LA-10284-MS

C.3

Los Alamos National Laboratory is operated by the University of California for the United States Department of Energy under contract W-7405-ENG-36.

A Review of the Literature on Bi-Level Mathematical Programming

Los Alamos
Los Alamos National Laboratory
Los Alamos, New Mexico 87545

LA-10284-MS

UC-32
Issued: October 1985

A Review of the Literature on Bi-Level Mathematical Programming

Charles D. Kolstad*

*Consultant at Los Alamos. University of Illinois at Urbana-Champaign, Urbana, IL 61801.

Los Alamos National Laboratory
Los Alamos, New Mexico 87545

CONTENTS

A REVIEW OF THE LITERATURE ON BI-LEVEL MATHEMATICAL PROGRAMMING

by

Charles D. Kolstad

ABSTRACT

This paper reviews the recent literature on applications and algorithms in bi-level programming. Bi-level programming involves two mathematical programs. One math program is concerned with minimizing w(x,t) over some region by varying the vector t. The variable x is actually a function x(t) and is defined implicitly as the solution vector to the second math program, which minimizes s(x,t) over some region by varying x. The review is divided into two main sections. One section covers applied problems that have been presented in the literature as bi-level math programs. Most such applications are in economics but some are in warfare planning. Another section of the paper concerns the many diverse algorithms that have been developed to solve the bi-level programming problem.

I. INTRODUCTION

Over the past decade there has been an increase in interest in multi-level mathematical programming, and in particular bi-level mathematical programming. The bi-level problem consists of two parts, an upper and lower part. Define the upper-level problem (denoted henceforth as "P1") as

(P1:) $$\min_t w(x,t) \qquad (1a)$$

$$\ni f(x,t) \le 0 \quad , \qquad (1b)$$

where x(t) is implicitly defined by the lower-level problem:

(B1:) $$x(t): \min_x s(x,t) \qquad (1c)$$

$$\ni g(x,t) \le 0 \quad , \qquad (1d)$$

where all variables and constraint functions may be vectors. A tremendous variety of applied problems, particularly economic problems, can be viewed as bi-level math programs. A Stackelberg duopoly or leader-follower continuous game (Chen and Cruz, 1972; Cruz, 1978; Papavassilopoulos, 1981) can be viewed as bi-level programming problems with the follower's problem corresponding to B1 and the leader's problem corresponding to P1. Many applications are in economic planning where the planner's problem is P1 and the economy responds according to B1. Related to this is the principal-agent problem where the principal (P1) tries to induce his agent (B1) to act in the principal's interest. Outside the economics literature, the max-min problem (Danskin, 1966) is that of maximizing the minimum of some function and is thus a special case of bi-level programming.

Unfortunately, good solution methods for the bi-level problem are not generally available. In fact, without significant restrictions on the sub-problem, the overall problem may well be nonconvex and thus difficult to solve for a global optimum.

The purpose of this paper is to provide a review of recent progress on bi-level programming (through 1982). The review covers both applications and algorithms. There has been a fair amount of work in both these areas with many algorithms springing from the need to solve specific applied problems. In the next section we review applications, some of which have appeared explicitly in the literature and others of which have only been suggested. This is followed by a section on algorithmic developments in multi-level programming. Most Soviet and Eastern European literature is not reviewed here (however, see Findeisen, 1982).

II. APPLICATIONS OF BI-LEVEL PROGRAMMING

In this section we provide a fairly comprehensive review of past applications of bi-level mathematical programming. The purpose of this section is to demonstrate the wide applicability of bi-level programming and thus its importance as a problem in mathematical programming. Unfortunately, because of the variety of disciplines in which applications occur, we have undoubtedly omitted some important work from our review.

The bulk of applications of bi-level programming that have appeared in the literature is in the economics realm, particularly central economic planning. The typical situation is that there is a planner with some social objective and a set of policy instruments to use for controlling one (or more) economic agents with different objectives.

In the context of the previously defined bi-level problem, the "policy problem" (P1) is given by Eqn. (1a-b), where the planner minimizes $w(x,t)$ subject to the constraints of Eqn. (1b). The planner can only effect his objective by adjusting the vector t, which may be a set of taxes, quotas, or some other instrument. The subordinate problem is given by Eqn. (1c-d) and, following Candler and Townsley (1982), is termed the "behavioral" problem (B1). Given a vector of policies, t, the subordinate agent must optimize his objective $s(x,t)$ by adjusting the vector x. Obviously, whatever x is chosen in the subordinate problem influences the planner's objective.

It is important to realize the distinction between the bi-level problem and the common decomposition of large planning problems into multi-level problems (e.g., Dantzig and Wolfe, 1961; Kornai and Liptak, 1965; Geoffrion, 1970). These methods are all concerned with breaking down a large math program into a number of smaller, more tractable units. An important aspect of these methods is a coincidence between the objectives of the multiple levels and the objective of the overall problem. The fact that the decomposed problem can be written as a single convex programming problem distinguishes decomposition from the general problem considered here.

