

Image Space Analysis of Generalized Fractional Programs

JAMES E. FALK¹ and SUSAN W. PALOCSAY²

¹*Department of Operations Research, The George Washington University, Washington, D.C. 20052, U.S.A.*; ²*Department of Information and Decision Sciences, James Madison University, Harrisonburg, VA, 22807, U.S.A.*

(Received: 20 October 1992; accepted: 24 March 1993)

Abstract. The solution of a particular nonconvex program is usually very dependent on the structure of the problem. In this paper we identify classes of nonconvex problems involving either sums or products of ratios of linear terms which may be treated by analysis in a transformed space. In each class, the image space is defined by a mapping which associates a new variable with each original ratio of linear terms. In the image space, optimization is easy in certain directions, and the overall solution may be realized by sequentially optimizing in these directions.

In addition to these ratio problems, we also show how to use image space analysis to treat the subclass of problems whose objective is to optimize a product of linear terms. For each class of nonconvex problems, we present an algorithm that locates global solutions by computing both upper and lower bounds on the solution and then solving a sequence of linear programming sub-problems. We also demonstrate the algorithms described in this paper by solving several example problems.

Key words. Nonconvex fractional problems, sums and products of ratios, global convergence.

1. Introduction

Nonconvex optimization problems occur naturally in a variety of economic, industrial, and engineering problems (see, e.g. Falk *et al.*, 1992; Konno and Inori, 1989; Almqvist and Levin, 1969). The difficulty associated with the existence of multiple local optima had led researchers to study special classes of nonconvex optimization problems (such as the minimization of concave functions) and to define a number of underlying strategies for global optimization, such as enumerative methods, branch and bound, and cutting plane methods (see Horst and Tuy, 1990).

In this paper, we describe a new general strategy that can be applied to nonconvex optimization problems where the objective function is either a sum or a product of two or more ratios of linear functions or a product of linear functions. In our approach, we first transform the problem into an “image space” by associating a new variable with each of the individual ratios or linear functions in the objective function. Then we proceed to solve the problem by analyzing it in the image space where the objective function is more tractable—in certain directions.

The first class of optimization problems which we address has the form

$$\mathcal{P}_1 : \underset{x \in S}{\text{maximize}} f(x) = \sum_{i=1}^m \frac{n_i(x)}{d_i(x)} = \sum_{i=1}^m \left(\frac{c_i^T x + \gamma_i}{h_i^T x + \delta_i} \right)$$

where $S = \{x \mid Ax \leq b, x \geq 0\}$. The numerator and denominator of each ratio in the objective function are linear functions where c_i and h_i are n -component vectors, γ_i and δ_i are constants, and $h_i^T x + \delta_i > 0$.

When problem \mathcal{P}_1 is transformed into its image space, each ratio is mapped into a single coordinate r_i and the objective function becomes one of maximizing the linear function $\sum_{i=1}^m r_i$ over the feasible region in the image space. By sequentially optimizing in the new coordinate directions, we derive an algorithm for problem \mathcal{P}_1 in the next section with an emphasis on illustrating the concept of image space analysis.

In Section 3, we extend our approach to the related class of problems

$$\mathcal{P}_2 : \underset{x \in S}{\text{minimize}} g(x) = \prod_{i=1}^m \frac{n_i(x)}{d_i(x)} = \prod_{i=1}^m \left(\frac{c_i^T x + \gamma_i}{h_i^T x + \delta_i} \right)$$

where $S = \{x \mid Ax \leq b, x \geq 0\}$, c_i and h_i are n -component vectors, γ_i and δ_i are constants, and $h_i^T x + \delta_i > 0$. We present an algorithm similar to the one developed for problem \mathcal{P}_1 where a new variable w_i is associated with each ratio and the transformed problem \mathcal{P}_2 is a problem of minimizing the hyperbolic function $\prod_{i=1}^m w_i$ in image space.

