# MATHEMATICAL PROGRAMMING AND THE COMPUTATION OF OPTIMAL TAXES FOR ENVIRONMENTAL POLLUTION CONTROL †

Stephen E. Jacobsen
Engineering Systems Department
School of Engineering and Applied Science
University of California, Los Angeles, Ca. 90024/USA

ABSTRACT

This paper considers some theoretical and computational problems that arise when trying to find optimal taxes for environmental pollution control. The paper takes cognizance of the reality of mixed-economy difficulties (and, therefore, Lagrangian decomposition is not appropriate), and also demonstrates that a "property-rights" approach to environmental quality control may not be appropriate. The paper presents a water quality control problem which highlights the difficulties. In addition, the resulting mathematical program is nonconvex and a solution algorithm is presented.

† This research was supported by the National Science Foundation, Grant ENG 74-02629

## 1. TAXATION AND DECENTRALIZABLE PROGRAMS

Consider the separable mathematical program, denoted by (P),

$$\begin{aligned} \max \quad & \sum_{j=1}^{k} f_j(x^j) \\ \text{s.t.} \quad & \sum_{j=1}^{k} g_j(x^j) \leq 0 \\ & x^j \in X^j, \quad j = 1,\ldots,k \ . \end{aligned} \qquad (P)$$

Here, $x^j$ represents the decision vector of the $j^{\underline{th}}$ subsector and $X^j$ represents the set from which decisions must be drawn. In addition, there is a set of constraints upon resources which link the subsectors. That is, each $g_j$ is a vector-valued function in $R^m$ and $g_j(x^j)$ represents the amounts of the m scarce resources consumed when the $j^{\underline{th}}$ subsector utilizes the decision vector $x^j$. The function

$$\sum_{j=1}^{k} f_j(x^j)$$

is understood to be the "benefit measure" of a central authority (C.A.) as a function of the decisions of the k subsectors.

Let $\lambda = (\lambda_1,\ldots,\lambda_m) \geq 0$ be an m-vector of per unit taxes upon the m scarce resources. Let

$$L_j(x^j, \lambda) = f_j(x^j) - \lambda g_j(x^j) \ .$$

Definition 1: The system (P) is centralized, with respect to $\lambda$, if the decision maker of the $j^{\underline{th}}$ subsector, $j = 1,\ldots,k$, seeks to solve

$$\begin{aligned} \max \quad & L_j(x^j, \lambda) \\ \text{s.t.} \quad & x^j \in X^j \end{aligned} \qquad (P_j(\lambda))$$

when confronted with the nonnegative tax vector $\lambda$. The system (P) is said to be mixed if this is not the case.

Definition 2: We say a centralized system is capable of decentralization if the Lagrangian of (P) possesses a saddle-point.

We now see the meaning of decentralization. Since $(\bar{x}, \bar{\lambda})$ is a saddle-point for (P), it must be the case that $\bar{x}$ is optimal for (P). Therefore, if the C.A. knows an appropriate per unit tax vector, $\bar{\lambda}$, the centralized system assumption implies that the $j^{\underline{th}}$ coordinate vector, $\bar{x}^j$, of $\bar{x}$ is an optimizer of the $j^{\underline{th}}$ subsector's problem $(P_j(\bar{\lambda}))$.

This implies that if $\bar{\lambda}$ presented to each of the subsectors then, if $P_j(\bar{\lambda})$ has a unique optimal solution, the optimal vector $\bar{x} = (\bar{x}^1,..,\bar{x}^k)$ will be recovered by individual subsystem optimization. If some of the subsectors have more than one optimal vector for $P_j(\bar{\lambda})$ it may occur that the vector presented to the C.A., by subsector individual optimization, will not produce an optimal vector for the problem facing the C.A. (i.e., problem P).

Centralized systems capable of decentralization have nice implications for optimization by economic iteration without complete information by the C.A. (e.g., see Dantzig [ 3]). This Walrasian tatonnement interpretation is the basis of dual algorithms for solving problems such as (P) (e.g., see Uzawa [13], Huard [ 6], Falk [ 4]).

## 2. THE BENEFIT MEASURE

Effective intervention in any sphere of economic activity must be based upon a comparison of some notion of benefits and losses associated with various policies. Such, of course, must also be the basis of efforts for environmental pollution control. A C.A.'s policies to achieve a better environment will generally be translated, in the form of higher prices, to the consumers of products produced by "polluting" firms. What constitutes an appropriate measure of the reduction in consumer welfare is a difficult question that is yet to be completely resolved. Traditionally, economists have relied on the consumer's surplus notion. Recently, Willig [14] has given a rather definitive treatment of the notion of consumer's surplus and has shown that under fairly general conditions the consumer's surplus closely approximates the so-called equivalent and compensating variations.