In the economics literature the subordinate problem (B1) often serves a very specific purpose, i.e., that of a simulator of a market economy. It has been known for some time that the operation of a portion of a competitive economy can be simulated using mathematical programming (Samuelson, 1952; Takayama and Judge, 1971). In short, in a market for a single good, if there are $i = 1,\ldots, I$ consumers each consuming $\hat{q}_i$ and $j = 1,\ldots, J$ producers each producing $\tilde{q}_j$, then a market equilibrium can be associated with the solution $(\hat{\underline{q}},\tilde{\underline{q}})$ to

$$\max_{\hat{\underline{q}},\tilde{\underline{q}}} \sum_{i=1}^{I} \int_0^{\hat{q}_i} P_i(x)dx - \sum_{j=1}^{J} \int_0^{\tilde{q}_j} s_j(x)dx \tag{2a}$$

$$\ni \sum_{i=1}^{I} \hat{q}_i - \sum_{j=1}^{J} \tilde{q}_j \leq 0 \tag{2b}$$

$$\hat{q}_i \geq 0, \quad \forall_i \tag{2c}$$

$$\tilde{q}_j \geq 0, \quad \forall_i \quad , \tag{2d}$$

where $P_i(x)$ is the inverse demand function for consumer i and $s_j(x)$ is the supply or marginal cost function for producer j. This suggests that very often the subordinate problem (B1) in bi-level math programs is a single math program simulating the decentralized market processes of a competitive economy.* The effect of a per-unit tax on such an economy can be simulated by subtracting a term for tax payments from Eqn. (2a). A quota system applied in an economically efficient manner can be simulated by adding appropriate constraints to Eqn. (2).** It is within this framework that most economic applications of bi-level mathematical programming occur: an overall social objective (the planning problem) subject to equilibrium in a market economy (the behavioral problem) with communication between the two levels through taxes, quotas, or some other set of policy instruments.

In spirit, the bi-level problem has a long history in economics -- social objectives vs objectives of individual economic agents. It is difficult to identify the earliest treatment of bi-level problems. Stackelberg's (1952) leader-follower duopoly model is fundamentally a bi-level problem. The leader's problem is P1 and the follower's problem is B1. Marschak (1953) considers the problem of governmental control of a monopolist with zero marginal costs facing linear demand for a single good. The policy problem is that of the government choosing a per-unit tax on the monopolist's output in order to optimize a governmental objective. Two objectives are considered. One is simply to maximize tax revenue. The other is to maximize output subject to a lower limit on tax revenue.

The earliest explicit discussion in the economics literature of bi-level math programming appears to be Candler and Norton (1977a). They consider a numerical example of a milk-producing monopoly in the Netherlands, regulated by the government. The behavioral problem represents the objectives of the

---

*The formulation of Eqn. (2) can obviously be made more complicated. The most common extension is to introduce multiple products and the notion of space where products are distinguished.

**A quota is a restriction on overall output from a particular sector of the economy. Within an optimization model of a competitive economy, it would be represented as a constraint on aggregate output. In practice, the quota would have to be translated to the firm level through a license system or some other mechanism. For an aggregate constraint to realistically represent the action of a quota, the licenses must be allocated to firms in an economically efficient manner. This can be assured by allowing private trading of licenses among firms.

monopoly that seeks to maximize revenue from sales of milk, butter, and cheeses. The Dutch government controls a milk subsidy and duties on imported butter. The government is assumed to have a composite objective consisting of consumer prices, government outlay (the less, the better), and farm income (the more, the better).

Other applications of bi-level programming have been suggested by Candler et al. (1981), principally in the area of development planning. They suggest that the market problem be a model of a sector of a developing economy (such as the agricultural sector), simulating the competitive interactions of economic agents in that sector in response to a number of governmental policies such as price supports or controls, taxes, subsidies, or production quotas. The policy problem can involve a variety of objectives including employment generation, economic growth, nutrition, or simply output. A specific problem examined by Candler et al. (1981) is that of irrigation policy. The behavioral model is of an agricultural region served by a single irrigation system. The behavioral model simulates the decisions of farmers as to how much water to use when subject to policies imposed by administrators of the irrigation system. Farmers are assumed to maximize profit given local prices. Policies considered are (a) system water allocations to be distributed efficiently among the farmers and (b) a cotton production quota. Two objectives are considered: (a) maximization of value-added tax at international prices (as opposed to local prices) and (b) maximization of employment. Although the data used in their analyses were hypothetical, the example illustrates a major area of application of bi-level programming.

Candler and Norton (1977b) have utilized a previously developed large-scale model of Mexican agriculture as the behavioral subproblem. For policy objectives, they examine employment, farm income, corn and wheat production (all to be maximized), and governmental expenses (to be minimized). Policy variables used to influence the subproblem include subsidies on fertilizer use, subsidies on irrigation investment loans, support prices on wheat and corn, and water taxes. The contribution of this work is not only in their realistic policy application but also in their computational experience. Although they came up with improved policies through bi-level optimization, they did discover some nonconvexities in their overall problem which made it difficult for them to find a global optimum (for some objectives). This has turned out to be a significant problem in bi-level programming.

Fortuny-Amat and McCarl (1981) consider the problem of a fertilizer supplier who monopolizes a specific region. Farmers in that region have an inelastic demand but can buy from distant sellers. Thus, the behavioral problem is that of the farmers' decision process. The behavioral problem is complicated by five variations on the basic product--fertilizer. These variations have to do with whether or not fertilizer application equipment is loaned with the fertilizer and whether or not prices are FOB the fertilizer plant or delivered to the farm. The policy problem is that of the monopolist who must decide how much to charge for his product and product variations in order to maximize monopoly rent subject to constraints on availability of capital and labor.

