AD 691841

Office of Naval Research
Contract N00014-67-A-0298-0006 NR-372-012

STEEPEST ASCENT FOR LARGE-SCALE LINEAR PROGRAMS

By

Richard C. Grinold

April 1969

Technical Report No. 567

This document has been approved for public release and sale; its distribution is unlimited. Reproduction in whole or in part is permitted by the U. S. Government

Division of Engineering and Applied Physics
Harvard University • Cambridge, Massachusetts

D D C
RECEIVED
AUG 20 1969
C

Reproduced by the
CLEARINGHOUSE
for Federal Scientific & Technical
Information Springfield Va. 22151

HS

Office of Naval Research

Contract N00014-67-A-0298-0006

NR-372-012

STEEPEST ASCENT FOR LARGE-SCALE LINEAR PROGRAMS

By

Richard C. Grinold

Technical Report No. 587

This document has been approved for public release and sale; its distribution is unlimited. Reproduction in whole or in part is permitted by the U. S. Government.

April 1969

The research reported in this document was made possible through support extended the Division of Engineering and Applied Physics, Harvard University by the U. S. Army Research Office, the U. S. Air Force Office of Scientific Research and the U. S. Office of Naval Research under the Joint Services Electronics Program by Contracts N00014-67-A-0298-0006, 0005, and 0008.

Division of Engineering and Applied Physics

Harvard University Cambridge, Massachusetts

STEEPEST ASCENT FOR LARGE-SCALE LINEAR PROGRAMS

By

Richard C. Grinold

Division of Engineering and Applied Physics

Harvard University Cambridge, Massachusetts

ABSTRACT

Many structured large-scale linear programming problems can be transformed into an equivalent problem of maximizing a piecewise linear, concave function subject to linear constraints. The equivalent problem can, in turn, be solved in a finite number of steps using a steepest ascent algorithm. This principle is applied to block diagonal systems yielding refinements of existing algorithms. An application to the multi-stage problem yields an entirely new algorithm.

INTRODUCTION

There are two distinct approaches for the solution of large-scale linear programs. The direct or compact basis technique uses special pivoting and storage rules to maintain an easily handled form of the basis. We shall consider the indirect approach. The key element in this method is the definition and solution of an equivalent concave programming problem. This equivalent problem is generally much smaller than the original problem. In many papers it is not mentioned explicitly.

This paper is related to and motivated by the work of Geoffrion [11] and Lasdon [16]. They discuss two methods, which Geoffrion has aptly termed price and resource directive, of defining equivalent problems for nonlinear, organizational problems. In addition, several possible solution techniques are described. Both authors propose the use of large step gradient or feasible direction algorithms. Our attention will be focused on this type of algorithm when the original problem is linear. While we are restricting our attention to linear programs, we do consider a wider class of problem. Thus multi-stage as well as block diagonal (organizational) problems are considered. In principle the theory is applicable to any large scale linear program that can be transformed into an equivalent problem of maximizing a piecewise linear, concave function. In this case we give a general method for constructing an algorithm. Also, by restricting our attention to linear problems, we are able to obtain sharper results, such as finite convergence.

We shall demonstrate that the equivalent problem can itself be transformed into a second linear program. This new linear program is then solved in a finite number of steps using the Primal-Dual (PD) algorithm of Dantzig, Ford and Fulkerson, [8]. If a Feasible Direction (FD) algorithm, [22], is designed for the concave program it yields a procedure identical† to the PD algorithm. This extends the equivalence from the linear case: [22], page 100. Three examples are explored and the algorithm is applied to the equivalent concave program in each instance. The three are: block diagonal problems, using the price directive approach, block diagonal problems, with the resource directive approach, and multi-stage or lower block trianglular problems. In each instance we obtain a finite algorithm for solving these problems.

Section I quickly reviews basic facts of linear programming and linear inequalities, while introducing necessary terminology. In section II we examine the fundamental problem, maximizing a piecewise linear, concave function subject to linear inequalities. The Primal-Dual (PD) and Feasible Direction (FD) algorithms are outlined and compared. The remaining three sections describe applications of the general principle.

The examples given in sections III-V are complementary. The first presents an application of the method in a familiar setting, i. e. the price directive approach to the block diagonal problem. This paves the way for a description of the resource directive method for the same, block diagonal problem. This application, in turn, employs many of the techniques needed in the multi-stage example.

---

†The two differ only in the selection of step size.

In the first two applications we obtain refinements of existing algorithms. The method and results in the multi-stage case are entirely new. This problem originally motivated the study.

A few items have been intentionally ommitted. There is no attempt to show direct applicability of the theory for each example. It is a straight forward exercise to transform any of the equivalent problems into the form of (II:1). In all the examples described, the question of finding a first feasible solution is ommitted. Information on this can be found in, [2], [6] and [23]. No detailed comparison with other algorithms has been made. There is no speculation on the relative efficiency of the procedures, and no consideration of the implementation problems. The purpose of the paper is to demonstrate a unified manner in which algorithms can be constructed and to show the results in three familiar cases.

An effort has been made to minimize notation. Superscripts are used to differentiate among vectors and matrices, while subscripts indicate elements of a particular vector or matrix. When no confusion is possible the notation $p^i \in P$ will indicate that p^i is a column of the matrix P. The vector e, prehaps superscripted, will indicate a vector of ones of appropriate length. Thus if λ is an n vector $e\lambda = \sum_{i=1}^{n} \lambda_i$, and if ρ is a scalar $\rho(e) = (\rho, \rho, \ldots, \rho)$. The norm of a vector is the absolute value of the sum of its coordinates: thus $\|f\| = \sum_{i=1}^{n} |f_i|$. Finally, the operator Min, when applied to a vector

takes on the value of the smallest coordinate. Therefore Min [uP] is equal to the minimum of the scalars up^i, for $p^i \in P$.

I. BACKGROUND

The facts and definitions used in the paper are summarized here.

First consider the convex polyhedral set X

$$X \equiv \{x \mid Ax = b,\ x \geqq 0\} \tag{1}$$

A vector p is an <u>extreme point</u> of X if p cannot be expressed as a proper convex combination of two distinct points in X. Vector r is an <u>extreme ray</u> of X if r is an extreme point of the set Y.

$$Y \equiv \{y \mid Ay = 0,\ ey = 1,\ y \geqq 0\} \tag{2}$$

Let the extreme points and rays be the columns of P and R, resp.. We rely heavily on the important representation theorem of Goldman.

Th: (Goldman) [13]

$$x \in X \quad \text{if and only if} \quad x = P\lambda + R\gamma, \quad \text{where} \quad e\lambda = 1,\ \lambda \geqq 0,\ \gamma \geqq 0 \tag{3}$$

Now consider the linear programming problem:

$$\begin{aligned} &\text{Min } (c - u)x \\ &\text{Subject to } x \in X \end{aligned} \tag{4}$$

Using (3), we can deduce the following about problem (4):

(i) The problem is feasible iff X has an extreme point.

(ii) For any $x \in X$

$(c - u)x = (c - u)P\lambda + (c - u)R\lambda \;:\; e\lambda = 1,\ \lambda \geqq 0,\ \gamma \geqq 0$

(iii) Problem (4) has a finite optimal solution iff X is nonvoid and $(c - u)R \geqq 0$.

(iv) If (4) has a finite optimal solution it has an extreme point optimal solution.

