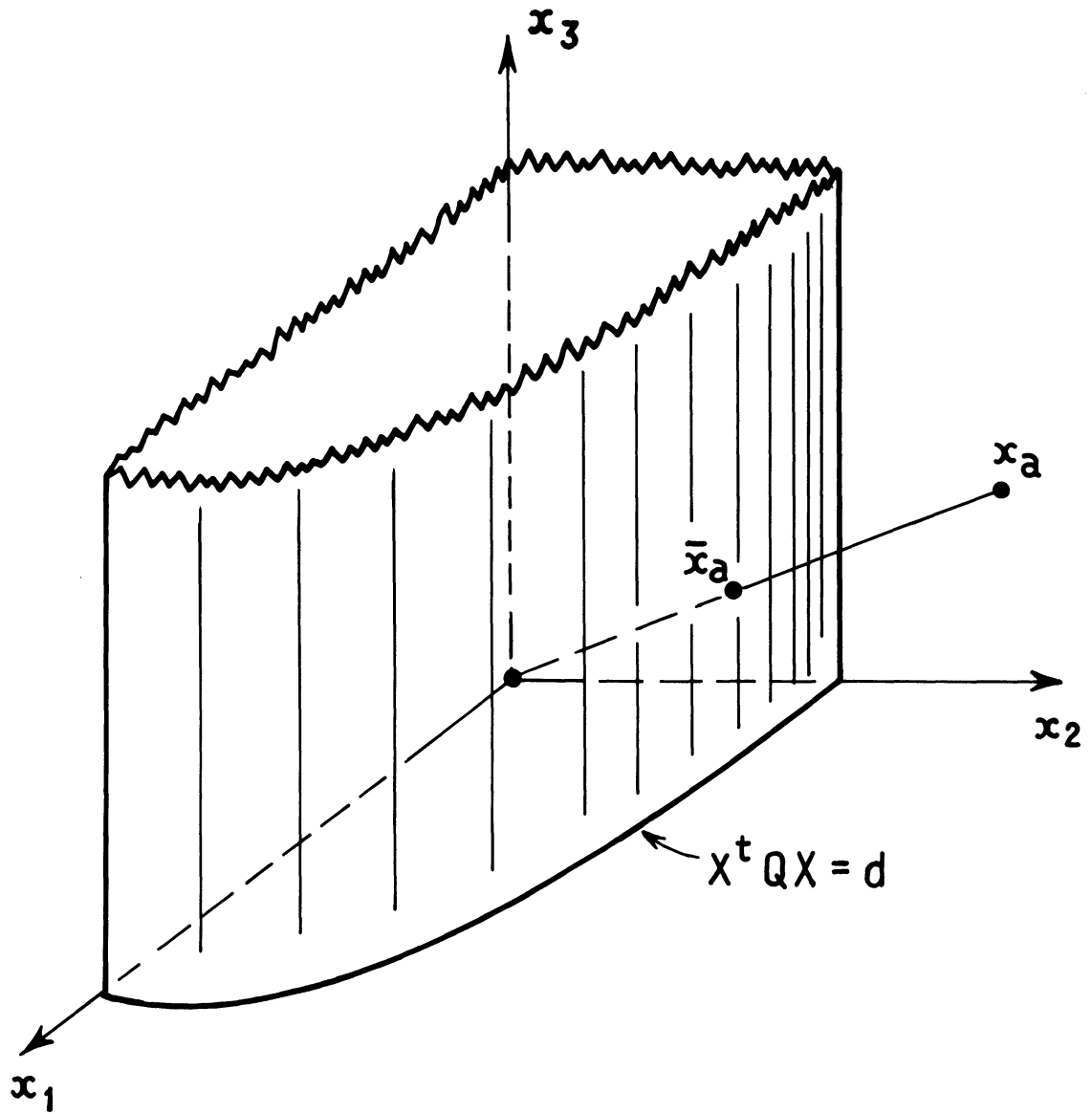


Fig. 2.--The quadratic program surface.

- (i) Let $\mathcal{Y} = \{X; AX \leq b, x_i \geq 0 \text{ integer}; i = 1, \dots, p\}$
- (ii) By definition, $\pi_a = \frac{2r_a X_a^t Q X_a}{d} = 2 \sqrt{\frac{X_a^t Q X_a}{d}}$
- (iii) From Lemma 1, $\pi_a \geq 2 X_a^t Q \bar{X} / d$, for any \bar{X} satisfying $\bar{X}^t Q \bar{X} = d$
- (iv) Suppose (π_a, X_a) do not solve (II). Then there exists

$$\text{some } X_b \text{ for which: } \pi_a > \pi_b \geq \frac{2 X_b^t Q \bar{X}_b}{d} = 2 \sqrt{\frac{X_b^t Q X_b}{d}}$$

- (v) By definition of π_a , we have:

$$\pi_a = 2 \sqrt{\frac{X_a^t Q X_a}{d}} > 2 \sqrt{\frac{X_b^t Q X_b}{d}}; X_a, X_b \in \mathcal{Y}$$

Since this implies that $X_a^t Q X_a > X_b^t Q X_b$, a contradiction of the minimum property of X_a in the hypothesis ($X_a^t Q X_a \leq X_b^t Q X_b$) the proof is complete.

Lemma 2: Let (X_b, π_b) solve (II). Then $\pi_b = 2/r_b$.

Proof:

- (i) Since (X_b, π_b) solve (II), $d\pi_b \geq 2\bar{X}^t Q X_b$ for any \bar{X} satisfying $\bar{X}^t Q \bar{X} = d$.
- (ii) Since π_b is minimum, and $2\bar{X}^t Q X_b$ is maximum at $2\bar{X}_b^t Q X_b$ (lemma 1) it must be true that: $d\pi_b = 2\bar{X}_b^t Q X_b = 2d/r_b$ so that $\pi_b = 2/r_b$

Theorem 2: If (X_b, Π_b) solves (II), then X_b solves (I).

Proof:

- (i) Let a solution to (I) be X_a . Then $\Pi_a = \Pi_b$ since (X_a, Π_a) solves (II).
- (ii) $f(X_b) = X_b^t Q X_b = (1/r_b^2)d = (1/r_a^2)d = f(X_a)$ so that X_b solves (I).

A constructive procedure for a solution to (II), and hence (I), is then described as proceeding in stages, the k^{th} stage represented by the mixed-integer program:

$$\begin{array}{ll}
 \text{Minimize} & \Pi \\
 \text{Subject to} & AX \leq b \\
 & X \geq 0, \text{ integer-valued} \\
 & \Pi \geq 2 \bar{X}_i^t Q X / d, \quad i = 1, \dots, k-1
 \end{array}$$

Let the solution to this stage be: (X_k, Π_k) . Suppose, for the first time, X_k also satisfies: $\Pi_k d \geq 2 \bar{X}_k^t Q X_k$. We then have, by the following lemma, that (X_k, Π_k) solves (II) and hence X_k is a solution to (I).

Lemma 3: Let the sequence $\{S(k)\}$ be defined as above in the constructive procedure, with K the smallest integer such that the solution (X_k, Π_k) to problem $S(k)$ also satisfies $d \Pi_k \geq 2 \bar{X}_k^t Q X_k$.

Then (X_k, π_k) solves problem (II).

Proof:

- (i) Suppose that (X_k, π_k) as defined in the hypothesis does not solve (II), but that (X_g, π_g) does.
- (ii) Then $\pi_g < \pi_k$, and further π_g is not a member of the solutions $P = \{\pi_1, \pi_2, \dots, \pi_k\}$ to the problems $S(1), S(2), \dots, S(k)$, since $d\pi_g \geq 2\bar{X}_g^t Q X_g$ and this property was true only for π_k in the sequence P .
- (iii) There exists a "v" such that $\pi_v \in P, \pi_{v+1} \in P$ and $\pi_v \leq \pi_g < \pi_{v+1}$ ($\pi_j \leq \pi_{j+1}$)
- (iv) By lemma 1, we have $d\pi_g = 2\bar{X}_g^t Q X_g \geq 2\bar{X}^t Q X_g$ for every \bar{X} that satisfies $\bar{X}^t Q \bar{X} = d$, so that (X_g, π_g) solves $S(v+1)$ rather than (X_{v+1}, π_{v+1}) . This contradiction completes the proof.

An example of the use of this method is shown by a simple two-factor, two-level additive model represented in tabular form:

		FACTOR 2	
		LEVEL 1	LEVEL 2
FACTOR 1	LEVEL 1	X_1	X_3
	LEVEL 2	X_2	X_4

The constraints are assumed to be:

$$AX=b \begin{cases} x_1+x_2 & = 2 \\ & x_3+x_4 = 2 \\ x_1 & +x_3 = 1 \end{cases}$$

$$C^t X \leq c_0 \{ x_1 + 2x_2 + 3x_3 + 4x_4 \leq 11$$

The matrix $2Q$ is:

$$2Q = \begin{bmatrix} 18 & 8 & 8 & 2 \\ 8 & 18 & 2 & 8 \\ 8 & 2 & 18 & 8 \\ 2 & 8 & 8 & 18 \end{bmatrix}$$

A solution to the noninteger quadratic program:

$$\begin{aligned} \text{Minimize} & \quad X^t Q X \\ \text{Subject to} & \quad AX = b \\ & \quad C^t X \leq c_0 \\ & \quad X \geq 0 \end{aligned}$$

is $(\frac{1}{2}, \frac{3}{2}, \frac{1}{2}, \frac{3}{2})$ which can be illustrated using methods of section 3.3. The first step in the Kunzi-Oettli algorithm, given this non-integer solution, is then to solve:

$$\begin{array}{ll}
\text{Minimize} & \pi \\
\text{Subject to} & AX = b \\
& C^t X \leq c_0 \\
& 80\pi \geq 28x_1 + 44x_2 + 28x_3 + 44x_4 \\
& X \geq 0, \text{ integer-valued.}
\end{array}$$

A solution to this problem is $\pi_1 = 2, X_1 = (0, 2, 1, 1)$, with $\bar{X}_1 = X_1 \sqrt{\frac{80}{82}}$. Since $\pi_1(d/2) = 80$ is less than $\bar{X}_1^t Q X_1 = 82 \sqrt{\frac{80}{82}} = 80.99$, we proceed to problem S(2):

$$\begin{array}{ll}
\text{Minimize} & \pi \\
\text{Subject to} & AX = b \\
& C^t X \leq c_0 \\
& 80\pi \geq 28x_1 + 44x_2 + 28x_3 + 44x_4 \\
& 80\pi \geq \sqrt{\frac{80}{82}} (26x_1 + 46x_2 + 30x_3 + 42x_4) \\
& X \geq 0, \text{ integer-valued.}
\end{array}$$

A solution to problem S(2) is $X_2 = (0, 2, 1, 1), \pi_2 = 2.025$. Further, $\pi_2(d/2) = 80.992 \bar{X}_2^t Q X_2 = 80.99$, satisfying the requirement that X_2 is optimal.