Another set of problems in the area of environmental regulation has motivated this author and apparently Wayne Bialas to research the question of bi-level programming. The problem is to drive polluters to efficient levels of emissions through an emissions tax. The same tax (per unit of emissions) applies to many different sources of pollution in a region even though each source contributes in a different way to concentrations of pollution in the environment, due to locational differences and transport of pollutants by the environment. Thus the subproblem (B1) simulates the market's response to a tax or set of taxes. The policy problem (P1) seeks to minimize real social costs while meeting pollution concentration standards (constraints). This problem was encountered by Bialas for the case of water pollution and Kolstad (1982) for air pollution.

A very different problem was explored by Falk and McCormick (1982): that of a cooperative game. The problem is that of an imperfect cartel of several countries in the international coal market. Since in an imperfect cartel, side-payments are not permitted, cartel objectives may not be to maximize joint profits. Falk and McCormick utilize Nash's solution to this bargaining problem. If u_i is the i^{th} cartel member's gain from joining the cartel (relative to his profit in a noncooperative setting), then the Nash solution is to maximize $\prod_i u_i$, the product of the u_i's. Falk and McCormick formulate this as a bi-level problem, utilizing a very simple competitive model of coal trade as the subproblem B1. The upper-level problem (P1) is Nash's product of individual gains from cartelization, $\prod_i u_i$. Using a numerical example with a two-member cartel, Falk and McCormick demonstrate that two relative maxima exist for the overall problem, only one of which is a global maximum. Kolstad and Lasdon (1985) have examined a similar problem in the same market.

De Silva (1978) has examined the regulation of the oil industry in the United States. The problem is to choose optimal price ceilings for oil discovered before and after a base point in time. The subproblem (B1) is that of a profit-maximizing oil company faced with price ceilings. The policy problem (P1) is that of the federal government choosing price ceilings in order to maximize a composite objective, including the value of oil discovered and produced during the planning period.

Cassidy et al. (1971) analyze the problem of bi-level planning where states (of the US) develop an optimal set of public works projects using money from the federal government. The subproblem (B1) is that of a state deciding on a set of projects which maximize a linear welfare function. The policy problem (P1) is that of allocating resources to each state to optimize a Federal objective couched in terms of the equity of the resource allocation.

There has been a variety of other research concerned with topics closely related to bi-level programming. In the early 1970s a series of articles appeared concerning programs involving the optimal value function of a secondary math program (Bracken and McGill, 1973a, b, 1974a, b; Bracken et al., 1977). All of the applications cited in these papers are in the area of warfare, principally the optimal structure and location of strategic nuclear forces--submarines, bombers, and missiles. If the problem is couched as a two-person Stackelberg game, the subordinate problem (B1) concerns one's opponent's objective (i.e., reducing one's war-making capability and causing other damage) while the upper-level problem (P1) concerns one's own objective (damage to your opponent). For their example of bomber basing, your opponent's goal is to use its submarines to destroy as many of your bombers as possible (problem B1). Your problem is to determine a least-cost bomber location pattern which assures that a given number of bombers survive.

Applications in the area of dynamic Stackelberg games are more remotely related to bi-level programming. Luh et al. (1982) propose constantly varying time-of-day pricing for electric power. They propose the customer as the subordinate agent responding to prices and influenced by a variety of stochastic variables such as the weather. The upper-level decision-maker is the electric power producer who chooses a price at an instant in time so as to clear the market in a least-cost manner.

The large literature on max-min problems is not considered here (see e.g., Danskin, 1966). As will become apparent in the next section, the max-min problem is a special case of a bi-level math program.

III. ALGORITHMIC DEVELOPMENTS

Most of the applications reviewed in the previous section are accompanied by algorithms for solving the particular problem considered. At least a dozen different algorithms appear in the literature, most of which will be discussed in this section. There are three classes into which most algorithms fall. One class of algorithms is concerned exclusively with the linear bi-level problem. These algorithms are concerned with efficiently moving from one extreme point to another until an optimum is found. Another set of algorithms utilizes the Kuhn-Tucker-Karush conditions of the subproblem as constraints on the overall problem, thus turning the bi-level problem into a nonconvex single mathematical program. A third set of algorithms is based on various descent approaches for the policy problem with gradient information from the subproblem acquired in a variety of ways.

A. Extreme Point Search.

All of these methods are concerned with purely linear bi-level problems. All of the algorithms discussed in this section ignore constraints on the upper-level problem (P1). Consequently, we can write the upper-level problem as

$$\text{(P3:)} \qquad \min_{t} \; c_t t + c_x x \tag{3a}$$

$$\ni t \geq 0 \quad , \tag{3b}$$

and the subproblem, implicitly defining x (t), as

$$\text{(B3:)} \qquad \min_{x} \; d_x x \tag{3c}$$

$$\ni A_x x \leq b - A_t t \tag{3d}$$

$$x \geq 0 \quad . \tag{3e}$$

A basic result of Bialas and Karwan (1982) is utilized in most of these algorithms:

Theorem 1 (Bialas and Karwan): Any solution to problem P3-B3 occurs at an extreme point of the constraint set of problem B3.

