ON THE MINIMIZATION OF A CERTAIN CONVEX FUNCTION ARISING IN APPLIED DECISION THEORY

W. A. Ericson

*The University of Michigan**

ABSTRACT

The author, in an expository paper [4], has presented an algorithm for choosing a non-negative vector $\vec{n}$ to minimize the function $v(\vec{n}) = \vec{\pi}(\vec{N} + \vec{M})^{-1}\vec{\pi}^t$ subject to the constraint $\vec{c}\vec{n}^t = C > 0$, where $\vec{N} = \text{diag}(\vec{n})$, $\vec{\pi} \neq \vec{0}$, $\vec{c} > \vec{0}$ are given vectors and $\vec{M}$ is positive definite symmetric. In this paper a derivation of this algorithm is presented, including an exact solution in a degenerate case, only alluded to in [4]. Several applications, in addition to that of [4], are briefly indicated.

1. INTRODUCTION AND APPLICATIONS:

Let $\vec{M}$ be a $k \times k$ positive definite symmetric matrix, let $\vec{\pi}(\neq \vec{0})$ and $\vec{c} > 0$ be given k-dimensional row vectors, and let $C > 0$ be a given scalar. The problem discussed here is that of choosing $\vec{n} = (n_1, \ldots, n_k)$ to minimize

$$(1) \qquad v(\vec{n}) = \vec{\pi}(\vec{N} + \vec{M})^{-1}\vec{\pi}^t$$

subject to the constraints

$$(2) \qquad \vec{n} \geq \vec{0}$$

and

$$(3) \qquad \vec{c}\,\vec{n}^t \leq C ,$$

where $\vec{N} = \text{diag}(\vec{n}) = \begin{bmatrix} n_1 & & & 0 \\ & n_2 & & \\ & & \ddots & \\ 0 & & & n_k \end{bmatrix}$.

(A superscript "t" will throughout denote transpose.)

The author [4] has shown that the posterior variance of $\Sigma \pi_i \mu_i$ is of the form $v(\vec{n})$ when the μ_i's are assigned a k-dimensional normal prior distribution with variance-covariance matrix $\vec{M}^{-1}$, and where given μ_i the conditional distribution of relevant sample

*The research reported here was partially supported by the Division of Research, Graduate School of Business Administration, Harvard University. Final preparation of the paper was supported by the National Science Foundation, Grant No. NSF-GP-6008.

statistics $\bar{x}_i$ (i = 1, . . . , k) based on sample sizes n_i, respectively, are independently normal with mean μ_i and variance proportional to $1/n_i$. Also it is easy to verify, using results given in Raiffa and Schlaifer [8], that under a wide variety of models of k independent data generating processes and natural conjugate prior distributions, the posterior variance (or prior expectation of the posterior variance) of a linear combination of the unknown process parameters may be put into the form (1), where the n_i are essentially the to-be-determined sample sizes. Thus the problem indicated above is one of choosing sample sizes to minimize the posterior variance of a linear combination of unknown parameters subject to a cost constraint. This has immediate applicability to problems of optimal stratified sample allocation as well as to optimal design for inference regarding the difference (or contrasts) among treatment effects. These problems have been treated by the author in a series of papers [4], [5], and [6].

2. PRELIMINARIES

Let the ij-th elements of $\vec{M}$ and $(\vec{N} + \vec{M})^{-1}$ be denoted respectively by m_{ij} and $u_{ij}(\vec{n})$, further let

$$t_i(\vec{n}) = \sum_{j=1}^{k} \pi_j u_{ij}(\vec{n}), \qquad i = 1, \ldots, k .$$

Observe that it may be assumed that $\vec{\pi} \geq \vec{0}$, for if some one or more π_i's are negative they may, without altering the function $v(\vec{n})$, be replaced by their absolute values provided merely that the sign of every element in the corresponding rows and columns of $\vec{M}$ be reversed.

Also letting $R_n = \{\vec{n} \mid \vec{n} \geq \vec{0}, \vec{c}\vec{n}^t \leq C\}$, it is clear that R_n is a closed, bounded, convex set of k dimensional vectors $\vec{n}$. Several useful properties of $v(\vec{n})$ for $\vec{n} \in R_n$ are given in the following lemma.

LEMMA 1: For $\vec{n} \in R_n$, $v(\vec{n})$ possesses the following properties:

a) $v(\vec{n})$ is twice differentiable with respect to $\vec{n}$ having derivatives

$$\frac{\partial v(\vec{n})}{\partial \vec{n}} \equiv \left(\frac{\partial v(\vec{n})}{\partial n_1}, \ldots, \frac{\partial v(\vec{n})}{\partial n_k}\right) = \left(-t_1^2(\vec{n}), \ldots, -t_k^2(\vec{n})\right), \tag{4}$$

and

$$\frac{\partial^2 v(\vec{n})}{\partial \vec{n}^2} \equiv \left[\ldots \frac{\partial^2 v(\vec{n})}{\partial n_i \partial n_j} \ldots\right] = [\ldots 2t_i(\vec{n}) t_j(\vec{n}) u_{ij}(\vec{n}) \ldots] . \tag{5}$$

b) $v(\vec{n}) > 0$, for $\vec{n} \geq \vec{0}$.

c) $v(\vec{n})$ is a non-increasing function of $\vec{n}$.

d) $v(\vec{n})$ is a convex function.

e) For every $\epsilon > 0$, there is an $\vec{n}^* \geq \vec{0}$ such that $\vec{n} > \vec{n}^*$ implies $v(\vec{n}) \leq \epsilon$.

PROOF: Property b) follows immediately from the definition of $v(\vec{n})$, and the assumption that $\vec{\pi} \neq \vec{0}$. The derivatives of a) may be verified by direct calculation, or much more easily by using the matrix derivative theory of Dwyer [3]. Properties c) and d) follow immediately from the form taken by the partial derivatives (4), (5), and the positive definiteness of $[\ldots u_{ij}(\vec{n}) \ldots]$. To demonstrate e), it suffices, by c), to show that for every $\epsilon > 0$ there is an $\vec{n}^* > \vec{0}$ such that $v(\vec{n}^*) \leq \epsilon$. Assume the contrary, i.e., for all $\vec{n} \geq \vec{0}$, $v(\vec{n}) > \epsilon$. Let $\epsilon = \vec{\pi}\vec{R}\vec{\pi}^t$ for some arbitrary positive definite diagonal matrix $\vec{R}$. Then as shown in [1, p. 341] or in [2, p. 58], there exists a non-singular matrix $\vec{L}$ such that $\vec{L}(\vec{N} + \vec{M})^{-1}\vec{L}^t = \vec{D}$, a diagonal matrix, and $\vec{L}\vec{R}\vec{L}^t = \vec{I}$, the identity. Letting $\vec{\pi} = \vec{x}\vec{L}$, one finds the obvious contradiction that for <u>all</u> $\vec{n} > \vec{0}$, $\vec{x}(\vec{L}^t)^{-1}(\vec{N} + \vec{M} - \vec{R}^{-1})(\vec{L})^{-1}\vec{x}^t < 0$.