3.5 Witzgall's parabolic constraint algorithm

An algorithm which also handles the case when Q is positive definite or positive semidefinite is reported by Witzgall (55). Here,

problem (I) in section 5.2 is transformed to:

$$\begin{aligned}
 &\text{Minimize} && Z \\
 &\text{Subject to} && Z \geq X^t Q X \\
 &&& AX \leq b \\
 &&& X_i \geq 0, \text{ integer } (i=1, 2, \dots, P)
 \end{aligned}$$

with the objective function linear and the constraint region nonlinear. The technique is shown to resemble the integer linear programming algorithm of Gomory, except that the nonlinear constraints are handled slightly differently.

The method requires that the function $X^t Q X$ be transformed into: $a_0 - L_0(X) - L_1^2(X) - \dots - L_k^2(X)$, where each $L_i(X)$ is linear in X and all are linearly independent of $L_0(X)$. This transformation could be unattractive as evidenced by the following two-factor, two-level problem.

$$\begin{aligned}
 \text{Here, } \xi_1 &= (1, 1, 0, 1, 0), & \xi_2 &= (1, 0, 1, 1, 0), \\
 \xi_3 &= (1, 1, 0, 0, 1), & \xi_4 &= (1, 0, 1, 0, 1),
 \end{aligned}$$

so that:

$$S = \begin{bmatrix} x_1 + x_2 + x_3 + x_4 & x_1 + x_2 & x_3 + x_4 & x_1 + x_3 & x_2 + x_4 \\ x_1 + x_2 & x_1 + x_2 & 0 & x_1 & x_2 \\ x_3 + x_4 & 0 & x_3 + x_4 & x_3 & x_4 \\ x_1 + x_3 & x_1 & x_3 & x_1 + x_3 & 0 \\ x_2 + x_4 & x_2 & x_4 & 0 & x_2 + x_4 \end{bmatrix}$$

For this problem, the Q matrix is then:

$$\begin{bmatrix} 9 & 4 & 4 & 1 \\ 4 & 9 & 1 & 4 \\ 4 & 1 & 9 & 4 \\ 1 & 4 & 4 & 9 \end{bmatrix}$$

and $X^t Q X$ is equal to $9(x_1^2 + x_2^2 + x_3^2 + x_4^2) + 8(x_1 x_2 + x_1 x_3 + x_2 x_4 + x_3 x_4) + 2(x_1 x_4 + x_2 x_3)$.

Finally, "completing the square" as Witzgall suggests, results in:

$$X^t Q X = 9L_1^2(X) + (65/9)L_2^2(X) + (a/b)L_3^2(X) + (c/d)L_4^2(X)$$

where:

$$a = 4144$$

$$b = 585$$

$$c = 405,076$$

$$d = 151,515$$

$$L_1(X) = x_1 + (4/9)(x_2 + x_3 + x_4/4)$$

$$L_2(X) = x_2 + (9/65)x_3 + (36/65)x_4$$

$$L_3(X) = x_3 + (585/1036)x_4$$

$$L_4(X) = x_4$$

This transformation to the L_i (not necessarily unique) could also be accomplished by computing the eigenvectors of Q . The inner product of these vectors with X would provide the set $\{L_i\}$.

3.6 A partial enumeration method of Lawler and Bell; an example

The two previous algorithms utilized the quadratic character of the objective function in their solutions. An enumeration technique of Lawler and Bell (38) is now used which does not require such a form, but takes advantage of the monotone property of $X^t Q X, X \geq 0$. This algorithm could also be used for the eigenvalue formulation since that objective was also monotone in X . Because of the calculations involved, the quadratic form will be used to illustrate this method. The problem stated in the previous section is stated with a feasibility region:

$$\begin{aligned}
 \text{Minimize} \quad & \sum s_{ij}^2 = g_0(X) \\
 \text{Subject to} \quad & g_1(X) = x_1 + x_2 \leq 4 = b_1 \\
 & g_2(X) = x_3 + x_4 \leq 2 = b_2 \\
 & g_3(X) = x_1 + x_3 \leq 3 = b_3 \\
 & g_4(X) = x_1 + x_2 + x_3 + x_4 = m = b_4 \\
 & g_5(X) = x_1 + 2x_2 + 5x_3 + 3x_4 \leq 14 \\
 & g_6(X) = x_2 + x_4 \leq 3 = b_6 \\
 & x_i \geq 0, \text{ integer } (i=1,2,3,4)
 \end{aligned}$$

Since 6 is the maximum value of m consistent with the constraints, this is the value that will be assumed (not necessarily optimal over all m). Further, with this value of m , the inequalities become equalities (with the exception of the cost constraint). The objective function may be reduced so that the problem is stated:

$$\text{Minimize } x_1^2 + x_2^2 + x_3^2 + x_4^2 + \text{(constant, ignored here)}$$

$$\text{Subject to } g_1(x) = 4$$

$$g_2(x) = 2$$

$$g_3(x) = 3$$

$$g_4(x) = 6$$

$$g_5(x) \leq 14$$

$$g_6(x) = 3$$

$$x_i \geq 0, \text{ integer } (i = 1, 2, 3, 4)$$

Since $x_1 \leq 4$, $x_2 \leq 4$, $x_3 \leq 2$, $x_4 \leq 2$, following (38) we define:

$$x_1 = z_{10} + 2z_{11} + 2^2z_{12}, \quad x_2 = z_{20} + 2z_{21} + 2^2z_{22},$$

$$x_3 = z_{30} + 2z_{31}, \quad x_4 = z_{40} + 2z_{41}$$

$$\text{where } z_{ij} = 0, 1.$$

The algorithm is then as follows

$$(\text{Let } Z = (z_{41}, z_{40}, z_{31}, z_{30}, z_{22}, z_{21}, z_{20}, z_{12}, z_{11}, z_{10}) :$$

The vector Z is considered in increasing numerical order.¹ Define Z^* to be the first vector following Z in that order which has the property that Z , element by element, is not less than or equal to Z^* . Let $g_0(Z')$ be the minimum value of $g_0(Z)$ to date. The following rules are defined indicating which vectors may be "skipped" over (note the $g_i(X)$ are nondecreasing in X or Z):

Rule 1: If $g_0(Z) \geq g_0(Z')$ skip to Z^* since no vector between Z and Z^* can have a lesser value than $g_0(Z')$.

Rule 2: If Z is feasible, skip to Z^* since no vector between Z and Z^* can have a lesser value of the function g_0 than $g_0(Z)$.

Rule 3: If $g_i(Z^*-1) < b_i, i \neq 5$, skip to Z^* since if $g_i(Z^*-1) < b_i$ no vector less than Z^*-1 but greater than Z can satisfy $g_i(Z^*-1) = b_i, i \neq 5$, since g_i is nondecreasing.

Rule 4: If $g_i(Z) > b_i$, skip to Z^* since g_i is nondecreasing ($i \neq 0$).

Figure 3 is a partial history of the solution to this problem. The optimal solution is $X = (2,2,1,1)$, with $g_0 = 10$.

¹At this point, the numerical "value" of Z is obtained by identifying Z with the integer $n(Z) = z_{41} \cdot 2^8 + z_{40} \cdot 2^7 + \dots + z_{11} \cdot 2^1 + z_{10} \cdot 2^0$. In the rules 1-4, Z refers to this binary number.

X_4		X_3		X_2			X_1			RULE	INFEASIBLE?	SKIP TO Z^* ?
Z_{41}	Z_{40}	Z_{31}	Z_{30}	Z_{22}	Z_{21}	Z_{20}	Z_{12}	Z_{11}	Z_{10}			
0	0	0	0	0	0	0	0	0	1	3		X
0	0	0	0	0	0	0	0	1	0	3		X
0	0	0	0	0	0	0	1	0	0	3		X
0	0	0	0	0	0	1	0	0	0	3		X
0	0	0	0	0	1	0	0	0	0	3		X
0	0	0	0	1	0	0	0	0	0	3		X
0	0	0	1	0	0	0	0	0	0	3		X
<hr/>												
0	0	1	0	0	0	0	1	0	0	4		X
0	0	1	0	0	0	1	0	0	0		X	
0	0	1	0	0	0	1	0	0	1		X	
0	0	1	0	0	0	1	0	1	0	4		X
0	0	1	0	0	0	1	1	0	0	4		
<hr/>												
0	1	0	1	0	1	0	0	1	0	2	$g_0=10$	X
0	1	0	1	0	1	0	1	0	0	4		X
0	1	0	1	0	1	1	0	0	0	4		X
0	1	0	1	1	0	0	0	0	0	4		
0	1	1	0	0	0	0	0	0	0	4		
1	0	0	0	0	0	0	0	0	0		X	
1	0	0	0	0	0	0	0	0	1		X	
<hr/>												
1	0	0	0	0	0	1	0	0	1		X	
1	0	0	0	0	0	1	0	1	0			
1	0	0	0	0	0	1	0	1	1		$g_0=11$	X
1	0	0	0	0	0	1	1	0	0	4		

Fig. 3.--Illustration of the enumeration technique.

3.7 Use as a starting point for the minimum eigenvalue search problem

In section 3.1 it was pointed out that a relation does exist between the minimum eigenvalue and the sum of squares criterion. It has not been proved that both criteria yield identical results in all cases, however. The solution, \bar{X} , to the minimum sum of squares problem could provide an initial starting point for the minimum eigenvalue gradient search procedure, probably a very good one. If it is true for \bar{X} that $\lambda_r(\bar{X})$ is very close to an upper bound (say, the one provided in section 2.2), then \bar{X} might be considered adequate for practical purposes.

In this chapter, an alternate formulation to the problem of Chapter II was accomplished. This alternate approach was shown to be (a) computationally attractive, (b) heuristically related to the Chapter II formulation, and (c) related to improving an upper bound for the Chapter II objective function, with some constraint on the feasibility region. The two criteria are next shown to be more closely related when, in the next chapter, the constraint region is further restricted to one for which constant cost of experimentation (all c_i equal) exists, and the total number of observations is specified.

CHAPTER IV

FURTHER EXAMPLES OF THE SUM OF SQUARES CRITERION

4.1 The equal-cost problem

The sum of squares is further substantiated as a reasonable objective in this chapter for the case when all c_i (cost or time per unit increase in x_i) are equal so that the constraint $C^t X \leq c_0$ may be replaced by $\sum x_i \leq N$. It will be shown that in three important cases this approach yields precisely the solution that would have been derived by the minimum eigenvalue criterion. In a fourth example, the arrangement proposed by Webb (51) is also derived. To facilitate these results, three observations are now made, one which describes the optimal solution of a simple noninteger quadratic program, and the other two, sufficient conditions for optimality of the integer quadratic program for the additive linear model.

4.2 Three preliminary observations

Observation one.-- $X^* = k(1, 1, \dots, 1)$ for scalar $k > 0$ not necessarily integer, solves:

$$\begin{aligned} \text{Minimize } & \sum \lambda_i^2 = X^t Q X = f(X) \\ \text{Subject to } & \sum x_i = N \\ & X \geq 0 \end{aligned}$$

where the model from which Q is derived is balanced (section 3.2).

- (i) Define E to be a $(p \times 1)$ vector with all components unity, and let $F = f(X) + \sigma(\sum x_i - N)$. Since Q is balanced, let $QE = \gamma E$ for some scalar γ .
- (ii) Since $X^t Q X$ is convex, a sufficient condition for optimality of nonnegative X^* is the existence of σ^* (and let $\kappa = N/p$ to insure feasibility) so that:
- $$\nabla_X F(X^*, \sigma^*) = 0, \text{ or } 2QX^* + \sigma^* E = 0$$
- (iii) Now $X^* = (N/p)E$, and $\sigma^* = -2N\gamma/p$ satisfies the above condition, since $\sum_j q_{ij} = \gamma$ for $i = 1, \dots, p$, true for a balanced model.