Assume there are n consumer goods under consideration. Let $D_i(p_1,\ldots,p_i,\ldots,p_n)$ be the demand function for the $i^{\underline{th}}$ good as a function of the prices of the n consumer goods. Assume $p^o = (p_1^o,\ldots,p_n^o)$ is the initial price vector and the C.A. has implemented a policy which changes the consumer price vector to $p^1$. The loss in consumer welfare is defined to be

$$\sum_{i=1}^{n} \int_{p_i^o}^{p_i^1} D_i(\bar{p}_i)\, du_i$$

where

$$\bar{p}_i = (p_1^1,\ldots,p_{i-1}^1,u_i,p_{i+1}^o,\ldots,p_n^o) \quad .$$

To this loss we added producers' loss (i.e., change in profits) and C.A.

loss (i.e., change in tax take) to arrive at the expression

$$S(p^o,p^1) = \sum_{i=1}^{n} \int_{p_i^o}^{p_i^1} D_i(\bar{p}_i)du_i + \pi^o - \pi^1 + T^o - T^1$$

as a measure of loss due to the C.A.'s policy which causes a consumer good price vector change from $p^o$ to $p^1$ .

Under independence of demands (i.e., the quantity of the $i^{\underline{th}}$ good demanded depends only upon its own price) and some mild additional assumptions, it can be shown that

$$S(p^o, p^1) = - \sum_{i=1}^{n} \int_{o}^{q_i^1} [D_i^{-1}(u_i) - S_i^{-1}(u_i)]du_i$$

$$+ \sum_{i=1}^{n} \int_{o}^{q_i^o} [D_i^{-1}(u_i) - S_i^{-1}(u_i)]du_i \quad ,$$

where $q_i = D_i(p_i)$ and $S_i^{-1}$ is the marginal cost of production function for the $i^{\underline{th}}$ firm. Therefore, under these assumptions, we see that minimizing total loss is equivalent to maximizing total surplus (i.e., the sum of consumer and producer surplus).

## 3. A WATER POLLUTION EXAMPLE

In this section we present a simple example which highlights the theoretical and computational difficulties of computing optimal taxes. We assume there are k firms located on a one-directional stream and that each firm must dispose of a certain amount of waste which in turn, depends upon the factor usage vector of the firm. For simplicity of notation we assume independence of demands. Let $D_j^{-1}$ denote the inverse demand function for the $j^{\underline{th}}$ firm's product; let $y_j$ denote the product output level of the $j^{\underline{th}}$ firm and let $C_j(y_j, p^j)$ denote the production cost, to firm j, of producing output level $y_j$ when the per unit factor price is $p^j$. Let $x^j$ denote the factor usage vector of firm j and let $\phi_j$ denote the $j^{\underline{th}}$ firm's production function. Then, assuming each firm is a price taker on factor markets, we have that

$$C_j(y_j, p^j) = \min_{x^j \geq 0} \{p^j x^j \mid \phi_j(x^j) \geq y_j\} .$$

Let $W_j(x^j)$ denote the amount of waste which the $j^{\underline{th}}$ firm must dispose of when using factor vector $x^j$. Let $T_j(\alpha_j, W_j(x^j))$ denote the $j^{\underline{th}}$

firm's treatment cost function as a function of the total amount of waste and the percentage-of-treatment variable, $\alpha_j$. For simplicity, we assume waste discharged to the stream consumes only dissolved oxygen and that the C.A. has set maximum allowable dissolved oxygen deficits. Therefore, the mathematical program is

$$\text{maximize} \quad \sum_{j=1}^{k} \left[ \int_{o}^{y_j} D_j^{-1}(z_j)\,dz_j - C_j(y_j, p^j) - T_j(\alpha_j, W_j(x^j)) \right]$$

$$\text{subject to } D_{j\ell}(W_1(x^1), \ldots, W_j(x^j), \alpha_1, \ldots, \alpha_j) \leq \overline{D}_{j\ell}, \quad \ell=0,\ldots,m_j \quad j=1,\ldots,k \qquad \text{(PcA)}$$