Observe that property c) means, in effect, that the solution to the problem originally posed will be an $\vec{n}$ such that

$$\vec{c}\vec{n}^t = C, \tag{6}$$

so that in analyzing this problem we may replace the constraint (3) by the equality (6). This Lemma also establishes all the needed properties so that general convex programming algorithms may be utilized in finding specific numerical solutions. The specific algorithm and solution developed here will yield analytic details of the solution which have been found extremely useful in applications in decision theory. Finally the properties of $v(\vec{n})$ given in Lemma 1 establish the applicability of a special case of the fundamental result of Kuhn and Tucker [7]. (See also the presentation in [9].) Specifically we have, adapting the Kuhn-Tucker result to this problem:

LEMMA 2: The minimum of $v(\vec{n})$, (1), subject to the conditions (2) and (6) is at $\vec{n} = \vec{n}^0$ if and only if there exists a $\lambda_0 > 0$ such that:

$$-\lambda_j \equiv -t_j^2(\vec{n}^0) + \lambda_0 c_j \geqq 0 \quad \text{if} \quad n_j^0 = 0 \tag{7}$$

$$-t_j^2(\vec{n}^0) + \lambda_0 c_j = 0 \quad \text{if} \quad n_j^0 > 0 \tag{8}$$

and

$$\vec{n}^0\vec{c}^t = C. \tag{9}$$

Our goal is to use this basic result to find an algorithm whereby the solution to the problem can be mapped out for all $C > 0$. The solution which is obtained below may be described briefly as follows: the interval $C > 0$ is partitioned into sub-intervals within each of which some subset of the n_i^0's are non-zero, each such n_i^0 being a linear function of C in that interval. Over all $C > 0$ each n_i^0 is a continuous piecewise linear function of C. The algorithm below gives these linear functions explicit form and gives a method for determining the sub-intervals.

3. BASIC RESULTS

Let $K = \{1, 2, \ldots, k\}$, $S = \{i_1, \ldots, i_r\}$ be any subset of K, and $W = K - S$. Now by Lemma 2 it is clear that for every $C > 0$ there is some $S \subseteq K$ such that the solution for that C is characterized by

$$n_j^0 \begin{cases} = 0 & . \ j \in W \\ > 0 & j \in S . \end{cases} \tag{10}$$

A useful first step in obtaining the solution for all $C > 0$ is to characterize the set of C's (perhaps empty) for which $\vec{n}^0$ is of the form (10) for some given subset S of K. This set of C's is easily obtained using Lemma 2. Note that condition (8) may be satisfied for $j \in S$ by $t_j(\vec{n}) = \pm(\lambda_0 c_j)^{1/2}$; however, if S is specified and if, additionally, one specifies a set of signed ones s_j (+1 or -1) for $j \in S$, then the Kuhn-Tucker conditions may be solved for λ_0, n_j^0, $(j \in S)$ and λ_j $(j \in W)$ in terms of C. The resulting n_j^0 will then be the minimizing set for all C's such that $n_j^0 \geq 0$ $(j \in S)$, $\lambda_0 > 0$ and $\lambda_j \geq 0$ $(j \in W)$. Details of this initial step are given in Theorem 1 below.

Given S and s_j $(j \in S)$, we adopt the following definitions and conventions:

In general given any matrix $\vec{A}$ and subsets U and V of K, $\vec{A}_{UV}$ will denote the matrix formed from $\vec{A}$ by deleting all rows i for $i \notin U$ and all columns j for $j \notin V$. In particular if $\vec{M}$ is the $k \times k$ positive definite symmetric matrix of (1), we will need the following:

$$\vec{V}_{WW} = \vec{M}_{WW}^{-1} \tag{11}$$

$$\vec{B}_{WS} = \vec{V}_{WW}\vec{M}_{WS} \quad (= \vec{0} \text{ if } S = \phi \text{ or } K) \tag{12}$$

$$\vec{R}_{SS} = \vec{M}_{SS} - \vec{M}_{SW}\vec{V}_{WW}\vec{M}_{WS} . \tag{13}$$

Similarly, $\vec{\pi}_S$ and $\vec{c}_S$ are row vectors formed from $\vec{\pi}$ and $\vec{c}$, respectively, by deleting all entries in W (not in S). Let

$$\vec{\psi} = (\psi_1 \ldots \psi_k), \quad \text{where } \psi_j = \begin{cases} s_j c_j^{1/2} & j \in S \\ \left[c_j + \lambda_j(W)/\lambda_0(S)\right]^{1/2} & j \in W \end{cases} \tag{14}$$

where $\lambda_j(W)$ and $\lambda_0(S)$ will be the quantities λ_j and λ_0 of (7) and (8). $\vec{\psi}_S$ and $\vec{\psi}_W$ are defined in analogy with $\vec{\pi}_S$. Also we let

$$\vec{G}(S) = \begin{bmatrix} 1/\psi_{i_1} & & 0 \\ & \ddots & \\ 0 & & 1/\psi_{i_r} \end{bmatrix} . \tag{15}$$

Finally we adopt the convention that if $\vec{x} = (x_1 \ldots x_k)$ then $\vec{x}^2 = (x_1^2, \ldots, x_k^2)$. With these preliminaries we have:

THEOREM 1: For any given subset S of K and a given set of signed ones, s_j $(j \in S)$, and for all $C > 0$ satisfying

(16) a) $$\gamma(S) \equiv \vec{\psi}_S\left(\vec{\pi}_S - \vec{\pi}_W \vec{B}_{WS}\right)^t \neq 0$$