We extend the basic concept of image space analysis in Section 4 to the class of problems referred to as linear multiplicative programs (Konno, Yajima, and Matsui, 1991). These optimization problems have the form

$$\mathcal{P}_3 : \underset{x \in S}{\text{minimize}} h(x) = \prod_{i=1}^m (c_i^T x + \gamma_i)$$

where $S = \{x \mid Ax \leq b, x \geq 0\}$, c_i are n -component vectors, and γ_i are constants.

Under the image space transformation, each linear function in the objective function is treated as a coordinate y_i and problem \mathcal{P}_3 is equivalent to minimizing the hyperbolic function $\prod_{i=1}^m y_i$. We describe an algorithm for problem \mathcal{P}_3 which places upper and lower bounds on the optimal solution and then exploits the linearity of the coordinate functions y_i .

Finally, in Section 5, we summarize our work and briefly discuss additional possible applications of our approach for globally optimizing nonconvex programming problems.

2. Optimizing the Sum of Linear Fractional Functions

In this section, we describe a new algorithm for locating a global solution to problem \mathcal{P}_1 , the problem of maximizing a sum of ratios of linear functions over

linear polyhedra. Although a considerable amount of work has been done on the problem of optimizing a single ratio of linear functions, there have been very few results that address the more general case where the objective function consists of a sum of ratios (Ritter, 1967; Schaible, 1977, 1981).

Almogy and Levin (1971) presented theoretical results for problems where the objective function is a sum of ratios of linear or quasi-concave functions and algorithmic results for problems involving up to three separable linear ratios. However, they use a parametric approach based on the work (for the single ratio problem) of Dinkelbach (1967) and Jagannathan (1966). We have shown this approach to be in error (Falk and Palocsay, 1992).

In another approach to the problem, Cambini *et al.* (1989) developed an algorithm for the case of two ratios based on the concept of “optimal level solutions”, by which they generate a sequence of ever-improving local solutions which must, in a finite number of steps, terminate at the desired globally optimal point. More recently, Konno, Yajima, and Matsui (1991) have proposed a parametric linear programming algorithm for minimizing the sum of two linear fractional functions.

2.1. SOLUTION ALGORITHM

In our algorithm, problem \mathcal{P}_1 is transformed from “ X -space” into an “image space” by mapping each ratio in the objective function into one dimension of the new space. The image of the feasible region is defined as the (generally nonconvex) set

$$T^1 = \{(r_1, \dots, r_m) \mid r_i = n_i(x)/d_i(x), i = 1, \dots, m; \text{ for some } x \in S\}$$

where the functions $d_i(x)$ are assumed to be positive. Figure 1 illustrates this mapping concept for $m = 2$. In this section, we assume that m is equal to two only to simplify the presentation of the algorithm. We discuss the extension of the algorithm for m greater than two in the next section.

After its transformation into the image space, problem \mathcal{P}_1 becomes a problem of maximizing the linear function $r_1 + r_2$ over T^1 . We observe that the new objective function has linear isovalue contours and use this observation to get initial bounds on a subset of T^1 in which the solution is guaranteed to lie.

Initially, we determine the values

$$u_1^0 = \text{maximum}_{x \in S} \frac{n_1(x)}{d_1(x)} \text{ and } u_2^0 = \text{maximum}_{x \in S} \frac{n_2(x)}{d_2(x)}$$

to place upper bounds on r_1 and r_2 , respectively. These values serve as bounds on the image space as shown in Figure 2. Fortunately, these two problems are linear fractional programs which can easily be solved (e.g., by using the Charnes and