$$\phi_j(x^j) - y_j \geq 0, \quad j=1,\ldots,k$$

$$0 \leq \alpha_j \leq 1, \; x^j \geq 0, \; y_j \geq 0 \quad j=1,\ldots,k$$

where $\overline{D}_{j\ell}$ is the greatest allowable dissolved oxygen deficit in reach $j\ell$ between the $j^{\underline{th}}$ and $(j+1)^{\underline{st}}$ firms, $m_j$ is the number of reaches between the $j^{\underline{th}}$ and $(j+1)^{\underline{st}}$ firms, and the set of constraint functions, representing dissolved oxygen deficit, are derived from the Streeter-Phelps equations (e.g., see Revelle, Loucks, and Lynn [11]). To be specific, we are assuming

$$D_{j\ell}(W_1(x^1), \ldots, W_j(x^j), \alpha_1, \ldots, \alpha_1) = \sum_{i=1}^{j} a_{j\ell,i} W_i(x^i)(1-\alpha_i) \qquad (1)$$

where

$a_{j\ell,i}$ = number of pounds of oxygen demanded, in reach $j\ell$ between the $j^{\underline{th}}$ and $(j+1)^{\underline{st}}$ firms, due to one pound of BOD discharged to the stream at location io (the location of the $i^{\underline{th}}$ firm upstream of the $j^{\underline{th}}$ firm),

and

$W_i(x^i)$ = number of pounds of BOD that the $i^{\underline{th}}$ firm must dispose of (by treatment or discharge to stream).

It is clear that these functions crucially depend upon the waste load and treatment decision variables of the firms along the stream. The dissolved oxygen deficit constraint functions are generally <u>non-convex</u> functions of the firms' decision variables $x^j$ and $\alpha_j$, $j=1,\ldots,k$. In fact, even if $W_j$ is a linear function of $x^j$, then $W_j(x^j)(1-\alpha_j)$ is a pseudo-concave and nonconvex function of $(x^j, \alpha_j)$. Therefore, $D_{j\ell}$ given by (1) is not even, generally, a pseudo-concave function of $(x^1,\ldots,x^j, \alpha_1,\ldots,\alpha_j)$ since sums of pseudo-concave functions are not generally pseudo-concave. In fact, in many applications it is also the case that the treatment capital investment functions $T_j(\alpha_j, W_j(x^j))$

exhibit economies-of-scale in the argument $W_j$. That is, $-T_j$ is convex or has convex segments in the argument $W_j$. This implies that the objective function facing the C.A. (the objective function of $P_{cA}$) is not generally a concave function of the decision variables of the problem.

Note also that a significant amount of information appears to be required by the C.A. In particular, knowledge of $D_j^{-1}$, $C_j$, $\phi_j$, and $T_j$ is required. However, under our assumptions, knowledge of $\phi_j$ is equivalent to knowledge of $C_j$(e.g., see Jacobsen [7,8]).

Recall that the major advantage of the decentralization of a centralized system is that system optimization can be carried out without complete information on the part of the C.A. Since the information requirements for the C.A. appear to be quite great for $P_{cA}$, we now investigate whether or not $P_{cA}$ is a centralized system capable of decentralization. Because $P_{cA}$ generally possesses a nonconcave objective subject to numerous nonconvex constraints, no claim can generally be made regarding the existence of a saddle-point. That is, the system generally is not capable of decentralization.

Moreover, if the $j^{\underline{th}}$ firm is confronted with a tax vector $\lambda^*$ it is not the case that the $j^{\underline{th}}$ firm will seek to solve

$$\begin{aligned} &\text{maximize} \quad L_j(y_j,\ x^j,\ \alpha_j;\lambda^*) \\ &\text{subject to} \quad \phi_j(x^j) - y_j \geq 0 \qquad\qquad P_j(\lambda^*) \\ &\qquad\qquad 0 \leq \alpha_j \leq 1,\ x^j \geq 0,\ y_j \geq 0 \end{aligned}$$

where

$$L_j(y_j,\ x^j,\ \alpha_j,\ \lambda^*) =$$

$$= \int_0^{y_j} D_j^{-1}(z_j)dz_j - C_j(y_j,p^j) - T_j(\alpha_j,W_j(x^j)) - \lambda^* a^j(1-\alpha_j)W_j(x^j) \ ,$$

and where $a^j$ is the $j^{\underline{th}}$ column of the set of constraints (1). That is, a profit maximizing firm, say, will attempt to maximize the difference between revenue and cost (production cost plus treatment cost plus the tax cost of waste discharge to the waterway). Therefore, a profit mazimizing firm will, rather, attempt to optimize a function $\bar{L}_j$, subject to the same constraints as in $P_j(\lambda^*)$, where

$$\bar{L}_j(y_j,\ x^j,\ \alpha_j,\ \lambda^*) =$$

$$= D_j^{-1}(y_j)y_j - C_j(y_j,p^j) - T_j(\alpha_j,W_j(x^j)) - \lambda^* a^j(1-\alpha_j)W_j(x^j).$$