Observation two.--Assume the three-factor model, $E(Y) = Z\alpha$,

or relative to an element $y_{ijk} \in Y$,

$$E(y_{ijk}) = \mu + \tau_i + \delta_j + \tau_k \begin{cases} i = 1, \dots, n_1 \\ j = 1, \dots, n_2 \\ k = 1, \dots, n_3 \end{cases}$$

The extension of the following to an arbitrary number of factors is obvious. In this example, the total number of factor-level combinations

is $p = n_1 n_2 n_3$. It is desired to establish sufficient conditions for X^* to solve:

$$(4.1.1) \quad \begin{aligned} & \text{Minimize} && \sum_{i,j} S_{ij}^2 \\ & \text{Subject to} && \sum_{i,j,k} x_{ijk} = N \\ & && x_{ijk} \geq 0, \text{ integer for all } i, j, k \end{aligned}$$

where it is also assumed that $N < p$ and x_{ijk} denotes the number of observations to be taken of the random variable y_{ijk} . The (\cdot) subscript denotes summation over that index, so that $x_{i\cdot\cdot} = \sum_{j,k} x_{ijk}$, $x_{\cdot\cdot k} = \sum_{i,j} x_{ijk}$, etc. For a given $X = \{x_{ijk}\}$, the matrix S is represented:

$$S = \begin{bmatrix} x_{\cdot\cdot\cdot} & | & W_1 & | & W_2 & | & W_3 \\ W_1^t & | & D_1 & | & D_{12} & | & D_{13} \\ \hline W_2^t & | & D_{12}^t & | & D_2 & | & D_{23} \\ W_3^t & | & D_{13}^t & | & D_{23}^t & | & D_3 \end{bmatrix}$$

where:

$$\begin{aligned} W_1(1 \times n_1) &= (x_{1\cdot\cdot}, x_{2\cdot\cdot}, \dots, x_{n_1\cdot\cdot}) \\ W_2(1 \times n_2) &= (x_{\cdot 1\cdot}, x_{\cdot 2\cdot}, \dots, x_{\cdot n_2\cdot}) \\ W_3(1 \times n_3) &= (x_{\cdot\cdot 1}, x_{\cdot\cdot 2}, \dots, x_{\cdot\cdot n_3}) \\ D_1(n_1 \times n_1) &= \text{diag}(x_{1\cdot\cdot}, x_{2\cdot\cdot}, \dots, x_{n_1\cdot\cdot}) \\ D_2(n_2 \times n_2) &= \text{diag}(x_{\cdot 1\cdot}, x_{\cdot 2\cdot}, \dots, x_{\cdot n_2\cdot}) \\ D_3(n_3 \times n_3) &= \text{diag}(x_{\cdot\cdot 1}, x_{\cdot\cdot 2}, \dots, x_{\cdot\cdot n_3}) \end{aligned}$$

$$\left. \begin{aligned} D_{12}(n_1 \times n_2) &= \{x_{ij.}\} \\ D_{13}(n_1 \times n_3) &= \{x_{i.k}\} \\ D_{23}(n_2 \times n_3) &= \{x_{.jk}\} \end{aligned} \right\} \begin{aligned} i &= 1, \dots, n_1 \\ j &= 1, \dots, n_2 \\ k &= 1, \dots, n_3 \end{aligned}$$

From observation one, it is known that the optimal solution without the integer constraint is $X^* = \{x_{ijk}^*\} = N/p(1, 1, \dots, 1)$. It is desired to find sufficient conditions in order that an integer-valued $\bar{X} = \{\bar{x}_{ijk}\}$ also be a solution to this problem. These conditions are obviously:

$$(4.1.2) \quad \left. \begin{aligned} \bar{x}_{ij.} &= x_{ij.}^* = n_3 N/p \\ \bar{x}_{i.k} &= x_{i.k}^* = n_2 N/p \\ \bar{x}_{.jk} &= x_{.jk}^* = n_1 N/p \\ x_{ijk} &\geq 0, \text{ integer} \end{aligned} \right\} \begin{aligned} i &= 1, \dots, n_1 \\ j &= 1, \dots, n_2 \\ k &= 1, \dots, n_3 \end{aligned}$$

Observation three.--Finally, we suppose that such an optimal integer-valued \bar{X} does not exist which satisfies conditions (4.1.2).

Let:

$$I_1 = \left\{ (i, j); j=1, i=1, 2, \dots, 1+n_1+n_2+n_3; i=1, j=1, 2, \dots, 1+n_1+n_2+n_3; \right. \\ \left. i=j=2, 3, \dots, 1+n_1+n_2+n_3 \right\}$$

$$I_2 = \left\{ (i, j); (i, j) \notin I_1, i, j=1, 2, \dots, 1+n_1+n_2+n_3 \right\}$$

The matrix S then has components S_{ij} , with (i, j) either in I_1 or I_2 .

With $g(X) = \sum_{i,j \in I_2} s_{ij}^2$ and $h(X) = \sum_{i,j \in I_1} s_{ij}^2$, the problem (4.1.1) now becomes:

Minimize $g(X) + h(X)$

Subject to $\sum x_{ijk} = N$

$x_{ijk} \geq 0, \text{integer}$

We seek sufficient conditions for a vector X^0 to simultaneously minimize $g(X)$ and $h(X)$, and hence minimize the sum.

The function $h(X)$ may be written:

$$h(X) = N^2 + 3 \left\{ \sum_i x_{i..}^2 + \sum_j x_{.j.}^2 + \sum_k x_{..k}^2 \right\}$$

A sufficient condition for $X^0 = \{x_{ijk}^0\}$ to minimize $h(X)$ is that the components for each of the three sums be equal, i.e.:

$$(4.1.3) \quad \begin{cases} x_{i..}^0 = N/n_1 \\ x_{.j.}^0 = N/n_2 \\ x_{..k}^0 = N/n_3 \\ x_{ijk} \geq 0, \text{integer} \end{cases}$$

Since $g(X) = 2 \left[\sum_{i,j} x_{ij.}^2 + \sum_{i,k} x_{i.k}^2 + \sum_{j,k} x_{.jk}^2 \right]$, a sufficient condition that $g(X)$ be minimized subject to the same conditions is that:

$$(4.1.4) \quad \begin{cases} x_{ij.}^0 = N/n_1 n_2 = n_3 N/p \\ x_{i.k}^0 = N/n_1 n_3 = n_2 N/p \\ x_{.jk}^0 = N/n_2 n_3 = n_1 N/p \\ x_{ijk}^0 \geq 0, \text{ integer} \end{cases}$$

Suppose $N/n_1 n_2, N/n_1 n_3$ or $N/n_2 n_3$ are noninteger, less than one. Then no integer-valued X^0 can satisfy conditions (4.1.4). By taking any one of the three elements of $g(x) = 2\{\sum x_{ij.}^2 + \sum x_{i.k}^2 + \sum x_{.jk}^2\}$ the problem becomes, say:

$$\begin{aligned} & \text{Minimize} && \sum_{i,j} x_{ij.}^2 \\ & \text{Subject to} && x_{...} = N \\ & && x_{ijk} \geq 0, \text{ integer} \end{aligned}$$

Since we are requiring the sum of squares of n_1, n_2 nonnegative integer elements to be minimized while constraining their sum to a number less than $n_1 n_2$, an obvious sufficient condition for optimality is:

$$(4.1.6) \quad \begin{cases} x_{ij.}^0 = 0, 1 \\ x_{i.k}^0 = 0, 1 \\ x_{.jk}^0 = 0, 1 \end{cases}$$

Here, the last two equations are obtained in a manner exactly as the first.

These three observations are now used in further illustrations of the sum of squares criterion. For brevity, a solution using the sum of squares criterion will be called L-optimal if it is a member of the set of solutions to the minimum eigenvalue criterion with the same constraint region.

4.3 The one-factor problem

The problem to be solved according to the second criterion is (we may assume each x_i is equal to an eigenvalue of S in this case):

$$\text{Minimize} \quad \sum_i^p \lambda_i^2 = \sum_i^p x_i^2$$

$$\text{Subject to} \quad x_i \leq a_i \quad (i=1, \dots, p)$$

$$\sum x_i = m$$

$$x_i \geq 0, \text{ integer } (i=1, \dots, p)$$

It is assumed that $a_1 \leq a_2 \leq \dots \leq a_p$, and that a feasible solution exists, i.e., $\sum_i a_i \geq m$. It is desired to show that for a solution to this problem $X' = \{x'_i\}$, it is true that $x'_i = \text{Minimum} \{a_i, [m/p]\}$, $[b]$ denoting the largest integer less than or equal to b . In this case the Minimum $\{x_i\}$ has attained its largest possible value, and hence X' is L-optimal.

Let $\Gamma_m = \{X; |x_i - x_j| \leq 1; \sum x_i = m; x_i, x_j \text{ integer}, i, j = 1, \dots, p\}$.

If $X' \in \Gamma_m$, then X' is L-optimal since the value of minimum $\{x'_i\}$ cannot be increased subject to $\sum x'_i = m$. Suppose now that $X' \notin \Gamma_m$ and that $x_1 < a_1$. Two cases are considered (let $f = \text{maximum } \{x'_i\}$).

Case 1: $x_1 < f$. Then there exists a "j" such that

$$x'_j \geq x'_1 + 2 \text{ (since } X' \notin \Gamma_m \text{)}. \text{ Let } \hat{X} = \{\hat{x}_i\}$$

be defined as follows:

$$\hat{x}_1 = x'_1 + 1$$

$$\hat{x}_j = x'_j - 1$$

$$\hat{x}_k = x'_k, k \neq 1, k \neq j. (\hat{X} \text{ feasible})$$

By optimality of X' ,

$$\sum_1^p (x'_i)^2 \leq \sum_1^p (\hat{x}_i)^2 = \sum_{k \neq 1} (x'_k)^2 + (x'_1 + 1)^2 + (x'_j - 1)^2$$

or

$$(x'_1)^2 + (x'_j)^2 \leq (x'_1)^2 + (x'_j)^2 + 2(x'_1 - x'_j) + 2$$

so that $x'_j - x'_1 \leq 1$, a contradiction. Hence $X' \in \Gamma_m$ or

$X' \notin \Gamma_m$ but $x'_1 = a_1$. In either case, X' is L-optimal.

Case 2: $x_1 = f$. Here there must exist x'_j such that

$$x'_1 \geq x'_j + 2. \text{ Interchanging the role of } x_1 \text{ and}$$

x_j in case 1, the argument is identical, noting

that the X in this case is feasible since
 $a_1 \leq a_k, k=2, \dots, p$.