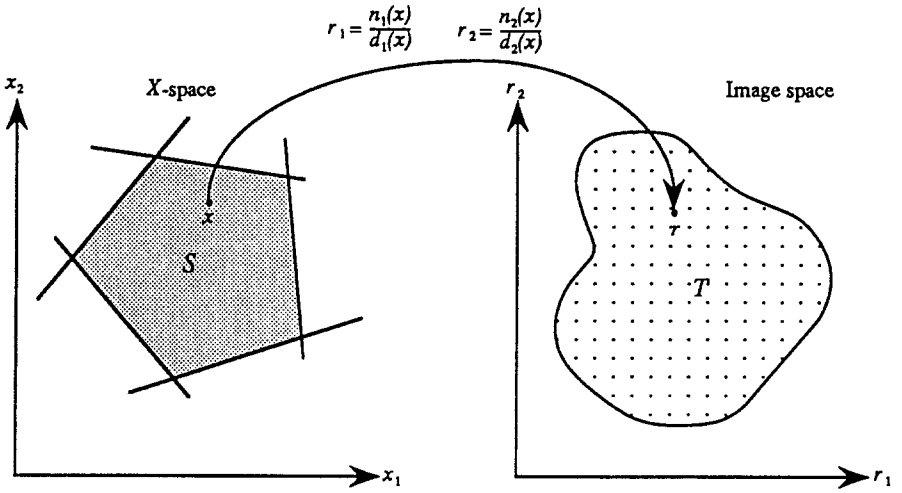


Fig. 1. "X-space" to "image space" mapping concept.

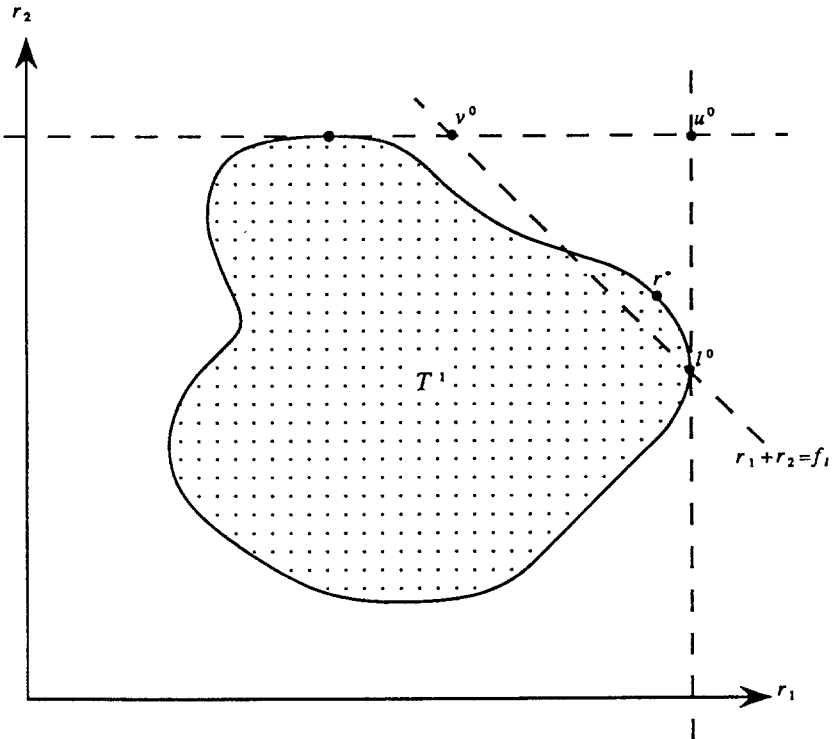


Fig. 2. Initial bounds on the optimal solutions of \mathcal{P}_1 in image space.

Cooper (1962) variable transformation which allows the solution of the linear fractional problem via an equivalent linear program).

