ORC 64-12 (RR)
22 JUNE 1964

NONLINEAR PROGRAMS WITH POSITIVELY BOUNDED JACOBIANS

by

Richard W. Cottle

OPERATIONS RESEARCH CENTER

INSTITUTE OF ENGINEERING RESEARCH

UNIVERSITY OF CALIFORNIA - BERKELEY

NONLINEAR PROGRAMS WITH POSITIVELY BOUNDED JACOBIANS

by

Richard W. Cottle
Operations Research Center
University of California, Berkeley

22 June 1964 ORC 64-12 (RR)

This research has been partially supported by the Office of Naval Research under Contract Nonr-222(83) and the National Science Foundation under Grant GP-2633 with the University of California. Reproduction in whole or in part is permitted for any purpose of the United States Government.

CHAPTER I

INTRODUCTION

I-1. The Mathematical Programming Problem

This report is concerned with properties and solution methods for inequality-constrained extremization problems. Generally, the problems can be put as follows:

(1) minimize $f_0(x)$ constrained by $f_i(x) \leq 0,\ 1 \leq i \leq m$,

where the $f_i (0 \leq i \leq m)$ are real-valued functions defined on real n-space.

Problem (1) is referred to in the literature as a <u>mathematical program.</u> In particular, (1) is a <u>linear program</u> if all the f_i are linear forms; otherwise it is a <u>nonlinear program.</u> The simplest kind of nonlinear program of interest is the <u>quadratic program,</u> in which the minimand, f_0 , is a convex quadratic function and the other f_i are linear forms. When all the functions in (1) are convex, it is called a <u>convex program</u>. A linear program may be regarded as a special type of quadratic program; both are examples of convex programs.

Differentiability of the f_i is not part of the statement of (1), although it has proved to be a useful assumption in many theoretical and computational investigations. Among studies in this field, the special case of convex programming has received the most attention, probably because the local extrema in such a problem are always global. In all the problems considered here, the extremand and the constraint functions will be assumed to possess at least continuous first partial derivatives; however, not all of them will be convex programs.

I-2. Results Presented in This Report

The mathematical programming problem is a generalization of the classical problem:

(2) minimize $f_0(x)$ constrained by $f_i(x) = 0, 1 \leq i \leq m$

which is usually handled by the method of Lagrange multipliers. This is always possible when the mapping $f = (f_1,\ldots,f_m)$ satisfies a regularity condition [7, p.198].

The extension of the Lagrange multiplier approach to the mathematical programming problem was first accomplished by John [28] who established necessary and sufficient conditions for a solution. Later, Kuhn and Tucker [32] presented their basic work giving necessary and sufficient conditions for a nonnegative saddle-point of a differentiable function and their relation to the mathematical programming problem. They obtained necessary conditions for a solution to (1) by introducing a regularity assumption which made it possible to apply the Minkowski-Farkas Lemma [18, p.5] and thereby produce multipliers like John's but with a leading multiplier of unity. John's Theorem and the Kuhn-Tucker Theorem on necessary conditions for a solution to (1) are discussed in Chapter II.

The well-known duality theory of linear programming was formulated in 1947 by von Neumann [34] in an unpublished manuscript (see[8, p.125]) and developed by Gale, Kuhn, and Tucker [23]. After a gap of several years, the notion of duality was extended to quadratic programming [12], [13], and to convex programming [42], [27]. For quadratic programming, the proof of the duality theorem can be obtained by applying the duality theorem of linear programming, whereas duality in convex programming has usually been handled by an application of the Kuhn-Tucker Theorem. The dual of a quadratic or convex program, as originally presented, differs from it so conspicuously that the pair of problems appears to lack the beautiful symmetry of the analogous problems in von Neumann's formulation for the linear case. In Chapter III, a symmetric formulation of duality for nonlinear programming is presented which includes those above as special cases. Part of this chapter is based on the author's paper [6] on "Symmetric dual quadratic programs." The extension of this idea to the nonlinear case is from a joint work [11] with G. B. Dantzig and E. Eisenberg.

The study of a dual pair of problems often leads to the consideration of a single system, the composite of the two. In Chapter IV, the composite problem is viewed as a special case of the general program

(3) minimize $zW(z)$ constrained by $W(z) \geq 0$, $z \geq 0$

where W is a mapping of N-space into itself. Necessary and sufficient conditions will be given for a solution to (3) when W is a differentiable mapping with a suitably restricted Jacobian matrix. Several important properties of the Jacobian matrix of the mapping W in (3) are left invariant by an operation known as principal pivoting which is an exchange of the dependence roles of certain variables. The positivity of the determinant of a (not-necessarily symmetric) positive definite matrix turns out to be a simple consequence of the results on principal pivoting.

Finally, the concept of a positively bounded Jacobian matrix is presented in Chapter V in response to the need for a sufficient condition to guarantee the existence of a solution to the problem (3). The existence theorem provides at the same time an algorithm in those cases where the functions and values it demands can be computed and actually inspires the algorithm of G. B. Dantzig and the author [9] for solving (3) when $W(z) = Mz + q$ and M is a positive semi-definite matrix. A new minimax theorem follows from the main existence theorem.

I-3. Notations and Terminology

There is a fairly standard vocabulary in mathematical programming which facilitates discussion; we shall use it freely here. Given the problem (1), the minimand, f_0, is called the objective function. A vector x is feasible for (1) if it satisfies the side conditions (or constraints) $f_i(x) \leq 0$, $1 \leq i \leq m$. The constraint set of the problem is the set of its feasible vectors. This may be empty, in which case (1) is infeasible. A feasible vector is optimal if it solves the problem.

All numerical quantities considered here belong to the reals, denoted R. Vectors will belong to finite-dimensional real vector spaces, R^n, and whether they are to be regarded as rows or columns will always be clear from the context in which they appear. Thus, for example, the expressions

$$x = (x_1, \ldots, x_n) \tag{4}$$

$$Ax = b \tag{5}$$

$$xy = \sum_{i=1}^{n} x_i y_i \tag{6}$$

are easily understood. In (4), x is a row vector, while in (5) it is a column vector. Equation (6)--in which x is a row and y is a column--defines the symbol xy. In short, no special notational provisions will be made for transposing vectors. We may not treat matrices so informally.

The transpose of a matrix A will be denoted A' .

Vector inequalities will be used extensively. Let $x = (x_1,\ldots,x_n)$ and $y = (y_1,\ldots,y_n)$. Then

$$x \geq y \quad \text{if} \quad x_i \geq y_i \ , \quad 1 \leq i \leq n$$

$$x > y \quad \text{if} \quad x_i > y_i \ , \quad 1 \leq i \leq n .$$

The reverse inequalities $\leq$ and $<$ are defined analogously. The same symbol, 0 , will be used to denote the zero vector and the ordinary scalar; no confusion should result from this. A vector x is called <u>nonnegative</u> or <u>positive</u> according as $x \geq 0$ or $x > 0$. A nonnegative vector which is not 0 is called <u>semi-positive</u>. The <u>nonnegative orthant</u> in R^n is the set R^n_+ consisting of all its nonnegative vectors.

If f is a differentiable real-valued function on an open subset of R^n, the <u>gradient</u> of f will be denoted

$$\nabla f = \left(\frac{\partial f}{\partial x_1},\ldots,\frac{\partial f}{\partial x_n}\right) ,$$

and if it is twice differentiable, the <u>Hessian</u> of f will be denoted

$$\nabla^2 f = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right) .$$

In dealing with differentiable functions of two vector arguments, it will be necessary to use partial gradients; if $F(x,y)$ is such a function,

$$\nabla_1 F = \left(\frac{\partial F}{\partial x_1},\ldots,\frac{\partial F}{\partial x_n}\right)$$

$$\nabla_2 F = \left(\frac{\partial F}{\partial y_1},\ldots,\frac{\partial F}{\partial y_m}\right) .$$

Also, for twice differentiable functions $F(x,y)$:

$$\nabla_{11} F = \left(\frac{\partial^2 F}{\partial x_i \partial x_j}\right) , \qquad (n \times n)$$

$$\nabla_{12}F = \left(\frac{\partial^2 F}{\partial x_i \partial y_j} \right), \qquad (n \times m)$$

$$\nabla_{21}F = \left(\frac{\partial^2 F}{\partial y_i \partial x_j} \right), \qquad (m \times n)$$

$$\nabla_{22}F = \left(\frac{\partial^2 F}{\partial y_i \partial y_j} \right), \qquad (m \times m) .$$

A square matrix M of order r will be called <u>positive semi-definite</u> if

(7) $$xMx \geq 0 \quad \text{all} \quad x \in R^n .$$

When equality in (7) holds only if $x = 0$, M is <u>positive definite</u>. It is <u>negative semi-definite</u> (<u>definite</u>) if, and only if, $-M$ is positive semi-definite (definite).

Although it is true that if

$$\tilde{M} = (1/2)(M + M'),$$

then $\tilde{M}$ is symmetric and

$$x\tilde{M}x = xMx \quad \text{all} \quad x \in R^n ,$$

it will not be possible to replace M by $\tilde{M}$ because M will represent a nonsymmetric Jacobian matrix needed elsewhere.

Finally, a word about the reference system. The manner in which chapters and their sections are numbered should be clear at this point. Equations are numbered consecutively within each chapter. Definitions, remarks, theorems, and the like are treated as equations. A reference to an equation outside a given chapter will be made by prefixing the chapter number to the equation number. This rule will not be followed when referring to an equation in the same chapter. In line with current practice, numbers in square brackets refer to books and papers listed at the end of the report.

CHAPTER II

NECESSARY CONDITIONS OF OPTIMALITY

II-1. John's Theorem

The earliest result on the necessary conditions of optimality in the problem I-(1) seems to be that of John [28, Theorem 1]. He was concerned with the problem

(1) minimize $f_0(x)$ constrained by $f(\sigma,x) \leq 0$, $(\sigma,x) \in S \times X$

where S is a compact metric space and X is a subset of R^n ; the partial derivatives of f_0 and f with respect to each component of x are assumed to be continuous on X and $S \times X$, respectively.

When $X = R^n$ and S is the compact metric space $\{1,\ldots,m\}$, I-(1) can be viewed as a special case of II-(1). It is convenient to give John's proof for the necessary conditions of optimality in (1), specializing it to the mathematical programming problem. It will rest partly on the following statement, which is due to Gordan [25].

(2) <u>Theorem.</u> Let A be an $m \times n$ matrix. Exactly one of the following alternatives holds. Either

$$uA = 0$$

has a semi-positive solution or

$$Av < 0$$

has a solution.

Gordan's Theorem, which may more conveniently be found in [8, p.136] and [21, p.48], pre-dates, by about twenty-five years, the more widely known Minkowski-Farkas Lemma [18, p.5].

The proposition we want is

(3) Theorem (John). Let $\hat{x}$ solve the problem

(4) minimize $f_0(x)$ constrained by $f_i(x) \leq 0$, $1 \leq i \leq m$.

Then there exists a semi-positive vector $(\hat{u}_0, \hat{u}_1, \ldots, \hat{u}_m)$ such that

(5) $$\hat{u}_i f_i(\hat{x}) = 0, \; 1 \leq i \leq m$$

and the function

(6) $$\Phi(x) = \sum_{i=0}^{m} \hat{u}_i f_i(x)$$

has a critical point at $\hat{x}$, i.e.,

(7) $$\nabla \Phi(\hat{x}) = \sum_{i=0}^{m} \hat{u}_i \nabla f_i(\hat{x}) = 0 \; .$$

Proof. Define the sets[1]

(8) $$M = \{1, \ldots, m\}$$

and

(9) $$Z = \{i \mid i \in M, \; f_i(\hat{x}) = 0\} \; .$$

If Z is empty, let $\hat{u}_0 = 1$ and $\hat{u}_i = 0$ for all $i \in M$. Indeed, $\hat{x}$ must belong to the interior of the constraint set, and the vanishing of the gradient, ∇f_0, at the unconstrained local minimum implies (7). If Z is nonempty, it will suffice to show that there is no solution v to the system

(10) $$\nabla f_0(\hat{x}) v < 0$$

$$\nabla f_i(\hat{x}) v < 0 \; , \qquad i \in Z .$$

For if the system has no solution, then defining $\hat{u}_i = 0$ for all indices $i \in M - Z$, there exists, by Gordan's Theorem, a semi-positive vector $(\hat{u}_0, \hat{u}_1, \ldots, \hat{u}_m)$ satisfying (7). This vector will satisfy (5) by the definitions made above.