Now we will assume A has full row rank, and that (4) has been solved by the simplex method yielding a finite optimal solution. The final simplex tableau yields a new representation of X.

$$X = \{(w, y, z) \mid Iw + Fy + Dz = \bar{b},\ w \geqq 0,\ y \geqq 0,\ z \geqq 0\}$$

(w, y, z) reflects a partitioning of the columns of A into three sets, and $(I, F, D, \bar{b})$ are the coefficients found in the final simplex tableau. The reduced cost coefficients for the three vectors are $(0, 0, d)$, where $d > 0$. If δ is the optimal value of (4) then the value of any solution $x = (w, y, z)$ is given by

$$(c - u)x = 0w + 0y + dz + \delta$$

From this it is easy to see that:

Prop: (5)

$x = (w, y, z)$ is optimal if and only if $Iw + Fy = \bar{b},\ w \geqq 0,\ y \geqq 0$ and $z = 0$

Therefore the set of optimal solutions, $\bar{X}$, is also a convex polyhedral set and its defining relations may be obtained directly from the final simplex tableau†

† Of course this data can be obtained using the revised simplex method.

$$\bar{X} \equiv \{\bar{x} \mid \bar{A}\bar{x} = \bar{b},\ \bar{x} \geqq 0\}$$

where $\bar{A} = (I, F)$, $x = (w, y)$.

$\bar{X}$ can also be defined in terms of the extreme points and rays of X. Let $\bar{P}$ be the matrix of optimal extreme point solutions and $\bar{R}$ the matrix of <u>extreme tight rays</u>. If p is a column of $\bar{P}$, then $(c - u)p = \delta$; if r is a column of $\bar{R}$, then $(c - u)r = 0$

Prop: (6)

$x \in \bar{X}$ if and only if $x = \bar{P}\lambda + \bar{R}\gamma$ where $e\lambda = 1,\ \lambda \geqq 0,\ \gamma \geqq 0$

Proof:

For any feasible x,

$(c - u)x = (c - u)P\lambda + (c - u)R\gamma$; where $e\lambda = 1,\ \lambda \geqq 0,\ \gamma \geqq 0$

Since the optimal value is δ we must have

$$(c - u)p^i \geqq \delta \quad \text{for each } p^i \in P$$

$$(c - u)r^j \geqq 0 \quad \text{for each } r^j \in R$$

Thus it is easy to see that

$$(c - u)x = \delta \text{ if and only if}$$

$$(c - u)p^i > \delta \text{ implies } \lambda_i = 0$$

$$(c - u)r^j > 0 \text{ implies } \gamma_j = 0$$

This establishes (6). ||

Some information about $\bar{P}$ and $\bar{R}$ can be derived from the final tableau. $(w, y) = (\bar{b}, 0)$ is obviously an extreme optimal solution.

If some column, say f^1, of F is nonpositive; then

$$(w, y_1, y_2, \ldots\ldots) = (-f^1, 1, 0, \ldots\ldots)/(1 + \|f^1\|)$$

is a tight extreme ray. We will call these the <u>at hand</u>, extreme optimal solutions and tight rays.

A typical problem we shall have is: given a vector v, and scalar ρ: Does (ρ, v) satisfy;

$$\begin{aligned} \rho(e) + v\bar{P} &\leqq 0 \\ v\bar{R} &\leqq 0 \end{aligned} \tag{7}$$

This question can be answered by considering the linear program

$$\begin{aligned} &\text{Min } - v\bar{x} \\ &\text{s.t. } \bar{A}\bar{x} = \bar{b},\ \bar{x} \geqq 0 \end{aligned} \tag{8}$$

This type of problem will be refered to as a <u>tight</u> program, since we are looking among solutions that are optimal for another objective function. The tight problem has several important properties that we shall use in sections III-V.

Prop: (9)

(i) (ρ, v) solves (7) iff the optimal value of the tight program is $\geqq \rho$.

(ii) If (ρ, v) does not solve (7), then the tight program will generate and $p \in \bar{P}$, or $r \in \bar{R}$, such that

$$\rho + vp > 0$$

or

$$vr > 0$$

Proof:

The tight problem is assumed to be feasible, so only three things can occur.

Case 1. The tight problem is unbounded below. In this case the simplex method will generate an extreme ray such that $-vr < 0$.

Case 2. The tight problem has an optimal extreme point solution p, but $-vp < \rho$.

Case 3. The tight problem has an optimal extreme point solution p, and $-vp \geqq \rho$.

Since p is optimal we must have $-vp^i \geqq \rho$, for each column p^i of $\overline{P}$, and $-vr^j \geqq 0$, for each $r^j \in \overline{R}$. ||

II. THE PD AND FD ALGORITHMS

In this section we shall show how the PD and FD algorithms can be applied to solve a simple concave programming problem. It will be clear that the two approaches are nearly identical.

The simple problem is:

$$\begin{aligned} &\text{Maximize } h(u) \\ &\text{Subject to } u \in U \equiv \{u \mid uD \leqq g\} \end{aligned} \tag{1}$$

The objective, $h(u)$, is the optimal value of a linear program.

$$\begin{aligned} &h(u) = \text{Min } (c - u)x \\ &\text{Subject to } x \in X \equiv \{x \mid Ax = b,\ x \geqq 0\} \end{aligned} \tag{2}$$

We shall say, $h(u) = -\infty$ if X is void, and $+\infty$ if (2) is unbounded. If either U or X is void then (1) is infeasible. These cases are readily detected, so we will assume U and X are not empty.

As noted in I, we can write x as

$$x = P\lambda + R\gamma \;;\; e\lambda = 1,\ \lambda \geqq 0,\ \gamma \geqq 0$$

where P is the matrix of extreme points, and R the matrix of extreme rays. Suppose for some r^j, that $(c - u)r^j < 0$; then (2) has no lower bound, and $h(u) = -\infty$. Since we are trying to maximize h(u) this situation should be avoided. This is accomplished by adding the constraints $(c - u)R \geqq 0$ to the original problem. Under those restrictions

$$h(u) = \text{Min}[(c - u)P]$$

In other words, (2) has an extreme point optimal solution. Thus instead of (1) we could look at the equivalent problem

$$\text{Maximize Min } [(c - u)P] \tag{3}$$

$$\text{Subject to } uD \leqq g$$

$$uR \leqq cR$$

Problem (3) can be simplified further by introducing the scalar objective ρ and requiring $\rho \leqq (c - u)p^i$ for all $p^i \in P$. Thus (1) can be expressed as a linear program.

$$\begin{aligned} &\text{Maximize } \rho \\ &\text{Subject to } \rho(e) + uP \leqq cP \\ &\qquad\qquad\qquad uR \leqq cR \\ &\qquad\qquad\qquad uD \leqq g \end{aligned} \tag{4}$$

Unfortunately, we don't know P and R and it is computationally impractical to find them. The PD algorithm can work in this enviroment since it is only concerned with the active, or tight constraints. We shall assume the reader is familiar with the PD algorithm and briefly indicate its application to (4). Complete descriptions and proofs of finite convergence can be found in: [6], [8], [20].

Assume we have established that U and X are nonempty, our current solution (δ, u) satisfies the constaints of (4), and $\delta = h(u)$.

Let $\bar{P}$ be the matrix of p^i satisfying

$$\delta = h(u) = (c - u)p^i ;$$

$\bar{R}$ the matrix of r^j satisfying

$$0 = (c - u)r^j ;$$

and finally $\bar{D}$ the matrix of d^k satisfying

$$g_k = ud^k .$$

$\bar{P}$, $\bar{R}$, and $\bar{D}$ represent the optimal extreme points, tight extreme rays and tight constraints.