Solution points of these two problems, denoted $x^{1,0}$ and $x^{2,0}$ respectively, determine the points $r(x^{1,0})$ and $r(x^{2,0})$ in the image space where

$$r(x^{1,0}) = (r_1(x^{1,0}), r_2(x^{1,0})) = \left(\frac{n_1(x^{1,0})}{d_1(x^{1,0})}, \frac{n_2(x^{1,0})}{d_2(x^{1,0})} \right)$$

and

$$r(x^{2,0}) = (r_1(x^{2,0}), r_2(x^{2,0})) = \left(\frac{n_1(x^{2,0})}{d_1(x^{2,0})}, \frac{n_2(x^{2,0})}{d_2(x^{2,0})} \right).$$

Both of these points are feasible and we identify the one which provides the best lower bound for the subset of T^1 , i.e., we choose the point which gives us

$$f_i = \text{maximum} \{r_1(x^{1,0}) + r_2(x^{1,0}), r_1(x^{2,0}) + r_2(x^{2,0})\}.$$

Note that the isovalue contour $r_1 + r_2 = f_i$, illustrated in Figure 2, determines an initial triangular subset of the image space that will contain the solution. We denote the second and third points of the triangle as l^0 and v^0 where $l^0 = (l_1^0, l_2^0) = (u_1^0, f_i - u_1^0)$ and $v^0 = (v_1^0, v_2^0) = (f_i - u_2^0, u_2^0)$. Thus, we have reduced the size of the feasible region in image space which must be searched and determined both upper and lower bounds on the optimal global solution value.

After locating the three points u^0 , l^0 , and v^0 , we can, in some cases, immediately determine if either l^0 or v^0 (whichever point is feasible) is an optimal solution to the transformed problem \mathcal{P}_1 by applying the following theorem (for the proof, see Falk and Palocsay, 1992).

THEOREM 2.1. *Let $r^1 \in T^1$ and $r^2, r^3 \in E^2$ such that r^1, r^2 , and r^3 are the extreme points of a triangular region in image space. Let $H(r) = \text{maximum}_{x \in S} \{ \sum_{i=1}^2 [n_i(x) - r_i \cdot d_i(x)] \}$. If $H(r^1) = 0$, $H(r^2) < 0$, and $H(r^3) < 0$, then $r^* = r^1$ where*

$$r_i^* = \frac{n_i(x^*)}{d_i(x^*)} \text{ and } x^* \text{ is an optimal solution to problem } \mathcal{P}_1.$$

We can use this result to check for optimality when the parametric function H is zero at either l^0 or v^0 . For example, in Figure 2 l^0 is a feasible point in the image space. If $H(l^0)$ is zero and both $H(v^0)$ and $H(u^0)$ are strictly negative, then we know that r^* is the point l^0 .

If the optimality check fails, then we are interested in efficiently searching for the optimal solution in the bounded subset of T^1 determined during the initialization phase of the algorithm. The iterative method we use is one of

alternately solving two linear fractional programs which maximize r_1 and r_2 , respectively, in the defined region. In each of these fractional programs, we have replaced one of the ratios in the objective function of problem \mathcal{P}_1 with a parameter t and added a constraint on t based on the replaced ratio to obtain the problems

$$\begin{aligned} &\text{maximize } \frac{n_1(x)}{d_1(x)} + t_2 \\ &\text{subject to} \\ &\quad x \in S \\ &\quad t_2 \leq \frac{n_2(x)}{d_2(x)} \quad \text{for fixed } t_2 \end{aligned}$$

and

$$\begin{aligned} &\text{maximize } t_1 + \frac{n_2(x)}{d_2(x)} \\ &\text{subject to} \\ &\quad x \in S \\ &\quad t_1 \leq \frac{n_1(x)}{d_1(x)} \quad \text{for fixed } t_1. \end{aligned}$$

The parameters t_1 and t_2 are chosen at each iteration k so that r_1 and r_2 remain within the limits of the rectangular region defined by the points u^k , v^k , and l^k of the current bounded region and the unlabeled point (v_1^k, l_2^k) . Note that at least one of these two problems is already solved since one of the two points l^k and v^k is both feasible and a lower bound partially defined by either u_1^0 or u_2^0 .