1 The sets M and Z will be used again.

Now, on the contrary, suppose (10) had a solution $\hat{v}$. Let $S(\hat{x},\epsilon)$ denote the closed ball with center $\hat{x}$ and radius $\epsilon > 0$. There exists an $\epsilon > 0$ and a $\delta > 0$ such that for all $x \in S(\hat{x},\epsilon)$

$$\nabla f_0(x)\hat{v} + \delta \leq 0$$

$$\nabla f_i(x)\hat{v} + \delta \leq 0, \qquad i \in \mathcal{Z} .$$

For some positive γ,

$$f_i(\hat{x}) + \gamma \leq 0, \qquad i \in \mathcal{M} - \mathcal{Z} .$$

Choose $\alpha > 0$ satisfying

$$\alpha\Big(\sum_{j=1}^{n} \hat{v}_j^2\Big)^{1/2} < \epsilon$$

$$\alpha\max\{|\nabla f_i(x)\hat{v}| \quad i \in \mathcal{M}, \quad x \in S(\hat{x},\epsilon)\} < \gamma .$$

The compactness of $S(\hat{x},\epsilon)$ and the continuity of the partial derivatives of the f_i imply the existence of the above maximum. For some θ_i, $0 < \theta_i < 1$,

$$f_0(\hat{x} + \alpha\hat{v}) = f_0(\hat{x}) + \alpha\nabla f_0(\hat{x} + \theta_0\alpha\hat{v})\hat{v}$$

$$f_i(\hat{x} + \alpha\hat{v}) = f_i(\hat{x}) + \alpha\nabla f_i(\hat{x} + \theta_i\alpha\hat{v})\hat{v} , \qquad i \in \mathcal{M} .$$

Hence

$$f_0(\hat{x} + \alpha\hat{v}) \leq f_0(\hat{x}) - \alpha\delta < f_0(\hat{x})$$

$$f_i(\hat{x} + \alpha\hat{v}) \leq f_i(\hat{x}) - \alpha\delta < 0 , \qquad i \in \mathcal{Z}$$

$$f_i(\hat{x} + \alpha\hat{v}) \leq -\gamma + \alpha|\nabla f_i(\hat{x} + \theta_i\alpha\hat{v})\hat{v}| < 0 , \quad i \in \mathcal{M} - \mathcal{Z} .$$

These last inequalities contradict the optimality of $\hat{x}$ in (4), and so (10) can have no solution. As noted earlier, this implies (7).

(11) Remark. if $\hat{u}_0$ is positive, it may be assumed to equal 1. If the vectors

$\nabla f_i(\hat{x})$, $i \in \mathcal{Z}$, are positively independent [21, p. 62], i.e., linearly independent over the nonnegative orthant, it is clear from the theorem that $\hat{u}_0$ cannot equal 0.

It will be useful to record the conclusion of the theorem when x is required to be nonnegative.

(12) Corollary. Let $\hat{x}$ be an optimal solution of the program

(13) minimize $f_0(x)$ constrained by $x \geq 0$, $f_i(x) \leq 0$, $1 \leq i \leq m$.

Then there exists a semi-positive vector $(\hat{u}_0, \hat{u}_1, \ldots, \hat{u}_m)$ such that

(14) $$\hat{u}_i f_i(\hat{x}) = 0, \qquad 1 \leq i \leq m$$

(15) $$\sum_{i=0}^{m} \hat{u}_i \nabla f_i(\hat{x}) \geq 0$$

(16) $$[\sum_{i=0}^{m} \hat{u}_i \nabla f_i(\hat{x})]_j \hat{x}_j = 0, \qquad 1 \leq j \leq n.$$

If $f: R^n \to R^m$ is a differentiable mapping with component functions $f_1, \ldots, f_m$, let J_f be its $m \times n$ Jacobian matrix:

(17) $$J_f = (\frac{\partial f_i}{\partial x_j}) .$$

The following consequence of John's Theorem guarantees the existence of multipliers in (14) - (16) with $\hat{u}_0 = 1$.

(18) Corollary. Let $\hat{x}$ be an optimal solution of (13). Suppose there are p columns of $J_f(\hat{x})$, say $j_1, \ldots, j_p$, such that the system of inequalities

(19) $$u(\frac{\partial f_i}{\partial x_{j_k}}(\hat{x})) \geq 0, \qquad u \geq 0$$

has only the trivial solution $u = 0$. Then there exists a semi-positive vector $(\hat{u}_0, \hat{u}_1, \ldots, \hat{u}_m)$ satisfying (14) - (16), and such that $\hat{u}_0 = 1$.

Proof. Apart from the stipulation that $\hat{u}_0 = 1$, such a vector exists by the previous corollary. If $\hat{u}_0 = 0$, then the vector $(\hat{u}_1, \ldots, \hat{u}_m)$ is semi-positive. This and (15) imply that (19) has a nontrivial solution, contrary to our assumption. Hence $\hat{u}_0 > 0$, and therefore may be taken to equal 1.

II-2. The Kuhn-Tucker Theorem

The Kuhn-Tucker Theorem, like John's Theorem, is concerned with necessary conditions of optimality in the mathematical programming problem. It includes, however, an extra hypothesis which has the effect of assuring the existence of multipliers with $\hat{u}_0 = 1$.

The existence of the multipliers in John's Theorem, above, followed, via Gordan's Theorem, from the fact that a system of linear inequalities, (10), had no solution. Suppose we had been able to show that the less restrictive system

(20)
$$\nabla f_0(\hat{x})v < 0$$
$$\nabla f_i(\hat{x})v \leq 0, \qquad i \in \mathcal{Z}$$

also has no solution. By the Minkowski-Farkas Lemma, there would exist a vector $(\hat{u}_1, \ldots, \hat{u}_m)$ such that

$$\nabla f_0(\hat{x}) + \sum_{i=1}^{m} \hat{u}_i \nabla f_i(\hat{x}) = 0$$

and, by definition,

$$\hat{u}_i f_i(\hat{x}) = 0, \qquad i \in \mathcal{M} .$$

Thus, $(1, \hat{u}_1, \ldots, \hat{u}_m)$ would be semi-positive and would correspond to the vector of multipliers in John's Theorem. The difficulty lies in the fact that it is not always possible to show that (20) has no solution. The example [32, p.484] illustrates this. Let $f_0(x) = -x_1$, $f_1(x) = (x_1 - 1)^3 + x_2$, $f_2(x) = -x_1$, and $f_3(x) = -x_2$. It turns out that $\hat{x} = (1,0)$ is the unique optimal solution to problem (6), but $\hat{v} = (1,0)$ satisfies (20) where we note that $\mathcal{Z} = \{1,3\}$.

Some regularity condition must be imposed if we wish to assert that (20) has no solution when $\hat{x}$ is optimal.

(21) Definition. Let $\bar{x}$ be a boundary point of the constraint set

$$C_f = \{x \mid x \in R^n,\ f(x) \leq 0\}$$

where $f: R^n \to R^m$. In this case, the set $\mathcal{Z} = \{i \mid i \in \mathcal{M},\ f_i(\bar{x}) = 0\}$ is nonempty. The (Kuhn-Tucker) constraint qualification is satisfied at $\bar{x}$ if for every vector v satisfying the system of homogeneous linear inequalities

(22) $$\nabla f_i(\bar{x})v \leq 0, \quad i \in \mathcal{Z}$$

there exists a continuously differentiable arc $\alpha:[0,1] \to C_f$ such that $\alpha(0) = \bar{x}$, and $\alpha'(0) = \lambda v$ for some $\lambda > 0$. (See [32, p.483].)

It is easily shown, along the lines of [12, p.136], that the constraint qualification is always satisfied at the boundary points of C_f when f is composed of linear forms. Arrow and Hurwicz [1, p.2] make the interesting observation that the constraint qualification (21) is a property of f rather than of C_f. They show that two distinct mappings can induce the same constraint set and yet differ with respect to satisfaction of the constraint qualification.

(23) Theorem (Kuhn and Tucker). Let $\hat{x}$ be an optimal solution of (6). If $\hat{x}$ belongs to the boundary of C_f, assume that the constraint qualification is satisfied at $\hat{x}$; then there exists a nonnegative vector

$\hat{u} = (\hat{u}_1, \ldots, \hat{u}_m)$ such that

(24) $$\nabla f_0(\hat{x}) + \sum_{i=1}^{m} \hat{u}_i \nabla f_i(\hat{x}) = 0$$

and

(25) $$\hat{u}_i f_i(\hat{x}) = 0, \quad i \in \mathcal{M} .$$

Proof. If $\hat{x}$ is an interior point of C_f, we may let $\hat{u} = 0$, as in the proof of (3). If $\hat{x}$ is a boundary point of C_f, the optimality of $\hat{x}$ and the constraint qualification imply that (20) has no solution. For otherwise we may take α (and λ) as in (21) and deduce that

$$\frac{d}{d\theta}\Big(f_0(\alpha(0))\Big) = \nabla f_0(\alpha(0))\alpha'(0) = \nabla f_0(\hat{x})\lambda v < 0 .$$

Since α maps the unit interval into C_f, this inequality implies that there exist points x of C_f such that $f_0(x) < f_0(\hat{x})$ which contradicts the optimality of $\hat{x}$. Therefore (20) has no solution, the Minkowski-Farkas Lemma may be applied, and the required vector exists.

The equations (24) and (25) are usually called the <u>Kuhn-Tucker conditions</u> for (6). When (6) is a convex program, (24) and (25) for some $\hat{u} \geq 0$ are sufficient conditions for optimality, regardless of any constraint qualification. (See [32, p.485].)

In some cases, the Kuhn-Tucker constraint qualification may be difficult to verify. But satisfaction of this regularity condition is merely a means to an end, namely, the Kuhn-Tucker conditions. We shall say that a differentiable mapping $f: R^n \to R^m$ is <u>Lagrange regular</u> at $\hat{x}$ if for some nonnegative $\hat{u}$, (24) and (25) are the necessary conditions of optimality in the program (6). This term has been used by Arrow, Hurwicz, and Uzawa [2, p.181] in their study of the interrelationships between some of the various constraint qualifications which have been proposed as substitutes for that of Kuhn and Tucker.

The property assumed in the corollary (18) made the mapping Lagrange regular at the point $\hat{x}$. All the programs we consider in this paper have either linear constraints or else constraints whose Jacobian matrix at the optimal solution has the property (19); in either case, Lagrange regularity is at hand.

We note, in passing, that the necessary conditions of optimality in (4) reduce to the Kuhn-Tucker conditions when the multiplier $\hat{u}_0$ is positive. For a related study, in which the positivity of the multiplier associated with the objective function is crucial, see [4, p.227].

Finally, let us point out why the Kuhn-Tucker conditions are preferable to those of John. The answer lies in the Equivalence Theorem [32, p.486]. When (13) is a convex program for which the constraint qualification (21) is satisfied, $\hat{x}$ is an optimal solution if, and only if, there exists a nonnegative vector $\hat{u}$ such that $(\hat{x}, \hat{u})$ provides a nonnegative saddle-point for the <u>Lagrangian function</u>

$$\Phi(x,u) = f_0(x) + \sum_{i=1}^{m} u_i f_i(x) .$$

CHAPTER III

SYMMETRIC DUAL PROGRAMS

III-1. Duality in Linear Programming

In order to put our subject in perspective, we recall briefly the symmetric duality theory of linear programming. This involves a pair of problems such as[1]

(1) minimize cx constrained by $Ax + b \geq 0,\ x \geq 0$

and

(2) maximize $-by$ constrained by $-A'y + c \geq 0,\ y \geq 0$

where A is an $m \times n$ matrix, b is an m-vector, and c is an n-vector; all three are constants, whereas x and y represent vector variables.

The symmetry of this pair stems from the fact that negation and transposition are involutory operations. (1) is called the <u>primal problem</u> and (2), the <u>dual problem</u>. This terminology, due to von Neumann, is somewhat arbitrary but is traditional and will be used here.

Letting $\mathcal{P}$ and $\mathcal{D}$ denote the constraint sets of (1) and (2), respectively, the major theorems on duality in linear programming are:

(3) <u>Weak Duality Theorem.</u> $\sup_{\mathcal{D}} -by \leq \inf_{\mathcal{P}} cx$

(4) <u>Duality Theorem.</u> If either program in the dual pair has an optimal solution, then so does the other, and when this is so,

$$\max_{\mathcal{D}} -by = \min_{\mathcal{P}} cx .$$

---

1 The reader already familiar with duality in linear programming will detect that (1) and (2) are an equivalent, but unorthodox, statement of the dual pair. The motivation for this will become clear later.

(5) Complementary Slackness Theorem. $\hat{x} \in \mathcal{P}$ and $\hat{y} \in \mathcal{D}$ are optimal solutions of the primal and dual problems, respectively, if, and only if,

(6) $$\hat{y}_i(A\hat{x} + b)_i = 0, \quad 1 \leq i \leq m$$

(7) $$\hat{x}_j(-A'\hat{y} + c)_j = 0, \quad 1 \leq j \leq n \; .$$

(8) Existence Theorem. If $\mathcal{P}$ and $\mathcal{D}$ are nonempty, then both problems have optimal solutions.