To obtain a new dual solution we solve the restricted problem.

$$\text{Maximize } \rho \tag{5}$$

$$\text{Subject to } \rho(e) + v\bar{P} \leqq 0$$

$$v\bar{R} \leqq 0$$

$$v\bar{D} \leqq 0$$

$$(-1, -1, \ldots, -1) \leqq v \leqq (1, 1, \ldots, 1)$$

The restricted problem always has an optimal solution (ρ, v) with $\rho \geqq 0$. If $\rho = 0$, then (δ, u) is the optimal solution to problem (4); other wise we compute the following numbers:

$$\lambda^1 = \underset{\{j \mid vr^j > 0\}}{\text{Minimum}} \left[\frac{(c - u)r^j}{vr^j} \right]$$

$$\lambda^2 = \underset{\{k \mid vd^k > 0\}}{\text{Minimum}} \left[\frac{g_k - ud^k}{vd^k} \right]$$

$$\lambda^3 = \underset{\{i \mid \rho + vp^i > 0\}}{\text{Minimum}} \left[\frac{(c - u)p^i - \delta}{\rho + vp^i} \right]$$

$$\bar{\lambda} = \text{Min}[\lambda^1, \lambda^2, \lambda^3] > 0$$

If $\bar{\lambda}$ is infinite: (4) is unbounded, otherwise the new solution is $(\delta, u) + \bar{\lambda}(\rho, v)$. There is a strict improvement at each iteration and the optimal solution is attained in a finite number of iterations.

Now we shall describe the application of a FD algorithm to (1). The objective, $h(u)$, is concave and piecewise linear, but not differentiable. We will show below, however, that it does have finite directional

derivatives in all the interesting directions. The directional derivative of h(u) in the direction v is defined as:

$$\nabla h(u:v) = \lim_{\alpha \to 0^+} \frac{h(u + \alpha v) - h(u)}{\alpha}$$

Suppose we have a $u \in U$ and h(u) is finite. A direction v is <u>feasible</u> if $(u + \alpha v) \in U$ for some $\alpha > 0$. This implies that $vd^k \leqq 0$, for each k such that $ud^k = g_k$. We will look for the feasible direction that maximizes the rate of increase in h. That is

$$\text{Maximize } \nabla h(u:v)$$

$$\text{Subject to } v\bar{D} \leqq 0$$

$$(-1, -1, \ldots, -1) \leqq v \leqq (1, 1, \ldots, 1)$$

The last constraint bounds the direction, since $\nabla h(u:v)$ is homogeneous in v. $\bar{D}$ is the matrix of tight constraints: $d^k \in \bar{D}$ iff $ud^k = g_k$.

Suppose $(c - u)r^j = 0$, and $vr^j > 0$, then for every $\alpha > 0$ $(c - u - \alpha v)r^j < 0$, so $h(u + \alpha v) = -\infty$ and $\nabla h(u:v) = -\infty$. Since we are trying to maximize $\nabla h(u:v)$ we won't consider moving in directions which give an infinite decrease. If we define $\bar{R}$ as before, we can write the direction finding problem as;

$$\text{Maximize } \nabla h(u:v)$$

$$\text{Subject to } v\bar{D} \leqq 0$$

$$v\bar{R} \leqq 0$$

$$(-1, -1, \ldots, -1) \leqq v \leqq (1, 1, \ldots, 1)$$

To show the direction finding problem is equivalent to the restricted problem, we will use a theorem of Danskin, [5]. A proof is included since it is slightly different from the original. First, let's define $\bar{P}$ as the matrix whose columns, p, satisfy

$$(c - u)p = h(u)$$

Th: [5], page 22 (7)

If $v\bar{R} \leqq 0$, then

$$\nabla h(u:v) = \mathrm{Min}[-v\bar{P}]$$

Proof:

Since $v\bar{R} \leqq 0$, there exists $\bar{\alpha} > 0$ such that for $\alpha \in [0, \bar{\alpha}]$

$$(c - u - \alpha v)R \leqq 0$$

and

$$h(u + \alpha v) = \mathrm{Min}[(c - u - \alpha v)P]$$

For any $p^i \in \bar{P}$, $(c - u - \alpha v)p^i \geqq h(u + \alpha v)$ for $\alpha \geqq 0$, and $(c - u)p^i = h(u)$

Thus

$$-vp^i = \frac{(c - u - \alpha v)p^i - (c - u)p^i}{\alpha} \geqq \frac{h(u + \alpha v) - h(u)}{\alpha}$$

as $\alpha \to 0^+$ we get

For any $p^i \in \bar{P}$,

$$-vp^i \geqq \limsup_{\alpha \to 0^+} \frac{h(u + \alpha v) - h(u)}{\alpha}$$

Thus

$$\mathrm{Min}[-v\bar{P}] \geqq \limsup_{\alpha \to 0^+} \frac{h(u + \alpha v) - h(u)}{\alpha}$$

Now suppose some sequence α^n satisfies

$$\alpha^n \longrightarrow 0$$

$$\frac{h(u + \alpha^n v) - h(u)}{\alpha^n} > \beta$$

For each n, there exists a $p^{i_n} \in P$ such that

$$h(u + \alpha^n v) = (c - u - \alpha^n v)p^{i_n}$$

Since P has a finite number of columns, one must occur infinitely often. Denote this column as p^{i*} and the subsequence on which it occurs as α^m.

For each m and any $p \in P$

$$h(u + \alpha^m v) = (c - u - \alpha^m v)p^{i*} \leqq (c - u - \alpha^m v)p$$

In the limit we have:

$$h(u) = (c - u)p^{i*} \leqq \mathrm{Min}[(c - u)P]$$

This implies that $p^{i*} \in \overline{P}$: so

$$\beta = -vp^{i*} \geqq \mathrm{Min}[-vP]$$

Thus

$$\liminf_{\alpha \to 0^+} \frac{h(u + \alpha v) - h(u)}{\alpha} \geqq \mathrm{Min}[-v\overline{P}]$$

which proves the theorem ||

Using (7), we can write the direction finding problem as:

$$\text{Maximize } \{\text{Min}[-v\overline{P}]\}$$

$$\text{Subject to } vR \leqq 0$$

$$vD \leqq 0$$

$$(-1, -1, \ldots, -1) \leqq v \leqq (1, 1, \ldots, 1)$$

Define a new variable ρ, requiring that $\rho \leqq -vp^i$ for $p^i \in \overline{P}$, and transform the direction finding problem into a linear program.

$$\text{Maximize } \rho \tag{8}$$

$$\text{Subject to } \rho\ (e) + v\overline{P} \leqq 0$$

$$v\overline{R} \leqq 0$$

$$v\overline{D} \leqq 0$$

$$(-1, -1, \ldots, -1) \leqq v \leqq (1, 1, \ldots, 1)$$

This demonstrates the equivalence of the direction finding and the restricted problems.

As before, (8) will have an optimal solution (ρ, v) with $\rho \geqq 0$. If $\rho = 0$, then the current solution, u, is optimal for for (1). Otherwise we can make a strict improvement by changing u in the direction v. The next step in the FD algorithm is to move in the direction v and maximize the objective, $h(u + \alpha v)$, while maintaining feasibility: $(u + \alpha v) \in U$.