The solutions of these problems are used to slice off pieces of the bounded region either vertically or horizontally and thereby reduce the size of the search space. If we use the Charnes and Cooper transformation for linear fractional programs, we are actually solving a sequence of linear programs. As these iterations continue, the best upper bound does not increase and the best lower bound does not decrease. If the best upper and lower bounds converge to a common value, it must be optimal.

However, in general, the procedure could "stall" and not be able to further reduce the size of the current bounded region. This situation is illustrated in Figure 3. Note that stalling occurs at iteration k only when the two points v^k and l^k are feasible and on the same isovalue contour.

If the procedure stalls, we then (arbitrarily) divide the rectangle $[(v_1^k, l_2^k), l^k, u^k, v^k]$ into two rectangles

$$[(v_1^k, l_1^k), \frac{1}{2}((v_1^k, l_2^k) + l^k), \frac{1}{2}(v^k + l^k), v^k]$$

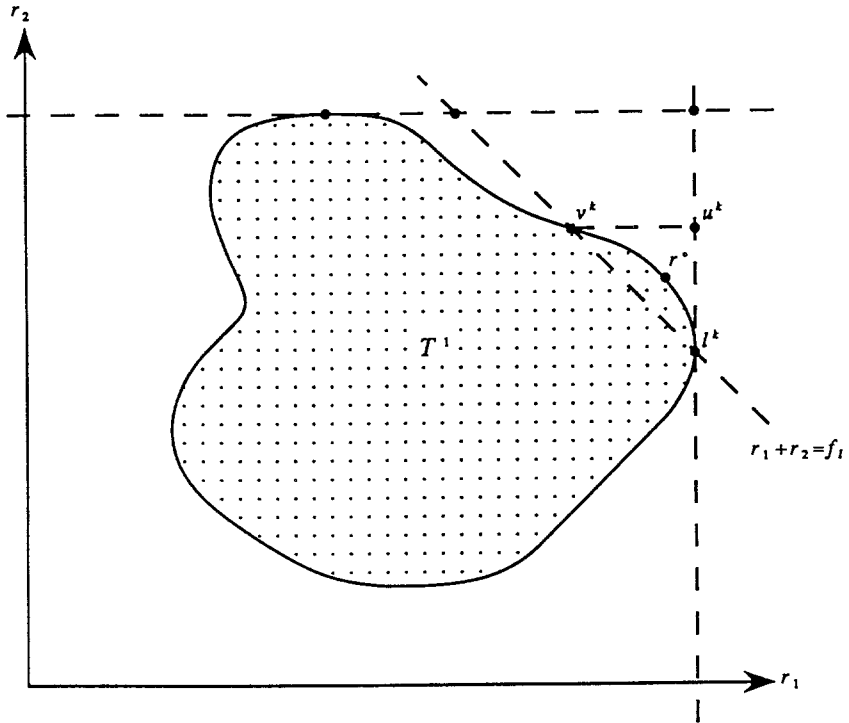


Fig. 3. Stalling in the iterative procedure for \mathcal{P}_1 .

and

$$[\frac{1}{2}((v_1^k, l_2^k) + l^k), l^k, u^k, \frac{1}{2}(v^k + l^k)].$$

We next consider two subproblems defined by the bounds on the two newly-created rectangles. Note that if the best solution of either of these subproblems is better than the feasible points v^k and l^k , the basic procedure will continue with this better point defining a single new triangle (based on a new isovalue contour) which is known to contain the optimal solution. Note also that if the solutions of the new subproblems are both worse than v^k and l^k , then the basic procedure applied to each subproblem will continue.

The only difficulty that remains is in the (unlikely) case that the best solution of one or both of the subproblems gives a value to $r_1 + r_2$ identical to v^k and l^k . In such a case, we would continue to split the rectangle(s) until either one happens on a point which gives a different value to $r_1 + r_2$ than the common value of v^k and l^k , or the area of the region known to contain an optimal ratio is below some

given tolerance. Note that the “area of uncertainty” is, in fact, halved in the worse case scenario where the new split fails to restart a stalled problem.