(9) Unboundedness Theorem. If exactly one of $\mathcal{P}$ and $\mathcal{D}$ is nonempty, either $\sup \; -by = +\infty$ or $\inf \; cx = -\infty$ according as $\mathcal{P}$ or $\mathcal{D}$ is empty.

III-2. Symmetric Dual Nonlinear Programs

Since 1959, the notion of duality has been extended to quadratic programming by Dennis [12] and Dorn [13] and to convex programming by Wolfe [42], Huard [27], Mangasarian [33], and others. In each case, the dual pair of programs lacks symmetry.

The object of this section is to present a treatment of symmetric dual programs [11] which will cover (1) and (2) and for which analogues of at least (3), (4), and (5) can be proved. In the case of quadratic programming, all five of the corresponding theorems are true. In the general nonlinear case, (9) is false.

Suppose $K: R^n_+ \times R^m_+ \to R$ is a continuously differentiable function and consider the programs[2]:

(10) $$\text{minimize} \quad F(x,y) = K(x,y) - y\,\nabla_2 K(x,y)$$

$$\text{constrained by} \quad \nabla_2 K(x,y) \leq 0$$

$$(x,y) \geq 0$$

---

2 In (10) only x (in (11) only y) need be nonnegative.

(11) $$\text{maximize } G(x,y) = K(x,y) - x\nabla_1 K(x,y)$$
$$\text{constrained by } \nabla_1 K(x,y) \geq 0$$
$$(x,y) \geq 0$$

designated primal and dual, respectively. Neither one is necessarily a convex program. Under some additional hypotheses, they are a dual pair, and for suitable choices of K, the dual pairs previously mentioned are special cases.

As in III-1, we denote the constraint sets of these problems by $\mathcal{P}$ and $\mathcal{D}$.

Let $X \subseteq R^n$ and $Y \subseteq R^m$ be nonempty convex sets. A function $K:X\times Y \to R$ will be called convex-concave [38] if $K(\ ,y)$ is a convex function on X for each $y \in Y$ and $K(x,\ )$ is a concave function on Y for each $x \in X$.

(12) Weak Duality Theorem. Let $K:R^n_+ \times R^m_+ \to R$ be a differentiable convex-concave function. Then

(13) $$\sup_{\mathcal{D}} G(x,y) \leq \inf_{\mathcal{P}} F(x,y)$$

Proof. Let $\emptyset$ denote the empty set. We adopt the convention that

(14) $\sup G(x,y) = -\infty$ if $\mathcal{D} = \emptyset$, $\inf F(x,y) = +\infty$ if $\mathcal{P} = \emptyset$.

It therefore suffices to assume that both $\mathcal{P}$ and $\mathcal{D}$ are nonempty. Let $(x,y) \in \mathcal{P}$ and $(\tilde{x},\tilde{y}) \in \mathcal{D}$. Since K is a differentiable convex-concave function, we have [29, p. 405]

(15) $$K(x,\tilde{y}) - K(\tilde{x},\tilde{y}) \geq (x - \tilde{x})\nabla_1 K(\tilde{x},\tilde{y})$$

(16) $$K(x,\tilde{y}) - K(x,y) \leq (\tilde{y} - y)\nabla_2 K(x,y).$$

Subtracting (16) from (15) and rearranging terms, we get

(17) $$F(x,y) - G(\tilde{x},\tilde{y}) \geq x\nabla_1 K(\tilde{x},\tilde{y}) - \tilde{y}\nabla_2 K(x,y) \geq 0.$$

(18) Remark. The obvious significance of the weak duality theorem is that when K is differentiable and convex-concave, $(\bar{x},\bar{y}) \in \mathcal{P}$ and $(\tilde{x},\tilde{y}) \in \mathcal{D}$ would be optimal solutions if $F(x,y)$ and $G(\tilde{x},\tilde{y})$ were equal. We note from (17) that[3] if this were the case, the following would hold:

$$\tilde{y}_i(\nabla_2 K(x,y))_i = 0, \quad 1 \leq i \leq m$$

$$x_j(\nabla_1 K(\tilde{x},\tilde{y}))_j = 0, \quad 1 \leq j \leq n.$$

(19) Duality Theorem. Let $K: R^n_+ x R^m_+ \to R$ be a twice continuously differentiable function. If $(\hat{x},\hat{y})$ is an optimal solution of the primal problem, (10), and $\nabla_{22}K(\hat{x},\hat{y})$ is negative definite, then $(\hat{x},\hat{y}) \in \mathcal{D}$ and $F(\hat{x},\hat{y}) = G(\hat{x},\hat{y})$. If K is convex-concave, $(\hat{x},\hat{y})$ solves the dual problem, (11).

Proof. The constraints of the primal problem are Lagrange regular at $(\hat{x},\hat{y})$ because the hypotheses of II-(18) are satisfied. There exists an m-vector $\hat{v} \geq 0$ such that $\hat{v}\nabla_2 K(\hat{x},\hat{y}) = 0$ and

$$\nabla_1 F(\hat{x},\hat{y}) + \nabla_{12}K(\hat{x},\hat{y})\hat{v} \geq 0$$

$$\hat{x}[\nabla_1 F(\hat{x},\hat{y}) + \nabla_{12}K(\hat{x},\hat{y})\hat{v}] = 0$$

$$\nabla_2 F(\hat{x},\hat{y}) + \nabla_{22}K(\hat{x},\hat{y})\hat{v} \geq 0$$

$$\hat{y}[\nabla_2 F(\hat{x},\hat{y}) + \nabla_{22}K(\hat{x},\hat{y})\hat{v}] = 0.$$

When F is replaced by its definition, these reduce to

(20) $$\nabla_1 K(\hat{x},\hat{y}) + \nabla_{12}K(\hat{x},\hat{y})(\hat{v} - \hat{y}) \geq 0$$

(21) $$\hat{x}\nabla_1 K(\hat{x},\hat{y}) + \hat{x}\nabla_{12}K(\hat{x},\hat{y})(\hat{v} - \hat{y}) = 0$$

---

[3] If the scalar product of two nonnegative vectors is zero, then each of the summands in the expression is zero.

(22) $$\nabla_{22}K(\hat{x},\hat{y})(\hat{v} - \hat{y}) \geq 0$$

(23) $$\hat{y}\nabla_{22}K(\hat{x},\hat{y})(\hat{v} - \hat{y}) = 0$$

It follows from $\hat{v} \geq 0$, (22), and (23) that

$$(\hat{v} - \hat{y})\nabla_{22}K(\hat{x},\hat{y})(\hat{v} - \hat{y}) \geq 0.$$

This inequality and the negative definiteness of $\nabla_{22}K(\hat{x},\hat{y})$ imply $\hat{v} = \hat{y}$. From (20)and (21), we now get

(24) $$\nabla_1 K(\hat{x},\hat{y}) \geq 0$$

(25) $$\hat{x}\nabla_1 K(\hat{x},\hat{y}) = 0.$$

Since $\hat{v} = \hat{y}$ and $\hat{v}\nabla_2 K(\hat{x},\hat{y}) = 0$,

(26) $$\hat{y}\nabla_2 K(\hat{x},\hat{y}) = 0.$$

Consequently, $(\hat{x},\hat{y}) \in \mathcal{D}$ and $F(\hat{x},\hat{y}) = G(\hat{x},\hat{y}) = K(\hat{x},\hat{y})$. When K is convex-concave, $(\hat{x},\hat{y})$ solves the dual problem by the weak duality theorem.

(27). Remark. A similar result obtains when $(\hat{x},\hat{y})$ solves the dual problem and $\nabla_{11}K(\hat{x},\hat{y})$ is positive definite. In both theorems, the *complementary slackness conditions* are visible in equations such as (25) and (26).

For suitable choices of K, the symmetric programs (10) and (11) reflect the *form* of certain known dual pairs of the programming problems. For instance if

(28) $$K(x,y) = cx - by - yAx$$

the symmetric dual linear programs (1) and (2) result from the operations indicated by (10) and (11), respectively. For this definition of K, the hypotheses of the weak duality theorem are satisfied. On the other hand, the definiteness assumption of our duality theorem (19) cannot be satisfied by such a function. However, the contribution of the linearity of K in each of its variables is sufficient to allow weaker hypotheses for the duality theorem.

Likewise, the primal convex program

$$\text{(29)} \qquad \begin{aligned} &\text{minimize} && f_0(x) \\ &\text{constrained by} && f_i(x) \leq 0, \quad 1 \leq i \leq m \\ &&& x \geq 0 \end{aligned}$$

has the dual, according to Wolfe [42],

$$\text{(30)} \qquad \begin{aligned} &\text{maximize} && f_0(x) + \sum_{i=1}^{m} y_i f_i(x) - \sum_{j=1}^{n} x_j [\nabla f_0(x) + \sum_{i=1}^{m} y_i \nabla f_i(x)]_j \\ &\text{constrained by} && \nabla f_0(x) + \sum_{i=1}^{m} y_i \nabla f_i(x) \geq 0 \\ &&& y \geq 0. \end{aligned}$$

Using (10) and (11), this pair of programs can be obtained from the differentiable convex-concave function

$$\text{(31)} \qquad K(x,y) = f_0(x) + \sum_{i=1}^{m} y_i f_i(x).$$

Again, the derivation of the dual pair (29), (30) in this manner is formal. The proof that if $\hat{x}$ is an optimal solution of (29), then there exists a $\hat{y}$ such that $(\hat{x},\hat{y})$ is an optimal solution of (30) and such that the extremal values of the objective functions are equal can be obtained from the Kuhn-Tucker Theorem, in which the constraint qualification must be assumed. Strong assumptions, like those of (19), are made by Huard [27] and Mangasarian [33] in the converse proof of the duality theorem.[4]

III-3. Symmetric Dual Quadratic Programs

Let A, b, and c have the meanings assigned in III-1; let D and E be symmetric, positive semi-definite matrices of orders n and m, respectively. The function K defined by

$$K(x,y) = \tfrac{1}{2}xDx + cx - \tfrac{1}{2}yEy - by - yAx \tag{32}$$

induces symmetric dual quadratic programs [6] in (10) and (11). When D and E are both zero matrices, the function K in (32) becomes identical to that defined in (28), and the related problems reduce to symmetric dual linear programs. When D or E is a zero matrix, one gets Dorn's dual quadratic programs [13]. See Dorn's paper [14] for an earlier effort toward symmetrizing quadratic programming.

The function defined in (32) is convex-concave [16, p. 51] and has continuous partial derivatives of all orders. Therefore the weak duality theorem holds for this K. However, $\nabla_{11}K(x,y) = D$ and $\nabla_{22}K(x,y) = -E$; these matrices are positive and negative semi-definite, respectively, and therefore the duality theorem (19) does not apply to this K. Nevertheless, a duality theorem can be established and several other interesting generalizations of the duality theory of linear programming can be made. This will be done in the present section.

---

[4] It is incorrectly assumed in [33] that local strict convexity for a twice continuously differentiable function implies the positive definiteness of its Hessian matrix in that neighborhood. This is false, though its converse is true [29, p. 406].

Explicitly, the dual programs are

$$\text{(33)} \qquad \begin{aligned} &\text{minimize} \quad F(x,y) = \tfrac{1}{2}xDx + \tfrac{1}{2}yEy + cx \\ &\text{constrained by} \quad Ax + Ey + b \geq 0 \\ &\qquad\qquad (x,y) \geq 0 \end{aligned}$$

and

$$\text{(34)} \qquad \begin{aligned} &\text{maximize} \quad G(x,y) = -\tfrac{1}{2}xDx - \tfrac{1}{2}yEy - by \\ &\text{constrained by} \quad Dx - A'y + c \geq 0 \\ &\qquad\qquad (x,y) \geq 0. \end{aligned}$$

The duality theorem for this pair runs as follows:

(35) Duality Theorem. If $(\hat{x},\hat{y})$ solves the primal problem, (33), then there exists a vector $\hat{v}$ such that $(\hat{x},\hat{v})$ solves the dual problem, (34), $E\hat{y} = E\hat{v}$, and $F(\hat{x},\hat{y}) = G(\hat{x},\hat{v})$. Conversely, if $(\hat{x},\hat{y})$ solves the dual problem, then there exists a vector $\hat{u}$ such that $(\hat{u},\hat{y})$ solves the primal problem, $D\hat{x} - D\hat{u}$, and $F(\hat{u}, \hat{y}) = G(\hat{x},\hat{y})$. Moreover, the pairs $(\hat{x},\hat{v})$ and $(\hat{u},\hat{y})$ solve both problems.