To this end we define:

$$\lambda^2 \equiv \text{Max}\ \{\lambda \mid (u + \lambda v) \in U\}$$

$$g(\lambda) \equiv h(u + \lambda v)$$

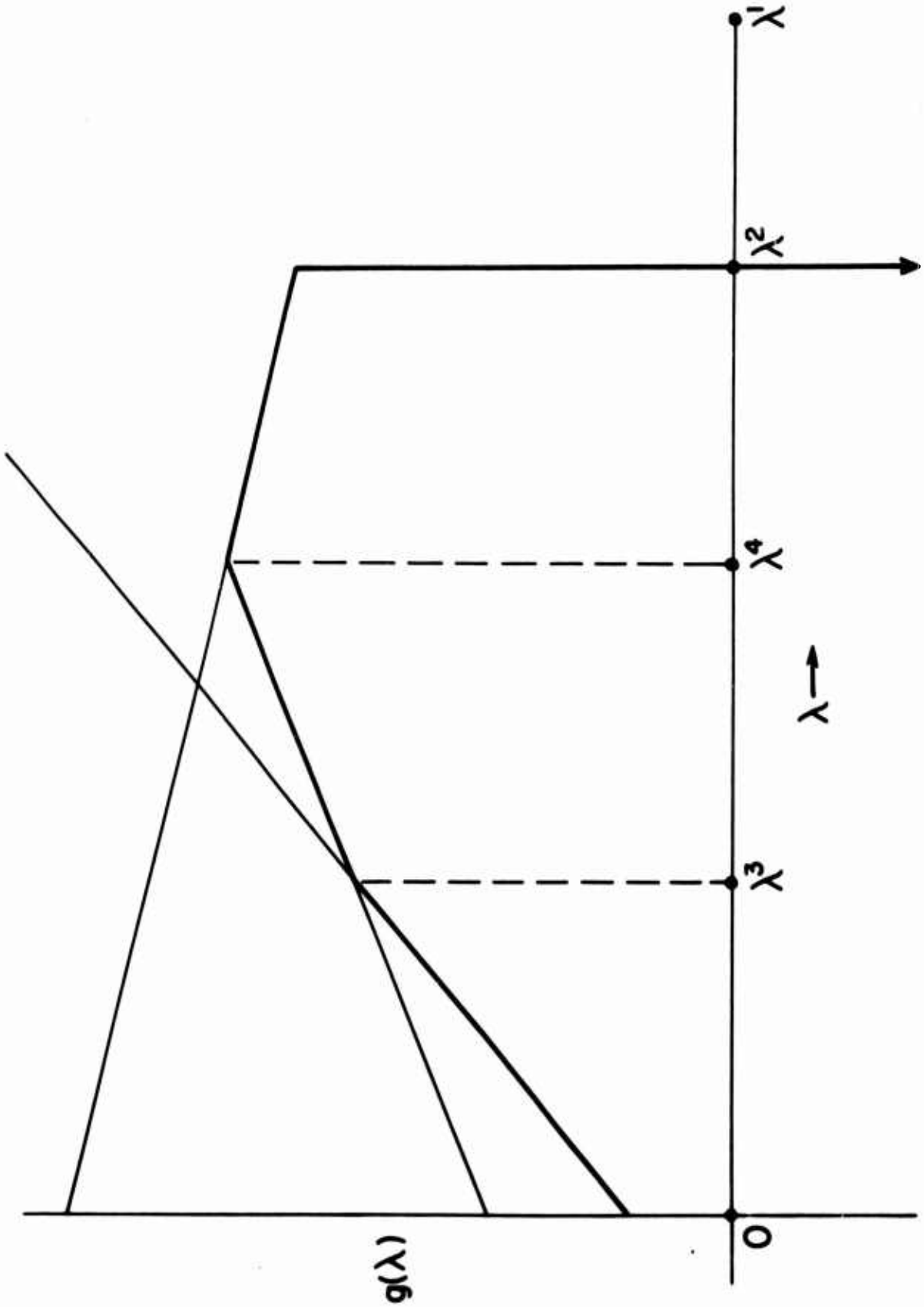
$$g(\lambda^4) \equiv \text{Max}\ \{g(\lambda) \mid \lambda \geqq 0\}$$

$$\bar{\lambda} = \text{Min}\ \{\lambda^2, \lambda^4\} > 0$$

If $\bar{\lambda}$ is infinite then (1) is unbounded, otherwise we define $(\bar{\delta}, \bar{u}) = (\delta, u) + \bar{\lambda}(\rho, v)$ as our new solution.

It is easy to see that the λ^2 defined in (9) is the same as (6). Also, note that $\lambda^4 \leqq \lambda^1$; since for $\lambda^1 < \lambda$, $g(\lambda)$ is equal to $-\infty$. If we graph $g(\lambda)$ we see it is piecewise linear and concave. It increases strictly until $\lambda = \lambda^3$, as defined by (6). At this point the rate of increase is reduced but the function may continue to increase. Thus $\lambda^3 \leqq \lambda^4$ and in some case, see below, $\lambda^3 < \lambda^4$.

Therefore the two algorithms are not precisely equivalent. In some instances they will choose different new solutions. The FD algorithm will always find a solution with a higher value. On the other hand, the PD algorithm converges in a finite number of steps, while the convergence properties of this FD algorithm are uncertain. Also, the PD algorithm examines more options. After stopping at λ^3, the algorithm is still free to go to λ^4. Thus the step lengths of the PD algorithm seem to be theoretically superior. In terms of the FD algorithm, taking suboptimal steps can be interpreted as an anti-zigzagging procedure.


λ^1
λ^2
λ^4
λ
λ^3
0
$g(\lambda)$

III. EXAMPLE: BLOCK DIAGONAL LINEAR PROGRAMS-THE PRICE DIRECTIVE APPROACH

Our first example will be an application to the block diagonal linear program:

$$\text{Minimize} \sum_{t=1}^{T} c^t x^t \tag{1}$$

$$\text{Subject to} \sum_{t=1}^{T} B^t x^t = d$$

$$A^t x^t = b^t, \; x^t \geqq 0 \quad t = 1, 2, \ldots, T$$

This problem is familiar to all those who have studied Dantzig and Wolfe Decomposition, [10]. Other methods have been proposed by, Balas [2], Bell [3], and Abadie and Williams [1]. All of these methods are price directive, that is they are designed to solve the concave programming problem (3). We will apply the algorithm outlined in section II to problem (3). The resulting algorithm is a slight generalization of Balas's. The generalization allows us to handle the case of unbounded subproblems in a direct manner. In addition, we have described a convenient method for solving the restricted problem.

In what follows we will describe how the algorithm applies to this particular problem. Imbedded in this discussion is the outline of a separate algorithm for solving the restricted problem, using the Dantzig and Wolfe technique.

The subproblems are

$$\text{Minimize } (c^t - uB^t)x^t \tag{2}$$

$$\text{Subject to } A^t x^t = b^t,\ x^t \geqq 0$$

Define $h^t(u)$ as the optimal value (2). If (2) is infeasible $h^t(u) = +\infty$; and $h^t(u) = -\infty$ if (2) is unbounded. We will assume that each subproblem is feasible, if not, then the original problem, (1), is infeasible. It is well known that solving the dual of (1) is equivalent to solving

$$\text{Maximize } \sum_{t=1}^{T} h^t(u) \tag{30}$$

Subject to u unrestricted

Suppose we have found a u such that each subproblem has a finite optimal solution. Let $\bar{P}^t$ and $\bar{R}^t$ denote the matrices of optimal extreme points and tight extreme rays for each subproblem. Thus

$$(c^t - uB^t)\bar{P}^t = h^t(u)(e^t)$$

and

$$(c^t - uB^t)\bar{R}^t = 0$$

If we are considering changing u in the direction v, we must insure that $-vB^t\bar{R}^t \geqq 0$ for each t: otherwise any slight change will make one of the subproblems unbounded. Using Theorem II:7, we can see that

$$\nabla h^t(u:v) = \text{Min}[-vB^t\bar{P}^t]$$

The restricted problem is

$$\text{Maximize} \sum_{t=1}^{T} \text{Min}\,[-vB^t\bar{P}^t]$$

$$\text{Subject to } vB^t\bar{R}^t \leq 0, \quad t = 1, 2, \ldots, T$$

$$(-1, -1, \ldots, -1) \leq v \leq (1, 1, \ldots, 1)$$

This is equivalent to

$$\text{Maximize } \rho^1 + \rho^2 + \ldots\ldots + \rho^T \tag{4}$$

$$\text{Subject to } \rho^t(e^t) + vB^t\bar{P}^t \leq 0$$

$$vB^t\bar{R}^t \leq 0 \quad t = 1, 2, \ldots, T$$

$$(-1, -1, \ldots, -1) \leq v \leq (1, 1, \ldots, 1)$$

Finally, the dual of (4) is

$$\text{Minimize } e(\eta + \sigma) \tag{5}$$

$$\text{Subject to } \sum_{t=1}^{T} B^t(\bar{P}^t\lambda^t + \bar{R}^t\gamma^t) + \eta - \sigma = 0$$

$$e^t\lambda^t = 1 \quad t = 1, 2, \ldots, T$$

$$\lambda^t \geq 0, \ \gamma^t \geq 0, \ \eta \geq 0, \ \sigma \geq 0$$

Problem (5) can be solved by the Dantzig and Wolfe method, to avoid the task of enumerating all the columns of $\bar{P}^t$ and $\bar{R}^t$. We will describe one major cycle of the algorithm.