Therefore,

$$L_j(y_j, x^j, \alpha j; \lambda^*) - \bar{L}_j(y_j, x^j, \alpha_j; \lambda^*) =$$

$$= \int_o^{y_j} D_j^{-1}(z_j)dz_j - D_j^{-1}(y_j)y_j,$$

the consumers' surplus associated with output level $y_j$. That is, $P_{cA}$ is not a centralized system.

The fact that such systems are neither centralized nor capable of decentralization has, unfortunately, serious consequences for both the Walrasian tatonnement interpretation and algorithmic effectiveness of the various dual algorithms. That is, the rather elegant connection between decentralization and the amount of information needed (by the C.A.) to optimize is lost for mixed economies. Secondly, the dual price procedures, viewed as just computer algorithms for optimization, require the existence of saddle-points. Note that, in principle, the C.A. can centralize the system $P_{cA}$ by offering, to the firms, schedules of the consumers' surplus

$$S_j(y_j) = \int_o^{y_j} D_j^{-1}(u_j)du_j - D_j^{-1}(y_j)y_j$$

However, this raises other distributional problems beyond the scope of the present paper.

## 4. PROPERTY RIGHTS

In an attempt to deal with information requirements, economists have often suggested a "property rights" approach. Briefly, the idea is that a C.A. will provide a fixed number of "pollution permits," each of which allows the owner to dump a fixed number of units (e.g., one) of polluting material. The firms are then free to buy and sell these permits on an open market and, therefore, an equilibrium price will be arrived at. That is, the market mechanism itself will provide the appropriate environmental charge so that environmental services (for instance, the assimilative capacity of a stream) will not be under-valued. This strategy is, perhaps, best articulated by Dales [ 2]. In this section we demonstrate that such a strategy is generally not valid.

We begin by answering the following question:

*If a market for disposal rights provides an equilibrium price, then what measure (if any) of benefits is being maximized by the creation of such a market and price?*

To answer this question, we proceed as follows. Assume there are n firms indexed $i = 1,..,n$. Let $x^i$ be the decision vector for the $i^{\underline{th}}$ firm and let $f_i(x^i)$ be the profit accrued to the $i^{\underline{th}}$ firm when it decides upon the input vector $x^i$. Let $W_i(x^i)$ be the solid waste produced for disposal by the $i^{\underline{th}}$ firm when the $i^{\underline{th}}$ firm is using the input vector $x^i$. Let w be the regional limit (per unit time) upon the quantity of waste that the authority will accept for disposal services.

If there is a price $\lambda$ per unit of waste, then the $i^{\underline{th}}$ firm will seek to solve its own profit maximization problem. That is, the firm seeks to compute its modified optimal profits.

$$\mathscr{L}_i(\lambda) = \text{maximize } f_i(x^i) - \lambda W_i(x^i)$$
$$\text{subject to } x^i \varepsilon X^i ,$$

where $X^i$ is the set of feasible decisions available to the $i^{\underline{th}}$ firm. Let $x^i(\lambda)$ be an optimizing vector for the above optimization problem of the $i^{\underline{th}}$ firm and let

$$w_i(\lambda) = W_i(x^i(\lambda))$$

denote the resulting level of waste offered for disposal. Then,

$$w(\lambda) = \sum_{i=1}^{n} w_i(\lambda)$$

denotes the total quantity demanded for waste disposal services when $\lambda$ is the per unit price for such services. It is well known that the function $\mathscr{L}_i(\lambda)$ is convex in $\lambda$.

Also, by non-negativity of the function $W_i(x^i)$, it can easily be shown that $\mathscr{L}_i(\lambda) \geq \mathscr{L}_i(\gamma)$ and $w_i(\lambda) \geq w_i(\gamma)$ if $\lambda < \gamma$. This implies that $w(\lambda)$ is a non-increasing function of the waste disposal price $\lambda$.