4.4 The latin square

The notation of section 4.2 will be used with the model

$$E(Y) = Z\alpha \text{ being written: } E(y_{ijk}) = \mu + \tau_i + \delta_j + \tau_k \quad (i, j, k=1, \dots, d).$$

The number of observations of the random variable y_{ijk} is again given by x_{ijk} . The problem is stated:

$$\begin{aligned} \text{Minimize} \quad & \sum_{i,j} s_{ij}^2 \\ \text{Subject to} \quad & \sum_{i,j,k} x_{ijk} = d^2 \\ & x_{ijk} \geq 0, \text{ integer } (i, j, k=1, \dots, d) \end{aligned}$$

In section 4.2, we have the sufficient conditions (4.1.2) for $\{x_{ijk}^0\} = X^0$ to solve this problem (here $p=d^3, n_1=n_2=n_3=d, N=d^3$):

$$\left. \begin{aligned} x_{ij.}^0 &= 1 \\ x_{i.k}^0 &= 1 \\ x_{.jk}^0 &= 1 \end{aligned} \right\} i, j, k=1, 2, \dots, d$$

(the three implying)

$$x_{ijk}^0 = 1 \text{ or } 0$$

These conditions precisely define a latin square arrangement, and since Ehrenfeld (19) has shown that $\lambda_r(X)$ is maximized for this factor arrangement, X^0 , this solution is L-optimal.

4.5 The balanced incomplete block

The model for this problem is:

$$E(y_{ij}) = \mu + \tau_i + \delta_j \quad (i=1, \dots, t; j=1, \dots, b)$$

The element x_{ij} is the number of observations to be taken with factor one (usually called treatments) at level i , and factor two (blocks) at level j . The problem is written:

$$\text{Minimize} \quad \sum_{i,j} s_{ij}^2$$

$$\text{Subject to} \quad x_{..} = N < bt$$

$$x_{ij} \geq 0, \text{ integer} \quad (i=1, \dots, t; j=1, \dots, b)$$

It is supposed that a balanced incomplete block exists so that N/b and N/t are integers. It is now desired to examine the sufficient conditions for X^0 to be a solution. We utilize the results of section 4.2, observation three, specialized to this two-factor case. Since $N/bt < 1$ (here $n_1=b, n_2=t$ in notation of section 4.2), sufficient conditions for optimality of $\{x_{ij}^0\} = X^0$ are obtained from (4.1.3) and (4.1.5):

$$x_{i.}^{\circ} = N/b \text{ (assumed integer), } i = 1, \dots, t$$

$$x_{.j}^{\circ} = N/t \text{ (assumed integer), } j = 1, \dots, b$$

$$x_{ij}^{\circ} = 0 \text{ or } 1$$

These conditions define a balanced incomplete block and hence X° is L-optimal since Ehrenfeld (19) has shown that $\lambda_r(X)$ is maximized with such an arrangement.

4.6 Webb's problem

Webb (51) reported a factor arrangement for a three-factor problem with 4, 4, and 3 levels. This problem was referred to in section 2.7 (Figure 1), and it was noted there that the gradient search technique failed to obtain the design proposed by Webb. The number of observations was limited to twelve (here $\rho = n_1 n_2 n_3 = 48$) or one-fourth the number of factor-level combinations. We shall see that the arrangement proposed by Webb (the encircled elements of Figure 1) satisfies the sufficient conditions for optimality relative to the sum of squares criterion.

The model is a three-factor additive one so that the notation of observation three, section 4.2, may be used where now $n_1 = 4$, $n_2 = 4$, $n_3 = 3$, and $N = 12$. The minimum sum of squares problem is:

$$\text{Minimize} \quad \sum_{i,j} s_{ij}^2$$

$$\text{Subject to} \quad X \dots = 12$$

$$x_{ijk} \geq 0, \text{integer}; \quad i, j = 1, \dots, 4; \quad j = 1, 2, 3$$

Let X° be an optimal solution to this problem. Since $N/n_1 n_3 = 1$ and $N/n_2 n_3 = 1$, (4.1.5) yields the conditions:

$$\left. \begin{array}{l} x_{i \cdot k}^\circ = 1 \\ x_{\cdot j k}^\circ = 1 \end{array} \right\} \begin{array}{l} i, j = 1, \dots, 4 \\ k = 1, 2, 3 \end{array}$$

Since $N/n_1 n_2 < 1$, 4.1.6 yields:

$$x_{ij \cdot}^\circ = 0, 1$$

Property (4.1.3) finally yields:

$$x_{i \cdot \cdot}^\circ = 3, \quad i = 1, \dots, 4$$

$$x_{\cdot j \cdot}^\circ = 3, \quad j = 1, \dots, 4$$

$$x_{\cdot \cdot k}^\circ = 4, \quad k = 1, 2, 3$$

We now observe that Webb's arrangement clearly satisfies all these criteria as shown by the matrix S :

S

=

12	3	3	3	3	3	3	3	3	3	4	4	4
3	3	0	0	0	1	0	1	1	1	1	1	1
3	0	3	0	0	0	1	1	1	1	1	1	1
3	0	0	3	0	1	1	1	0	1	1	1	1
3	0	0	0	3	1	1	0	1	1	1	1	1
3	1	0	1	1	3	0	0	0	1	1	1	1
3	0	1	1	1	0	3	0	0	1	1	1	1
3	1	1	1	0	0	0	3	0	1	1	1	1
3	1	1	0	1	0	0	0	3	1	1	1	1
4	1	1	1	1	1	1	1	1	4	0	0	0
4	1	1	1	1	1	1	1	1	0	4	0	0
4	1	1	1	1	1	1	1	1	0	0	4	4

CHAPTER V

SUMMARY AND AREAS OF FUTURE INVESTIGATION

5.1 Summary

In structuring the problem of experimental factor arrangement as a formal mathematical optimization problem, two objective functions were used. The first, a minimum eigenvalue criterion, was a conservative approach described possibly as a MAX-MIN method in that the minimum informational content (over normalized linear combinations of parameters to be estimated) was maximized. This criterion led to a multidimensional lattice search problem, useful only when the method of "rounding off" noninteger solutions could be justified allowing conventional continuous parameter search methods to be used. It was empirically determined that such approaches might be valid for problems which allowed the total number of observations to be at least as great as the number of factor combinations.

An alternate criterion was introduced which was related to the minimum eigenvalue criterion. The second objective function,

the sum of squares of the eigenvalues of the information matrix, was seen to yield identical solutions to the minimum eigenvalue problem in three important cases: (a) the one-factor problem, (b) the latin square, and (c) the incomplete block arrangement. Furthermore, this second criterion yielded a quadratic form in the decision variable so that existing algorithms for the integer quadratic program could be utilized. Computational feasibility of both criteria is marginal (Appendix C), with problems on the order of five factors and four to eight levels per factor probably requiring minutes to hours of computation time on the 7090. Investigative effort in programming efficiency could possibly cut these times by a factor of ten.

Originally, work on this dissertation began with a cost-minimization problem. That is, with the environmental constraints and a lower bound for information, the objective was to seek a minimum cost factor arrangement. With information defined as the minimum positive eigenvalue, λ_r , the problem became an exceedingly difficult one since the constraint region was no longer defined by linear relations. It should be noted that with the sum of squares definition of information, this leads to a quadratic form in the constraints, the problem that Witzgall's algorithm (55) does handle. It was eventually decided, however, that probably most sizable experimental factor arrangement problems have a fixed upper bound on

total cost, and it is desired to find the arrangement of maximum information content; hence the formulation of Chapter I.

5.2 Areas of future investigation

The areas most suggestive for future effort are most easily categorized as (a) theoretical and (b) operational. Both are now discussed separately.

Theoretical areas of investigation.--One extension of this dissertation would be to include random effects models and mixed models. Another would be to alter the conservative objective function by considering the minimization of a particular variance with other variances being constrained. Effects of "decomposing" the larger problems into smaller ones and heuristically "tying" them together might prove fruitful.

There remains the question as to what sufficient conditions on the constraint region must be in order that the quadratic formulation yield the same solutions as the minimum eigenvalue problem. Finally, the question of improving existing partial enumeration techniques for this specific problem has possibility.

Operational questions.--The first operational question to be answered would seem to be whether "typical" problems which involve

constraints of the type mentioned can really be solved. A particularly promising area could be that of marketing. "The high cost of retail experimentation discourages the consideration of more than three or four alternatives" (25). Because the determination of a feasible solution involves existing computer programs, the question of how "close" such methods do get to optimal solutions would be very useful information.

Finally, due consideration should be given to the vast existing knowledge in experimental layout. If the constraints readily admit a factor arrangement of known optimal properties (presuming such information could be stored in a computer, such as characteristics for an incomplete block), then a computer program should test for such a condition. Such investigations could obviously become quite exhaustive.

APPENDIXES

APPENDIX A

EIGENVALUE PROPERTIES

Below we prove properties b, d, and f of section 2.2, Chapter II.

Property b: $\lambda_r(X) \leq \text{Minimum}_i \{s_{ii}\}$

- (i) Let $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r$ be the nonzero characteristic roots of the real, symmetric ($m \times m$) matrix S of rank r . Let R ($m \times m$) be such that

$$R^t S R = [\text{diag} \{ \lambda_i \}], \quad R^t R = I$$

- (ii) Then $1/\lambda_r = \text{Max}_X \left\{ \frac{1}{X^t S X} ; SX \neq 0, X^t X = 1 \right\} = B$, since, for $X = RY$,

$$\begin{aligned} B &= \text{Max}_Y \left[\frac{1}{Y^t R^t S R Y} ; X^t X = 1 \Rightarrow Y^t Y = 1 \right] \\ &= \text{Max}_Y \left[1 / \sum_i \lambda_i y_i^2 ; \sum_i y_i^2 = 1 \right] = 1/\lambda_r. \end{aligned}$$

- (iii) Therefore, letting X_i be a vector of all zeros except unity in the i^{th} position, it must be true that

$$\lambda_r \leq X_i^t S X_i = s_{ii} \text{ for all } i, \text{ or } \lambda_r \leq \text{minimum}_i \{s_{ii}\}.$$

Property d: $\sum \lambda_i = v \sum x_i$, where v is the (constant) number of nonzero (one) entries in any ξ_i .

Recall, $S = \sum x_i \xi_i \xi_i^t$ where elements of ξ_i are 0 or 1 (and hence so are $\xi_i \xi_i^t$) and each first component of ξ_i is 1.

We have

$$\begin{aligned} \sum \lambda_i &= \text{trace } S = \sum x_i [\text{trace } \xi_i \xi_i^t] \\ &= \sum x_i [\text{trace } \xi_i^t \xi_i] = v \sum x_i \end{aligned}$$

Property f: For the additive model with g factors, each factor having n_i levels, $\lambda_r(X) \leq N/\bar{n}$, where $N = \sum x_i$, and \bar{n} is the maximum of the n_i .

The additive model introduces a constraint on the ξ_i to the effect that the components of ξ_i relating to factor k (n_k in all, one component for each of the n_k levels) may have only one non-zero entry. The singular S model assumes all the ξ_i have "1" in their first position. Then S_{ii} represents the total number of times parameter "i" (pertaining to a level of one of the factors) is observed. With $\sum x_i = N$, we have:

$$\begin{aligned} S_{11} &= N \\ \sum_{i=2}^{n_1+1} S_{ii} &= N \\ \sum_{i=n_1+2}^{2+n_1+n_2} S_{ii} &= N \\ &\dots \\ &\text{etc.} \end{aligned}$$

Therefore, with the definitions $\bar{n} = \text{maximum}_j \{n_j\}$, $\underline{s} = \text{minimum}_i \{s_{ii}\}$, the largest value that \underline{s} can assume is N/\bar{n} . From property (b) above, $\lambda_r(X) \leq \underline{s}$, so that the quantity N/\bar{n} represents the maximum value attainable for $\lambda_r(X)$ with fixed N .