(17) b) $$\vec{\lambda}(W) = \left[\vec{\pi}_W \vec{V}_{WW} - \lambda_0^{1/2}(S)\vec{\psi}_S \vec{B}^t_{WS}\right]^2 - \lambda_0(S)\vec{c}_W \leq \vec{0}$$

and

(18) c) $$\vec{\nu}(S) \equiv \left[\lambda_0^{-1/2}(S)\left(\vec{\pi}_S - \vec{\pi}_W \vec{B}_{WS}\right) - \vec{\psi}_S \vec{R}_{SS}\right]\vec{G}(S) \geq \vec{0},$$

where

(19) $$\lambda_0^{1/2}(S) = \frac{\gamma(S)}{C + \vec{\psi}_S \vec{R}_{SS} \vec{\psi}^t_S},$$

the solution to the problem of minimizing (1) subject to (2) and (3) is given by

(20) $$\vec{n}_W^{\,0} = \vec{0}$$

and

(21) $$\vec{n}_S^{\,0} = \vec{\nu}(S).$$

PROOF: By the definition of $t_j(\vec{n})$ the conditions (7) and (8) become

$$\vec{\pi}(\vec{N} + \vec{M})^{-1} = \lambda_0^{1/2}(S)\vec{\psi}.$$

Using the fact that $\vec{N}$ is diagonal and the assumption that $\vec{N}_{WW} = \vec{0}$, this system becomes

(22) $$\lambda_0^{-1/2}(S)\vec{\pi}_S = \vec{\psi}_S\left(\vec{N}_{SS} + \vec{M}_{SS}\right) + \vec{\psi}_W \vec{M}_{WS}$$

and

(23) $$\lambda_0^{-1/2}(S)\vec{\pi}_W = \vec{\psi}_S \vec{M}_{SW} + \vec{\psi}_W \vec{M}_{WW}.$$

Solving (23) for $\vec{\psi}_W$ one finds

$$\vec{\psi}_W = \lambda_0^{-1/2}(S)\vec{\pi}_W \vec{V}_{WW} - \vec{\psi}_S \vec{B}^t_{WS},$$

and by the definition of $\vec{\psi}_W$, (14); and by letting $\vec{\lambda}(W)$ be the vector of $\lambda_j(W)$'s, one has immediately

(24) $$\vec{\lambda}(W) = \left(\vec{\pi}_W \vec{V}_{WW} - \lambda_0^{1/2}(S)\,\vec{\psi}_S \vec{B}^t_{WS}\right)^2 - \lambda_0(S)\,\vec{c}_W .$$

Further, by substituting the above expression for $\vec{\psi}_W$ in (22) one finds

$$\vec{\psi}_S \vec{N}_{SS} = \lambda_0^{-1/2}(S)\left(\vec{\pi}_S - \vec{\pi}_W \vec{B}_{WS}\right) - \vec{\psi}_S \vec{R}_{SS} .$$

But the elements of $\vec{\psi}_S \vec{N}_{SS}$ are just $s_j c_j^{1/2} n_j^0$ for $j \in S$. Hence multiplying this expression by $\vec{G}(S)$ one has

(25) $$\vec{n}_S^{\,0} = \left[\lambda_0^{-1/2}(S)\left(\vec{\pi}_S - \vec{\pi}_W \vec{B}_{WS}\right) - \vec{\psi}_S \vec{R}_{SS}\right] \vec{G}(S) .$$

Further, by multiplying the expression above for $\vec{\psi}_S \vec{N}_{SS}$ by $\vec{\psi}_S^{\,t}$ and solving this scalar equation for $\lambda_0^{-1/2}(S)$, one finds that

$$\lambda_0^{-1/2}(S) = \frac{\vec{\psi}_S \vec{N}_{SS} \vec{\psi}_S^{\,t} + \vec{\psi}_S \vec{R}_{SS} \vec{\psi}_S^{\,t}}{\left(\vec{\pi}_S - \vec{\pi}_W \vec{B}_{WS}\right)\vec{\psi}_S^{\,t}} ;$$

and by the use of condition (9), that $\vec{\psi}_S \vec{N}_{SS} \vec{\psi}_S^{\,t} = \sum_{j \in S} c_j n_j^0 = C$ yields Eq. (19). When these solutions for $\vec{\lambda}(W)$, $\vec{n}^0(S)$, and $\lambda_0(S)$ are used, the conditions a) - c) of the theorem are mere re-expressions of Lemma 2, establishing the theorem.

Observe that if the subset S and signs s_j satisfy the conditions of this theorem then so do S and the signs $-s_j$. Thus it may be assumed that given any subset S we will always take signs s_j such that $\gamma(S) > 0$.

It is also important to note that for a given subset S of K and signed ones s_j ($j \in S$) each of the k conditions (17) and (18) defines an interval on C where the corresponding inequality obtains. It is then for all C in the interval, possibly empty, found by taking the intersection of these k intervals that $\vec{n}^0$ is given by (20) and (21). Clearly to check all possible inputs to Theorem 1, i.e., all subsets S of K and all sign assignments, s_j, is not feasible. Fortunately there is a simple procedure for determining the sequence of meaningful S and s_j's which define the successive intervals on C mentioned earlier. As a first step in this procedure we have the following special case of Theorem 1 for the situation where all π_j are positive, i.e., no π_j is zero. Taking $S = K$ and $s_j = 1$ ($j = 1, \ldots, k$) the following corollary is an immediate consequence of the preceding theorem:

COROLLARY: If $\vec{\pi} > \vec{0}$ and

(26) $$C \geq C_0 = \max_{\ell}\left\{\frac{\vec{\psi}\vec{M}_\ell^{\,t}}{\pi_\ell}\right\} \vec{\pi}\vec{\psi}^t - \vec{\psi}\vec{M}\vec{\psi}^t ,$$

where $\vec{M}_\ell$ is the ℓth row of $\vec{M}$, $\vec{\psi} = \left(c_1^{1/2}, \ldots, c_k^{1/2}\right)$, then the minimizing $\vec{n}$ is given by

(27) $$\vec{n}^0 = \left[\left(\frac{C + \vec{\psi}\vec{M}\vec{\psi}^t}{\vec{\pi}\vec{\psi}^t}\right)\vec{\pi} - \vec{\psi}\vec{M}\right]\vec{G},$$

where $\vec{G} = \begin{bmatrix} c_1^{-1/2} & & 0 \\ & \ddots & \\ 0 & & c_k^{-1/2} \end{bmatrix}$.

Thus if $\vec{\pi} > \vec{0}$ the first (open) interval $C \geq C_0$ and the solution vector $\vec{n}^0$ for C's in that interval are established by this corollary. What remains is only to provide a mechanism by which one can, given any meaningful interval on C (set of inputs to Theorem 1 yielding $\vec{n}^0$ for a non-empty interval on C), find the next set of inputs to that theorem which yield the solution in the next adjacent interval on C. A special important case of such a result is given in Theorem 2 below. This result can be used in practice to find the solution (or at least a good approximation) for any problem of the type under discussion. This is the so-called non-degenerate case.