Now assume $\lambda^* \geq 0$ is an equilibrium price. That is, $\lambda^*$ is such that

$$w(\lambda^*) = w$$

or, equivalently, $\lambda^*$ is the price which equates the fixed supply w with demand $w(\lambda)$. Now, when the $i^{\underline{th}}$ decision unit is faced with waste disposal price $\lambda^*$, it will solve its own optimization problem and decide upon vector $x^i(\lambda^*)$. Therefore, for each $i = 1,..,n$, the following condition holds:

$$x^i(\lambda^*) \text{ solves } \max f_i(x^i) - \lambda^* W_i(x^i) \text{ subject to } x^i \in X^i \quad .$$

Also, by assumption, the following two conditions hold:

$$w = w(\lambda^*) = \sum_{i=1}^{n} W_i(x^i(\lambda^*)) \quad ,$$

and

$$0 = \lambda^* \left[ \sum_{i=1}^{n} W_i(x^i(\lambda^*)) - w \right] \quad .$$

Of course, these conditions comprise the statement that $(x^*, \lambda^*) = (x^1(\lambda^*),...,x^n(\lambda^*), \lambda^*)$ is a saddle-point for the optimization problem

$$\max \sum_{i=1}^{n} f_i(x^i)$$

$$\text{subject to} \sum_{i=1}^{n} W_i(x^i) \leq w \qquad (2)$$

$$x^i \in X^i \ , \quad i = 1,..,n \quad .$$

We therefore have an answer to the above question.

> *If a waste disposal authority sets a quantity limit w and if an efficient private market acts to equilibrate supply and demand for the limited disposal services, then the resulting actions of the n decision makers are such that the sum of the individual objective functions of the n decision makers is maximized subject to the authority's quantity constraint.*

The attractiveness of using property rights to generate a market for environmental services is somewhat mitigated in certain situations. For instance, if there are several firms discharging wastes to a waterway and a market is created for which these firms are to purchase limited quantities of the stream's assimilative capacity, our result implies that the sum of the firms' individual objectives will be maximized. However, if some of the firms have considerable monopoly or oligopoly power in their respective product markets, the result will then be that the product output levels of each of the firms will diverge from the optimal levels which would occur if we were to use the benefit measure developed above. Broadly speaking, there seems to be no justification for maximizing the sum of the polluting firms' profits subject to a stream assimilative capacity constraint. This methodological problem of welfare economics seems to have gone unnoticed by Dales.

While the above argument demonstrates that a "property rights" approach may be methodologically incorrect, it is also the case that such an approach may also be technically incorrect. In particular, suppose we approve of maximizing the sum of firms' profits. Then the above argument demonstrates that an equilibrium price implies the existence of a saddle-point for optimization problem (2). But suppose such a problem does not have a saddle-point as is likely to be the case for problems such as $P_{cA}$(i.e., there are economies-of-scale or nonconvex constraints). Then it is clear that an equilibrium price cannot be found by any ordinary market mechanism. For instance, consider the following simple example

$$\begin{aligned} \max \quad & x^2 + y^2 \\ \text{s.t.} \quad & \tfrac{1}{2}x + 2y \leq 3 \\ & 0 \leq x \leq 2 \ , \ 0 \leq y \leq 2 \ . \end{aligned}$$

Then

$$x(\lambda) = \begin{Bmatrix} 2 \ , & \lambda \in [0,4] \\ 0 \ , & \lambda \geq 4 \end{Bmatrix} .$$

Therefore, $w_1(\lambda) = \frac{1}{2}x(\lambda)$ (note that $x(\lambda)$ and, therefore, $w_1(\lambda)$ are not single-valued). Also,

$$y(\lambda) = \begin{Bmatrix} 2 \ , & \lambda \in [0,1] \\ 0 \ , & \lambda \geq 1 \end{Bmatrix}$$

and $w_2(\lambda) = 2y(\lambda)$. Therefore

$$w(\lambda) = \begin{Bmatrix} 5 \ , & \lambda \in [0,1] \\ 1 \ , & \lambda \in [1,4] \\ 0 \ , & \lambda \geq 4 \end{Bmatrix}$$

and hence there is no intersection of the demand curve $w(\lambda)$ with the fixed supply of 3 units. What market behavior would occur, when an equilibrium price does not exist, is not clear.