The various algorithms are concerned with efficient searches of these extreme points. Three algorithms have been discussed in the literature. The algorithm due to Candler and Townsley (1982) is the most widely discussed, principally because of the large number of papers on bi-level programming of which Candler is a coauthor. Other algorithms of this type are due to Bialas and Karwan (1982) and Papavassilopoulos (1982).

1. Candler-Townsley. The Candler-Townsley algorithm is described in some depth in Candler and Townsley (1982) and with more brevity in Bard and Falk (1982). The algorithm focuses on the relationship between P3-B3 and the following LP:

$$\text{(P4:)} \qquad \min_{\bar{x},t} c_t t + c_{\bar{x}} \bar{x} \tag{4a}$$

$$\ni B \bar{x} \leq b - A_t t \tag{4b}$$

$$\bar{x} \geq 0 \tag{4c}$$

$$t \geq 0 \quad . \tag{4d}$$

In P4, B is an "optimal" basis from A_x; i.e., B satisfies optimality conditions for B3 (nonnegative reduced costs). In P4, the vector x has been restricted to $\bar{x}$, corresponding to the columns of A_x in B. Note that with B so defined, any solution of P4 is feasible for P3-B3 (i.e., an optimal solution of B3 that is feasible for P3). The algorithm thus involves moving from one such "optimal" basis B to another, solving P4 each time. If one ensures that the objective of P4 improves, and thus there is no cycling, then the following theorem assures that P3-B3 will eventually be solved.

Theorem 2 (Candler and Townsley): If there exists an optimal solution to P3-B3 (t^*,x^*), then there exists a basis B^* of A_x with nonnegative reduced costs with respect to B3 such that (t^*,x^*,B^*) solves P4.

Thus their algorithm focuses on searching the bases of A_x until a solution of P3-B3 is found. We describe the process intuitively since the details of the search process are quite elaborate. Given a feasible solution to P3-B3 (t_ℓ,x_ℓ) and a corresponding "optimal" basis B_ℓ, solve P4. The nonbasic columns of A_x which have negative reduced costs (with respect to the objective function of P4) are candidates for pivoting into a new basis; denote the set of these columns by T_ℓ. Candler and Townsley prove that any basis $B_{\ell+1}$ which improves the optimal value of P4 (and thus moves closer to an optimum of P3-B3) must contain an element of each of the T_k, $k = 1,\ldots,\ell$. They further define a supplemental set of nonbasic columns of A_x so that one is guaranteed to find a basis which is feasible for P4. Thus, one sequentially changes B in P4 and then solves P4 until a solution to P3-B3 is found.

2. K-th Best Algorithm. Bialas and Karwan (1982) take a slightly different approach focusing on the relationship between P3-B3 and P5:

$$\text{(P5:)} \quad \min_{t,x} c_t t + c_x x \tag{5a}$$

$$\ni A_x x + A_t t \leq b \tag{5b}$$

$$t,x \geq 0 \quad . \tag{5c}$$

Theorem 1 above indicates that a solution of P3-B3 will occur at an extreme point of the constraint set of P5. The "K-th best" algorithm is an efficient way of searching these extreme points. Suppose that the entire set of M extreme points of the constraint set of P5 is enumerated in ascending order of objective function value $[(\hat{t}_1,\hat{x}_1), (\hat{t}_2,\hat{x}_2),\ldots,(\hat{t}_M,\hat{x}_M)]$; i.e., $c_t\hat{t}_i + c_x\hat{x}_i \leq c_t\hat{t}_{i+1} + c_x\hat{x}_{i+1}$. We know one of the extreme points will solve P3-B3. The algorithm moves sequentially through these ordered extreme points until one is found which is an optimal solution to B3. One seeks the index K* where

$$K^* = \min[i\varepsilon(1,\ldots,M)\,|\,(\hat{t}_i,\hat{x}_i) \text{ solves B4}] \quad . \tag{6}$$

Obviously the first of the sequence of extreme points can be found by solving P5 directly. The mechanism for moving from $(\hat{t}_i,\hat{x}_i)$ to $(\hat{t}_{i+1},\hat{x}_{i+1})$ is straightforward. Define $T_i = \{(\hat{t}_k,\hat{x}_k)|k \leq i\}$. Define $W_i = \{(t,x)|\exists(\hat{t},\hat{x})\varepsilon T_i\}$, such that (t,x) is an adjacent extreme point to $(\hat{t},x)\}$. Let $V_i = W_i \,\Omega\, T_i^c$, where T_i^c is the complement of T_i. In other words, W_i is the set of extreme points adjacent* to one of the previously examined extreme points. The set V_i is that set less the previously examined extreme points. It is easy to see that $(\hat{t}_{i+1},\hat{x}_{i+1})$ is the solution to $\min_{x,t}[c_x x + c_t t|(x,t)\varepsilon V_i]$. The algorithm terminates when $(\hat{t}_i,\hat{x}_i)$ solves B3. Since this algorithm approaches the optimal

*At an optimal solution, adjacent extreme points are obtained by pivoting into the basis each of the non-basic variables.

solution from a region of infeasibility, the solution will be a global solution even if P3-B3 is not convex.

3. Papavassilopoulos. In a recent paper, Papavassilopoulos (1982) presents several algorithms for solving the linear bi-level program. Unfortunately it is not clear whether any computational experience with these algorithms exists. We focus on the first of his algorithms.