DEFINITION: By non-degeneracy is meant that $\vec{\pi} > \vec{0}$ and also for every interval on C for which the conditions b) and c) of Theorem 1 are met, these conditions are satisfied with strict inequality in each, except at the end points of the interval where one and only one of the k conditions b) and c) is met with strict equality.

Most real problems of the form (1), (2), and (3) are of the non-degenerate type and, if not, may be subsumed under this general case by randomly perturbing the π_i's by adding to them arbitrarily small and unequal ϵ_i's. In this non-degenerate case an algorithm for obtaining the solution for all $C > 0$ is completed by the following theorem:

THEOREM 2: In the non-degenerate case if for $0 \leq C_L \leq C \leq C_U \leq \infty$ the conditions a) through c) of Theorem 1 are satisfied for some given subset S^* of K and signs s_j^* $(j \in S^*)$ and further if

(a) for $C < C_L$ the condition c) is violated for $j^* \in S^*$ then for a finite interval on C $(C \leq C_L)$, $\vec{n}^0$ is given by Theorem 1 by taking $S = S^* - \{j^*\}$ and signs $s_j = s_j^*$ for $j \in S$; and

(b) for $C < C_L$ the condition b) of Theorem 1 is violated for $j^* \in W^* \equiv K - S^*$, then for a finite interval on $C < C_L$ the optimum $\vec{n}^0$ is given by Theorem 1 taking $S = S^* \cup \{j^*\}$ and $s_j = s_j^*$, $j \neq j^*$, and s_j^* determined so that for $\lambda_0(S^*)$ evaluated at $C = C_L$, s_j^* takes the sign of the j*th element of

(28) $$\vec{\pi}_{W^*}\vec{V}_{W^*W^*} - \lambda_0^{1/2}(S^*)\vec{\psi}_{S^*}\vec{B}^t_{W^*S^*}.$$

This result is a simple special case of Theorem 3 below, and its proof follows from the proof of the latter theorem.

Thus in the non-degenerate case the solution for all $C > 0$ is mapped out by using the Corollary to Theorem 1, then Theorem 2a, Theorem 1, Theorem 2a or 2b, etc., until the resulting sequence of non-overlapping intervals covers $C \geq 0$. It should also be pointed out that at each step in this process no matrix inversion is necessary, for explicit formulae are given in

[4] (special cases of Lemma 3 below) for changing the expressions of Theorem 1 from the inputs S* to those for S. A numerical example of the algorithm resulting from the preceding theory is also given in [4].

4. DEGENERACY

Although in most practical applications the expedient is to perturb the π_i's and thereby eliminate any possible degeneracy, in this section the necessary general theory for finding exact solutions is given. Recall that degeneracy may arise either because one or more (but clearly not all) π_i's may be zero or where at the lower endpoint, C_L, of one or more of the sub-intervals on C several of the conditions (17) or (18) may hold simultaneously with equality at $C = C_L$ and are violated for $C < C_L$. We handle this second type of degeneracy first. The problem arising is solved by a suitable generalization of Theorem 2. The following preliminaries are needed.

Let S, J, and W' form a partition of $K = \{1, 2, \ldots, k\}$ and $\vec{M}$ be the $k \times k$ positive definite symmetric matrix of (1). Let $\vec{M}^*$ consist of permuted rows and columns of $\vec{M}$ so that

$$\vec{M}^* = \begin{bmatrix} \vec{M}_{SS} & \vec{M}_{SJ} & \vec{M}_{SW'} \\ \vec{M}_{JS} & \vec{M}_{JJ} & \vec{M}_{JW'} \\ \vec{M}_{W'S} & \vec{M}_{W'J} & \vec{M}_{W'W'} \end{bmatrix},$$

where the matrix elements of $\vec{M}^*$ are as defined earlier. It is assumed that the relevant $\vec{\psi}$, $\vec{\pi}$, $\vec{c}$, and so on, are conformably permuted and partitioned. In Lemma 3 below the quantities defined in (11), (12), and (13) for S and its complement $W = W' \cup J$ are related to the same quantities taking $S' = S \cup J$ and its complement W'. We first redefine these six quantities in partitioned form as follows:

for S and W

$$\text{(29)} \qquad \vec{V}_{WW} \equiv \begin{bmatrix} \vec{V}_{JJ} & \vec{V}_{JW'} \\ \vec{V}_{W'J} & \vec{V}_{W'W'} \end{bmatrix} = \begin{bmatrix} \vec{M}_{JJ} & \vec{M}_{JW'} \\ \vec{M}_{W'J} & \vec{M}_{W'W'} \end{bmatrix}^{-1} = \vec{M}_{WW}^{-1},$$

$$\text{(30)} \qquad \vec{B}_{WS} \equiv \begin{bmatrix} \vec{B}_{JS} \\ \vec{B}_{W'S} \end{bmatrix} = \vec{M}_{WW}^{-1} \vec{M}_{WS} = \begin{bmatrix} \vec{M}_{JJ} & \vec{M}_{JW'} \\ \vec{M}_{W'S} & \vec{M}_{W'W'} \end{bmatrix}^{-1} \begin{bmatrix} \vec{M}_{JS} \\ \vec{M}_{W'S} \end{bmatrix},$$

and

$$\text{(31)} \qquad \vec{R}_{SS} = \vec{M}_{SS} - \vec{M}_{SW} \vec{M}_{WW}^{-1} \vec{M}_{WS};$$

while for S' and W' we have

(32) $$V'_{W'W'} = \left[M_{W'W'}\right]^{-1},$$

(33) $$\vec{B}'_{W'S'} \equiv \left[\vec{B}'_{W'S} \quad \vec{B}'_{W'J}\right] = \vec{M}^{-1}_{W'W'}\left[\vec{M}_{W'S} \quad \vec{M}_{W'J}\right],$$

and

$$\vec{R}'_{S'S'} \equiv \begin{bmatrix} \vec{R}'_{SS} & \vec{R}'_{SJ} \\ \vec{R}'_{JS} & \vec{R}'_{JJ} \end{bmatrix}$$

$$= \begin{bmatrix} \vec{M}_{SS} - \vec{M}_{SW'}\vec{M}^{-1}_{W'W'}\vec{M}_{W'S} & \vec{M}_{SJ} - \vec{M}_{SW'}\vec{M}^{-1}_{W'W'}\vec{M}_{W'J} \\ \vec{M}_{JS} - \vec{M}_{JW'}\vec{M}^{-1}_{W'W'}\vec{M}_{W'S} & \vec{M}_{JJ} - \vec{M}_{JW'}\vec{M}^{-1}_{W'W'}\vec{M}_{W'J} \end{bmatrix}.$$