## 5. A SOLUTION PROCEDURE

The above sections demonstrate that a Walrasian price adjustment process, for the purpose of arriving at prices which will induce firms to act optimally, is virtually useless. In particular, many environmental problems, such as $P_{cA}$, do not correspond to centralized systems capable of decentralization. This implies that a great deal of economic information is required by the C.A. in order to solve $P_{cA}$ and to then set some policy (i.e., taxes) to encourage firms to behave optimally. Nevertheless, it behooves the C.A. to do the best it can as far as a solution of $P_{cA}$ is concerned. The following discussion is based upon that of Hillestad and Jacobsen [5] .

Under some fairly general additional assumptions we sketch an algorithm which is based upon the fact that the constraints of $P_{cA}$ can be converted to a set of convex constraints plus a set of reverse convex constraints (a single constraint $g(x) \leq 0$ is said to be a reverse convex constraint if g is a concave function). The term "reverse convex" is taken from Meyer [10]. Once this conversion is accomplished, a combination of Kelley [ 9] cuts for the convex constraints and Tui [12] cuts for the reverse convex constraints can be used to develop the algorithm.

The constraints of $P_{cA}$ can be converted to the above mentioned form as follows. Let $z_j = W_j(x^j)$, $d_j = 1 - \alpha_j$, $j=1,\ldots,k$. Let $D_{j\ell}$ denote the j x j diagonal matrix whose $i^{\underline{th}}$ diagonal term is $a_{j\ell,i}$. Let

$$Q_{j\ell} = \begin{bmatrix} 0 & D_{j\ell}/2 \\ D_{j\ell}/2 & 0 \end{bmatrix}$$

and let $z^j = (z_1,\ldots,z_j)$, $d^j = (d_1,\ldots,d_j)$. Then the constraint (1) can be rewritten as

$$(z^j,d^j)Q_{j\ell}(z^j,d^j)' \leq \bar{D}_{j\ell} \quad . \qquad (3)$$

Now, by the principle of diagonal dominance, we can express (3) as the difference of convex quadratics. That is, (3) is equivalent to

$$(z^j,d^j)Q^1_{j\ell}(z^j,d^j)' - (z^j,d^j)Q^2_{j\ell}(z^j,d^j)' \leq \bar{D}_{j\ell} \qquad (4)$$

where

$$Q^1_{j\ell} = Q_{j\ell} + \varepsilon_{j\ell}\, I \quad ,$$

$$Q^2_{j\ell} = \varepsilon_{j\ell}\, I \quad ,$$

and $\varepsilon_{j\ell}$ is a number chosen so that

$$\varepsilon_{j\ell} > \max_{i=1,\ldots,j} \left\{ a_{j\ell,i}/2 \right\} \quad .$$

Now, by the introduction of another variable, we can express (4) as

$$\begin{aligned} (z^j,d^j)Q^1_{j\ell}(z^j,d^j)' - u_{j\ell} &\leq \bar{D}_{j\ell} \\ (z^j,d^j)Q^2_{j\ell}(z^j,d^j)' - u_{j\ell} &\leq 0 \qquad (5) \\ -(z^j,d^j)Q^2_{j\ell}(z^j,d^j)' + u_{j\ell} &\leq 0 \end{aligned}$$

and where the first and second constraints of (5) are convex and the third constraint of (5) is of the reverse convex type.

Now, under the assumptions that each production function $\phi_j$ is concave and each waste load function $W_j$ is either convex or concave, the constraints of $P_{cA}$ can be equivalently rewritten as a set of convex and a set of reverse convex constraints as follows:

$$\begin{aligned} (z^j,d^j)Q^1_{j\ell}(z^j,d^j)' - u_{j\ell} &\leq \bar{D}_{j\ell} \\ (z^j,d^j)Q^2_{j\ell}(z^j,d^j)' - u_{j\ell} &\leq 0 \qquad \ell=0,\ldots,m_j \\ & \qquad\qquad j=1,\ldots,k \\ -(z^j,d^j)Q^2_{j\ell}(z^j,d^j)' + u_{j\ell} &\leq 0 \qquad\qquad\qquad (6) \\ z_j - W_j(x^j) &\leq 0 \qquad j=1,\ldots,k \\ -z_j + W_j(x^j) &\leq 0 \\ \phi_j(x^j) - y_j &\geq 0 \qquad j=1,\ldots,k \\ 0 \leq d_j \leq 1,\ x^j \geq 0,\ y_j &\geq 0 \quad . \end{aligned}$$

Similarly, it is easy to show that if $T_j(\alpha_j,z_j)$ is quadratic or is of the form

$$T_j(\alpha_j,z_j) = T_{j1}(\alpha_j)T_{j2}(z_j)$$

then also the treatment cost function terms of the objective of $P_{cA}$ can be incorporated into the constraints in the form of convex and reverse convex functions. Then, by introducing a new variable, the objective can be placed in the constraints and by then bounding the feasible region by a bounded convex polyhedron our problem $P_{cA}$ is converted to the form

$$\begin{aligned} &\max\ cx \\ \text{subject to}\ & h(x) \leq 0 \\ & g(x) \leq 0 \\ & A_o x \leq b_o \\ & x \geq 0 \end{aligned}$$

where h is a vector of convex functions and g is a vector of concave functions.