As in the previous algorithms a sequence of extreme points is generated, each of which will be feasible* for P3-B3. For each extreme point $(\hat{t}_i, \hat{x}_i)$, the next extreme point of the sequence is chosen by examining all extreme points adjacent to $(\hat{t}_i, \hat{x}_i)$. An adjacent extreme point is chosen, $(\hat{t}_{i+1}, \hat{x}_{i+1})$, which (a) strictly reduces the objective of P3 while (b) maintaining optimality of B3. Since this algorithm approaches the solution from a region of feasibility, global optimality is not assured.

B. Kuhn-Tucker-Karush Methods.

A number of algorithms involve transforming the behavioral problem B1 into Kuhn-Tucker-Karush necessary conditions for optimality and then rewriting P1-B1 as

$$(P7:) \quad \min_{x,t,\mu} \quad w(x,t) \tag{7a}$$

$$\ni f(x,t) \leq 0 \tag{7b}$$

$$\nabla_x s(x,t) + \mu' \nabla_x g(x,t) = 0 \tag{7c}$$

$$\mu' g(x,t) = 0 \tag{7d}$$

$$g(x,t) \leq 0 \tag{7e}$$

$$\mu \geq 0 \quad . \tag{7f}$$

Problem P7 is of course equivalent to P1-B1 since any solution to P7 will satisfy Eqns. (7c-f) and thus solve B1 (providing s is strictly quasi-convex with respect to x and g is quasi-convex with respect to x). The difficulty is that the constraint set of P7 is not convex, principally because of Eqn. (7d).

*For (t,x) to be feasible for P3-B3, (t,x) must solve B3 while $t \geq 0$.

Thus most conventional descent algorithms cannot be applied to P7. Also, if the subproblem B1 is large, then the number of constraints and variables associated with the Kuhn-Tucker-Karush conditions will be large. Thus these techniques do not seem well-suited to problems involving large subproblems. The three algorithms presented below solve P7 in different ways. All algorithms consider the case of linear Kuhn-Tucker-Karush conditions for the subproblem.*

1. Bard and Falk. Bard and Falk (1982) consider the linear version of P7 for which the only problematic constraint is Eqn. (7d). The core of their algorithm is to rewrite P7 as a separable convex program; i.e., for $x \in R^n$, all functions can be written as

$$f(x) = \sum_{j=1}^{n} f_j(x_j) \quad .$$

Introducing the variables λ (equal in dimension to g), constraint (7d) is equivalent to

$$\sum_i [\min(0,\lambda_i) + \mu_i] = 0 \tag{8a}$$

$$\lambda_i + g_i + \mu_i = 0, \ \forall_i \tag{8b}$$

$$\lambda_i \geq 0, \ \forall_i \quad . \tag{8c}$$

Although Eqn. (8a) is not a nice smooth function, it does have the separability characteristic which Bard and Falk need to apply an existing algorithm for separable nonconvex programs. The algorithm uses a branch-and-bound technique and involves a partition of the feasible region. Computational tests applied to small problems have produced good results although computations increase rapidly with the size of the constraint region. Thus, the technique may be time consuming when applied to problems involving large subproblems.

2. Fortuny-Amat and McCarl. As did Bard and Falk, Fortuny-Amat and McCarl (1981) focus on the complementary slackness condition, Eqn. (7d). They examine

---

*Bard (1983) has recently proposed an algorithm for solving the general problem P7. His method involves a grid search between estimated upper and lower bounds on w(x,t) in Eqn. 7a.

the case where P1 and B1 are each quadratic programs.* If we assume that each objective function is convex, then if constraint Eqn. (7d) is ignored, problem P7 is a convex program which can be easily solved. Introducing the variable η (with the same dimension as g) such that each η_i is either 0 or 1, P7 can be transformed into

$$\text{(P9:)} \quad \min_{x,t,\mu} \; w(x,t) \tag{9a}$$

$$\ni f(x,t) \leq 0 \tag{9b}$$

$$\nabla_x s(x,t) + \mu' \nabla_x g(x,t) = 0 \tag{9c}$$

$$\mu \leq M\eta \tag{9d}$$

$$g(x,t) \geq - M(1 - \eta) \tag{9e}$$

$$g(x,t) \leq 0 \tag{9f}$$

$$\mu \geq 0 \tag{9g}$$

$$\eta_i = 0 \text{ or } 1 \quad , \tag{9h}$$

where M is a fixed, large positive number. For a fixed η, problem P9 is convex and can be readily solved (since in our example s is quadratic and g linear) for a global optimum. The Fortuny-Amat and McCarl algorithm uses a branch-and-bound technique to enumerate the possibilities for η, solving P9 at each iteration. In commenting on their computational experience, the authors seem to suggest that for large subproblems (i.e., η of large dimension), their algorithm is not very satisfactory.

3. <u>Parametric Complementary Pivot</u>. Problem P7 involves finding x, t, and μ which optimizes the objective function, w (Eqn. 7a). Bialas and Karwan reformulate P7 as that of finding a <u>feasible</u> x, t, and μ such that the objective is less than some upper bound α. By solving the problem with successively smaller upper bounds until no feasible solution can be found, a solution to P7 will obviously be obtained. Thus the reformulated problem is

---

*A quadratic program involves a quadratic objective and linear constraints.