These expressions follow immediately from (11) - (13). With these definitions one has:

LEMMA 3: Under the above definitions with $\vec{M}^*$ positive definite symmetric:

(35) $$\vec{V}_{WW} = \begin{bmatrix} \left(\vec{R}'_{JJ}\right)^{-1}, & -\left(\vec{R}'_{JJ}\right)^{-1}\left(\vec{B}'_{W'J}\right)^{t} \\ -\vec{B}'_{W'J}\left(\vec{R}'_{JJ}\right)^{-1}, & \vec{V}'_{W'W'} + \vec{B}'_{W'J}\left(\vec{R}'_{JJ}\right)^{-1}\vec{B}'_{JW'} \end{bmatrix},$$

(36) $$\vec{B}_{WS} = \begin{bmatrix} \left(\vec{R}'_{JJ}\right)^{-1}\vec{R}'_{JS} \\ \vec{B}'_{W'S} - \vec{B}'_{W'J}\left(\vec{R}'_{JJ}\right)^{-1}\vec{R}'_{JS} \end{bmatrix},$$

and

(37) $$\vec{R}_{SS} = \vec{R}'_{SS} - \vec{R}'_{SJ}\left(\vec{R}'_{JJ}\right)^{-1}\vec{R}'_{JS};$$

while conversely

(38) $$\vec{V}'_{W'W'} = \vec{V}_{W'W'} - \vec{V}_{W'J}\left(\vec{V}_{JJ}\right)^{-1}\vec{V}_{JW'},$$

(39) $$\vec{B}'_{W'S'} = \left[\vec{B}_{W'S} - \vec{V}_{W'J}\vec{V}^{-1}_{JJ}\vec{B}_{JS}, \quad -\vec{V}_{W'J}\vec{V}^{-1}_{JJ}\right],$$

and

(40) $$\vec{R}'_{S'S'} = \begin{bmatrix} \vec{R}_{SS} + \vec{B}_{SJ}\vec{V}^{-1}_{JJ}\vec{B}_{JS}, & \vec{B}_{SJ}\vec{V}^{-1}_{JJ} \\ \vec{V}^{-1}_{JJ}\vec{B}_{JS}, & \vec{V}^{-1}_{JJ} \end{bmatrix} .$$

PROOF: It is a commonly known and easily verifiable result that

$$\vec{V}_{WW} = \begin{bmatrix} \vec{M}_{JJ} & \vec{M}_{JW'} \\ \vec{M}_{W'J} & \vec{M}_{W'W'} \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} \left(\vec{R}'_{JJ}\right)^{-1}, & -\left(\vec{R}'_{JJ}\right)^{-1}\vec{M}_{JW'}\vec{M}^{-1}_{W'W'} \\ -\vec{M}^{-1}_{W'W'}\vec{M}_{W'J}\left(\vec{R}'_{JJ}\right)^{-1}, & \vec{M}^{-1}_{W'W'} + \vec{M}^{-1}_{W'W'}\vec{M}_{W'J}\left(\vec{R}'_{JJ}\right)^{-1}\vec{M}_{JW'}\vec{M}^{-1}_{W'W'} \end{bmatrix},$$

where $\vec{R}'_{JJ} = (\vec{M}_{JJ} - \vec{M}_{JW'}\vec{M}^{-1}_{W'W'}\vec{M}_{W'J})$, as per (34). The result (35) then follows from this form using (32) and (33). The expression (36) is readily obtained by substituting the abo expression in (30) and then using definitions (32) - (34). By substituting (36) in (31), one finds

$$\vec{R}_{SS} = \vec{M}_{SS} - \vec{M}_{SW'}\vec{B}'_{W'S} - \vec{M}_{SJ}\left(\vec{R}'_{JJ}\right)^{-1}\vec{R}'_{JS} + \vec{M}_{SW'}\vec{B}'_{W'J}\left(\vec{R}'_{JJ}\right)^{-1}\vec{R}'_{JS} ,$$

from which (37) follows using (34). Finally, expressions (38) - (40) are easily obtained by solving the system of equations (35) - (37) for the primed quantities.

Finally, since we are aiming at relating the successive inputs to Theorem 1, we restate, for convenience, the quantities used in that theorem for the subsets $S' = S \cup J$ and S in the partitioned notation used above. We have first for the subset S':

(41) $$\gamma(S') = \vec{\psi}_S\left(\vec{\pi}_S - \vec{\pi}_{W'}\vec{B}'_{W'S}\right)^t + \vec{\psi}_J\left(\vec{\pi}_J - \vec{\pi}_{W'}\vec{B}'_{W'J}\right)^t ,$$

(42) $$\vec{\lambda}(W') = \left[\vec{\pi}_{W'}\vec{V}'_{W'W'} - \lambda_0^{1/2}(S')\left(\vec{\psi}_S\vec{B}'^t_{W'S} + \vec{\psi}_J\vec{B}'^t_{W'J}\right)\right]^2 - \lambda_0(S')\vec{c}_{W'} ,$$

and

(43) $$\vec{\nu}_S(S') = \left[\lambda_0^{-1/2}(S')\left(\vec{\pi}_S - \vec{\pi}_{W'}\vec{B}'_{W'S}\right) - \vec{\psi}_S\vec{R}'_{SS} - \vec{\psi}_J\vec{R}'_{JS}\right]\vec{G}_S(S') ,$$

(44) $$\vec{\nu}_J(S') = \left[\lambda_0^{-1/2}(S')\left(\vec{\pi}_J - \vec{\pi}_{W'}\vec{B}'_{W'J}\right) - \vec{\psi}_S\vec{R}'_{SJ} - \vec{\psi}_J\vec{R}'_{JJ}\right]\vec{G}_J(S') ,$$

where

(45) $$\lambda_0^{1/2}(S') = \frac{\gamma(S')}{C + \vec{\psi}_S \vec{R}'_{SS} \vec{\psi}_S^t + \vec{\psi}_S \vec{R}'_{SJ} \vec{\psi}_J^t + \vec{\psi}_J \vec{R}'_{JS} \vec{\psi}_S^t + \vec{\psi}_J \vec{R}'_{JJ} \vec{\psi}_J^t},$$

and $\vec{G}_S(S')$ and $\vec{G}_J(S')$ are diagonal matrices having elements $1/\psi_j$ for $j \in S$ and $j \in J$, respectively.