2.2. EXTENSIONS OF THE ALGORITHM

The extension of our algorithm to the problem of maximizing a sum of m linear ratios when m is greater than two is straightforward although it becomes more difficult to visualize. We are now transforming problem \mathcal{P}_1 into the m -dimensional space where each of the ratios in the objective function is mapped into one of the m dimensions. The transformed problem consists of maximizing the sum of the r_i over T^1 so that the isovalue contours of the objective function are hyperplanes.

We determine an initial upper bound $u_i^0 = \text{maximum}_{x \in S} n_i(x)/d_i(x)$ for $i = 1, \dots, m$ and use the solution points $x^{i,0}$ to determine

$$f_i = \text{maximum}_{1 \leq i \leq m} \left\{ \sum_{j=1}^m \frac{n_j(x^{i,0})}{d_j(x^{i,0})} \right\}.$$

The points where this initial isovalue contour intersects the upper bounds, together with the upper bound point u^0 , define an m -simplex subset of the image space which contains the optimal solution r^* . Theorem 2.1, which gives us a criteria for optimality, immediately extends to the case of m ratios by the convexity property of H proven in Falk and Palocsay (1992).

We extend the iterative steps in the algorithm by replacing all of the ratios in the objective function of \mathcal{P}_1 except one with the parameters t_i and adding the appropriate constraints on these parameters. Thus, each of these problems remains a linear fractional program, and so is equivalent to a linear program. Convergence of the algorithm occurs in the same manner it did for the case of two ratios.

If the algorithm stalls, we apply the same approach described for two ratios. In the extension to the case of m ratios, we use a hyperplane to divide the m -simplex subset into two equal-sized subregions. Then we identify an isovalue contour in each subregion using the initialization steps and choose one of them to restart the iterations if it has a better isovalue contour value than the current one. If not, we continue by executing the algorithm separately in each of the two subregions until we obtain a global optimal solution in each subregion or we find a new improved isovalue contour for continued iterations.

As a final note, we mention that we can easily use our algorithm to address the form of problem \mathcal{P}_1 with an objective to be minimized by converting the objective function into its equivalent maximization form ($\min f$ occurs at the \max of $-f$). Also we point out that while our figures illustrate the algorithm in the first quadrant, in general there is no sign restriction on the ratios in \mathcal{P}_1 .

2.3. EXAMPLE PROBLEMS

The first example problem is

$$\text{maximize } \left\{ \frac{-x_1 + 2x_2 + 2}{3x_1 - 4x_2 + 5} + \frac{4x_1 - 3x_2 + 4}{-2x_1 + x_2 + 3} \right\}$$

subject to

$$x_1 + x_2 \leq 1.5$$

$$x_1 \leq x_2$$

$$0 \leq x_1 \leq 1$$

$$0 \leq x_2 \leq 1.$$

Figure 4 illustrates the feasible region in image space for this problem. The initial upper bounds on r_1 and r_2 are $u_1^0 = 4$ and $u_2^0 = 2.1111$, respectively. Using the solution points $x^{1,0} = (0, 1)$ and $x^{2,0} = (0.75, 0.75)$ obtained from computing these bounds, we identify the feasible points $r(x^{1,0}) = (4, 0.25)$ and $r(x^{2,0}) = (0.6471, 2.1111)$. Thus, the initial linear isovalue contour $r_1 + r_2$ is $v^0 = (2.1389, 2.1111)$, and $u^0 = (4, 2.1111)$.

Computing $H(r)$ at each of these points yields $H(l^0) = 0.0$, $H(v^0) = -5.1806$, and $H(u^0) = -7.4444$. Since $H(l^0)$ is equal to zero and both $H(v^0)$ and $H(u^0)$ are negative, the criteria in Theorem 2.1 are satisfied and we can immediately identify l^0 as an optimal solution point r^* in image space and $x^* = (0, 1)$ as the corresponding optimal solution in X -space.