At the start of each major cycle a basic feasible solution of (5) will be available, and selection of the new price u, will generate new extreme points and extreme rays from at least one subproblem.

For each subproblem we have available a matrix representation of the optimal solutions.

$$\bar{X}^t = \{\bar{x}^t \mid \bar{A}^t\bar{x}^t = \bar{b}^t,\ \bar{x}^t \geq 0\} \quad , \quad \bar{A}^t = (I, F^t),\ \bar{x}^t = (w^t, y^t)$$

In this particular procedure we are going to be solving an incomplete inversion of (5). Thus all of the optimal extreme points and extreme rays are not known, just the few we have left from former restricted problems and the new ones we find at hand. We will call the problem the <u>partial</u> or <u>cut down</u> version of (5). Thus we have a cut down version of the restricted problem, which ignores some of the columns.

Suppose we have solved a cut down version of (5) and $(\rho^1, \rho^2, \ldots\ldots, \rho^T, v)$ are the optimal multipliers. If $(\rho^1, \rho^2, \ldots, \rho^T, v)$ is feasible for (4), then it is optimal. Feasibility can be tested by solving the <u>tight subproblems</u> (6).

$$\text{Min} \quad -vB^t\bar{x}^t \tag{6}$$

$$\text{Subject to } \bar{A}^t\bar{x}^t = \bar{b}^t,\ \bar{x}^t \geq 0$$

If, for every t, the optimal value of (6) is no less then ρ^t, then $(\rho^1, \rho^2, \ldots, \rho^T, v)$ is feasible for (4), and thus optimal. In any other case the tight subproblems will generate new columns for (5). This procedure continues until (5) has been solved, yielding a set of optimal multipliers $(\bar{\rho}^1, \rho^2, \ldots, \bar{\rho}^T, v)$. The columns of (5) for which

$$\bar{\rho}^t + \bar{v}B^t p = 0$$

or

$$vB^t r = 0$$

can be retained. These extreme points and extreme rays will be optimal and tight for the next price u'. Thus for the next cycle we will at least retain a basic feasible solution of (5). The other extreme points and rays in (5) can be deleted . They will not be optimal or tight in the subproblems (2) for the new price u'.

To find the new price we consider the parametric objectives $(c^t - uB^t - \alpha B^t)\, x^t$ in the subproblems (2). α is increased until some new activity has a zero reduced cost coefficient. Let α^t be the value of α at which this occurs.

Then let:

$$\alpha^* = \text{Min}\,[\alpha^t |\ t = 1, 2, \ldots, T] \tag{7}$$

If α^* is infinite, then (3) is unbounded: otherwise let

$$u' = u + \alpha^* \bar{v}$$

and for every t such that $\alpha^t = \alpha^*$, determine the new optimal solution or tight ray and introduce them into (5). This completes a major cycle.

The algorithm is summarized below, but first a few passing comments.

If T is very large we might want to use a compact basis technique in solving (5). The generalized upper bounding approach, [9], of Dantzig and Van Slyke can be employed.

Finally, as an alternative approach we note that problem (5) is equivalent to (8) which can, in turn, be solved using Rosen's technique

$$\text{Minimize} \quad e(\pi + \sigma) \tag{8}$$

$$\text{Subject to} \quad \sum_{t=1}^{T} \bar{B}^t \begin{bmatrix} w^t \\ y^t \end{bmatrix} + I\pi - I\sigma = 0$$

$$Iw^t + F^t y^t = \bar{b}^t \qquad t = 1, 2, \ldots, T$$

$$w^t \geqq 0, \; y^t \geqq 0, \; \pi \geqq 0, \; \sigma \geqq 0$$

$\bar{B}^t$ is the restriction of B^t to those activities that can be positive in an optimal basis.

BLOCK DIAGONAL-PRICE DIRECTIVE

Step 0: Obtain a u such that each subproblem,(2), has a finite optimal solution. If no such u exist, then the dual of (1) is infeasible. If any subproblem is infeasible, then (1) is infeasible.

Step 1: Problem (5) is solved using the Dantzig and Wolfe Decomposition technique. Columns are generated from the tight subproblems (6). u is optimal if the value of (5) is 0.

Step 2: Given the optimal multipliers of (5), analyze the subproblems. (2), parametrically to determine a new u. If α^* in (7) is infinite, then (1) is infeasible: otherwise generate the new columns and send them to (5). Go to step 1.

IV. EXAMPLE: BLOCK DIAGONAL LINEAR PROGRAMS-THE RESOURCE DIRECTIVE APPROACH

This example outlines another method for solving (1).

$$\text{Minimize} \sum_{t=1}^{T} c^t x^t \tag{1}$$

$$\text{Subject to} \sum_{t=1}^{T} B^t x^t = d$$

$$A^t x^t = b^t, \; x^t \geq 0 \quad t = 1, 2, \ldots, T$$

As before we shall apply our method to an equivalent, this time convex, programming problem [14]. The fundamental idea is partitioning the vector d among the subproblems, and letting the sectors compete for scarce resources. This approach was originally considered by Benders [4], and Kornai and Liptak [15]. It has recently been discussed in much greater generality by Geoffrion [11], and Silverman [19]. For complete background and motivation see [11].

When the algorithm of section II is applied to (4) we obtain an algorithm similar to Zschau's [23]; but a great deal easier to implement. Actually, it is closer to Geoffrion [11], although we get some simplification and finite convergence by considering the all linear case.

The subproblems are:

$$\text{Minimize } c^t x^t \tag{2}$$

$$\text{Subject to } B^t x^t = y^t$$

$$A^t x^t = b^t, \; x^t \geq 0$$

These are obtained by partitioning the vector d, $\sum_{t=1}^{T} y^t = d$. We will find it more convenient to work with the dual of (2).

$$\begin{aligned} &\text{Maximize} && u^t y^t + v^t b^t \\ &\text{Subject to} && u^t A^t + v^t B^t \leq c^t \end{aligned} \tag{3}$$

These are the problems we propose to solve, but we'll resist the temptation to transform them into "standard form". Define $h^t(y^t)$ as the optimal value of (3). As usual, $h^t(y^t)$ is $+\infty$ if (3) is unbounded and $-\infty$ if (3) is infeasible. Note that if any subproblem,(3),is infeasible, then the dual of (1) is also infeasible. The convex programming problem equivalent to (1) is:

$$\begin{aligned} &\text{Minimize} && \sum_{t=1}^{T} h^t(y^t) \\ &\text{Subject to} && \sum_{t=1}^{T} y^t = d \end{aligned} \tag{4}$$

Before describing the restricted problem, we will look at (3) in greater detail, omitting superscripts. Assume for some y, h(y) is finite, and let $(\bar{U}, \bar{V})$ denote the matrix of optimal extreme solutions. Thus each row (u, v) of $(\bar{U}, \bar{V})$ is an extreme point satisfying:

$$uy + vb = h(y)$$

Similarly, let $(\bar{R}, \bar{Q})$ be the matrix of tight extreme rays. If an extreme ray (r, q) is a row of $(\bar{R}, \bar{Q})$ then

$$ry + qb = 0$$

We are interested in changing y in some direction z. Unless $\bar{R}z \leq 0$, we will have $h(y + \alpha z) = +\infty$, for any positive α. In addition, we can show that

$$\nabla h(y:z) = \text{Max } [\bar{U}z]$$

if $\bar{R}z \leq 0$.