For the subset $S = S' - J$, the same quantities are given by:

(46) $$\gamma(S) = \vec{\psi}_S \left(\vec{\pi}_S - \vec{\pi}_{W'} \vec{B}_{W'S} - \vec{\pi}_J \vec{B}_{JS}\right)^t,$$

(47) $$\vec{\lambda}_{W'}(W) = \left[\vec{\pi}_{W'} \vec{V}_{W'W'} + \vec{\pi}_J \vec{V}_{JW'} - \lambda_0^{1/2}(S) \vec{\psi}_S \vec{B}^t_{W'S}\right]^2 - \lambda_0(S) \vec{c}_{W'},$$

(48) $$\vec{\lambda}_J(W) = \left[\vec{\pi}_{W'} \vec{V}_{W'J} + \vec{\pi}_J \vec{V}_{JJ} - \lambda_0^{1/2}(S) \vec{\psi}_S \vec{B}^t_{JS}\right]^2 - \lambda_0(S) \vec{c}_J,$$

and

(49) $$\vec{\nu}(S) = \left[\lambda_0^{-1/2}(S) \left(\vec{\pi}_S - \vec{\pi}_{W'} \vec{B}_{W'S} - \vec{\pi}_J \vec{B}_{JS}\right) - \vec{\psi}_S \vec{R}_{SS}\right] \vec{G}(S),$$

where

(50) $$\lambda_0^{1/2}(S) = \frac{\gamma(S)}{C + \vec{\psi}_S \vec{R}_{SS} \vec{\psi}_S^t}.$$

With these preliminaries we may state and prove the following generalization of Theorem 2:

THEOREM 3: (a) Consider the inputs to Theorem 1 consisting of the subset $S' = S \cup J$ of K and signs s'_j $(j \in S')$. Suppose for such inputs $\gamma(S') > 0$ and further that at the point $C = C_L$, $\vec{\nu}_J(S') = \vec{0}$ and $\vec{\nu}_J(S') < \vec{0}$ for $C < C_L$, then taking the new subset $S \subseteq K$ and signs $s_j = s'_j$ for $j \in S$ one finds $\gamma(S) > 0$; at $C = C_L$, $\lambda_0^{1/2}(S) = \lambda_0^{1/2}(S')$, $\vec{\nu}(S) = \vec{\nu}_S(S')$, $\vec{\lambda}_{W'}(W) = \vec{\lambda}(W')$, and $\vec{\lambda}_J(W) = \vec{0}$, while for $C < C_L$, $\vec{\lambda}_J(W) \not> \vec{0}$. Also if additionally at $C = C_L$, $\vec{\lambda}(W') < \vec{0}$ and $\vec{\nu}_S(S') > \vec{0}$ then in an interval on $C < C_L$, $\vec{\lambda}_{W'}(W) < \vec{0}$ and $\vec{\nu}(S) > \vec{0}$.

(b) Consider the inputs to Theorem 1 comprising the subset S of K (with complement $W = W' \cup J$) and signs s_j $(j \in S)$. Suppose for such inputs $\gamma(S) > 0$ and also at the point $C = C_L$, $\vec{\lambda}_J(W) = \vec{0}$ and $\vec{\lambda}_J(W) < \vec{0}$ for $C < C_L$, then letting the new subset be $S' = S \cup J$ and signs $s'_j = s_j$ for $j \in S$ and s_j for $j \in J$ (or equivalently $\vec{G}_J(S')$) determined so that at $C = C_L$

(51) $$\left[\vec{\pi}_{W'} \vec{V}_{W'J} + \vec{\pi}_J \vec{V}_{JJ} - \lambda_0^{1/2}(S) \vec{\psi}_S \vec{B}^t_{JS}\right] = \lambda_0^{1/2} \vec{G}_J(S'),$$

then $\gamma(S') > 0$ and at $C = C_L$, $\lambda_0^{1/2}(S') = \lambda_0^{1/2}(S)$, $\vec{\nu}(S') = \vec{\nu}(S)$, $\vec{\lambda}(W') = \vec{\lambda}_{W'}(W)$, and $\vec{\nu}_J(S') = \vec{0}$, while for $C < C_L$, $\vec{\nu}_J(S') \not< \vec{0}$. Also if additionally at $C = C_L$, $\vec{\lambda}_{W'}(W) < \vec{0}$ and $\vec{\nu}(S) > \vec{0}$, then in an interval on $C < C_L$, $\vec{\lambda}(W') < \vec{0}$ and $\vec{\nu}_S(S') > \vec{0}$.

PROOF: To demonstrate (a) note by hypothesis that for $C = C_L$, $\vec{\nu}_J(S') = \vec{0}$ which by (44) implies

$$\lambda_0^{1/2}(S') = \frac{\left(\vec{\pi}_J - \vec{\pi}_{W'}\vec{B}'_{W'J}\right)\vec{\psi}_J^{\,t}}{\vec{\psi}_S\vec{R}'_{SJ}\vec{\psi}_J^{\,t} + \vec{\psi}_J\vec{R}'_{JJ}\vec{\psi}_J^{\,t}}\,. \tag{52}$$

Equating this to (45) one finds that

$$C_L = \frac{\gamma(S')}{\left(\vec{\pi}_J - \vec{\pi}_{W'}\vec{B}'_{W'J}\right)\vec{\psi}_J^{\,t}}\left(\vec{\psi}_S\vec{R}'_{SJ}\vec{\psi}_J^{\,t} + \vec{\psi}_J\vec{R}'_{JJ}\vec{\psi}_J^{\,t}\right) - \left(\vec{\psi}_S, \vec{\psi}_J\right)\vec{R}'_{S'S'}\left(\vec{\psi}_S, \vec{\psi}_J\right)^t. \tag{53}$$

Taking the inputs $S \subseteq K$ and $s_j = s'_j$ for $j \in S$ one has by definition

$$\lambda_0^{1/2}(S) = \frac{\gamma(S)}{C + \vec{\psi}_S\vec{R}_{SS}\vec{\psi}_S^{\,t}} = \frac{\vec{\psi}_S\left(\vec{\pi}_S - \vec{\pi}_{W'}\vec{B}_{W'S} - \vec{\pi}_J\vec{B}_{JS}\right)^t}{C + \vec{\psi}_S\vec{R}_{SS}\vec{\psi}_S^{\,t}}. \tag{54}$$