We let

$$\begin{aligned} Ax &\leq b \\ x &\geq 0 \end{aligned}$$

denote the set of linear constraints at a generic iteration of the algorithm. The algorithm is composed of the following steps:

1) Let $x^o$ be a basic optimal solution for

$$\begin{aligned} &\max\ cx \\ \text{subject to}\ & Ax \leq b \\ & x \geq 0\ . \end{aligned}$$

2) If $h(x^o) \leq 0$ and $g(x^o) \leq 0$, then $x^o$ is optimal. Otherwise, let $h_k(x^o) = \max_i \{h_i(x^o)\}$, $g_\ell(x^o) = \max_j \{g_j(x^o)\}$ and generate a cut for the function $h_k$ or $g_\ell$ which solves

$$\max \{h_k(x^o), g_\ell(x^o)\}\ .$$

3) Append the cut to the linear constraints and return to Step 1. When the cut is to be made for $h_k$, the usual Kelley cut

$$h_k(x^o) + \nabla h_k(x^o)(x-x^o) \leq 0$$

is used. To describe the Tui cut, when the cut is to be made for $g_\ell$, we proceed as follows. Assume A is m x n and that the basic optimal solution $x^o$ (including slacks) is nondegenerate. This implies, since the feasible region for the linear constraints is assumed to be bounded, that the vertex $x^o$ has n neighboring vertices, denoted by $x^{oi}$, i=1,...,n. Let $d^{oi} = x^{oi} - x^o$, i=1,...,n. Let $\bar{\alpha}_j$, i=1,...,n, denote optimal solutions for the n one-dimensional problems

$$\max \ \alpha_i$$
$$\text{subject to} \quad g_\ell(x^o + \alpha_i d^{oi}) \geq 0 \ .$$

The Tui cut is defined by the half-space (not containing $x^o$) whose bounding hyperplane passes through the n linearly independent points

$$z^{oi} = x^o + \bar{\alpha}_i d^{oi}, \quad i=1,\ldots,n \ .$$

That is, compute the hyperplane passing through the points $z^{oi}$ and denote this hyperplane by $ax = \gamma$. If $ax^o < \gamma$, the Tui cut is $ax \geq \gamma$. If $ax^o > \gamma$, the Tui cut is $ax \leq \gamma$. Of course, it is well known that the direction vectors $d^{oi}$ are proportional to vectors of the form

$$(-\bar{a}_{1,m+i},\ldots,-\bar{a}_{m,m+i}, 0,\ldots,0,\underset{\uparrow\, i^{th}\text{ nonbasic coordinate}}{1}, 0,\ldots,0), \quad i=1,\ldots,n,$$

and the latter are readily available in the final simplex tableau which produced $x^o$, the basic optimal solution. Therefore, these latter vectors are to be used as direction vectors rather than pivoting to actually compute the neighboring vertices $x^{oi}$, $i=1,\ldots,n$.

Also, the sequence of cuts does not cut away any portion of the feasible region. In particular, it is well known that a Kelley cut (for a convex constraint) cannot delete any portion of the feasible region. To see that the same is true for a Tui cut we proceed as follows. Let $g_j$ be any concave constraint such that $g_j(x^o) > 0$. Observe that if

$$x \in \text{conv}\,\{x^o, z^{o1},\ldots,z^{on}\}$$

then there exist nonnegative weights whose sum is one and

$$g_j(x) \geq \lambda_o\, g_j(x^o) + \lambda_{o1} g_j(z^{o1}) + \ldots + \lambda_{on} g_j(x^{on}) = \lambda_o g_j(x^o) \quad .$$

Assume the Tui cut is given by $ax \leq \gamma$ (i.e., $ax^o > \gamma$) and let x be any feasible solution. Therefore, $g_j(x) \leq 0$. If $g_j(x) < 0$, then

$$x \notin \text{conv}\,\{x^o, z^{o1},\ldots,z^{on}\}$$

and therefore $ax \leq \gamma$. If $g_j(x) = 0$ and

$$x \in \text{conv}\,\{x^o, z^{o1},\ldots,z^{on}\},$$

it must be the case that x can be written as a convex combination of only $z^{o1},\ldots,z^{on}$ and, therefore, $ax = \gamma$.