Proof. By the symmetry of the problems, the converse need not be proved. The Kuhn-Tucker conditions (after some simplification and elimination of redundant information) read: There exists an m-vector $\hat{v}$ such that

$$\text{(36)} \qquad D\hat{x} - A'\hat{v} + c \geq 0$$

$$\text{(37)} \qquad E\hat{y} - E\hat{v} \geq 0$$

$$\text{(38)} \qquad \hat{x}(D\hat{x} - A'\hat{v} + c) = 0$$

$$\text{(39)} \qquad \hat{y}(E\hat{y} - E\hat{v}) = 0$$

$$\text{(40)} \qquad \hat{v}(A\hat{x} + E\hat{y} + b) = 0$$

$$\text{(41)} \qquad \hat{v} \geq 0.$$

From (36) and (41), we know that $(\hat{x},\hat{v})$ is a feasible pair for the dual problem. To show $F(\hat{x},\hat{y}) = G(\hat{x},\hat{v})$, it suffices to show $F(\hat{x},\hat{y}) \leq G(\hat{x},\hat{v})$. By (38) and (40), this amounts to showing

$$\hat{v}E\hat{y} \geq \tfrac{1}{2}\hat{v}E\hat{v} + \tfrac{1}{2}\hat{y}E\hat{y}. \tag{42}$$

In virtue of (39), this is equivalent to showing

$$\hat{y}E\hat{y} \geq \hat{v}E\hat{v} \tag{43}$$

which, however, is true since (39), (37), (41), and the symmetry of E imply

$$\hat{y}E\hat{y} = \hat{y}E\hat{v} = \hat{v}E\hat{y} \geq \hat{v}E\hat{v}.$$

Since E is a symmetric, positive semi-definite matrix, the reverse of (42) is always true, hence $(\hat{y} - \hat{v})E(\hat{y} - \hat{v}) = 0$, and this, in turn, implies [19, p. 108] $E\hat{y} = E\hat{v}$. This last fact shows that $(\hat{x},\hat{v})$ does indeed solve both problems.

We note that with $(\hat{x},\hat{v})$ as above, $E\hat{y} = E\hat{v}$, and the <u>complementary slackness conditions</u> appear in (38) and (40).

(44) <u>Remark</u>. This duality theorem was first proved by showing that if $(\hat{x},\hat{y})$ solves the primal problem, then there is a certain <u>linear</u> program of which it is also an optimal solution. The duality theorem of linear programming was then invoked and the required vector $\hat{v}$ was obtained. See [6] and [13].

The duality theorem (35) shows that if either problem, (33) or (34), can be solved, then $\mathcal{P}$ and $\mathcal{D}$ (their respective constraint sets) are nonempty, and moreover $\mathcal{P} \cap \mathcal{D}$ is nonempty. The converse of this is of considerable practical significance. The proof, given below, will depend on a result due to Frank and Wolfe [19, p. 108], but modified here for the situation at hand.

(45) Theorem. Let $f(x) = px + xCx$ be bounded below on a polyhedral convex set $\mathcal{R}$ in R^n. Then f assumes its infimum on $\mathcal{R}$.

Proof. (Since polyhedral convex sets are closed, the result is of particular interest when $\mathcal{R}$ is unbounded.) There is no loss of generality in assuming that C is a symmetric matrix. The demonstration is by induction on the dimension of $\mathcal{R}$; there is nothing to prove if $\mathcal{R}$ is of dimension zero. Suppose that the dimension of $\mathcal{R}$ is $k + 1$. It is possible to write

$$\mathcal{R} = \{s + \gamma t \mid s \in \mathcal{S}, t \in \mathcal{T}, \gamma \in R_+\}$$

where $\mathcal{S}$ is a bounded polyhedral convex set and $\mathcal{T}$ is the intersection of a certain polyhedral convex cone with the unit sphere [24, p. 44]. Note that for all triples $(r,t,\gamma) \in \mathcal{R} \times \mathcal{T} \times R_+$, $r + \gamma t \in \mathcal{R}$ and

(46) $$f(r + \gamma t) = f(r) + \gamma(p + 2rC)t + \gamma^2 tCt.$$

Since f is bounded below on $\mathcal{R}$, it follows from (46) that $tCt \geq 0$ for all $t \in \mathcal{T}$.

If $tCt > 0$ for all $t \in \mathcal{T}$, then $tCt > \alpha > 0$ for some α ($\mathcal{T}$ is compact); and $(p + 2sC)t > \beta$ for some β which may be assumed to be negative. For any $r = s + \gamma t \in \mathcal{R}$, we have

(47) $$f(r) = f(s) + \gamma(p + 2sC)t + \gamma^2 tCt > f(s) + \gamma\beta + \gamma^2\alpha.$$

The minimum of the quadratic function on the right-hand side of (47) is taken on when $\gamma = -\beta/2\alpha$. The infimum of f over $\mathcal{R}$ is therefore attained on the compact subset $\{s + \gamma t \mid s \in \mathcal{S}, t \in \mathcal{T}, 0 \leq \gamma \leq -\beta/2\alpha\}$.

Suppose $\bar{t}C\bar{t} = 0$ for some $\bar{t} \in \mathcal{T}$. Since f is bounded below on $\mathcal{R}$, $(p + 2rC)t \geq 0$ for all $t \in \mathcal{T}$. If for all $r \in \mathcal{R}$, $r + \gamma\bar{t} \in \mathcal{R}$ for all real γ (positive or negative), then $(p + 2rC)\bar{t} = 0$. In this case $f(r + \gamma\bar{t}) = f(r)$ for all $r \in \mathcal{R}$, and the range of f is unaffected by projecting $\mathcal{R}$ into the k-dimensional subspace orthogonal to $\bar{t}$. The inductive hypothesis will then apply and yield the result.

If there exists $(r,\gamma) \in \mathcal{R} \times R$ such that $r + \gamma\bar{t} \notin \mathcal{R}$, consider the sets

$$\mathcal{R}_1 = \{r \mid r \in \mathcal{R}, (p + 2rC)\bar{t} = 0\}$$

$$\mathcal{R}_2 = \{r \mid r \in \mathcal{R}, (p + 2rC)\bar{t} > 0\}.$$

By an earlier comment, $\mathcal{R} = \mathcal{R}_1 \cup \mathcal{R}_2$. By the argument just given, the infimum of f on $\mathcal{R}_1$ is attained there. For $r \in \mathcal{R}_2$, define

$$\gamma_r = \inf\{\gamma \mid r + \gamma\bar{t} \in \mathcal{R}\}.$$

For each $r \in \mathcal{R}_2$, $0 \geq \gamma_r > -\infty$. Thus,

$$b_r = r + \gamma_r\bar{t}$$

belongs to the boundary of $\mathcal{R}$. Now, for each $r \in \mathcal{R}_2$,

$$f(b_r) = f(r) + \gamma_r(p + 2rC)\bar{t} \leq f(r);$$

which means that the minimum of f on $\mathcal{R}_2$ may be sought on the boundary of $\mathcal{R}$. But on each of the k-dimensional bounding hyperplanes of $\mathcal{R}$, f attains its infimum. This completes the proof.

(48) Existence Theorem. If both $\mathcal{P}$ and $\mathcal{D}$ are nonempty, then the primal and dual quadratic programs (33) and (34) have optimal solutions.

Proof. By the weak duality theorem, the primal objective function (which is quadratic) is bounded below over $\mathcal{P}$. Since $\mathcal{P}$ is a polyhedral convex set, the theorem (45) applies. The remainder of the assertion is a consequence of the duality theorem.

To complete this list of extensions of the five theorems in III-1, we mention the

(49) Unboundedness Theorem. If exactly one of $\mathcal{P}$ and $\mathcal{D}$ is nonempty, either $\sup G(x,y) = +\infty$ or $\inf F(x,y) = -\infty$, according as $\mathcal{P}$ or $\mathcal{D}$ is empty.

Proof. This is a consequence of a more general result on programs with linear constraints. See [42, Theorem 3].

III-4. An Equivalence Theorem

Once more, let $K:R^n_+ x R^m_+ \to R$ be an arbitrary differentiable convex-concave function. A point $(\hat{x},\hat{y}) \in R^n_+ x R^m_+$ is called a nonnegative saddle-point of K if

$$K(x,\hat{y}) \leq K(\hat{x},\hat{y}) \leq K(x,\hat{y})$$

for all $(x,y) \in R^n_+ x R^m_+$. Kuhn and Tucker [32] gave necessary and sufficient conditions for a nonegative saddle-point of such a function:

$$\nabla_1 K(\hat{x},\hat{y}) \geq 0\ , \quad \hat{x}\nabla_1 K(\hat{x},\hat{y}) = 0\ , \quad \hat{x} \geq 0 \tag{50}$$

$$\nabla_2 K(\hat{x},\hat{y}) \leq 0\ , \quad \hat{y}\nabla_2 K(\hat{x},\hat{y}) = 0\ , \quad \hat{y} \geq 0\ . \tag{51}$$

Letting $\mathcal{P}$ and $\mathcal{D}$ be the constraint sets of the programs (10) and (11), respectively, we get the following

(52) Equivalence Theorem. The point $(\hat{x},\hat{y}) \in \mathcal{P} \cap \mathcal{D}$ satisfies the equation $F(\hat{x},\hat{y}) = G(\hat{x},\hat{y})$ if, and only if, it is a nonegative saddle-point of K.

Proof. For the necessity, refer to the remark (18). The sufficiency part is immediate.

CHAPTER IV

COMPOSITE PROGRAMS

IV-1. Optimality Criteria

In each of the duality theorems established above, there existed a point common to the intersection of the primal and dual constraint sets which gave equal values to their objective functions. It is natural to form a program whose optimal solutions (if any) are such points. This is done by means of the <u>composite program</u>:

$$\text{(1)} \qquad \begin{aligned} &\text{minimize} \quad H(x,y) = x\nabla_1 K(x,y) - y\nabla_2 K(x,y) \\ &\text{constrained by} \quad \nabla_1 K(x,y) \geq 0 \\ &\qquad\qquad\qquad -\nabla_2 K(x,y) \geq 0 \\ &\qquad\qquad\qquad (x,y) \geq 0. \end{aligned}$$

It consists of minimizing the difference between the primal and dual objective functions over the intersection of the primal and dual constraint sets.

When K is defined by III-(32), the composite program for quadratic programming is

$$\text{(2)} \qquad \begin{aligned} &\text{minimize} \quad xDx + yEy + cx + by \\ &\text{constrained by} \quad Dx - A'y + c \geq 0 \\ &\qquad\qquad\qquad Ax + Ey + b \geq 0 \\ &\qquad\qquad\qquad (x,y) \geq 0. \end{aligned}$$

We shall need (2) for future reference.

By defining $z = (x,y)$ and $W(z) = (\nabla_1 K(x,y), -\nabla_2 K(x,y))$, it is clear that (1) has the <u>form</u>

(3) $$\text{minimize } zW(z)$$
$$\text{constrained by } W(z) \geq 0$$
$$z \geq 0.$$

With $N = n + m$, W can be regarded as a mapping of R_+^N into R^N. When K is convex-concave and twice continuously differentiable, the mapping $W = (\nabla_1 K, -\nabla_2 K)$ is continuously differentiable and its Jacobian matrix,

$$J_W(z) = \begin{pmatrix} \nabla_{11}K(x,y) & \nabla_{12}K(x,y) \\ -\nabla_{21}K(x,y) & -\nabla_{22}K(x,y) \end{pmatrix},$$

can be shown to be positive semi-definite for all $(x,y) \in R_+^n x R_+^m$.

In this section, we study optimality criteria for programs of the form (3) where $W: R_+^N \to R^N$ is an arbitrary continuously differentiable map.

In (3), the objective function $zW(z)$ is obviously nonnegative over the constraint set, provided it is nonempty. Hence any feasible vector which makes the objective function vanish must be optimal. This much is true regardless of any assumptions on W. The necessity of the vanishing of the objective function at an optimal solution is the substance of the next result.

(4) Theorem. Let $\hat{z}$ be an optimal solution of (3) where W is a continuously differentiable map with a positive semi-definite Jacobian matrix at $\hat{z}$. Suppose furthermore that the constraints are Lagrange regular at $\hat{z}$. Then

(5) $$\hat{z}W(\hat{z}) = 0.$$

Proof. For convenience, let $M = J_W(\hat{z})$. By the assumption of

Lagrange regularity, there exists a vector $\hat{u} \geq 0$ such that

$$W(\hat{z}) + M'(\hat{z} - \hat{u}) \geq 0$$
$$\hat{z}[W(\hat{z}) + M'(\hat{z} - \hat{u})] = 0$$
$$\hat{u}W(\hat{z}) = 0.$$

These facts, the feasibility of $\hat{z}$, and the positive semi-definiteness of M' allow us to construct the following string of inequalities:

(6) $$0 \leq \hat{z}W(\hat{z}) = \hat{z}M'(\hat{u} - \hat{z}) \leq (\hat{z} - \hat{u})M'(\hat{u} - \hat{z}) \leq 0.$$

But (6) implies (5).