Evaluating this expression for $C = C_L$, as given in (53), and using (52), (41), together with (36) and (37) of Lemma 3, one has for $C = C_L$

$$\bullet\ \lambda_0^{1/2}(S) = \frac{\vec{\psi}_S\left(\vec{\pi}_S - \vec{\pi}_{W'}\vec{B}'_{W'S}\right)^t - \vec{\psi}_S\vec{R}'_{SJ}\left(\vec{R}'_{JJ}\right)^{-1}\left(\vec{\pi}_J - \vec{\pi}_{W'}\vec{B}'_{W'J}\right)^t}{\lambda_0^{-1/2}(S')\,\vec{\psi}_S\left(\vec{\pi}_S - \vec{\pi}_{W'}\vec{B}'_{W'S}\right)^t - \left(\vec{\psi}_J\vec{R}'_{JS}\vec{\psi}_S^{\,t} + \vec{\psi}_S\vec{R}'_{SJ}\left(\vec{R}'_{JJ}\right)^{-1}\vec{R}'_{JS}\vec{\psi}_S^{\,t}\right)}.$$

Hence using the hypothesis that $\vec{\nu}_J(S') = \vec{0}$ one has

$$\lambda_0^{-1/2}(S')\left(\vec{\pi}_J - \vec{\pi}_{W'}\vec{B}'_{W'J}\right)\left(\vec{R}'_{JJ}\right)^{-1}\vec{R}'_{JS}\vec{\psi}_S^{\,t} = \left(\vec{\psi}_S\vec{R}'_{SJ} + \vec{\psi}_J\vec{R}'_{JJ}\right)\left(\vec{R}'_{JJ}\right)^{-1}\vec{R}'_{JS}\vec{\psi}_S^{\,t}\,,$$

from which it follows that at $C = C_L$, $\lambda_0^{1/2}(S) = \lambda_0^{1/2}(S')$. This, in turn, establishes the fact that $\gamma(S) > 0$.

By similar substitutions it is readily verified that at $C = C_L$, $\vec{\nu}(S) = \vec{\nu}_S(S')$ and $\vec{\lambda}_{W'}(W) = \vec{\lambda}(W')$. And since $\vec{\nu}(S)$ and $\vec{\lambda}_{W'}(W)$ are continuous functions of C it follows that if at C_L $\vec{\nu}_S(S') > \vec{0}$ and $\vec{\lambda}(W') < \vec{0}$ then for some interval below C_L $\vec{\nu}(S) > \vec{0}$ and $\vec{\lambda}_{W'}(W) < \vec{0}$. There remains only consideration of $\vec{\lambda}_J(W)$. By definition.

$$\vec{\lambda}_J(W) = \left[\vec{\pi}_{W'}\vec{V}_{W'J} + \vec{\pi}_J\vec{V}_{JJ} - \lambda_0^{1/2}(S)\vec{\psi}_S\vec{B}_{JS}^{\,t}\right]^2 - \lambda_0(S)\vec{c}_J\,,$$

and by Lemma 3 this may be expressed as

$$\vec{\lambda}_J(W) = \left\{\left[\left(\vec{\pi}_J - \vec{\pi}_{W'}\vec{B}'_{W'J}\right) - \lambda_0^{1/2}(S)\vec{\psi}_S\vec{R}'_{SJ}\right]\left(\vec{R}'_{JJ}\right)^{-1}\right\}^2 - \left[\lambda_0^{1/2}(S)\vec{\psi}_J\right]^2 .$$

Again by the hypothesis that at C_L, $\vec{\nu}_J(S') = \vec{0}$, it follows that at that point $\vec{\lambda}_J(W) = \vec{0}$ for the quantity in curly brackets is just $\lambda_0^{1/2}(S)\vec{\psi}_J$. Thus the absolute value of that quantity at C_L is $\lambda_0^{1/2}\vec{\psi}_J\vec{S}_J > \vec{0}$ where S_J is a diagonal matrix whose elements are s_j for $j \in J$. Suppose that $\vec{\lambda}_J(W) > 0$ for some interval on C below C_L then

$$\text{(55)} \qquad \left[\left(\vec{\pi}_J - \vec{\pi}_{W'}\vec{B}'_{W'J}\right) - \lambda_0^{1/2}(S)\vec{\psi}_S\vec{R}'_{SJ}\right]\left(\vec{R}'_{JJ}\right)^{-1}\vec{S}_J - \lambda_0^{1/2}(S)\vec{\psi}_J\vec{S}_J > \vec{0} ,$$

for, by continuity, the absolute value of the quantity in curly brackets above is itself times S_J near C_L. By hypothesis $\vec{\nu}_J(S') < \vec{0}$ for $C < C_L$, or by (44) and the definition of $\vec{G}_J(S')$

$$\text{(56)} \qquad \left[\left(\vec{\pi}_J - \vec{\pi}_{W'}\vec{B}'_{W'J}\right) - \lambda_0^{1/2}(S')\left(\vec{\psi}_S\vec{R}'_{SJ} + \vec{\psi}_J\vec{R}'_{JJ}\right)\right]\vec{S}_J < \vec{0} .$$

By replacing the decreasing function of C, $\lambda_0^{1/2}(S')$ by the similarly behaved function $\lambda_0^{1/2}(S)$ the inequality in (56) continues to hold. Multiplying (55) by the transpose of (56) thus results in a negative scalar and since $\vec{S}_J\vec{S}_J^{\,t} = \vec{I}$, the identity, one has a contradiction of the positive definiteness of $\left(\vec{R}'_{JJ}\right)^{-1}$, we conclude that for an interval below C_L, $\vec{\lambda}_J(W) \not> \vec{0}$, establishing (a).

Part (b) of the theorem is established in strictly analogous fashion.

Several comments on this result are in order. First observe that if for the interval $C_L \le C \le C_U$ the solution, $\vec{n}^0$, is given by Theorem 1 using the inputs $S' = S \cup J$ and signs s'_j where all the conditions of that theorem hold with strict inequality except at C_L (and C_U) where $\vec{\nu}_J(S') = \vec{0}$ and all these components are negative for smaller C's then for the new subset S and signs s_j, as per (a) of Theorem 3, all conditions of Theorem 1 are satisfied for a contiguous interval on C below C_L with one exception. This exception being that for $C < C_L$ the only statement regarding $\vec{\lambda}_J(W)$ is that not all its components are positive. A similar interpretation is immediate from (b) of this result. Second, observe that if $J = \{j^*\}$ say, then this theorem clearly yields the next meaningful inputs to Theorem 1 and thus the solution for the adjacent subinterval on C. This is precisely the special case singled out as Theorem 2 earlier.