(P10:) Find $x(\alpha)$, $t(\alpha)$, $\mu(\alpha)$

$$\ni w(x,t) \leq \alpha \tag{10a}$$

$$f(x,t) \leq 0 \tag{10b}$$

$$\nabla_x s(x,t) + \mu' \nabla_x g(x,t) = 0 \tag{10c}$$

$$\mu' g(x,t) = 0 \tag{10d}$$

$$g(x,t) \leq 0 \tag{10e}$$

$$\mu \geq 0 \quad . \tag{10f}$$

For fixed α, it is relatively easy to write P10 as the problem of finding $z \geq 0 \ni F(z) \geq 0$, $\langle z,F(z)\rangle = 0$, the complementarity problem (see Cottle and Dantzig, 1974), for which algorithms are available. Although x and t are not explicitly restricted to be nonnegative in P10, they can be easily written as the difference between two nonnegative variables. For convenience, assume $t \geq 0$, $x \geq 0$. Then P10 plus these restrictions on t and x can be written in complementarity form as

(P11:)

$$\langle \alpha - w(x,t) - \nu \geq 0,\ \nu \geq 0\rangle = 0 \tag{11a}$$

$$\langle -\lambda - f(x,t) \geq 0,\ \lambda \geq 0\rangle = 0 \tag{11b}$$

$$\langle \chi \geq 0,\ t \geq 0\rangle = 0 \tag{11c}$$

$$\langle t - \chi \geq 0,\ \chi \geq 0\rangle = 0 \tag{11d}$$

$$\langle -g(x,t) \geq 0,\ \mu \geq 0\rangle = 0 \tag{11e}$$

$$\langle -\nabla_x s(x,t) - \mu' \nabla_x g(x,t) \geq 0,\ x \geq 0\rangle = 0 \quad . \tag{11f}$$

Equations (11a) and (11b) are restatements of Eqns. (10a) and (10b) where "dummy" variables, ν and λ have been introduced to be consistent with complementarity format. Equations (11c) and (11d) (with the dummy χ) are complicated ways of writing $t \geq 0$. Equations (11e-f) correspond to Eqns. (10c-f). Thus for a given α, a solution to P11 $(x,t,\mu,\gamma,\lambda,\chi)$ is feasible for P10. The minimum α for which P11 has a solution will yield the optimal solution to P7 and thus the optimal solution to P1-B1.

Bialas and Karwan apparently only examine the linear version of P7 and use their own algorithm to solve the resulting P11 and to choose successive α. They indicate that their algorithm has worked quite well for the small problem they have examined.

C. Descent Methods.

The workhorses of nonlinear programming have to be the descent methods where first derivative information is used to smoothly approach an optimum. There are two probable reasons these methods have not been more widely used for bi-level programming. One reason is the potential for multiple local optima. Another, possibly more fundamental problem, is the computation of derivatives associated with the subproblem B1. Although techniques for computing derivatives of solutions to mathematical programs with respect to parameters of those programs have been known for some time (see Fiacco and McCormick, 1968), they are not widely used.

Referring back to P1-B1, the basic approach is to apply one of the many descent methods to P1. In P1, x is viewed as a function of t, defined implicitly by B1. Gradients of w and f can be computed if $\nabla_t x^*$ is known. ($\nabla_t x^*$ reflects changes in the solution to B1, x^*, from infinitesimal changes in t.) Of course $x^*(t)$ may not be uniquely defined nor be differentiable at all t, and $\nabla_t x^*$ is unlikely to be continuous. These are potential problems.

1. Penalty/Barrier Function Methods. Shimizu and Aiyoshi (1981) propose rewriting the subproblem B1 as an unconstrained minimization problem through the use of a barrier function. A solution to B1 can then satisfy a stationarity condition of this unconstrained function.

Rewrite B1 as

$$\text{(P12:)} \quad \min_x \{P^r(x,t) \equiv s(x,t) + r\,\phi\,[g(x,t)]\} \quad , \qquad (12)$$

where $r > 0$ and ϕ is continuous and becomes infinite for (x,t) outside the feasible region. Thus if $x^r(t)$ solves P12, then under suitable conditions $\lim_{r\to 0} x^r(t)$ solves B1. Assuming P^r is strictly convex in x, then necessary and sufficient conditions for a solution to P12 are the stationarity conditions

$$\nabla_x P^r(x,t) = 0 \quad . \tag{13}$$

If x is regarded as an implicit function of t, this can be solved for $x^r(t)$ providing $\nabla_{xx}P^r$ is nonsingular. Furthermore,

$$\nabla x^r(t) = -[\nabla^2_{xx}P^r(x,t)]^{-1}\nabla^2_{xt}P^r(x,t) \quad . \tag{14}$$

Problem P1-B1 can now be rewritten as

$$\text{(P15:)} \qquad \min_t \; w(x^r(t),t) \tag{15a}$$

$$\ni f[x^r(t),t] \le 0 \quad , \tag{15b}$$

where $x^r(t)$ and $\nabla x^r(t)$ are given by Eqns. (13) and (14). Many methods are available for solving P15 since derivative information on $x^r(t)$ is available. Shimizu and Aiyoshi (1981) show that if (x^r,t^r) solves P15 then $\lim_{r\to 0}(x^r,t^r)$ solves P1-B1.

This method has been successfully applied to small problems. One difficulty not addressed by the authors is that only local solutions are found using this method.