Thus, under the assumptions of (4), the equation (5) is the criterion for optimality in (3).

Another version of the same theorem involves the assumption that the Jacobian matrix $J_W(\hat{z})$ has <u>positive principal minors</u>, i.e., that the square submatrices of $J_W(\hat{z})$ along its main diagonal have positive determinants.

(7) <u>Theorem</u>. Let $\hat{z}$ be an optimal solution of (3) where W is a continuously differentiable map whose Jacobian matrix has positive principal minors at $\hat{z}$. Then its constraints are Lagrange regular at $\hat{z}$, and the objective function vanishes there.

Proof. To see that the constraints are Lagrange regular at $\hat{z}$, we use the hypothesis that $J_W(\hat{z})$ has positive principal minors. In general, if M is a square matrix with positive principal minors, the system of homogeneous linear inequalities

$$uM \leq 0, \qquad u \geq 0$$

has only the trivial solution [22, Theorem 1]. The Lagrange regularity now follows by a straightforward application of II-(18) in which $f = -W$, $x = z$, and $p = N$.

For the second assertion, we again let $M = J_W(\hat{z})$, and let m^i denote its i-th column. The Kuhn-Tucker conditions componentwise are

$$\begin{aligned} W_i(\hat{z}) + m^i(\hat{z} - \hat{u}) &\geq 0 \\ \hat{z}_i[W_i(\hat{z}) + m^i(\hat{z} - \hat{u})] &= 0 \\ \hat{u}_i W_i(\hat{z}) &= 0 \qquad\qquad 1 \leq i \leq N \\ \hat{u}_i &\geq 0. \end{aligned}$$

Hence for each i,

$$0 \leq \hat{z}_i W_i(\hat{z}) = -\hat{z}_i m^i(\hat{z} - \hat{u}) \leq (\hat{u}_i - \hat{z}_i) m^i(\hat{z} - \hat{u}).$$

That is,

$$(\hat{z}_i - \hat{u}_i) m^i(\hat{z} - \hat{u}) \leq 0 \qquad\qquad 1 \leq i \leq N.$$

But since M (and therefore M') has positive principal minors, $\hat{z} - \hat{u} = 0$ by [22, Theorem 2]. The result now follows.

(8) Remark. It is known (see for example [37]) that positive definite matrices have positive determinants and consequently have positive principal minors. (We offer an alternate proof of this as a by-product of some results in section IV-3.) Of course, the converse of this proposition is false; symmetry is required as an extra hypothesis.

IV-2. Quadratic Programs

This section is devoted to an existence theorem for problem (3) when it is essentially a quadratic program. The composite problem for quadratic programming was stated as (2). If we set

$$M = \begin{pmatrix} D & -A' \\ A & E \end{pmatrix}, \qquad q = (c,b), \quad \text{and} \quad z = (x,y)$$

problem (2) becomes

(9) $$\text{minimize} \quad z(Mz + q)$$
$$\text{constrained by} \quad Mz + q \geq 0$$
$$z \geq 0.$$

This is an example of (3) where $W(z) = Mz + q$ and $J_W(z) = M$.

We study the general problem (9) under the assumption that either M is positive semi-definite or M has positive principal minors. With linear constraints, Lagrange regularity is present, and the optimality criteria, (4) and (7), are in full force: $\hat{z}$ solves (9) if, and only if, $\hat{z}(M\hat{z} + q) = 0$.

(10) Theorem. If M has positive principal minors, problem (9) has an optimal solution.

Proof. By [22, Corollary 2], there exists a vector satisfying the inequalities $Mz > 0$, $z > 0$. A suitably large positive scalar multiple of any such z will be feasible for (9). Since $z(Mz + q) \geq 0$ for all z in the nonempty polyhedral convex constraint set of (9), it has an optimal solution by III-(45).

(11) Remark. If (9) has a nonempty constraint set, it has an optimal solution, regardless of how M is qualified. But in general, the optimal solution need not make the objective function vanish.

We emphasize that when (9) represents the composition (2) of dual quadratic (or in the extreme case, linear) programs, an optimal solution of (9) solves both the primal and the dual problems Conversely, if an optimal solution to either the primal or the dual problem exists, then a solution to (9) exists. In this sense, they are equivalent. However, the composite problem has certain features which lend themselves to a constructive solution technique, i.e., an algorithm, which will be presented in Chapter V.

IV-3. Principal Pivoting

In problems such as (3), a vector $\hat{z}$ is sought which satisfies the conditions

$$W(\hat{z}) \geq 0 \tag{12}$$

$$\hat{z} \geq 0 \tag{13}$$

$$\hat{z}W(\hat{z}) = 0. \tag{14}$$

We now let w denote an N-vector and form the system of <u>equations</u>

$$w_i - W_i(z) = 0 \qquad 1 \leq i \leq N. \tag{15}$$

The problem can be viewed as one of finding a pair of nonnegative, orthogonal vectors, $\hat{w}$ and $\hat{z}$, satisfying the system (15).

Let us assume that W is defined on all of R^N. The N variables w_i in (15) are <u>dependent</u>, and the other N variables z_j are <u>independent</u>. In the language of linear programming, the w_i are <u>basic variables</u> and the z_j are <u>nonbasic variables</u>. (See [8].) This terminology will be used here. A <u>basic solution</u> to (15) is one in which the nonbasic variables all equal zero.

For the remainder of this paper, we denote the set of integers $\{1,\ldots,N\}$ by $\mathcal{N}$. For $i,j \in \mathcal{N}$, we shall define $W_{ij} = \partial W_i/\partial z_j$. This notation is suggested by the analogy with the special case $W(z) = Mz + q$.

If $M = (m_{ij})$, then $W_{ij} = m_{ij}$.

In the next chapter, we shall wish to reverse the dependence roles of certain pairs of variables in the system (15). The idea will be to select a certain equation, say k, in (15) and "solve it" for one of its nonbasic variables z_j as a function of w_k and the remaining $N - 1$ nonbasic variables. When this is done, we shall substitute for z_j in the remaining $N - 1$ equations and obtain a new system. The aim of this section is to make these notions more precise and to describe the relation between the Jacobian matrix of the original mapping and that of the derived mapping.

We recall, first, an obvious, but useful, fact about functions of one real variable. If $f:R \to R$ is a continuously differentiable function with derivative f' satisfying $f'(x) \geq \delta > 0$ $(f'(x) \leq -\delta < 0)$ for some δ and all $x \in R$, then f is a strictly increasing (decreasing) map of R onto R.

Now suppose that for fixed $k,j \in \mathcal{N}$, there exists $\delta > 0$ such that $W_{kj}(z) \geq \delta$ or $W_{kj}(z) \leq -\delta$ for all $z \in R^N$. Then the k-th component function W_k maps R^N onto R. In particular, for each $\bar{w}_k \in R$ and each $(N-1)$-tuple $(\bar{z}_1,\ldots,\bar{z}_{j-1},\bar{z}_{j+1},\ldots,\bar{z}_N) \in R^{N-1}$ there exists a unique $\bar{z}_j \in R$ such that

$$\bar{w}_k = W_k(\bar{z}_1,\ldots,\bar{z}_N).$$

Hence there exists a well-defined function $\tilde{W}_k$ on R^N such that

$$z_j = \tilde{W}_k(z_1,\ldots,z_{j-1}, w_k, z_{j+1},\ldots,z_N) \tag{16}$$

if, and only if,

$$w_k = W_k(z_1,\ldots,z_{j-1}, z_j, z_{j+1},\ldots,z_N) .$$

Since we pr... to denote basic and nonbasic variables generically by w and z, respectively, we alter the notation of the entire set of $2N$ variables as follows:

$$\tilde{w}_k = z_j, \quad \tilde{z}_j = w_k, \quad \tilde{w}_i = w_i \ (i \in \mathcal{N} - \{k\}), \quad \tilde{z}_i = z_i \ (i \in \mathcal{N} - \{j\}).$$

With the new notation, (16) becomes

$$(17) \qquad \tilde{w}_k - \tilde{W}_k(\tilde{z}) = 0.$$

After we define

$$(18) \qquad \tilde{W}_i(\tilde{z}) = W_i(\tilde{z}_1, \ldots, \tilde{z}_{j-1}, \tilde{W}_k(\tilde{z}), \tilde{z}_{j+1}, \ldots, \tilde{z}_N), \qquad i \in \mathcal{N} - \{k\},$$

a new, equivalent, system emerges:

$$(19) \qquad \tilde{w}_i - \tilde{W}_i(\tilde{z}) = 0, \qquad i \in \mathcal{N}.$$

The system (19) is said to be obtained from (15) by a <u>simple pivot</u> $\langle k,j \rangle$. The mapping $\tilde{W}$ is continuously differentiable.

(20) <u>Definition</u>. A <u>principal pivot</u> in (15) consists of a finite sequence $\{\langle k,j \rangle\}$ of simple pivots such that both projections of $\{\langle k,j \rangle\}$ into $\mathcal{N}$ equal $\mathcal{S}$, for some $\mathcal{S} \subseteq \mathcal{N}$. If $\mathcal{S} = \{k\}$, we shall speak of $\langle k,k \rangle$ as a <u>principal pivot on</u> k. We shall write

$$(21) \qquad \tilde{W} = P(W; \{\langle k,j \rangle\})$$

if the system $\tilde{w} - \tilde{W}(z) = 0$ is induced by the principal pivot $\{\langle k,j \rangle\}$ in the system $w - W(z) = 0$. For a principal pivot on k we use the simpler notation $\tilde{W} = P(W;k)$.

If $\tilde{W} = P(W; \langle k,j \rangle)$, the Jacobian matrix $J_{\tilde{W}}$ is given by the formulas (see [7, p. 118])

(22) $$\begin{aligned} \tilde{W}_{kj} &= 1/W_{kj} \\ \tilde{W}_{ij} &= W_{ij}/W_{kj} \\ \tilde{W}_{ki} &= W_{ki}/W_{kj} \qquad\qquad i,\ell \in \mathcal{N} - \{j,k\} \\ \tilde{W}_{i\ell} &= W_{i\ell} - (W_{ij}W_{k\ell}/W_{kj}). \end{aligned}$$

Now let $\tilde{W} = P(W;k)$. We shall need a relationship between the principal minors of $J_{\tilde{W}}$ and those of J_W, for which we introduce another notation. If M is an arbitrary matrix of order N and $\mathcal{I} \subseteq \mathcal{N}$, we denote by $(M)_{\mathcal{I}}$ the principal submatrix of M formed by deleting row and column i from M for each $i \in \mathcal{I}$. The empty matrix, $(M)_{\mathcal{I}}$, will be admitted, and its determinant will be defined to be 1. Of course, $M = (M)_{\emptyset}$.

(23) <u>Theorem.</u> If $\tilde{W} = P(W;k)$, then

(24) $$\left|(J_{\tilde{W}})_{\mathcal{I}}\right| = \left|(J_W)_{\mathcal{I}\,\Delta\{k\}}\right|/W_{kk} \qquad \text{for all } \mathcal{I} \subseteq \mathcal{N}$$

where vertical bars denote determinant and the connective Δ denotes the symmetric difference.[1]

Proof. The statement is obvious when $N = 1$. Assume $N > 1$ and (24) holds for all smaller positive integers. It will suffice to prove (24) in the cases $\mathcal{I} = \emptyset$ and $\mathcal{I} = \{k\}$. Now

$$\left|(J_{\tilde{W}})_{\emptyset}\right| = \begin{vmatrix} W_{11} & \cdots & W_{1,k-1} & 0 & W_{1,k+1} & \cdots & W_{1N} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ W_{k-1,1} & \cdots & W_{k-1,k-1} & 0 & W_{k-1,k+1} & \cdots & W_{k-1,N} \\ \dfrac{-W_{k1}}{W_{kk}} & \cdots & \dfrac{-W_{k,k-1}}{W_{kk}} & \dfrac{1}{W_{kk}} & \dfrac{-W_{k,k+1}}{W_{kk}} & \cdots & \dfrac{-W_{k,N}}{W_{kk}} \\ W_{k+1,1} & \cdots & W_{k+1,k-1} & 0 & W_{k+1,k+1} & \cdots & W_{k+1,N} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ W_{N1} & \cdots & W_{N,k-1} & 0 & W_{N,k+1} & \cdots & W_{NN} \end{vmatrix}$$

$$= \left|(J_W)_{\{k\}}\right|/W_{kk}.$$

---

[1] For a more comprehensive treatment of this and related problems, see Tucker's paper [40].