As before, the final simplex tableau will contain a matrix representation of the set of optimal solution. Rather than describing these equations we will denote the set as $\bar{W}^t$.

The reader can easily establish that (5) is the correct form for the restricted problem.

$$\text{Minimize } \sum_{t=1}^{T} \rho^t \tag{5}$$

$$\text{Subject to } \rho^t(e^t) - \bar{U}^t z^t \geq 0$$

$$- \bar{R}^t z^t \geq 0$$

$$(-1, -1, \ldots, -1) \leq z^t \leq (1, 1, \ldots, 1) \text{ for } t = 1, 2, \ldots, T$$

and

$$\sum_{t=1}^{T} z^t = 0$$

We propose to solve (5) using the dual simplex method with upper bounded variables, [21]. As before, after the subproblems have been solved certain new columns from $\bar{U}^t$ and $\bar{R}^t$ will be available. These are added as constraints to (5), which is solved by doing dual simplex steps. Thus we have solved a version of (5) in which some of the constraints have been ignored. Suppose $\{\rho^t, z^t\}_{t=1}^T$ is optimal for this <u>cut down</u> version of (5). We can show that $\{\rho^t, z^t\}_{t=1}^T$ is optimal for (5), with all constraints considered, by checking to see if $\{\rho^t, z^t\}_{t=1}^T$ is feasible for (5). This is accomplished by solving the <u>tight</u> subproblems:

$$\text{Maximize } u^t z^t \tag{6}$$

$$\text{Subject to } (u^t, v^t) \in \bar{W}^t$$

The current solution $\{\rho^t, z^t\}_{t=1}^T$ will be feasible for (5) if the optimal value of each tight subproblem is no greater than ρ^t. Otherwise the tight subproblem will generate a $\bar{u}^t$ or $\bar{r}^t$ such that;

$$\rho^t - \bar{u}^t z^t < 0$$

or

$$-\bar{r}^t z^t < 0$$

These rows are added to (5) and we do dual simplex steps until a new optimum is attained. This procedure is repeated until an optimal solution of (5), $\{\rho^t, z^t\}_{t=1}^T$ is available. If $\sum_{t=1}^T \bar{\rho}^t = 0$, then the current

partition solves (4): otherwise we can make a strict improvement by changing the current partition in the direction of $\{\bar{z}^t\}_{t=1}^{T}$.

For each subproblem (3), consider the parametric objective function $u^t(y^t + \alpha\bar{z}^t) + v^t b^t$. Increase α until some <u>new column</u> is able to enter the basis or an unbounded solution is detected. Call the value of α for which this occurs α^t. Let

$$\alpha^* = \mathrm{Min}[\alpha^t \mid t = 1, 2, \ldots, T] \tag{7}$$

If α^* is infinite then we have an unbounded solution of (4). Otherwise, the new partition is

$$(y^t)' = y^t + \alpha^* \bar{z}^t .$$

For each t, such that $\alpha^t = \alpha^*$, we send the new extreme points and rays to (5), where they will be infeasible under the old solution $\{\bar{\rho}^t, \bar{z}^t\}_{t=1}^{T}$. In addition, all the rows of (5) that were slack under the old optimal solution can be dropped, while the tight rows are retained.

The principle is exactly the same as the one employed in the price directive case, but we have seen fit to solve the restricted problem rather than its dual. The solution of the restricted problem involves the use of a Dantzig and Wolfe scheme for row generation. At any iteration in the solution of (5), the cut down or partial version of (5), is solved using the dual simplex method with upper-bounded variables.

BLOCK DIAGONAL-RESOURCE DIRECTIVE APPROACH

Step 0: Find a partition of d such that each subproblem, (3), has a finite optimal solution. If no such partition exists, then (1) is infeasible. If any subproblem is infeasible, then the dual of (1) is infeasible. Send the optimal extreme solutions and tight rays, at hand, to the restricted problem.

Step 1: Solve the partial version of (5) by the dual simplex method with upper bounded variables, and generate new rows by solving the tight subproblems, (6). If the optimal value of (5) is zero, then the current solution is optimal.

Step 2: Analyze the subproblems, (3), parametrically to determine the new partition. If α^*, (7), is infinite, then (1) has an unbounded solution: otherwise generate the new rows for (5) and go to step 1.

V. EXAMPLE: MULTI-STAGE LINEAR PROGRAMS-TRAJECTORY OPTIMIZATION

In this example we shall examine the multi-stage or lower block triangular linear program:

$$\text{Min} \sum_{t=1}^{T} c^t x^t \tag{1}$$

$$\text{Subject to } A^t x^t = d^t + \sum_{s=1}^{t-1} H^{ts} x^s \;\dagger$$

$$x^t \geq 0 \qquad t = 1, 2, \ldots, T$$

---

† $\sum_{1}^{-1} = \sum_{1}^{0} = 0$

Like the block diagonal problem, which is a special case, this problem has been studied extensively. A host of special techniques are described in, [17] and two compact basis methods are outlined in [14], and [18]. The problem can be hammered into block diagonal form, but this is awkward. A recent report, [12], by Glassey discusses some of these points and proposes an interesting price-resource communication technique which solves a special case of (2).

First, we will make an assumption about (1), in order to transform it into a more manageble form, (2). Then we'll describe (2) in terms of an equivalent convex programming problem. This equivalent problem is solved using the algorithm of section II.

Although the methods we are describing apply to (1), we will make an assumption which allows us to discuss a special case. The general case is described briefly in an appendix.

Our assumption is:

$$H^{ts} = H^{s} \text{ for all } s \text{ and } t = s+1,\ s+2, \ldots, T$$

Using this assumption and subtracting rows we can transform (1) into an equivalent problem:

$$\text{Minimize} \sum_{t=1}^{T} c^t x^t \tag{2}$$

$$\text{Subject to } A^t x^t = b^t + K^{t-1} x^{t-1}$$

$$x^t \geq 0 \qquad t = 1, 2, \ldots, T$$

Here, $K^t = H^t + A^t$ for $t = 1, \ldots, T-1$: and $K^0 = 0$. Also, $b^t = d^t - d^{t-1}$ for $t = 2, 3, \ldots, T$; and $b^1 = d^1$.

The dual of (2) is

$$\text{Maximize} \sum_{t=1}^{T} w^t b^t \tag{3}$$

$$\text{Subject to } w^t A^t \leqq c^t + w^{t+1} K^t \quad t = 1, 2, \ldots, T$$

If (2) is considered as a sequential problem, then the right hand side, $\{b^t + K^{t-1}x^{t-1}\}_{t=2}^{T}$ describes a trajectory. Our procedure is called trajectory optimization since we are going to fix the trajectory and then find the best solution that follows it. This computation yields information useful in finding a better trajectory or indicates that the present trajectory is optimal. As with price and resource directive approaches, selection of a trajectory allows us to decompose the problem into easily solvable subproblems.