APPENDIX B**OBTAINING A FEASIBLE SOLUTION****B.1 General approach**

If the gradient procedure of Chapter II is used, an initial feasible solution is needed. Further, every attempt to minimize costly computation of eigenvalues of that method should be investigated by looking for less costly methods of obtaining "good" initial feasible solutions. It is also possible that such a solution would provide a better algorithm for the integer quadratic programming problem of Chapter III. Additional reasons for such investigations are:

- (a) This approach could yield an optimal solution in itself.
- (b) Smaller problems might be solved heuristically.
- (c) Greater insight into the entire problem might be gained.

The general approach will be to ignore certain constraints so as to take advantage of the resulting problem structure, and if the solution satisfies the previously ignored constraints, a feasible solution has been found. By judiciously choosing the constraints to ignore, this strategy could prove successful.

The problem of obtaining a feasible solution is to find an X_0 so that the following four constraints are satisfied:

- (a) $AX^0 \leq b, a_{ij} = 0, 1 \ (i=1, \dots, h; j=1, \dots, p), X^0 \geq 0$
- (b) $C^t X^0 \leq c_0, c_j \geq 0 \ (j=1, \dots, p), X^0 \geq 0$
- (c) x_i^0 integer, $(i=1, \dots, p)$
- (d) $\lambda_r(X^0) > 0$

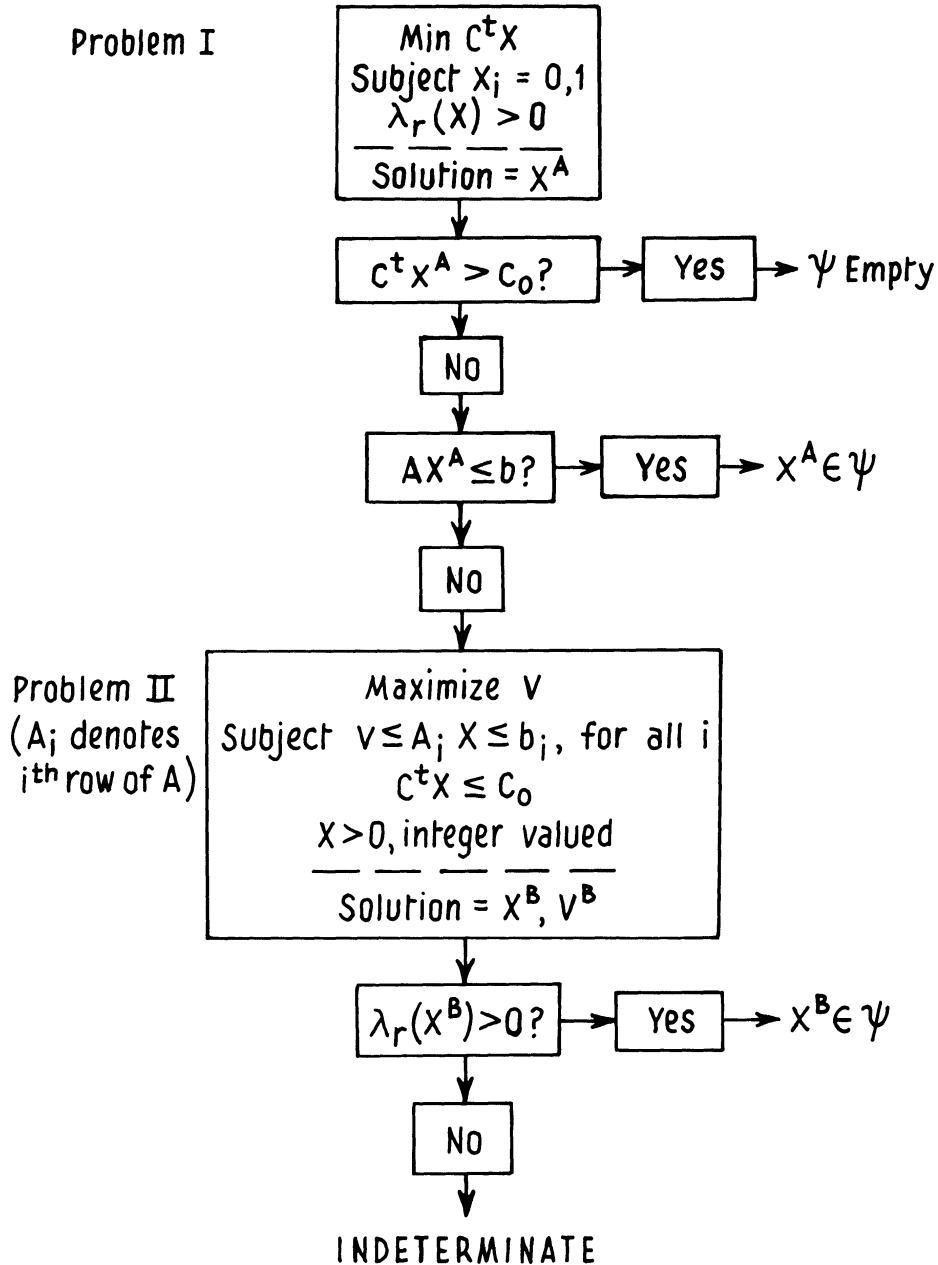
As in Chapter II, requirement (d) is added. In the general problem of the maximization of $\lambda_r(X)$, the requirement that this function be positive is superfluous and this restriction further introduces the undesirable property that the feasibility region is an open, rather than closed, set. For purposes of this section, however, the requirement that λ_r be positive has the advantage that should an initial solution actually be used, $\lambda_r > 0$ insures the maximum number of estimable linear combinations of the unknown parameters. Recall that the matrix A has elements zero or one and each row of AX^0 corresponds to the total number of observations available for a particular factor level. Requirement (d) is satisfied only if the space generated by the vectors $\{x_i^0; x_i^0 > 0\}$ is of dimension r , i.e., there must be r linearly independent vectors corresponding to the nonzero x_i^0 .

The problem of obtaining a feasible solution is approached by solving Problem I (Figure 4), or, if the solution to that problem

Feasible Solution Problem: Find X so that

- (1) $AX \leq b, X \geq 0$
- (2) $c^t X \leq C_0, X \geq 0$
- (3) X integer valued
- (4) $\lambda_r(X) > 0$

Let $\Psi =$ Set of feasible solutions



produces indefinite results, Problem II is then solved (Figure 4).

Problem I essentially seeks a minimum cost factor arrangement subject to $\lambda_r > 0$. If, for the solution, the (minimum) cost does not exceed the allowable upper bound, c_0 , but the resource constraints $AX \leq b$ are violated, Problem II is solved. This latter problem essentially ignores the constraint $\lambda_r > 0$, and seeks to maintain the cost and resource constraints. The function λ_r is not completely ignored in Problem II, however, since there is a set $\{K\}$ such that

$$s_{ii} = \{A_u X; u \in K\} \quad \text{for every diagonal element } \Delta_{ii}.$$

Here, A_u denotes the u^{th} row of A . Since v is maximized in Problem II, and $\lambda_r(X) \leq \underset{i}{\text{minimum}} \{s_{ii}\}$ (property b, section 2.2), Problem II maximizes this upper bound of $\lambda_r(X)$. This procedure does not guarantee that $\lambda_r(X^B) > 0$ for a solution, X^B to Problem II, and the result of solving the two problems may still be an indeterminate situation. The problems are now examined in more detail.

B.2 Solution of Problem I, linear model

Ignoring constraint (a), the following procedure is defined to obtain a "least cost" set of r linearly independent ξ_i which in turn define a set of least cost observations satisfying (d): Proceed by stages, considering vectors ξ_i , in order of nondecreasing cost,

C_i . At each stage choose the vector if it is linearly independent of the previously chosen ones; otherwise discard it from future consideration. This process is ended when r vectors are chosen. Corresponding to a chosen ξ_i , set x_i equal to one. Otherwise set x_i equal to zero. Let the resultant vector be denoted by X^A .

As a result of this process, three cases are possible. If $C^t X^A > c_0$, the problem has no solution since $C^t X^A$ is the minimum value of $C^t X$ subject to $\lambda_r(X) > 0$. The proof, due to E. Lawler, is found in section B.7. If $C^t X^A \leq c_0$, and for $X^A = (x_1^A, \dots, x_p^A)$ it is true that $\sum_j a_{ij} x_j^A \leq b_i, i = 1, \dots, p$, then X^A is an initial feasible solution. Otherwise one or more resource constraints are violated. This last condition is now considered.

B.3 Problem II, linear model

The problem is altered slightly to accommodate the last case above. Find $x_i \geq 0, y_j \geq 0, w \geq 0$ so that:

$$\sum_j a_{ij} x_j + \gamma_i = b_i \quad (i = 1, \dots, p)$$

$$\sum_j c_j x_j + w = c_0$$

$$\lambda_r(X) > 0$$

The third case in the previous example corresponds to a set $\{x_i^A \geq 0, w \geq 0\}$, but for which one or more γ_i are negative.

Ignoring the constraint $\lambda_r(x) > 0$, one could seek to "drive" those negative elements of γ_i toward positive values while maintaining the already feasible elements by solving the problem (equivalent to Problem II):

$$\begin{aligned}
 &\text{Maximize} && v \\
 &\text{Subject to} && v \leq \gamma_i \quad (i=1, \dots, h) \\
 & && \sum_j a_{ij} x_j + \gamma_i = b_i \quad (i=1, \dots, h) \\
 & && \sum_j c_j x_j + w = c_0 \\
 & && x_j, \gamma_i, w \geq 0 \quad (i=1, \dots, h; j=1, \dots, p)
 \end{aligned}$$

If, for a solution X^B , it is true that $\lambda_r(X^B) > 0$, X^B is a feasible solution. Otherwise, we have an indeterminate situation.

We now consider the special case of an additive model, with the general approach depicted by Figure 5 and discussed in the next section.