This is proof of the first part. For the second we have

$$\left|(J_{\tilde{W}})_{(k)}\right| = \begin{vmatrix} \tilde{W}_{11} & \cdots & \tilde{W}_{1,k-1} & 0 & \tilde{W}_{1,k+1} & \cdots & \tilde{W}_{1N} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \tilde{W}_{k-1,1} & \cdots & \tilde{W}_{k-1,k-1} & 0 & \tilde{W}_{k-1,k+1} & \cdots & \tilde{W}_{k-1,N} \\ \frac{W_{k1}}{W_{kk}} & \cdots & \frac{W_{k,k-1}}{W_{kk}} & 1 & \frac{W_{k,k+1}}{W_{kk}} & \cdots & \frac{W_{kN}}{W_{kk}} \\ \tilde{W}_{k+1,1} & \cdots & \tilde{W}_{k+1,k-1} & 0 & \tilde{W}_{k+1,k+1} & \cdots & \tilde{W}_{k+1,N} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \tilde{W}_{N1} & \cdots & \tilde{W}_{N,k-1} & 0 & \tilde{W}_{N,k+1} & \cdots & \tilde{W}_{NN} \end{vmatrix}$$

$$= \begin{vmatrix} W_{11} & \cdots & W_{1,k-1} & W_{1k} & W_{1,k+1} & \cdots & W_{1N} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ W_{k-1,1} & \cdots & W_{k-1,k-1} & W_{k-1,k} & W_{k-1,k+1} & \cdots & W_{k-1,N} \\ \frac{W_{k1}}{W_{kk}} & \cdots & \frac{W_{k,k-1}}{W_{kk}} & 1 & \frac{W_{k,k+1}}{W_{kk}} & \cdots & \frac{W_{kN}}{W_{kk}} \\ W_{k+1,1} & \cdots & W_{k+1,k-1} & W_{k+1,k} & W_{k+1,k+1} & \cdots & W_{k+1,N} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ W_{N1} & \cdots & W_{N,k-1} & W_{Nk} & W_{N,k+1} & \cdots & W_{NN} \end{vmatrix}$$

$$= \left|(J_W)_\emptyset\right| / W_{kk}.$$

These equations are based on the well-known effects that elementary row operations have on determinants. (See, for example, [5, pp. 301-302].)

As a direct consequence of (24), we have

(25) <u>Corollary</u>. If $\tilde{W} = P(W;k)$ and J_W has positive principal minors, then so does $J_{\tilde{W}}$.

This is one of the several invariants under (simple) principal pivoting. Another, which we state without its simple proof, will appear in a forthcoming paper by Tucker and Wolfe [41].

(26) Theorem. Let $W(z) = Mz + q$. Suppose a principal pivot in the system $w - (Mz + q) = 0$ yields the system $\tilde{w} - (\tilde{M}\tilde{z} + \tilde{q}) = 0$. Then $\tilde{M}$ is positive definite (semi-definite) if, and only if, M is positive definite (semi-definite).

Using this invariance theorem, we can prove an assertion made above.

(27) Theorem. The determinant of a positive definite matrix M is positive.

Proof. Every principal submatrix of a positive definite matrix is positive definite, and its main diagonal consists of positive elements. We use induction on the order of M. If M is of order 1, there is nothing to prove. If M is of order $N > 1$, a principal pivot on the index 1 yields a matrix $\tilde{M}$ which is also positive definite. By the inductive hypothesis, $|(\tilde{M})_{\{1\}}| > 0$, and by (24), $|M| = m_{11} |(\tilde{M})_{\{1\}}| > 0$, which completes the proof.

The class of matrices with positive principal minors includes all positive definite matrices, whether or not they are symmetric.

CHAPTER V

NONLINEAR PROGRAMS WITH POSITIVELY BOUNDED JACOBIANS

V-1. The Main Theorem

According to IV-(10), the program

$$\text{(1)} \qquad \begin{array}{ll} \text{minimize} & z(Mz+q) \\ \text{constrained by} & Mz + q \geq 0 \\ & z \geq 0 \end{array}$$

has an optimal solution if M has positive principal minors. (This has also been established constructively, that is, by an algorithm [9].) It is natural to ask whether the same kind of theorem can be proved about the program

$$\text{(2)} \qquad \begin{array}{ll} \text{minimize} & zW(z) \\ \text{constrained by} & W(z) \geq 0 \\ & z \geq 0 \end{array}$$

where it is assumed that $W: R^N \to R^N$ is a continuously differentiable mapping and has a Jacobian matrix $J_W(z)$ with positive principal minors for all $z \in R^N$. It cannot.

In fact, such hypotheses do not even assure the existence of feasible vectors. This can be seen in the one-dimensional problem with $W(z) = -e^{-z}$.

If we postulate feasibility, there is still no hope. Recall that our optimality criterion says that a feasible vector $\hat{z}$ in (2) is optimal if, and only if, $\hat{z}W(\hat{z}) = 0$. Now let

$$W_1(z_1, z_2) = -e^{-z_1} + z_2$$
$$W_2(z_1, z_2) = z_2.$$

Then the Jacobian matrix is

$$J_W(z) = \begin{pmatrix} e^{-z_1} & 1 \\ 0 & 1 \end{pmatrix},$$

and it has positive principal minors for all $z \in R^2$. The program is clearly feasible. If $\hat{z} = (\hat{z}_1, \hat{z}_2)$ solves it, then

$$\hat{z}_1(-e^{-\hat{z}_1} + \hat{z}_2) + \hat{z}_2^2 = 0.$$

Therefore $\hat{z}_2 = 0$; but the only <u>finite</u> value of z_1 for which $-z_1e^{-z_1}$ can vanish is 0. The vector $(\hat{z}_1, \hat{z}_2) = (0,0)$ is not feasible, since it violates the constraint $W_1(z_1, z_2) \geq 0$.

The search for a condition on the mapping W which would guarantee the existence of a solution to (2) has led to the following notion.

(3) <u>Definition</u>. A continuously differentiable mapping $W: R^N \to R^N$ has a <u>positively bounded Jacobian matrix</u> $J_W(z)$ if there exists a number such that $0 < \delta < 1$ and such that for all $z \in R^N$ every principal minor of $J_W(z)$ lies between δ and δ^{-1}.

In (2), if the mapping W has a positively bounded Jacobian matrix, the optimality criterion is the same since the Jacobian has positive principal minors. Furthermore, if $W(z) = Mz + q$, its Jacobian matrix is identically M and is positively bounded if, and only if, M has positive principal minors. Thus, the class of mappings with this rather strong property is nonempty. By considering what the condition means for real-valued functions of one real variable, we note that multivariate mappings can be constructed which also have this property.

We can now state an invariance theorem which will later play an important role.

(4) <u>Theorem</u>. If $W: R^N \to R^N$ has a positively bounded Jacobian matrix and $\tilde{W} = P(W;k)$, then $\tilde{W}$ also has a positively bounded Jacobian matrix.

Proof. There exists a real number δ such that $0 < \delta < 1$ and for all $z \in R^N$ and all $I \subseteq \mathcal{N}$,

$$\delta \leq |(J_W(z))_I| \leq \delta^{-1}.$$

It follows from IV-(24) that

$$\delta^2 \leq |(J_{\tilde{W}}(\tilde{z}))_I| \leq \delta^{-2}$$

for all $\tilde{z} \in R^N$ and $I \subseteq \mathcal{N}$, hence $\tilde{W}$ has a positively bounded Jacobian matrix.

The desirability of a positive lower bound for the partial derivatives W_{kk} $(k \in \mathcal{N})$ was suggested by the examples above. This property need not be preserved in the absence of an upper bound property. For example, if

$$W_1(z_1,z_2) = e^{z_1+z_2} + z_1$$

$$W_2(z_1,z_2) = z_1 + z_2 ,$$

we may choose any δ between 0 and 1 as a lower bound for the principal minors of $J_W(z)$. (Their actual lower bound is 1.) But they have no upper bound. If $W = P(W;1)$, we find that $W_{11} = 1/(e^{z_1+z_2} + 1)$ which is positive but not bounded below by any positive real number. Thus, the positive lower bound property is not invariant under principal pivoting unless there is also an upper bound property. This example helps to motivate the definition (3).

(5) <u>Definition</u>. A solution $(\bar{w},\bar{z})$ of the system

(6) $$w_i - W_i(z) = 0 \qquad i \in \mathcal{N}$$

is <u>nondegenerate</u> if at most N of its 2N components are zero. Otherwise it is degenerate.

(7) Remark. There are perturbation techniques in mathematical programming for ensuring nondegeneracy. One of them involves the replacement of the real variables by lexicographically ordered vectors.[1] (See [10] and [8, Chapter 10].) It will henceforth be assumed that all solutions to (6) are nondegenerate. Nondegeneracy implies (a) that in no basic solution of (6) can there be a basic variable with value zero, and (b) that for each value of z_r, $w_i(0,\dots,0,z_r,0,\dots,0) = 0$ for at most one $i \in N$.

Recall that to solve (2), we need a nonnegative solution $(\hat{w},\hat{z})$ of the system (6) such that $\hat{z}\hat{w} = 0$. The nonnegativity of $(\hat{w},\hat{z})$ corresponds to the feasibility of $\hat{z}$.

(8) Definition. The number of negative coordinates in a solution of (6) is called its infeasibility index.

A basic solution to (6) with an infeasibility index of zero solves problem (2). It is such a solution we shall show to exist. This will be accomplished by an iterative procedure which does not employ nonnegative solutions to (6) until the last step. (Therein lies part of its novelty.) But with a scheme of this kind, the "improvement" in successive iterations must be measured by something other that the change in the value of the objective function, $zW(z)$. We do this by obtaining a (finite) sequence of solutions to (6) such that the corresponding infeasibility indices form a monotonically decreasing sequence. The main result, then, is the following existence theorem.

(9) Theorem. Let $W:R^N \to R^N$ be a continuously differentiable mapping with a positively bounded Jacobian matrix. Then there exists a vector

---

[1] In (1), e.g., q is replaced by (q,I), each component of w,z by a vector.

$\hat{z} \geq 0$ such that $W(\hat{z}) \geq 0$ and $\hat{z}W(\hat{z}) = 0$.

Proof. It will be convenient to make the notation reflect the iterative nature of the method. Therefore we replace the original system (6) by

$$w_i^0 - W_i^0(z^0) = 0 \qquad i \in N$$

Every system we encounter in this proof will be derived from its predecessor by a simple principal pivot. In general, the system after the ν-th iteration will be denoted

$$w_i^\nu - W_i^\nu(z^\nu) = 0 \qquad i \in N$$

where

$$W^\nu = P(W^{\nu-1};k_\nu) \qquad \nu = 1,2,\ldots.$$

At any stage, if $(W^\nu(0),0) \geq 0$, we have the desired result. If this (basic) solution is not nonnegative, it has a positive infeasibility index. It suffices to show that it can be lowered by at least one through a finite number of principal pivots, for the infeasibility index of the initial solution $(W^0(0),0)$ is at most N.

Assume $W^\nu(0)$ has a negative component, say r. Since J_{W^ν} is positively bounded (recall the invariance theorem, (4)), there exists a positive number δ such that $W_{rr}^\nu(z^\nu) \geq \delta$. Hence $W_r^\nu(0,\ldots,0,Z_r^\nu,0,\ldots,0)$ $= 0$ for some positive value Z_r^ν of the variable z_r^ν. Consider the set of values of z_r^ν such that $0 \leq z_r^\nu \leq Z_r^\nu$ and

$$W_i^\nu(0,\ldots,0,z_r^\nu,0,\ldots,0) \geq 0 \quad \text{if} \quad W_i^\nu(0) \geq 0.$$

This bounded set is closed and hence compact. Let its maximum element be denoted ζ_r^ν. We shall call ζ_r^ν the <u>critical</u> <u>value</u> of z_r^ν for this iteration. By the nondegeneracy assumption, it is positive and uniquely determines an index s such that $W_s^\nu(0,\ldots,0,\zeta_r^\nu,0,\ldots,0) = 0$. Let $W^{\nu+1} = P(W^\nu;s)$. If $s = r$, that is $\zeta_r^\nu = Z_r^\nu$, then $(W^{\nu+1}(0),0)$ has a lower infeasibility index than $(W^\nu(0),0)$. If $s \neq r$, then the infeasibility index of the solution $(W^{\nu+1}(0,\ldots,0,\zeta_r^\nu,0,\ldots,0),$ $(0,\ldots,0,\zeta_r^\nu,0,\ldots,0))$ does not exceed that of the previous solution. We now consider the compact set of $z_r^{\nu+1}$ (alias z_r^ν) such that $\zeta_r^\nu \leq z_r^{\nu+1} \leq Z_r^{\nu+1}$ satisfying

$$W_i^{\nu+1}(0,\ldots,0,z_r^{\nu+1},0,\ldots,0) \geq 0 \text{ if } W_i^\nu(0,\ldots,0,\zeta_r^\nu,0,\ldots,0) > 0.$$

Let $\zeta_r^{\nu+1}$ be its maximum element. Then, by our assumption of nondegeneracy, it follows that $\zeta_r^{\nu+1} > \zeta_r^\nu$. Now we repeat these steps.