For our purpose the vector $z = (z^1, z^2, \ldots, z^{T-1})$ defines a trajectory $\{b^t + K^{t-1}z^{t-1}\}_{t=2}^{T}$. The subproblems optimize along this path.

$$\text{Minimize } c^t x^t \tag{4}$$

$$\text{Subject to } A^t x^t = b^t + K^{t-1} z^{t-1}$$

$$K^t x^t = K^t z^t$$

$$x^t \geqq 0$$

where $K^0 = K^T = 0$

The subproblems we will actually solve are the duals of (4).

$$\text{Maximize } u^t(b^t + K^{t-1}z^{t-1}) + v^tK^tz^t \tag{5}$$

$$\text{Subject to } u^tA^t + v^tK^t \leqq c^t$$

The optimal value of (5) is given by the function $h^t(z)$. As usual, $h^t(z)$ is $+\infty$ if (5) is unbounded and $-\infty$ if (5) has no feasible solution.

The equivalent convex programming problem is

$$\text{Minimize } \sum_{t=1}^{T} h^t(z) \tag{6}$$

$$\text{Subject to } z \geqq 0$$

The equivalence is almost tautalogical. It is based on two facts. If $\{x^t\}_{t=1}^{T}$ is any feasible solution of (2) then $\{x^t\}_{t=1}^{T-1}$ determines a trajectory along which $\{x^t\}_{t=1}^{T}$ is feasible. Also, if for some z, $\{x^t\}_{t=1}^{T}$ solves the subproblems (4), then $\{x^t\}_{t=1}^{T-1}$ traces out the same trajectory as z and $\{x^t\}_{t=1}^{T}$ is a feasible solution of (2).

Theorem: (7)

(i) If $\{x^t\}_{t=1}^{T}$ solves (2), then $\{x^t\}_{t=1}^{T-1}$ solves (6).

(ii) If $\{z^t\}_{t=1}^{T-1}$ solves (6), then any solution along the trajectory

$\{b^t + K^{t-1}z^{t-1}\}_{t=1}^{T}$ solves (2).

Proof:

Suppose $\{x^t\}_{t=1}^T$ solves (2).

Consider $x = (x^1, x^2, \ldots, x^{T-1})$ as a solution of (6). It is easy to see that $h^t(x) \leqq c^t x^t$, since x^t is a feasible solution of (4). Suppose there is some other z, and $\{y^t\}_{t=1}^T$ such that $\sum_{t=1}^T c^t y^t = \sum_{t=1}^T h^t(z) < \sum_{t=1}^T h^t(x) \leqq \sum_{t=1}^T c^t x^t$. Then $\{y^t\}_{t=1}^T$ is feasible for (2), contradicting the optimality of $\{x^t\}_{t=1}^T$.

Now assume z solves (6) and $\{x^t\}_{t=1}^T$ is optimal along that trajectory. If $\{x^t\}_{t=1}^T$ does not solve (2) there exists a $\{y^t\}_{t=1}^T$ such that $\sum_{t=1}^T c^t y^t < \sum_{t=1}^T c^t x^t$. $y = (y^1, y^2, \ldots, y^{T-1})$ determines a trajectory along which it is feasible, and $h^t(y) \leqq c^t y^t$ for each t. Thus

$$\sum_{t=1}^T h^t(y) \leqq \sum_{t=1}^T c^t y^t < \sum_{t=1}^T c^t x^t = \sum_{t=1}^T h^t(z)$$

contradicting the optimality of z. ||

Note the theorem implies the existence of an optimal solution that is an extreme point optimal solution of each subproblem.

Before stating the restricted problem, we will inspect the subproblems, (5), in greater detail. Assume we have found a trajectory z such that each subproblem (5) has a finite optimal solution. Let $(\bar{U}^t, \bar{V}^t)$ be the matrix of extreme optimal solutions. If (u^t, v^t) is a row of $(\bar{U}^t, \bar{V}^t)$ then it is an extreme point of (5) and

$$u^t(b^t + K^{t-1}z^{t-1}) + v^tK^tz^t = h^t(z)$$

Similarly let $(\bar{R}^t, \bar{Q}^t)$ be the matrix of extreme tight rays. If (r^t, q^t) is a row of $(\bar{R}^t\bar{Q}^t)$ then (r^t, q^t) is an extreme ray and

$$r^t(b^t + K^{t-1}z^{t-1}) + q^tK^tz^t = 0$$

If we are going to change $z = (z^1, z^2, \ldots, z^{T-1})$ in the direction $y = (y^1, y^2, \ldots, y^{T-1})$, we must have

$$\bar{R}^tK^{t-1}y^{t-1} + \bar{Q}^tK^ty^t \leqq 0$$

or the subproblems, (5), will become unbounded. We can also show using Theorem II:7, that

$$\nabla h^t(z:y) = \text{Max}\ [\bar{U}^tK^{t-1}y^{t-1} + \bar{V}^tK^ty^t]$$

Finally, we must have $y_i^t \geqq 0$, if $z_i^t = 0$.

Using all this information, we can write the restricted problem as:

$$\text{Minimize} \sum_{t=1}^{T} \rho^t \tag{6}$$

$$\begin{array}{llll}
\text{Subject to } \rho^1(e^1) & -\bar{V}^1K^1y^1 & \geqq 0 & \\
 & -\bar{Q}^1K^1y^1 & \geqq 0 & \\
\rho^t(e^t) & -\bar{U}^tK^{t-1}y^{t-1} - \bar{V}^tK^ty^t & \geqq 0 & \\
 & -\bar{R}^tK^{t-1}y^{t-1} - \bar{Q}^tK^ty^t & \geqq 0 & \text{for } t = 2, 3, \ldots, T-1 \\
\rho^T(e^T) & -\bar{U}^TK^{T-1}y^{T-1} & \geqq 0 & \\
 & -\bar{R}^TK^{T-1}y^{T-1} & \geqq 0 & \\
 & y_i^t & \geqq 0 & \text{if } z_i^t = 0, \\
 & & & t = 1, 2, \ldots, T-1 \\
(-1, -1, \ldots, -1) \leqq y^t \leqq (1, 1, \ldots, 1) & & &
\end{array}$$

The restricted problem has the same linked structure as the original problem. Note that some of the lower bounds on the y^t are redundant. The method suggested for solving (6) is almost identical with the resource directive case.

After (5) has been solved, for some z, a matrix representation of the optimal solutions is available. This is denoted $\bar{W}^t$. From this set of linear inequalities, we can immediately discover a few optimal extreme solutions and tight extreme rays. These are the solutions at hand. With this and past information we solve a partial or cut down version of (6), using the dual simplex method with upperbounds. This generates a solution $y = (y^1, y^2, \ldots, y^{T-1}), \{\rho^t\}_{t=1}^{T}$. This solution is optimal for (6) it it is feasible. Feasibility is checked by solving the tight problems

$$\text{Max } u^t(K^{t-1}y^{t-1}) + v^t(K^t y^t) \tag{7}$$

$$\text{Subject to } (u^t, v^t) \in \bar{W}^t$$

The solution is feasible if, for each t, the optimal value of the tight program is no greater than ρ^t. Otherwise the tight programs generate an infeasible constraint of (6).

Suppose $\bar{y}$ is optimal for (6). Now we will change z in the direction $\bar{y}$. The new z must be non negative, so we can proceed no further than.

$$\beta = \text{Min}\left[\frac{z_i^t}{-\bar{y}_i^t} \;\middle|\; \bar{y}_i^t < 0,\ t = 1, 2, \ldots, T-1\right]$$

In addition, for each subproblem, (5), we consider the parametric objective

$$u^t b^t + u^t K^{t-1} z^{t-1} + \alpha u^t K^{t-1} \bar{y}^{t-1} + v^t K^t z^t + \alpha v^t K^t \bar{y}^t$$

The parameter α is increased until some new reduced cost coefficent is zero, indicating an unbounded solution or a new extreme optimal solution. Call that value α^t and let

$$\begin{aligned} \alpha^* &= \text{Min } [\alpha^t \mid t = 1, 2, \ldots, t] \\ \theta &= \text{Min } [\alpha^*, \beta] > 0 \end{aligned} \tag{8}$$

If θ is infinite then there is an unbounded solution. Otherwise the new solution is

$$\bar{z}^t = z^t + \theta \bar{y}^t \qquad t = 1, 2, \ldots, T - 1$$

As before the constraints in that are slack after (6) has been solved can be eliminated, while the tight constraints should be retained. For each t such that $\alpha^t = \theta$, there is a new constraint generated for (6). Also, for each (t, i) such that $\theta = \left[\frac{z_i^t}{-\bar{y}_i^t}\right]$, the constraint $y_i^t \geqq 0$ is added to (6).