Therefore, the cutting plane algorithm above produces a sequence of nonincreasing upper bounds for the optimal value.

It may occur that $x^o$ is degenerate and there will generally be more than n neighboring vertices and, therefore, the Tui cut cannot be unambiguously executed. In fact, it can be shown that if one arbitrarily chooses n directions to form the cut, some part of the feasible region may be deleted. However, in such situations we may use the cut proposed by Carvajal-Moreno [1]. Carvajal's cut is motivated by the Tui cut and it can be shown that both are the same when $x^o$ is nondegenerate.

Even though we have not established a priori convergence conditions, the algorithm does produce a sequence of nonincreasing upper bounds. Moreover, the algorithm has been found to solve complex nonconvex programs (of the type of this paper) and, in the absence of other methods for such problems, appears to be a very useful tool.

Moreover, often lower bounds for $P_{cA}$ can be obtained. For instance, if $T_j$ is linear in $W_j$ and $W_j$ is convex in $x^j$ then $P_{cA}$ is a concave program for a fixed set of $\alpha$'s and is, similarly, a concave program for a fixed set of x's. Therefore, to perhaps obtain a good lower bound, we can fix the $\alpha$'s and optimize over the y's and x's. Then, fixing the y's and x's at the previous optimal values, we optimize over the $\alpha$'s, etc. Such a procedure, in conjunction with the cutting plane algorithm, produces upper and lower bounds for $P_{cA}$, an otherwise relatively intractable problem.

Of course, when the C.A. has obtained optimal or near optimal solutions for $P_{cA}$ it is still faced with the problem of how to set taxes so as to encourage firms to act optimally. About all that can be said at this point is that it may be possible to perform this "coordination" in the following manner. Given all the information needed to solve $P_{cA}$ the C.A. can then find each firm's optimal factor mix and treatment response as a function of tax parameters in the firms' profit functions. One would then hope that there are tax parameters, for each firm, which would force the firms to choose the corresponding optimal values of $P_{cA}$. However, it can be shown that this is generally not the case when considering environmental taxes and therefore the "coordination" problem for such systems is still generally unsolved.

## REFERENCES


1. R. Carvajal-Moreno, "Minimization of Concave Functions Subject to Linear Constraints," Operations Research Center ORC 72-3, University of California, Berkeley, (1972).

2. J. H. Dales, Pollution, Property and Prices, University of Toronto Press, 1968.

3. G. B. Dantzig, Linear Programming and Extensions, Princeton University Press, (1963).

4. J. E. Falk, "Lagrange Multipliers and Nonlinear Programming," Journal of Math Analysis and Applications, 19, 141-159, 1967.

5. R. J. Hillestad and S.E. Jacobsen, "A Cutting Plane Algorithm for Water Quality Control," submitted to J. ORSA.

6. P. Huard, "Convex Programming-Dual Algorithm," Operations Research Center ORC 63-20, University of California, Berkeley, 1963.

7. S. E. Jacobsen, "Production Correspondences," Econometrica, 38 (5), 754-771, 1970.

8. S. E. Jacobsen, "On Shephard's Duality Theorem," Journal of Economic Theory, 4 (3), 458-464, 1972.

9. J. E. Kelley, Jr., "The Cutting-Plane Method for Solving Convex Programs," J. Soc. Indust. Appl. Math., 8, (4), 703-712, (1960).

10. R. Meyer, "The Validity of a Family of Optimization Methods," SIAM J. Control, 8, 41-54, (1970).

11. C. S. Revelle, D.P. Loucks, and W.R. Lynn, "Linear Programming Applied to Water Quality Management," Water Resour. Res., 4, 1-9, (1968).

12. H. Tui, "Concave Programming Under Linear Constraints," Soviet Mathematics, 5, (6), (1964).

13. H. Uzawa, "Iterative Methods for Concave Programming," Chapter 10 of Studies in Linear and Nonlinear Programming, K. Arrow, L. Hurwicz, H. Uzawa (eds.), Stanford University Press, 1958.

14. R. Willig, "Consumer's Surplus: A Rigorous Cookbook," Institute for Mathematical Studies in the Social Sciences, Tech. Report No. 98, Stanford University, (1973).