B.4 The additive linear model; analogy to the transportation problem

At this point, for reasons presently obvious, we briefly consider the characteristics of the solution to the generalized transportation problem of linear programming. The primal and dual problems (28) consist in finding a solution (X^0, U^0) so that the following conditions hold:

Primal	Dual
$C^t X^0$ is a minimum	$b^t U^0$ is a maximum
$A X^0 = b$	$A^t U^0 \leq c_0$
$X_i^0 \geq 0$	

The matrix A has a special structure, and the equations $A X = b$ are illustrated for the case when elements of X have three subscripts:

$$\sum_{i,j,k} x_{ijk} = b^{(0)}$$

$$\sum_{j,k} x_{ijk} = b_i^{(1)} \quad (i=1, \dots, n_1)$$

$$\sum_{i,k} x_{ijk} = b_j^{(2)} \quad (j=1, \dots, n_2)$$

$$\sum_{i,j} x_{ijk} = b_k^{(3)} \quad (k=1, \dots, n_3)$$

In the transportation analogy, one would have three shipping "stages" labeled I, J, K , each with n_1, n_2, n_3 substages respectively. Then x_{ijk} refers to the amount of goods shipped from the i^{th} substage of I , through the j^{th} substage of J , to the k^{th} substage of K . An optimal solution to the noninteger problem (the solution is automatically integer-valued for two shipping stages, but for no greater such number) which minimizes $C^t X$ consists of a "basic" set $\{x_{ijk}^0\}$ of nonzero elements (the remaining being zero) and further, the columns of A corresponding to the basic elements are linearly independent. The vectors X^0 and U^0 are orthogonal, i.e., $U^0 \cdot X^0 = 0$,

and for $X_{ijk}^0 = 0$, the corresponding value of u_{ijk}^0 inserted in $C_{ijk}^1 = C_{ijk} - u_{ijk}^0$ gives the marginal cost C_{ijk}^1 in introducing a unit increase in the variable X_{ijk}^0 (zero, if X_{ijk}^0 is in the basic set). The additive experimental design model $E(Y) = Z\alpha$ is considered, and assume for simplicity that a three-factor model exists so that X_{ijk} must be determined for $i = 1, \dots, n_1, j = 1, \dots, n_2$, and $k = 1, \dots, n_3$. The vectors ξ_i^t (or rows of Z) are of the form $e_i^{(0)} + e_i^{(1)} + e_j^{(n_1+1)} + e_k^{(n_1+n_2+1)}$, where the vector $e_a^{(b)}$ is defined with $n_1 + n_2 + n_3$ components all zero except component $a+b$.

It may be seen that every ξ_i^t corresponds to exactly one column of A in the transportation problem structure so that specification of a basic solution X_0 , and a linearly independent set of columns of A then determines a linearly independent set of r (here $r = n_1 + n_2 + n_3 - 1$) vectors, $\{\xi_i^t\}$. The above analogy for the additive model enables a different, possibly faster and better approach to the problem of finding a "good" initial feasible solution.

B.5 An approach to the three-factor additive model

The problem now considered is again for three factors, the generalization to n factors being obvious. With the three factors at n_1, n_2 , and n_3 levels (or with this number of parameters associated with each, plus one constant parameter) let X_{ijk} again

represent the number of observations taken with factor one at level i , factor two at level j , and factor three at level K . Now define:

$$\Psi_1 = \{X; \lambda_r(X) > 0\}$$

$$\Psi_2 = \{X; AX \leq b, X \geq 0\}$$

$$\Psi_3(d) = \{X; AX = d, \text{ for } 0 < d \leq b, X \geq 0\}$$

$$\Psi_4 = \{X; C^T X \leq c_0\}$$

NOTE: $\Psi_3(d) \subset \Psi_2$

The inequality $AX \leq b$ is defined in further detail:

$$x_{\dots} \leq b^{(0)}$$

$$x_{i\dots} \leq b_i^{(1)} \quad (i = 1, \dots, n_1)$$

$$x_{\dots j} \leq b_j^{(2)} \quad (j = 1, \dots, n_2)$$

$$x_{\dots k} \leq b_k^{(3)} \quad (k = 1, \dots, n_3)$$

The subscript (\cdot) again indicates a summation over that index, i.e.,

$$x_{\dots j} = \sum_{i, k} x_{ijk}, \quad x_{ij\dots} = \sum_k x_{ijk}, \text{ etc.}$$

Two approaches are now considered, one preceding the other. They attempt to find a "good" initial feasible solution, taking advantage of existing linear programming techniques for their solution.

Recall that $\lambda_r(X) \leq \text{Minimum}_i \{S_{ii}\} = v$. Here, $v = \text{minimum}_{i, j, k} \{x_{i\dots}, x_{\dots j}, x_{\dots k}\}$. An attempt to obtain an initially good solution

is to make v as large as possible by solving the following integer

linear program (Problem II, Figure 5):

$$\begin{array}{ll}
 \text{Find } X, \text{ so as to Maximize} & v \\
 \text{Subject to} & v \leq x_{i..} \quad (i = 1, \dots, n_1) \\
 & v \leq x_{.j.} \quad (j = 1, \dots, n_2) \\
 & v \leq x_{..k} \quad (k = 1, \dots, n_3) \\
 & X \in \Psi_2 \cap \Psi_4 \\
 & X \text{ integer valued}
 \end{array}$$

If, for a solution (v_0, X_0) it is true that $v_0 = 0$, the original problem has no solution since it follows that $\lambda_r(X_0) = 0$. If $v_0 > 0$, and $\lambda_r(X_0) > 0$, then X_0 is the initial solution sought. Otherwise, the next problem is suggested.

Let the elements of the previous X_0 be denoted by x_{ijk}^0 , and define a vector d with elements:

$$\begin{array}{l}
 d^{(0)} = x_{...}^0 \\
 d_i^{(1)} = x_{i..}^0 \quad (i = 1, \dots, n_1) \\
 d_j^{(2)} = x_{.j.}^0 \quad (j = 1, \dots, n_2) \\
 d_k^{(3)} = x_{..k}^0 \quad (k = 1, \dots, n_3)
 \end{array}$$

Now consider the problem:

$$\begin{array}{ll}
 \text{Minimize} & c^t X \\
 \text{Subject to} & X \in \Psi_3(d) \\
 \text{With} & X \text{ integer-valued}
 \end{array}$$

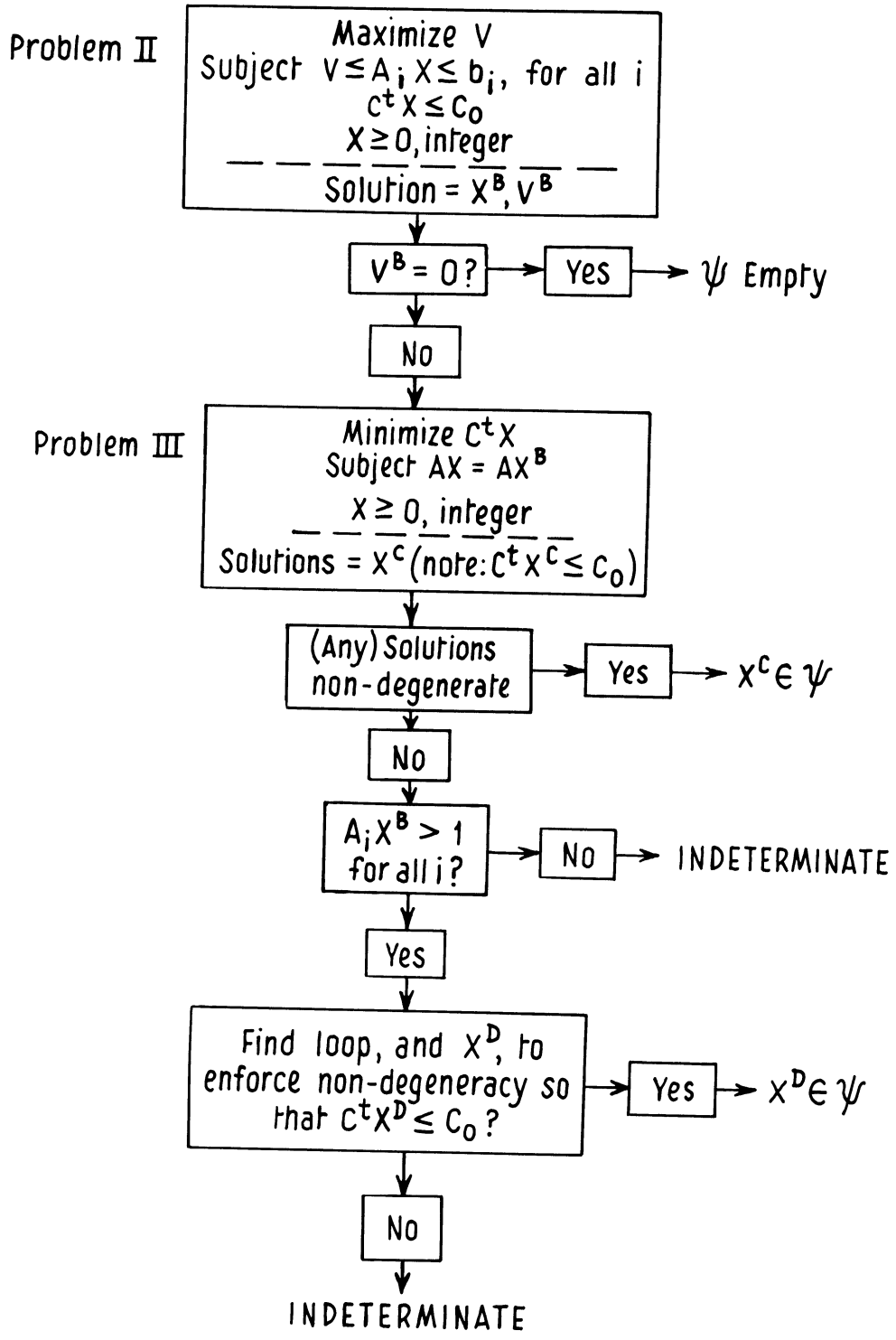


Fig. 5.--Additive model.

It is observed that since $X_0 \in \Psi_2$ and $v_0 > 0$, then $0 < d \leq b$, a requirement that $\Psi_3(d)$ be nonempty. This problem is precisely the three-dimensional transportation problem for which a procedure for approximating an integer solution exists (Dwyer, 18). If, for the solution (with the integer approximating technique) $C^t X_0 > c_0$, this approach has been fruitless. If $C^t X_0 \leq c_0$, and $\lambda_r(X_0) > 0$, an initial solution has been found. Finally, if $C^t X_0 \leq c_0$, but $\lambda_r(X_0) = 0$, some of the elements of the basic solution in X_0 are zero, or one has a degenerate solution (otherwise, the columns of A corresponding to nonzero x_j would span a space of dimension r).

Let these zero elements of the basic solution be the set J . One possible next step would be to increase each element of J to one, and if the resultant $\bar{x} \in \Psi_2 \cap \Psi_4$, it would be the feasible solution (clearly $\bar{x} \in \Psi_1$). Of course, it may be possible that even though a degenerate solution is obtained, one finds multiple solutions one of which is nondegenerate. Necessary conditions for such a situation could be easily developed.

B.6 The two-factor additive model: illustration

It is now assumed a two-factor problem ended in the first phase with $v_0 \geq 2$, but with the corresponding $X_0, \lambda_r(X_0) = 0$.