2. Direct Gradient Methods. De Silva (1978) has utilized a technique in which problem P1 is solved viewing x as a function of t. Given an estimate of t, problem B1 is solved to give both $x(t)$ and $\nabla x(t)$. In contrast to the barrier function approach of Shimizu and Aiyoshi (1981), in De Silva's method $x(t)$ can be computed using any nonlinear programming technique and $\nabla x(t)$ calculated directly using methods developed by Fiacco (1976) for sensitivity analysis. Thus one moves from one t to the next in P15 using any nonlinear programming algorithm that uses first derivative information on w and f. Given a t, any nonlinear programming method can be used to find $x(t)$ and thus $\nabla x(t)$.

A more efficient descent algorithm, particularly appropriate for large problems, has been developed by Kolstad and Lasdon (1985). They focus on the computation of $\nabla x(t)$. If B1 is very large, this can be very difficult to compute. Following Murtagh and Saunders (1981), they partition any solution vector $x^*(t)$ of B1 into components which are at bounds (nonbasic variables--$\tilde{x}^*$) and other components (basic and superbasic variables--$\hat{x}^*$): $x^* \equiv (\hat{x}^*, \tilde{x}^*)$. If strict complementary slackness is assumed, as t changes infinitesimally in B1, only $\hat{x}^*$ will change; the $\tilde{x}^*$ will remain at their bounds. This structuring of the problems greatly facilitates the computation of $\nabla x^*(t)$ since most components are generally nonbasic.

3. Optimal Value Functions. A subclass of the P1-B1 problem has been examined by several authors

(P16:)

$$\min_{t} w(\phi,t) \tag{16a}$$

$$\ni f(\phi,t) \leq 0 \quad , \tag{16b}$$

where

$$\phi(t) = \min_{x} s(x,t) \tag{16c}$$

$$\ni g(x,t) \leq 0 \quad . \tag{16d}$$

Since $\phi(t)$ is defined as the optimal value function of problem B1, we know in general that ϕ is convex (Mangasarian and Rosen, 1964). Thus, in many cases P16 is a strictly convex program which has a unique local optimum. Also, since ϕ is scalar-valued, $\nabla\phi$ is relatively easy to compute. Bracken and McGill (1974b) solve such problems, computing $\nabla\phi$ numerically. Geoffrion and Hogan (1972) examine a problem similar to P16 (actually a problem with multiple subproblems), focusing on calculating the directional derivatives of $\phi(t)$, since $\phi(t)$ is not everywhere differentiable even though it is usually continuous.

REFERENCES

1. Jonathan F. Bard, "An Algorithm for Solving the General Bi-Level Programming Problem," Math. of Ops. Res., 8 (2) 260-272 (1983).

2. J.F. Bard and J.E. Falk, "An Explicit Solution to the Multi-Level Programming Problem," Comput & Ops. Res., 9 (1), 77-100 (1982).

3. W.F. Bialas and M.H. Karwan, "On Two-Level Optimization," IEEE Trans. Auto. Cont., AC-27 (1), 211-214 (Feb. 1982).

4. W.F. Bialas and M.H. Karwan, "Two-Level Linear Programming," unpublished manuscript, Dept. of Industrial Engineering, State University at New York, Buffalo (undated).

5. J. Bracken and J.T. McGill, "A Convex Programming Model for Optimization Problems in the Constraints," Oper. Res., 21, 30-36 (1973).

6. J. Bracken and J.T. McGill, "Mathematical Programs With Optimization Problems in the Constraints," Oper. Res., 21, 37-44 (1973).

7. J. Bracken and J.T. McGill, "Defense Applications of Mathematical Programs with Optimization Problems in the Constraints," Oper. Res., 22, 1086-1096 (1974).

8. J. Bracken and J.T. McGill, "A Method for Solving Mathematical Programs With Nonlinear Programs in the Constraints," Oper. Res., 22 1097-1101 (1974).

9. J. Bracken, J.E. Falk, and F.A. Miercort, "A Strategic Weapons Exchange Allocation Model," Ops. Res., 25 (6), 968-976 (1977).

10. Wilfred Candler and Roger D. Norton, "Multi-Level Programming," World Bank Development Research Center Discussion Paper #20, Washington, DC (Jan. 1977).

11. Wilfred Candler and Roger Norton, "Multi-Level Programming and Development Policy," World Bank Staff Working Paper #258, Washington, DC (May 1977).

12. W. Candler, J. Fortuny-Amat, and B. McCarl, "The Potential Role of Multi-level Programming in Agricultural Economics," Amer. J. Agric. Econ., 63 (3), 521-531 (1981).

13. W. Candler and R. Townsley, "A Linear Two-Level Programming Problem," Comput. & Ops. Res., 9 (1), 59-76 (1982).

14. R.G. Cassidy, J.J.L. Kirby, and W.M. Raike, "Efficient Distribution of Resources Through Three Levels of Government," Mgmt. Sci., 17 (8), pp B462-B473 (April 1971).

15. C.I. Chen and J.B. Cruz, Jr., "Stackelberg Solution for Two-Person Games with Biased Information Patterns," IEEE Trans. Auto. Cont., AC-17 (6), 791-798 (1972).

16. R.W. Cottle and G.B. Dantzig, "Complementary Pivot Theory of Mathematical Programming," in Studies in Optimization, G.B. Dantzig and B.C. Eaves, Eds. (Mathematical Association of America, Buffalo, NY, 1974), pp. 27-51.