MULTI-STAGE PROGRAMS-TRAJECTORY OPTIMIZATION

Step 0: Find an initial trajectory such that each subproblem, (5), has a bounded optimal solution. If no such trajectory exists, then (2) is infeasible. If any of the subproblems are infeasible, then (3) is infeasible. Send the optimal solutions and extreme rays at hand to the restricted problem.

Step 1: Solve the partial version of the restricted problem by the dual simplex method with upper bounded variables, and generate new rows by solving the tight problems,(7). If the optimal value of (6) is zero, the current trajectory is optimal.

Step 2: Determine how far to proceed in the new direction, (8). If θ is infinite, the problem is unbounded: otherwise generate the new restricted problem and go to 1.

APPENDIX

For the general multi-stage problem, (V:I) the subproblems are for $t = 1, 2, \ldots, T$

$$\text{Min} \quad c^t x^t \tag{1}$$

$$\text{Subject to } A^t x^t = d^t + \sum_{s=1}^{t-1} H^{ts} z^s$$

$$H^{st} x^t = H^{st} z^t \qquad s = t+1,\ t+2, \ldots, T$$

$$x^t \geqq 0$$

Actually the dual of (1) will be solved

$$\text{Max} \quad u^t(d^t + \sum_{s=1}^{t-1} H^{ts} z^s) + \sum_{s=t+1}^{T} v^{st} H^{st} z^t \tag{2}$$

$$\text{Subject to } u^t A^t + \sum_{s=t+1}^{T} v^{st} H^{st} \leqq c^t$$

Although (2) has a large number of variables, it has the same number of constraints as the subproblems considered in the special case (V:5).

The construction of the restricted problem is straight forward. It will inherit the lower block triangular structure of the original problem (V:1).

REFERENCES

1. Abadie, J. M. and A. C. Williams, "Dual and Parametric Methods in Decomposition, "Recent Advances in Mathematical Programming (R. Graves and Ph. Wolfe, eds.) pp. 149-158, McGraw-Hill, New York, (1963).

2. Balas, E., "An Infeasibility-Pricing Decomposition Method for Linear Programs, "J. ORSA., Vol. 14, No. 5, (1966).

3. Bell, E. J., "Primal-Dual Decomposition Programming," ORC 65-23, Operations Research Center, University of California, Berkeley, (1965).

4. Benders, J. F., "Partitioning Procedures for solving Mixed Variable Programming Problems," Numerische Mathematik, 4, pp. 238-252, (1962).

5. Danskin, J. M., The Theory of Max-Min, Springer-Verlag, Inc. New York, (1967).

6. Dantzig, G. B., Linear Programming and Extensions, Princeton University Press, Princeton, N. J., (1963).

7. Dantzig, G. B., "On the Status of Multi-State Linear Programming Problems," Management Science, Vol. 6, No. 1, (1959).

8. Dantzig, G. B., Ford, L. R. and D. R. Fulkerson, "A Primal-Dual Algorithm for Linear Programs," in Linear Inequalities and Related Systems, (H. W. Kuhn and A. W. Tucker, eds), Annals of Mathematics Study No. 38, Princeton University Press, Princeton, N. J., (1956).

9. Dantzig, G. B. and R. M. Van Sylke, "Generalized Upper Bounding Techniques," J. of Computer and System Science, Vol. 1, No. 3, pp. 213-226, (1967).

10. Dantzig, G. B. and Ph. Wolfe, "Decomposition Principle for Linear Programs," J. ORSA., Vol. 8, No. 1, pp. 101-111, (1960).

11. Geoffrion, A. M., "Primal Resource-Directive Approaches for Optimizing Nonlinear Decomposable Systems," in Optimization Methods for Large Scale Systems, (D. A. Wismer ed.) McGraw-Hill, (forthcoming).

12. Glassey, C. R., "Dynamic Linear Programs for Production Scheduling," ORC 68-2, Operations Research Center, University of California, Berkeley, (1968).

13. Goldman, A. J., "Resolution and Separation Theorems for Polyhedral Convex Sets," Linear Inequalities and Related Systems, (H. W. Kuhn and A. W. Tucker, eds) Annals of Mathematics Study #38, Princeton University Press, (1956).

14. Heesterman, A. R. G., and J. Sandee, "Special Simplex Algorithm for Linked Problems," Management Science, Vol. 11, No. 3 pp. 420-428, (1965).

15. Kornai, J. and Liptak, T., "Two Level Planning," Econometrica, 33,1 pp. 141-169, (1965).

16. Lasdon, L. S., "Duality and Decomposition in Mathematical Programming," IEEE Trans. on System Science and Cybernetics, Vol. 4, No. 2, (1968).

17. Rosen, J. B. "Primal Partition Programming for Block Diagonal Matrices," Numerische Mathematik, 6, pp. 250-260, (1964).

18. Saigal, R., "Block Triangularization of Multi-Stage Linear Programs," ORC 66-9, Operations Research Center, University of California, Berekely, (1966).

19. Silverman, G. J., "Primal Decomposition of Mathematical Programs by Resource Allocation," Operations Research Department Tech., Mem. 116 Case Western Reserve University, (1968).

20. Simmonard, M., "Linear Programming," Prentice-Hall Inc., Englewood Clifts, New Jersey, (1966).

21. Wagner, H. M., "The Dual-Simplex Algorithm for Bounded Variables," Naval Research Logistics Quarterly, Vol. 5, No. 3, (1958).

22. Zoutendijk, G., "Methods of Feasible Directions," Elsevier Publishing Co., Amsterdam, (1960).

23. Zschau, E. V. W., "A Primal Decomposition Algorithm for Linear Programming," Working Paper # 91, Graduate School of Business Stanford University, (1967).

Unclassified

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Division of Engineering and Applied Physics Harvard University Cambridge, Massachusetts 02138	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

3. REPORT TITLE

STEEPEST ASCENT FOR LARGE-SCALE LINEAR PROGRAMS

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Interim technical report

5. AUTHOR(S) (First name, middle initial, last name)

Richard C. Grinold

6. REPORT DATE April 1969	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS 23
8a. CONTRACT OR GRANT NO. N00014-67-A-0298-0006 b. PROJECT NO. c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) Technical Report No. 587 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)	

10. DISTRIBUTION STATEMENT

This document has been approved for public release and sale; its distribution is unlimited. Reproduction in whole or in part is permitted by the U. S. Government.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY Office of Naval Research

13. ABSTRACT

Many structured large-scale linear programming problems can be transformed into an equivalent problem of maximizing a piecewise linear, concave function subject to linear constraints. The equivalent problem can, in turn, be solved in a finite number of steps using a steepest ascent algorithm. This principle is applied to block diagonal systems yielding refinements of existing algorithms. An application to the multi-stage problem yields an entirely new algorithm.

DD FORM 1 NOV 65 1473 (PAGE 1)

S/N 0101-807-6811

Unclassified

Security Classification

A-31408

14. KEY WORDS	LINK A ROLE	LINK A WT	LINK B ROLE	LINK B WT	LINK C ROLE	LINK C WT
Large Scale						
Linear Programming						
Multi-Stage						
Multi-Divisional						
Steepest Ascent						