The second problem above would then be stated:

$$\begin{array}{ll}
\text{Minimize} & C^t X \\
\text{Subject to} & x_{..} = d^{(0)} \\
& x_{i.} = d_i^{(1)} \quad (i = 1, \dots, n_1) \\
& x_{.j} = d_j^{(2)} \quad (j = 1, \dots, n_2) \\
& x_{ij} \geq 0
\end{array}$$

It is known that solutions to this problem, the two-dimensional transportation problem, are integer-valued. The solution, X_0 , is assumed to satisfy $C^t X_0 = c_0 - w_0$, for $w_0 > 0$, but that $\lambda_r(X_0) = 0$. The latter means that X_0 is a degenerate solution, but the cost constraint to the original problem is inactive. We seek a solution X_1 , such that $c_0 \geq C^t X_1 \geq C^t X_0$ for which the "row" and "column" sums, $X_{i.}$ and $X_{.j}$, remain unchanged. The matrix $C' = \{c'_{ij}\}$ is defined as the dual solution of X_0 (see for instance, 49, 31). It represents the marginal cost of introducing an element into the solution which is not in the present optimal solution. The vector $X_0 = \{x_{ij}^0\}$ has $n_1 + n_2 - 1$ members of the basic solution, the indices of which may be identified by elements of the "tree": $K = [(i_1, j_1), (i_2, j_1), (i_2, j_2), (i_3, j_2), \dots, (i_r, j_r)]$ where one or more $\{x_{ij} ; (i, j) \in K\}$ are zero. Notice that consecutive elements of K cannot correspond to zero x_{ij} for otherwise an $x_{i.}$ or $x_{.j}$ would be zero, contrary to the assumption that v_0 was at least two. The elements (i_1, j_1) ,

$(i_2, j_2), \dots$, etc., are sometimes called "odd" with the remaining elements of K called "even" (28). Next consider all elements (i, j) not in the basic solution and from these set $G = \{x_{ij} ; c'_{ij} \leq w_0\}$. Since c'_{ij} represents the cost of adding x_{ij} , we could not introduce this element if the resultant total cost were greater than C_0 . For every $x_{ij} \in G$, it is known (49) there exists a "loop" or sequence of positions $L_{ij} = (i, j), (i, j_1), (k, j_1), \dots$, with initial and final elements (i, j) and intermediate elements those of the basic solution where each such entry (a, b) is preceded by (a, k) or (k, b) for some k . For any $x_{ij} \in G$, a sufficient condition for a solution to our original problem (that of feasibility) is that for some L_{ij} , all degenerate basic elements are in the "odd" group. The resulting X_1 is to increase all odd elements of L_{ij} by one, with the even elements decreased by one. Then it follows that the columns A^i (and hence the ξ_i) corresponding to the nonzero x_i have r linearly independent vectors among them. If $C^t X_1 \leq C_0$ with $A X_1 = d \leq b$, we then have an initial feasible solution.

An extremely simple example of this two-factor problem is now given. Let the costs, C_{ij} , be represented in the following array with the original resource constraints indicated at the side and bottom of this array:

Problem (1)Maximize v Subject to $v \leq x_i, (i=1, \dots, n_1)$ $v \leq x_j, (j=1, \dots, n_2)$ $\Delta X \leq b$ $C^t X \leq 2I$ $x \geq 0, \text{ integer}$

1	11	6	≤ 6
3	2	14	≤ 6
12	4	5	≤ 6

 $\leq 2 \quad \leq 2 \quad \leq 2$ where $b^t = (6, 6, 6, 2, 2, 2)$ $C^t = (1, 11, 6, 3, 2, 14, 12, 4, 5).$ A solution to this problem (V_0, W_0) is indicated:

2	0	0
0	2	0
0	0	2

i.e., $V_0 = 2, C^t X_0 = 16, x_{11}^0 = x_{22}^0 = x_{33}^0 = 2$, with $x_{ij}^0 = 0, i \neq j$.Since the value of $\lambda_r(X_0)$ is zero in this case, the problem is:

Problem (2)

$$\begin{aligned} &\text{Minimize } C^t X \\ &\text{Subject to } AX = d \\ &\quad X \geq 0 \end{aligned}$$

where $d = (6, 2, 2, 2, 2, 2, 2)$. The transportation algorithm then yields the following solution, the exponents corresponding to the basic solution, and the main elements representing the dual variables or marginal costs.

0^2	11	5
0^0	0^2	11
7	0^0	0^2

Cost = 16.

Since there are two elements which are zero in this basic solution, $(2, 1)$ and $(3, 2)$, we seek an entry which has $c'_{ij} \leq 21 - 16 = 5 = w_0$ and for which a loop can be constructed as described above. It is seen that $(1, 3)$ conforms to these requirements with $L_{1,3} = [(1, 3), (3, 3), (3, 2), (2, 2), (2, 1), (1, 1), (1, 3)]$. Making the adjustment indicated above, the initial feasible solution now becomes:

1	0	1
1	0	0
0	1	1

Cost = 16 + 5 = 21.

It is interesting to note that, with the proper interpretation of the factors, this is a simple case of an incomplete block arrangement, one of optimal $\lambda_r(\lambda)$, provided the constraint $\sum x_{ij} = G$ is added (15) (also a solution to the problem of Chapter III).

Of course, full advantage may be taken of this two-factor, transportation problem analogy by utilizing existing algorithms for the solution to the latter. The additional constraint that exists is that the solution be nondegenerate. Such a constraint might be ignored initially to obtain what might be an excellent first step toward a solution to the factor arrangement problem. This would allow more complex cost functions and resource constraints. For instance, if each cell, or factor combination has an upper bound and if a convex piecewise linear cost function, rather than strictly linear, exist, the capacitated transportation algorithm of Graves and Thrall (28) could be used. It might be quite realistic that increased marginal costs occur for certain factor combinations, and such an effect should be included.

B.7 Proof of the minimum cost algorithm of section B.2

The proof of the algorithm attaining a minimum cost solution in section B.2 is embraced by the following Lemma, due to E. Lawler.¹

¹Private communication.

Lemma: Given a set of vectors $\{\alpha_i \in E^{(N)}; i=1, \dots, p > N\}$, with associated costs $c_i > 0$, the minimum cost algorithm of section B.2 results in a minimum $Z = \sum c_i$, the sum taken over all vectors chosen for which the chosen set is a basis for $E^{(N)}$ (the original set spans $E^{(N)}$).

Proof:

(i) Let $U_1 = \{\alpha_1, \alpha_2, \dots, \alpha_N\}$ be a solution using the algorithm, with associated cost C_1 .

Let $U_2 = \{\beta_1, \beta_2, \dots, \beta_N\}$ be a set of the original vectors which is also a basis for $E^{(N)}$, and has optimal cost C_2 (any optimal set has precisely N vectors).

Suppose $C_2 < C_1$.

(ii) Let $\hat{\alpha}$ be the minimum cost vector, not necessarily unique, in U_1 but not in U_2 . One exists since $U_1 \neq U_2$. $\hat{\alpha}$ may be represented as a unique linear combination of vectors in U_2 . This nontrivial combination must include at least one vector not in U_1 since that set is a basis for $E^{(N)}$. Let V_1 be the set of vectors in this unique linear combination which are in U_1 , and V_2 be the set of vectors also in this linear combination in U_2 , but not U_1 .

- (iii) Let β_k be the vector in V_2 of maximal cost, C_m . If C_0 is the cost of $\hat{\alpha}$, C_0 cannot be less than C_m , for otherwise $\hat{\alpha}$ could replace β_k in U_2 resulting in a basis at lower cost than optimal. But C_m cannot be less than C_0 since by construction, $\hat{\alpha}$ was the lowest cost vector, independent of V_1 . Therefore $C_m = C_0$.
- (iv) Replace β_k by $\hat{\alpha}$, denoting the set by W_1 . Repeating this process until no vector exists in U_1 which is not in W_k , we have $U_1 = W_k$, each of which have costs C_1, C_2 , unequal by assumption. By this contradiction it follows that all sets of optimal cost must have cost equal to C_1 , and hence the algorithm results in one such solution.

APPENDIX C

COMPUTATIONAL CONSIDERATIONS

C.1 Problem sizes

In order to assess present-day computational capability in solving the two problems of Chapters II and III, and the set proposed in Appendix B, the problem size relations are first defined. Let a linear model have "r" factors at n_1, n_2, \dots, n_r levels. Define $p_1 = 1 + \sum^r n_i$, and $p_2 = p_1 + \frac{r(r-1)}{2}$. Then $(p_1 \times p_1)$ represents the dimension of the information matrix in the additive model, and $(p_2 \times p_2)$ that size with the two-factor interaction terms included. No higher order interaction terms will be considered for purposes of computation calculation. Thus, in a 2^k factorial arrangement, with $k=6$ (six factors each at two levels), the information matrix (and also the Q matrix of the quadratic program) is (13×13) , or written as $(13, 13)$, for the additive case and $(28, 28)$ if the two-factor interactions are included.

It will be assumed that the number of inequality constraints is equal to $p_1 + 1$ (one per factor level, one for the total observations, and one for the cost constraint). Furthermore, there are p_1 (p_2 with the two-factor interactions) variables and all but the cost constraint coefficients will be $(0, 1)$. The percent of nonzeros in

the coefficient constraint matrix is approximately $P = \frac{\sum n_i^2}{N}$, where N is the total entries in the matrix, equal to $p_1(p_1+1)$ or $p_2(p_1+1)$. All the computation times reported refer to a 7090 processor unless otherwise stated.

C.2 The eigenvalue search problem

In the search technique of Chapter II, each step required (a) calculation of the minimum eigenvalue, and (b) calculation of a "step" direction. In the following, it will be assumed that the integer requirement may be ignored.

Calculation of the eigenvalues of a real symmetric matrix reportedly depends very much on the technique used. Estrin and Viswanathan (22) predict several computation times for various matrix dimensions. Included below in Table 5 is the maximum time of a method by Corbato (faster than Jacobi's method) labeled column 2, and the minimum times, predicted by a special-purpose computer configuration and algorithm, column 3. Here, n refers to the number of rows in the matrix.

The second phase in section 2.4 is the solution of a convex program. Actually, the objective function is linear, and all the constraints are linear with the exception of the simple form, $\sum y_i^2 \leq \kappa$. Being optimistic in that this program could be solved

TABLE 5

COMPARISON OF EIGENVALUE CALCULATIONS FOR
VARIOUS MATRIX DIMENSIONS

(1) n	(2) Corbato Method (sec.)	(3) Special (sec.)
10	1	0.3
25	18	2.8
50	145	12.9
100	1150	123.0

in approximately the same time as a quadratic program, Wolfe (56) reports that the latter involves about the same computing time as a linear program of equivalent dimension. On this basis, Smith (48) quotes such times for various dimensions and density of zero elements in the constraint matrix (for linear programs) as shown in Table 6.

Using the most optimistic times above, Table 7 shows calculations for one step in the gradient search technique for typical problem configurations noted. Here, T_1 is the eigenvalue calculation time estimate, and T_2 is the step calculation time estimate.

TABLE 6
EXPERIENCE IN LINEAR PROGRAM
CALCULATION TIMES

Dimension	Percent Zeroes in Constraint Matrix	Time (sec.)
5x10	-	1.8
27x20	50	24
118x225	95	258

TABLE 7
OPTIMISTIC COMPUTATION TIME PER ONE
STEP IN GRADIENT SEARCH

Factor- Levels	Description	Dimen- sion of S	Dimen- sion of Con- straint Matrix	Pct. 0's	T ₁ (sec.)	T ₂ (sec.)
2^3	additive	7x7	8x7	79	1	2
2^3	2-factor interaction	10x10	8x10	85	1	2
2^6	additive	13x13	14x13	87	1	2-24
2^6	2-factor interaction	28x28	14x28	94	4	12-24
$3^3 2^3$	additive	18x18	19x18	88	2	12-24
4x5x6	additive	16x16	17x16	72	2	12-24
4x5x6x7x8	2-factor interaction	40x40	17x40	70	13	24-48

It is seen that with the best possible present-day specialized programming, computing, and assumptions about applicability of other reported results, the time per step is several seconds to one minute for the problems listed. If a 7090 were used and the Corbato method programmed, a factor of eight to ten increase could easily result. Actually, the latter effect could possibly be offset by requiring only the minimum eigenvalue and eigenvector to be calculated. The reduction in computing time for this case is not clear. A problem involving sixty steps could take an hour of computing time in the best of conditions.