At every iteration, there is a basic set of variables. Since the critical values increase strictly, the repetition of a basis would imply a contradiction in the critical value of z_r^ν for that iteration. Therefore, after finitely many principal pivots, the r^{th} basic variable reaches the value 0; at this juncture, a principal pivot on the index r lowers the infeasibility index, and we say that a <u>major cycle</u> has ended. As we remarked earlier, at most N major cycles are required to obtain a nonnegative basic solution with infeasibility index zero.

V-2. An Algorithm for Quadratic Programming

The existence theorem (9) provides an effective algorithm for solving problem (2) when the required values are readily computed. This is particularly true when W is of the form $W(z) = Mz + q$ and M has positive principal minors.

In this section, we present an algorithm for solving the quadratic program (1) under the assumption that M is positive semi-definite. The algorithm, adapted from [9], is analogous to the method of existence theorem just proved. It will involve principal pivoting and share the characteristic that never does a feasible solution appear unless it is optimal.

We know, IV-(11), that if (1) is feasible, it has an optimal solution. But it is clear that (1) need not be feasible if M is merely positive semi-definite; our algorithm must be able to handle this contingency. We therefore begin our discussion with a useful observation.

(10) Remark. If the matrix

(11) $$\begin{pmatrix} m_{11} & m_{12} \\ m_{21} & 0 \end{pmatrix}$$

is positive semi-definite, then

(12) $$m_{12} + m_{21} = 0.$$

This is true because

(13) $$m_{11}x_1^2 + (m_{12} + m_{21})x_1x_2 \geq 0, \quad \text{all} \quad (x_1,x_2) \in R^2.$$

However, (13) can hold only if (12) holds.

This remark helps us to prove

(14) Lemma. Let $(w,z) = (\bar{w},\bar{z})$ be a solution of the system $Iw - (Mz + q) = 0$ where M is a positive semi-definite matrix of order N. If for some r

$$\bar{z} \leq 0, \quad \bar{w}_r < 0, \quad m_{rr} = 0, \quad \text{and} \quad m_{ir} \geq 0 \quad \text{all} \quad i \in N,$$

then the system has no nonnegative solution.

Proof. Since $m_{rr} = 0$ and $m_{ir} \geq 0$, for all $i \in N$, it follows from the remark above that $m_{ri} \leq 0$ for all $i \in N$. Now

$$\bar{w}_r = \sum_{i=1}^{N} m_{ri}\bar{z}_i + q_r < 0$$

implies that $q_r < 0$ since $\sum_{i=1}^{N} m_{ri}\bar{z}_i \geq 0$. But the equation

$$w_r - \sum_{i=1}^{N} m_{ri}z_i = q_r$$

can have no nonnegative solution; therefore the entire system has no nonnegative solution.

The reason for assuming $\bar{z} \leq 0$ will become clear in the sequel. Looking ahead to the algorithm, we perceive that if the conditions of the Lemma [illegible], no critical value of the nonbasic variable z_r can be found. As matters stand, the converse of this proposition is not true. What may happen is that the entries of the r-th column of M possess signs which agree with those of the corresponding basic variables. These entries, being the partial derivatives of the basic variables w_i with respect to z_r, make it impossible to ascertain a critical value for z_r. Under such circumstances, z_r could be increased without limit. However, some solvable programs have this property. As a very simple example, let

$$M = \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix} \quad \text{and} \quad q = (-1,-2). \tag{15}$$

The vector $\hat{z} = (0,2)$ is optimal in the program (1) formed with M and q.

To circumvent this difficulty, we introduce an artificial lower bound for negative basic variables. Let β be any number such that

(16) $$-\infty < \beta < \min \{q_i \mid i \in \mathcal{N}\}.$$

If the minimum on the right-hand side is nonnegative, then the zero vector is optimal and there is nothing more to be done. The other possibility implies that β is negative. Throughout the algorithm, β will be used as a lower bound for all basic variables whose values are negative. The lower bound for all nonnegative basic variables is still zero.

At this point, we must revise our notion of nondegeneracy.

(17) <u>Definition</u>. A solution $(\bar{w},\bar{z})$ of the system $Iw - (Mz + q) = 0$ is nondegenerate if at most N of its $2N$ components have the value 0 or β. Otherwise it is degenerate.

This new definition extends the earlier one (5). We shall assume that all solutions to the system $Iw - (Mz + q) = 0$ are nondegenerate. This will imply that if $(\bar{w},\bar{z})$ is a solution and $\bar{z} \in \{0,\beta\}^N$, then for all $i \in \mathcal{N}$, $\bar{w}_i \notin \{0,\beta\}$. If for $i \in \mathcal{N} - \{r\}$, $\bar{z}_i \in \{0,\beta\}$, then for all values of z_r, $W_i(\bar{z}_1,\ldots,\bar{z}_{r-1},z_r,\bar{z}_{r+1},\ldots,\bar{z}_N) \in \{0,\beta\}$ for at most one $i \in \mathcal{N}$.

The notation will be analogous to that of (9). We replace the original system by $Iw^0 - (M^0 z^0 + q^0) = 0$. In general, either the system $Iw^{\nu+1} - (M^{\nu+1} z^{\nu+1} + q^{\nu+1}) = 0$ will be obtained from $Iw^\nu - (M^\nu z^\nu + q^\nu) = 0$ by a principal pivot operation, or else it will be the identical system.

The steps of the algorithm are listed below; a more detailed discussion follows.

Step 0) Set $\nu = 0$; define $(\bar{w}^0,\bar{z}^0) = (q^0,0)$ and define β as any number satisfying (16)

Step 1) If $(\bar{w}^\nu,\bar{z}^\nu) \geq 0$, stop. The solution is optimal.

Step 2) Determine $r \in \mathcal{N}$ such that $\bar{w}^\nu_r < 0$ or $\bar{z}^\nu_r = \beta$.

Step 3) Let ζ_r^ν be the largest value of $z_r^\nu \geq \bar{z}_r^\nu$ satisfying the following conditions:

(i) $z_r^\nu \leq 0$ if $\bar{z}_r^\nu = \beta$;

(ii) $W_r^\nu(z_1,\ldots,\bar{z}_{r-1}^\nu, z_r^\nu, \bar{z}_{r+1}^\nu,\ldots,\bar{z}_N^\nu) \leq 0$ if $\bar{w}_r^\nu < 0$;

(iii) $W_i^\nu(\bar{z}_1^\nu,\ldots,\bar{z}_{r-1}^\nu, z_r^\nu, \bar{z}_{r+1}^\nu,\ldots,\bar{z}_N^\nu) \geq 0$ if $\bar{w}_i^\nu > 0$;

(iv) $W_i^\nu(\bar{z}_1^\nu,\ldots,\bar{z}_{r-1}^\nu, z_r^\nu, \bar{z}_{r+1}^\nu,\ldots,z_N) \geq \beta$ if $\bar{w}_i^\nu < 0$.

Step 4) If $\zeta_r^\nu = +\infty$, stop. No feasible solution exists.

Step 5) If $\zeta_r^\nu = 0$, let $\bar{z}_r^{\nu+1} = 0$, $\bar{z}_i^{\nu+1} = \bar{z}_i^\nu$, for all $i \in \mathcal{N} - \{r\}$, and let $\bar{w}^{\nu+1} = W^\nu(\bar{z}^{\nu+1}) = W^{\nu+1}(\bar{z}^{\nu+1})$. Return to step 1 with ν replaced by $\nu + 1$.

Step 6) Let s be the unique index determined by the conditions (ii), (iii), and (iv) in step 3.

Step 7) If $m_{rr}^\nu = 0$, perform the principal pivot $\{\langle r,s\rangle,\langle s,r\rangle\}$; put $\bar{w}_r^{\nu+1} = \bar{z}_s^\nu$, $\bar{w}_s^{\nu+1} = \zeta_r^\nu$, $\bar{z}_i^{\nu+1} = \bar{z}_i^\nu$ for all $i \in \mathcal{N} - \{r,s\}$ and then $\bar{w}_i^{\nu+1} = W_i^{\nu+1}(\bar{z}^{\nu+1})$ for all $i \in \mathcal{N} - \{r,s\}$. Return to step 3 with ν replaced by $\nu + 1$ and r replaced by s.

Step 8) If $m_{rr}^\nu > 0$, perform a principal pivot on s. Let $\bar{z}_s^{\nu+1} = W_s^\nu(\bar{z}_1^\nu,\ldots,\bar{z}_{r-1}^\nu,\zeta_r^\nu,\bar{z}_{r+1}^\nu,\ldots,\bar{z}_N^\nu)$ and $\bar{w}_i^{\nu+1} = W_i^{\nu+1}(\bar{z}^{\nu+1})$.

Step 9) If $s = r$, return to step 1 with ν replaced by $\nu + 1$.

Step 10) If $s \neq r$, return to step 3 with ν replaced by $\nu + 1$.

This completes the statement of the algorithm. Now for a few explanatory remarks. We point out first that the procedure is designed to rule out any increase in the infeasibility index from one iteration to the next. Every return to step 1 is accompanied by a decrease in the infeasibility index of at least one. The only possible values of the nonbasic variables at a return to step 1 are 0 and β. Consequently, $(\bar{w}^{\nu}, \bar{z}^{\nu}) \geq 0$ (in step 1) implies $\bar{z}^{\nu} = 0$ and hence the orthogonality of the pair.

In step 2, such an r exists if step 1 does not cause termination. Not both $\bar{w}_r^{\nu}$ and $\bar{z}_r^{\nu}$ are negative. This is a property of the initial solution and it is preserved throughout the process.

Step 3 is clear.

The assertion in step 4 follows from (14) and explains its hypothesis, $\bar{z} \leq 0$.

The condition of step 5 implies that $\bar{z}_r^{\nu} = \beta$. No pivoting is required in this case. We simply increase the value of the r-th nonbasic variable and adjust the values of the basic variables accordingly. Then we increase ν to $\nu + 1$.

Step 6 makes sense because the conditions (ii), (iii), and (iv) determine some s; it is unique by nondegeneracy.

If $m_{ss}^{\nu} = 0$, then $r \neq s$ in step 7, for the definition of s and the nondegeneracy assumption imply that $\partial w_s^{\nu}/\partial z_r^{\nu}$ is not zero. Under these circumstances, we cannot make a principal pivot on s (as we would have done in (9)). We have $\partial w_s^{\nu}/\partial z_r^{\nu} = m_{sr}^{\nu} < 0$ (either (iii) or (iv) determines s). Hence, by (10), $m_{rs}^{\nu} > 0$. The principal pivot $\{\langle r,s\rangle, \langle s,r\rangle\}$ can be achieved. Recall that

$$w_r^{\nu+1} = z_s^{\nu}, \quad w_s^{\nu+1} = z_r^{\nu}, \quad w_i^{\nu+1} = w_i^{\nu} \quad \text{all } i \in N - \{r,s\}$$

$$z_r^{\nu+1} = w_s^{\nu}, \quad z_s^{\nu+1} = w_r^{\nu}, \quad z_i^{\nu+1} = z_i^{\nu} \quad \text{all } i \in N - \{r,s\}.$$

The values they acquire in the new solution are given as follows:

$$\bar{z}_r^{\nu+1} = W_s^\nu(\bar{z}_1^\nu,\ldots,\bar{z}_{r-1}^\nu\ ,\ \zeta_r^\nu,\bar{z}_{r+1}^\nu,\ldots,\bar{z}_N^\nu)$$

$$\bar{z}_s^{\nu+1} = W_r^\nu(\bar{z}_1^\nu,\ldots,\bar{z}_{r-1}^\nu\ ,\ \zeta_r^\nu,\bar{z}_{r+1}^\nu,\ldots,\bar{z}_N^\nu)$$

$$\bar{z}_i^{\nu+1} = \bar{z}_i^\nu \quad \text{for all} \quad i \in \mathcal{N} - \{r,s\}$$

$$\bar{w}_i^{\nu+1} = W_i^{\nu+1}(\bar{z}^{\nu+1}) \quad \text{for all} \quad i \in \mathcal{N}.$$

After the principal pivot $\{\langle r,s\rangle,\langle s,r\rangle\}$, the matrix $M^{\nu+1}$ has among its entries

$$m_{rr}^{\nu+1} = \frac{-m_{rr}^\nu}{m_{rs}^\nu m_{sr}^\nu} \geq 0$$

$$m_{rs}^{\nu+1} = \frac{1}{m_{rs}^\nu} > 0$$

$$m_{sr}^{\nu+1} = \frac{1}{m_{sr}^\nu} < 0$$

$$m_{ss}^{\nu+1} = 0.$$

The instruction, in step 7, to return to step 3, with ν replaced by $\nu + 1$ and r replaced by s, suggests an attempt to increase $z_s^{\nu+1}$ (alias w_r^ν) from its negative value. It was this variable which we originally wanted to increase toward zero. Notice that $m_{ss}^{\nu+1} = 0$, so that $w_s^{\nu+1}$ (alias z_r^ν) will not be affected by the forthcoming increase.