Finally and more generally, Theorem 3 may be used as follows. Suppose that for some inputs to Theorem 1 the minimizing $\vec{n}$ is given for all $C \in [C_L, C_U]$, but at C_L several of the ν_i's and/or λ_i's are zero and are respectively less than zero and greater than zero for C's below C_L. One can then apply (a) and (b) of the theorem sequentially, resulting in a finite sequence of subsets of K and corresponding signs, the last of which must define the inputs yielding $\vec{n}^0$ for the interval $[C_{L'}, C_L]$ for some $C_{L'} < C_L$. That such a process terminates follows immediately from the theorem, for at each application the number of conditions of Theorem 1 which are violated for $C < C_L$ by the new subset and signs must be reduced by at least one.

Exact solutions under the degeneracy of the form discussed thus far may also be obtained by perturbation. For example, if for some solution interval several conditions on the ν_i's and/or λ_i's are met with equality at the lower endpoint C_L and violated below, then by randomly perturbing the π_i's, say, the degeneracy may be removed and Theorem 2 used to map out the solution for the perturbed problem. This solution will consist of a sequence of inputs to Theorem 1 defining the solution for a corresponding sequence of non-overlapping intervals on C. These inputs may be checked in sequence for the unperturbed problem, most often the inputs yielding the first interval of non-trivial length will provide the solution for the next interval below C_L for the original problem.

Finally, these same ideas can be used to find solutions when degeneracy in the form of zero π_i's occurs. Note that if some of the π_i's are zero all of the previous theory holds with the exception of the corollary to Theorem 1. Thus zero π_i's raise only one difficulty, viz., how to obtain a start. That is, the first interval, $[C_0, \infty]$, is not given by the preceding theory. As pointed out earlier, a solution in this case may be obtained arbitrarily close to the exact solution by perturbing the π_i's by adding to each a small positive ϵ_i. Moreover, such approximate solutions can be adjusted to yield exact results by checking the successive inputs to Theorem 1, obtained for the perturbed problem, for the unperturbed problem. Once some set of inputs yield the solution via Theorem 1 for the unperturbed problem for any non-empty interval on C, then the previous theory may be used to map out the exact solution for all other C. Note that although we have concentrated on going from one interval to the next on the left, the results above may also be used in reverse fashion to obtain the next adjacent interval to the right.

5. $v\left(\vec{n}^0\right)$ AS A FUNCTION OF C

Letting $\vec{n}^0(C)$ be the solution to the problem of minimizing (1) subject to (2) and (3), it is useful then to examine

$$v^*(C) \equiv v\left(\vec{n}^0(C)\right) . \tag{57}$$

Let $I_1, I_2, \dots, I_r$ be the sequence of non-overlapping intervals on C within each of which the solution vector $\vec{n}^0$ is given by Theorem 1 for some subset S_i of K and signs s_{ij} for $j \epsilon S_i$, $i = 1, \dots, r$. It is a known result in Lagrangian theory, easily verified directly here, that

$$\frac{dv^*(C)}{dC} = -\lambda_0\left(S_i\right) \quad \text{for} \quad C \epsilon I_i . \tag{58}$$

Further it follows from the preceding theory that for all i $\lambda_0(S_i)$ exists and is positive, hence v*(C) is a continuous decreasing function of C. Furthermore, its derivative, $-\lambda_0(S_i)$, is also a continuous function of C, for by (19) it is continuous within each I_i and by the result of Theorem 3 it is continuous at the endpoints of these intervals. Also from (58) and (19)

$$\frac{d^2 v^*(C)}{dC^2} = \frac{2\left\{\gamma\left(S_i\right)\right\}^2}{\left[C + \vec{\psi}_{S_i} R_{S_i S_i} \vec{\psi}^t_{S_i}\right]^3} > 0 \quad \text{for} \quad C \epsilon I_i ,$$

establishing convexity of v*(C). Finally it may be observed that as C increases v*(C) goes to zero, this follows immediately from Lemma 1. Summarizing these results we have:

LEMMA 4: $v^*(C) = v\left(\vec{n}^0(C)\right)$ is a continuous, strictly decreasing, convex function of C such that

a) $\lim_{C \to \infty} v^*(C) = 0$

and

b) v*(C) possesses a continuous first derivative given by

$$\frac{dv^*(C)}{dC} = -\lambda_0\left(S_i\right) = -\left\{\frac{\gamma\left(S_i\right)}{C + \vec{\psi}_S \vec{R}_{S_i S_i} \vec{\psi}^t_{S_i}}\right\}^2, \quad C \in I_i. \tag{59}$$

This lemma enables us to give a relatively simple expression for v*(C) as a function of $C \geq 0$.

THEOREM 4: For $C \in I_i$, $i = 1, \ldots, r$,

$$v^*(C) = K_i + \frac{\left[\gamma\left(S_i\right)\right]^2}{C + \vec{\psi}_{S_i} \vec{R}_{S_i S_i} \vec{\psi}^t_{S_i}} \tag{60}$$

where K_i is determined so that v*(C) is a continuous function of C for all $C \geq 0$. For the interval I_1 containing zero K_1 is determined so that $v^*(\vec{0}) = \vec{\pi}\vec{M}^{-1}\vec{\pi}^t$ or for the interval I_r ($C \geq C_0$) $K_r = 0$.

PROOF: Follows immediately from the lemma.

6. ACKNOWLEDGMENT

The author wishes to express his gratitude to Professor Howard Raiffa for suggesting this problem and for his guidance and encouragement.

REFERENCES

[1] Anderson, T. W., Introduction to Multivariate Statistical Analysis (John Wiley and Sons, Inc., New York, 1958).

[2] Bellman, R., Introduction to Matrix Analysis (McGraw-Hill Book Co., Inc., New York, 1960).

[3] Dwyer, P. S., "Some Applications of Matrix Derivatives in Multivariate Analysis," J. Am. Statistical Assn. 62, 607 - 625 (June 1967).

[4] Ericson, W. A., "Optimum Stratified Sampling Using Prior Information," J. Am. Statistical Assn. 60, 750 - 771 (Sept. 1965).

[5] Ericson, W. A., "On the Economic Choice of Experiment Sizes for Decision Regarding Certain Linear Combinations," J. Roy. Statistical Soc. (B), Part III (1967).

[6] Ericson, W. A., "Optimum Allocation in Stratified and Multistage Samples Using Prior Information," to appear in the J. Am. Statistical Assn.

[7] Kuhn, H. W. and A. W. Tucker, "Non-Linear Programming," Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability (Ed. J. Neyman) (University of California Press, Berkeley, 1950).

[8] Raiffa, H. and R. O. Schlaifer, Applied Statistical Decision Theory (Division of Research, Harvard Business School, Boston, 1961).

[9] Vajda, S., Mathematical Programming (Addison-Wesley Publishing Co., Inc., Reading, Mass., 1961).

* * *