C.3 The quadratic integer program

The three algorithms are now investigated as to their computational feasibility. Only in the case of the partial enumeration technique is there such experience, so that estimates will have to be made in the case of the algorithms specifically involving the quadratic objective function.

The Kunzi-Oettli algorithm.--Problem II of section 3.2 is a mixed-integer program with possibly a noninteger valued objective function, which is pointed out by the authors (37) and Balinski (4), and convergence is not guaranteed with the mixed integer algorithm proposed by Gomory (27). The authors also point out that replacing

the objective function π by $\tilde{N}\pi$ for large \tilde{N} might suffice, where now $\pi\tilde{N}$ would be required to be integer so that an all-integer algorithm could be used. Assuming that this could be accomplished, computational experience on integer programs varies widely in performance, depending upon the code and problem structure. Balinski (4), in attempting to consolidate the widely varying reports, mentions that Martin, using a special algorithm (41), solves scheduling problems of the order $(70,160) \leq (m,n) \leq (1600,2200)$ regularly, approaching one hour each in computing time of the 7090 processor. In these problems, the constraint matrices had about 85 percent zeroes. The remaining nonzeros had about 85 percent ones. In Table 7, section C.2, the values there are not inconsistent with Martin's.

Witzgall's parabolic constraint algorithm.--Since the method of Witzgall (55) parallels that of Gomory (27) with the former method involving extra computation only in one row (corresponding to the single quadratic function) we might assume that this method would be approximately the same as Gomory's for similar size problems of a linear integer program character. Experience, however, is meager with Gomory's algorithm, although ten problems of the order $(10,50)$ have been reported solved in presumably "reasonable" lengths of time (27).

Partial enumeration technique.--More computational data are available for this algorithm, used in section 3.6. The authors (38) report times of 56 to 716 seconds for a problem with twenty-seven $(0,1)$ variables and a constraint matrix involving integers having values up to 9. "Covering" problems were solved for twenty-five $(0,1)$ variables in a maximum of 2.5 minutes, with a problem of thirty-five variables unsolved after 10 minutes. One interesting fact reported is that computation time may vary by 10 or 20 to 1, depending on the initial ordering of the variables. Also, the extreme simplicity of the algorithm allows rapid modification of the rules to suit the problem (Rule 3, section 3.6 was a trivial addition due to the equality constraint).

C.4 General remarks

Probably the only remark of substance that could be made concerning computational feasibility of the methods in previous chapters is that for moderate-sized problems ($4 \times 5 \times 6 \times 7 \times 8$, say), considerable computational effort and skill might result in operational methods of solution. Smaller problems (2^5) could undoubtedly be solved with existing algorithms in reasonable time (minutes, not hours) provided the variables did not exceed a value of one or two.

As in most specific application areas of integer programming, it would seem that most significant improvement would come in investigating and taking advantage of, the specific problem structure. For instance, if the cost constraints were not present, the entire constraint matrix would consist of 0's and 1's, a fact leading possibly to advantageous algorithm modification. Toward this end, it would also appear that a primary consideration would be simplicity of the initial algorithm which was being modified.

REFERENCES

1. Addleman, S. "Irregular Fractions of the 2^n Factorial Experiments," *Technometrics*, Vol. 3, No. 4, November 1961.
2. Anderson, T., An Introduction to Multivariate Statistical Analysis, J. Wiley and Sons, 1958.
3. Balinski, M., "Notes on Operations Research," 1962 Engineering Summer Conference, University of Michigan.
4. Balinski, M., "Integer Programming: Methods, Uses, Computation," *Management Science*, November 1965.
5. Barankin, W., and R. Dorfman, "On Quadratic Programming," University of California Publication in Statistics, Vol. 2, 1958.
6. Bellman, R., Introduction to Matrix Analysis, McGraw-Hill, 1960.
7. Benders, J. F., A. R. Catchpole, and C. Kuiken, "Discrete Variables Optimization Problems," Rand Symposium on Mathematical Programming, March 16-20, 1959.
8. Boot, J. C., Quadratic Programming, North-Holland Publishing Company, 1964.
9. Bose, R. C., I. M. Chakravarti, and D. E. Knuth, "On Methods of Constructing Sets of Mutually Orthogonal Latin Squares Using a Computer," *Technometrics*, Vol. 2, No. 4, November 1960.
10. Box, G. E. P., and D. W. Behnken, "Some New Three Level Designs for the Study of Quantitative Variables," *Technometrics*, Vol. 2, No. 4, November 1960.

11. Box, G. E. P., and J. S. Hunter, "The 2^{k-p} Fractional Factorial Designs," *Technometrics*, Vol. 3, No. 3, August 1961.
12. Chakrabarti, M. C., "On the C-Matrix in Design of Experiments," *Journal of the Indian Statistical Association*, 1959.
13. Charnes, A., and W. Cooper, Management Models and Industrial Applications of Linear Programming, Vol. 11, J. Wiley and Sons, 1961.
14. Connor, W., and S. Young, "Fractional Factorial Designs for Experiments with Factors and Two and Three Levels," National Bureau of Standards Applied Mathematics Series 58, U.S. Government Printing Office, Washington, D.C., September 1, 1961.
15. Cox, P., Planning of Experiments, J. Wiley and Sons, 1958.
16. Dantzig, G. B., "On the Significance of Solving Linear Programming Problems with Integer Variables," *Econometrika* (28), 1960.
17. Dantzig, G. B., "Quadratic Programming, a Variant of the Wolfe-Markowitz Algorithm," Research Report 2 of the Operations Research Center of the University of California, 1961.
18. Dwyer, P., "The Direct Solution of the Transportation Problem with Reduced Matrices," *Journal of Management Science*, Vol. 13, No. 1, September 1966.
19. Ehrenfeld, S., "On the Efficiency of Experimental Designs," *The Annals of Mathematical Statistics*, Vol. 26, No. 2, June 1955.
20. Ehrenfeld, S., "Complete Class Theorems in Experimental Design," Third Berkeley Symposium on Mathematical Statistics and Probability, Vol. 1, 1956.
21. Elston and Bush, "The Hypothesis That Can Be Tested When There Are Interactions in the Analysis of Variance Model," *Biometrics*, December 1964.

22. Estrin, G., and C. Viswanathan, "Organization of a 'Fixed-Plus Variable' Structure for Computation of Eigenvalues and Eigenvectors of Real Symmetric Matrices," Journal of the Association for Computing Machinery, January 1962.
23. Falk, J., "A Constrained LaGrangian Approach to Nonlinear Programming," Technical Report 06122-8-T, University of Michigan Office of Research Administration, March 1965.
24. Fisher, F. A., Statistical Methods for Research Workers, Hafner Publishing Company (13th Edition, 1963).
25. Gatty, R., "Statistical Models for Experiments in Merchandising," 1965 Proceedings, American Statistical Association, Business and Economic Statistics Section.
26. Goertzel, G., and N. Tralli, Some Mathematical Methods of Physics, McGraw-Hill, 1960.
27. Gomory, R., "An Algorithm for the Mixed Integer Problem," RM-2597 Rand Corporation, July 7, 1960.
28. Graves, G. W., and R. M. Thrall, "Capacitated Transportation Problem with Convex Polygonal Costs," Rand Memorandum RM-4941-PR, April 1966.
29. Graybill, F., An Introduction to Linear Statistical Models, I, McGraw-Hill, 1961.
30. Graybill, F., and W. Pruitt, "The Staircase Design: Theory," Annals of Mathematical Statistics, (29) 1958.
31. Hadley, G., Linear Programming, 1962, Addison-Wesley.
32. Hadley, G., Nonlinear and Dynamic Programming, 1964, Addison-Wesley.
33. Johnson, N., and F. Keone, Statistics and Experimental Design, Volume II, J. Wiley and Sons, 1964.
34. Kempthorne, O., Design and Analysis of Experiments, J. Wiley and Sons, 1952.

35. Kempthorne, O., "The Efficiency Factor of an Incomplete Block Design," *The Annals of Mathematical Statistics*, Vol. 27, 1956.
36. Kshirsagar, A., "A Note on Incomplete Block Designs," *The Annals of Mathematical Statistics*, Vol. 29, 1958.
37. Kunzi, H., and W. Oettli, "Integer Quadratic Programming," Recent Advances in Mathematical Programming, Edited by R. Graves and P. Wolfe, McGraw-Hill, 1963.
38. Lawler, E., and M. Bell, "A Method for Solving Discrete Optimization Problems," *Operations Research*, November-December 1966, Vol. 14, No. 6.
39. Lax, M., "Localized Perturbations," *Physical Review*, Vol. 94, 1954.
40. Mann, H., Analysis and Design of Experiments, Dover, 1949.
41. Martin, G., "An Accelerated Euclidean Algorithm for Integer Linear Programming" in (37).
42. Plackett, R., and J. Burman, "The Design of Optimum Multifactor Experiments," *Biometrika* (33), 1946.
43. Rao, C., "A Note on Balanced Designs," *Annals of Mathematical Statistics*, 1953.
44. Rao, C., Linear Statistical Inference and Its Applications, J. Wiley and Sons, 1965.
45. Rhode, C., and J. Harvey, "Unified Least Squares Analysis," *Journal of the American Statistical Association*, June 1965.
46. Roy, S., and Gnanadesikan, "Some Contributions to ANOVA in One or More Dimensions: I," *Annals of the American Statistical Association*, Vol. 30 (1959).
47. Shah, B., "On Balancing in Factorial Experiments," *Annals of Mathematical Statistics*, Vol. 29, No. 3, September 1958.

48. Smith, D., and W. Orchard-Hays, "Computational Efficiency in Product Form LP Codes," Recent Advances in Mathematical Programming, McGraw-Hill, 1963.
49. Thrall, R., "Lecture Notes on Operations Research," Engineering Summer Conferences 1966, University of Michigan.
50. Wald, A., "On the Efficient Design of Statistical Investigations," The Annals of Mathematical Statistics, Vol. 14, No. 2, June 1943.
51. Webb, S., "Construction and Comparison of Non-Orthogonal Incomplete Factorial Design," Proceedings of the Eleventh Conference on the Design of Experiments, Army Research Development and Testing ARD Report 66-2.
52. Webb, S., "Design, Testing and Estimation in Complex Experimentation," ARL 65-116, Part 1, June 1965, Aerospace Research Laboratories.
53. Wilde, D., Optimum Seeking Methods, Prentice-Hall, 1964.
54. Wilkinson, J., The Algebraic Eigenvalue Problem, Clarendon Press, 1965.
55. Witzgall, C., "An All-Integer Programming Algorithm with Parabolic Constraints," Journal Society Industrial and Applied Mathematics, Vol. 11, No. 4, December 1963.
56. Wolfe, P., "Methods of Nonlinear Programming," p. 67 in Recent Advances in Mathematical Programming, McGraw-Hill, 1963.