The principal pivot indicated in step 8 is possible because $m_{ss}^\nu > 0$.

If the alternative corresponding to step 9 is the case, the meaning of the previous step is that the negative variable w_r^ν has risen to 0, as we wished. In other words, s was determined by (ii). In the new solution, we shall have $\bar{z}_r^{\nu+1} = 0$ and $\bar{w}_r^{\nu+1} > 0$.

From step 10, we re-enter step 3 and increase the r-th nonbasic variable further. An important point is that after the principal pivot on s, we get

$$m_{sr}^{\nu+1} = \frac{-m_{sr}^{\nu}}{m_{ss}^{\nu}} > 0 .$$

Hence the s-th basic variable will <u>increase</u> as $z_r^{\nu+1}$ is increased.

We are now ready to argue the finiteness of the process. As we have already stated, every return to step 1 entails a reduction in the infeasibility index. Since it never increases during the process, only finitely many returns to step 1 can occur. (These correspond to major cycles in (9).) Only finitely many returns to step 3 can occur between returns to step 1. If step 2 is entered at iteration ν_0, a particular index r is determined, and step 3 asks that $z_r^{\nu_0}$ be increased. In each iteration preceding a return to step 1, the variables $w_r^{\nu_0}$ and $z_r^{\nu_0}$ each increase monotonically, and their sum increases in a strict sense. The $N - 1$ nonbasic variables distinct from $z_r^{\nu_0}$ (or $w_r^{\nu_0}$, whichever happens to be nonbasic) all have value 0 or β. These facts preclude the possibility of infinitely many iterations before a return to step 1. Consequently the process is finite.

V-3. An Application of the Main Theorem

The existence of theorem (9) is applicable to the noonegative saddle-point problem (see III-4).

(18) <u>Theorem</u>. Let $K: R^n x R^m \to R$ be a twice continuously differentiable function and suppose that $(\nabla_1 K, -\nabla_2 K)$ has a positively bounded Jacobian

matrix. Then K has a nonnegative saddle-point $(\hat{x},\hat{y})$.

Proof. We use the interpretation of (2) as the composite program IV-(1). By (9), there exists a point $(\hat{x},\hat{y})$ such that:

$$(19) \qquad \nabla_1 K(\hat{x},\hat{y}) \geq 0, \quad \hat{x}\nabla_1 K(\hat{x},\hat{y}) = 0, \quad \hat{x} \geq 0;$$

$$(20) \qquad \nabla_2 K(\hat{x},\hat{y}) \leq 0, \quad \hat{y}\nabla_2 K(\hat{x},\hat{y}) = 0, \quad \hat{y} \geq 0.$$

Since K is twice continuously differentiable, $\nabla_{11}K$ and $-\nabla_{22}K$ are symmetric; they both have positive principal minors. Hence both are positive definite. This implies that K is strictly convex-concave [29, p.406]. The proof is complete since (19) and (20) are the necessary and sufficient conditions for a nonnegative saddle-point of a differentiable convex-concave function.

We mention, in closing, that for this point

$$\max_{y\geq 0} \min_{x\geq 0} K(x,y) = \min_{x\geq 0} \max_{y\geq 0} K(x,y) = K(\hat{x},\hat{y})$$

Therefore (18) may be construed as a <u>minimax theorem</u>. Many such theorems may be found in the literature [17], [38], [39]. Although many involve convex-concave (but otherwise more general) functions K, compactness appears to be an essential ingredient in the hypotheses. It is felt that the noncompactness of R^n_+ and R^m_+ justifies the restrictive hypotheses of (18).

BIBLIOGRAPHY

1. K. J. Arrow and L. Hurwicz, Reduction of constrained maxima to saddle-point problems, in [36], 1-20.

2. K. J. Arrow, L. Hurwicz, and H. Uzawa, Constraint qualifications in maximization problems, Naval Res. Logist. Quart. 8 (1961), 175-191.

3. R. Bellman and M. Hall (eds.), Combinatorial Analysis, Proceedings of Symposia in Applied Mathematics 10, American Mathematical Society, 1960.

4. C. Berge, Topological Spaces (tr. E. M. Patterson), Oliver and Boyd, Edinburgh and London, 1963.

5. G. Birkhoff and S. MacLane, A Survey of Modern Algebra, rev. ed., The Macmillan Company, New York, 1953.

6. R. W. Cottle, Symmetric dual quadratic programs, Quart. Appl. Math. 21 (1963), 237-243.

7. R. Courant, Differential and Integral Calculus, 2nd ed., Vol. 2, Interscience Publishers, Inc., New York, 1953.

8. G. B. Dantzig, Linear Programming and Extensions, Princeton University Press, Princeton, New Jersey, 1963.

9. G. B. Dantzig and R. W. Cottle, Positive (semi-) definite matrices and mathematical programming, Operations Research Center, University of California, Berkeley, ORC 63-18 (RR) 13 May 1963.

10. G. B. Dantzig, A. Orden, and P. Wolfe, The generalized simplex method for minimizing a linear form under linear inequality restraints, Pacific J. Math. 5 (1955), 183-195.

11. G. B. Dantzig, E. Eisenberg, and R. W. Cottle, Symmetric dual nonlinear programs, submitted to Pacific J. Math.

12. J. B. Dennis, Mathematical Programming and Electrical Networks, The Technology Press of the Massachusetts Institute of Technology and John Wiley and Sons, Inc., New York, 1959.

13. W. S. Dorn, Duality in quadratic programming, Quart. Appl. Math. 18 (1960), 155-162.

14. W. S. Dorn, A symmetric dual theorem for quadratic programming, J. Operations Res. Soc. Japan 2 (1960), 93-97.

15. W. S. Dorn, Self-dual quadratic programs, J. Soc. Indust. Appl. Math. 9 (1961), 51-54.

16. H. G. Eggleston, Convexity, Cambridge Tracts in Mathematics and Mathematical Physics No. 47, Cambridge University Press, Cambridge, 1958.

17. K. Fan, Minimax theorems, Proc. Nat. Acad. Sci. U.S.A. 39 (1953), 42-47.

18. J. Farkas, Theorie der einfachen Ungleichungen, J. reine Angew. Math. 124 (1902), 1-27.

19. M. Frank and P. Wolfe, An Algorithm for quadratic programming, Naval Res. Logist. Quart. 3 (1956), 95-110.

20. K. O. Friedrichs, C. E. Neugebauer, and J. J. Stocker (eds.), Studies and Essays, Courant Anniversary Volume, Interscience Publishers, Inc., New York, 1948.

21. D. Gale, The Theory of Linear Economic Models, McGraw-Hill Book Company, Inc., New York, 1960.

22. D. Gale, Univalence theorems for differentiable mappings, Discussion Paper No. 28, Institute for Social and Economic Research, Osaka University, November, 1962.

23. D. Gale, H. W. Kuhn, and A. W. Tucker, Linear programming and the theory of games, Chap. 19 of [30].

24. A. J. Goldman, Resolution and separation theorems for polyhedral convex sets, Chap. 3 of [31].

25. P. Gordan, Über die Auflösung linearer Gleichungen mit reellen Coefficienten, Math. Ann. 6 (1873), 23-28.

26. R. L. Graves and P. Wolfe (eds.), Recent Advances in Mathematical Programming, McGraw-Hill Book Company, Inc., New York, 1963.

27. P. Huard, Dual programs, Paper No. 7 in [26].

28. F. John, Extremum problems with inequalities as side conditions, in [20], 187-204.

29. S. Karlin, Mathematical Methods and Theory in Games, Programming, and Economics, Vol. 1, Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1959.

30. T. C. Koopmans (ed.), Activity Analysis of Production and Allocation, John Wiley and Sons, Inc., New York, 1951.

31. H. W. Kuhn and A. W. Tucker (eds.), Linear Inequalities and Related Systems, Annals of Mathematics Studies Number 38, Princeton University Press, Princeton, New Jersey, 1956.

32. H. W. Kuhn and A. W. Tucker, Nonlinear programming, in [35], 481-492.

33. O. L. Mangasarian, Duality in nonlinear programming, Quart. Appl. Math. 20 (1962), 300-302.

34. J. von Neumann, On a maximization problem, Institute for Advanced Study, Princeton, New Jersey, November, 1947.

35. J. Neyman (ed.), Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, Berkeley and Los Angeles, 1951.

36. J. Neyman (ed.), Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, Berkeley and Los Angeles, 1955.

37. A. M. Ostrowski and O. Taussky, On the variation of the determinant of a positive definite matrix, Nederl. Akad. Wetensch. Proc. Ser. A 54 (1951), 383-385.

38. M. Shiffman, On the equality min max = max min, and the theory of games, The RAND Corporation, RM-243, 22 September 1949.

39. M. Sion, On general minimax theorems, Pacific J. Math 8 (1958), 171-176.

40. A. W. Tucker, _A combinatorial equivalence of matrices_, in [3], 129-140.

41. A. W. Tucker and P. Wolfe, (tentatively entitled) _On the theory of quadratic programming_, The RAND Corporation.

42. P. Wolfe, _A duality theorem for non-linear programming_, Quart. Appl. Math. 19 (1961), 239-244.

BASIC DISTRIBUTION LIST

FOR UNCLASSIFIED TECHNICAL REPORTS

Head, Logistics and Mathematical
Statistics Branch
Office of Naval Research
Washington 25, D. C

C. O., ONR Office
Navy No. 100, Box 39, F. P. O.
New York, New York

ASTIA Document Service Center
Building No. 5
Cameron Station
Alexandra, Virginia

Institute for Defense Analyses
Communications Research Division
von Neumann Hall
Princeton, New Jersey

Technical Information Officer
Naval Research Laboratory
Washington 25, D. C.

C. O., ONR Branch Office
1030 East Green Street
Pasadena 1, California
ATTN: Dr. A. R. Laufer

Bureau of Supplies and Accounts
Code OW, Department of the Navy
Washington 25, D. C.

Professor Russell Ackoff
Operations Research Group
Case Institute of Technology
Cleveland 6, Ohio

Professor Kenneth J. Arrow
Serra House, Stanford University
Stanford, California

Professor G. L. Bach
Carnegie Institute of Technology
Planning and Control of Industrial
Operations, Schenley Park
Pittsburgh 13, Pennsylvania

Professor L. W. Cohen
Mathematics Department
University of Maryland
College Park, Maryland

Professor Donald Eckman
Director, Systems Research Center
Case Institute of Technology
Cleveland, Ohio

Professor Lawrence E. Fouraker
Graduate School of Business
Harvard University
Cambridge, Massachusetts

Professor David Gale
Department of Mathematics
Brown University
Providence 12, Rhode Island

Dr. Murray Geisler
The RAND Corporation
1700 Main Street
Santa Monica, California

Professor L. Hurwicz
School of Business Administration
University of Minnesota
Minneapolis 15, Minnesota

Dr. James R. Jackson, Director
Western Management Sciences Institute
University of California
Los Angeles 24, California

Professor Samuel Karlin
Mathematics Department
Stanford University
Stanford, California

Professor C. E. Lemke
Department of Mathematics
Rensselaer Polytechnic Institute
Troy, New York

Professor W. H. Marlow
Logistics Research Project
The George Washington University
707 - 22nd Street, N. W.
Washington 7, D. C.

BASIC DISTRIBUTION LIST

FOR UNCLASSIFIED TECHNICAL REPORTS

Professor A. Charnes
The Technological Institute
Northwestern University
Evanston, Illinois

Professor R. Radner
Department of Economics
University of California
Berkeley 4, California

Professor Murray Rosenblatt
Department of Mathematics
Brown University
Providence 12, Rhode Island

Mr. J. R. Simpson
Bureau of Supplies and Accounts
Navy Department (Code W31)
Washington 25, D. C.

Professor A. W. Tucker
Department of Mathematics
Princeton University
Princeton, New Jersey

Professor J. Wolfowitz
Department of Mathematics
Lincoln Hall, Cornell University
Ithaca 1, New York

C. O. ONR Branch Office
346 Broadway
New York 13, New York
ATTN: J. Laderman

Professor Oskar Morgenstern
Economics Research Project
Princeton University
92 A Nassau Street
Princeton, New Jersey