17. J.B. Cruz, "Leader-Follower Strategies for Multi-level Systems," IEEE Trans. Auto Cont., AC-23 (2), 244-254 (1978).

18. J.W. Danskin, "The Theory of Max-Min with Applications," J. SIAM Appl. Math. 14 (4), 641-664 (1966).

19. G. Dantzig and P. Wolfe, "Decomposition Algorithm for Linear Programs," Econometrica, 29, 767-778 (1961).

20. James E. Falk and Garth P. McCormick, "Mathematical Structure of the International Coal Trade Model," US Department of Energy Report DOE/NBB-0025, Washington, DC (Sept. 1982).

21. A.V. Fiacco, "Sensitivity Analysis for Nonlinear Programming Using Penalty Methods," Math. Prog., 10, 287-311 (1976).

22. A.V. Fiacco and G.P. McCormick, Nonlinear Programming: Sequential Unconstrained Minimization Techniques (John Wiley, New York, 1968).

23. W. Findeisen, "Decentralized and Hierarchical Control Under Consistency or Disagreement of Interests," Automatica, 18 (6), 647-664 (Nov. 1982).

24. J. Fortuny-Amat and B. McCarl, "A Representation of a Two-Level Programming Problem," J. Ops. Res. Soc., 32, 783-792 (1981).

25. A.M. Geoffrion, "Primal Resource-Directive Approaches for Optimizing Nonlinear Decomposable Systems," Oper. Res., 18 (3), 375-403 (1970).

26. A.M. Geoffrion and W.W. Hogan, "Coordination of Two-Level Organizations with Multiple Objectives," in Techniques in Optimization, edited by A.V. Balakrishnan (Academic Press, London, 1972).

27. C.D. Kolstad, "An Empirical Analysis of Regulatory Efficiency in Air Pollution Control," Los Alamos National Laboratory document LA-UR-81-1727, Los Alamos, NM (Nov. 1982).

28. C.D. Kolstad and L.S. Lasdon, "A Descent Algorithm for Bi-Level Mathematical Programs," unpublished manuscript, Dept. Economics, University of Illinois at Urbana-Champaign (1985).

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

30. Peter B. Luh, Yu-Chi Ho, and Ramal Muralidharan, "Load Adaptive Pricing: An Emerging Tool for Electric Utilities," IEEE Trans. Auto. Cont., AC-27 (2), 320-329 (April 1982).

31. O.L. Mangasarian and J.B. Rosen, "Inequalities for Stochastic Nonlinear Programming Problems," Ops. Res. 12 (1), 143-154 (1964).

32. J. Marschak, "Economic Measurement for Policy and Prediction," in W.C. Hood and T.C. Koopmans (eds.) Studies in Econometric Method, Cowles Foundation Monograph #14 (Yale University Press, New Haven, 1953).

33. Bruce A. Murtagh and Michael A. Saunders, "A Projected Lagrangian Algorithm and its Implementation for Sparse Nonlinear Constraints," Systems Optimization Laboratory Report SOL 80-1R, Department of Operations Research, Stanford University, Stanford, CA (February 1981).

34. G.P. Papavassilopoulos, "Solution of Some Stochastic Nash and Leader-Follower Games," SIAM J. Cont. & Opt., 19 (5), 651-666 (Sept. 1981).

35. G.P. Papavassilopoulos, "Algorithms for Static Stackelberg Games with Linear Costs and Polyhefra Constraints," Proc., IEEE Conf. on Decision and Control, Atlanta, GA, pp 647-652 (Dec. 1982).

36. P.A. Samuelson, "Spatial Price Equilibrium and Linear Programming," Amer. Econ. Rev. 42, 283-303 (1952).

37. K. Shimizu and E. Aiyoshi, "A New Computational Method for Stackelberg and Min-Max Problems by Use of a Penalty Method," IEEE Trans. Auto. Cont., AC-26 (2), 460-466 (April 1981).

38. Anura H. de Silva, "Sensitivity Formulas for Nonlinear Factorable Programming and Their Application to the Solution of an Implicitly Defined Optimization Model of US Crude Oil Production," D. Sc. dissertation, George Washington University, Washington, DC (January 1978).

39. H. von Stackelberg, The Theory of the Market Economy (Oxford University Press, London, 1952).

40. T. Takayama and G. Judge, Spatial and Temporal Price and Allocation Models (North-Holland, Amsterdam, 1971).

☆ U.S. GOVERNMENT PRINTING OFFICE 1985—676-034/20148

Printed in the United States of America
Available from
National Technical Information Service
US Department of Commerce
5285 Port Royal Road
Springfield, VA 22161

Microfiche (A01)

Page Range	NTIS Price Code	Page Range	NTIS Price Code	Page Range	NTIS Price Code	Page Range	NTIS Price Code
001-025	A02	151-175	A08	301-325	A14	451-475	A20
026-050	A03	176-200	A09	326-350	A15	476-500	A21
051-075	A04	201-225	A10	351-375	A16	501-525	A22
076-100	A05	226-250	A11	376-400	A17	526-550	A23
101-125	A06	251-275	A12	401-425	A18	551-575	A24
126-150	A07	276-300	A13	426-450	A19	576-600	A25
						601-up*	A99

*Contact NTIS for a price quote.

Los Alamos