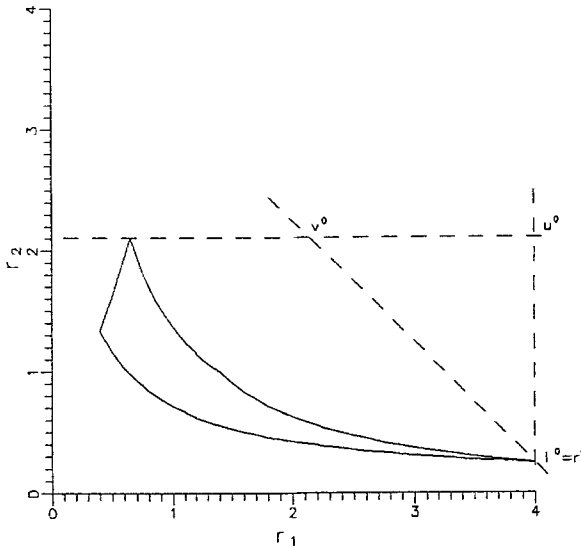


Fig. 4. Graphical representation of the first \mathcal{P}_1 example problem.

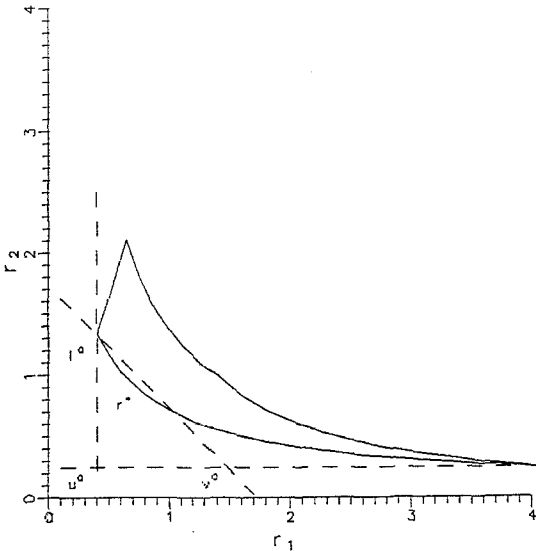


Fig. 5. Graphical representation of the second \mathcal{P}_1 example problem.

Now suppose that we want to minimize the objective function in this problem. After converting the problem into its equivalent maximization form, we find the initial upper bound on (r_1, r_2) is $u^0 = (-0.4, -0.25)$ with corresponding solutions $x^{1,0} = (0, 0)$ and $x^{2,0} = (0, 1)$. The feasible points $r(x^{1,0})$ and $r(x^{2,0})$ are computed as $(-0.4, -1.33)$ and $(-4, -0.25)$, respectively. The initial triangular region is defined by $u^0 = (-0.4, -0.25)$, $l^0 = (-0.4, -1.33)$, and $v^0 = (-1.48, -0.25)$ with isovalue contour $r_1 + r_2$ equal to -1.73 . These bounds are illustrated in Figure 5 for the minimization form of the problem.

Since the optimality check fails for this problem, we begin maximizing in the direction of r_2 to reduce the size of the bounded region containing the optimal solution. However, in this case, the algorithm stalls after approximately twelve iterations and we proceed by following the approach outlined in Section 2.1. The result is identification of a new isovalue contour which is used to determine a new triangular region to restart the steps of the algorithm. After a further reduction of the search space, the algorithm stalls again and we repeat the steps for the stalling procedure to identify the optimal solution at $r^* = (0.6654, 0.9586)$ and $x^* = (0, 0.2839)$.

3. Optimizing the Product of Linear Fractional Functions

In this section, we are also concerned with optimization problems involving ratios of linear functions. However, in problem \mathcal{P}_2 we are interested in optimizing the product of these ratios rather than their sum. We relate this problem to problem

\mathcal{P}_1 and show that we can also apply the image space transformation approach to solve this problem.

Motivated by bond portfolio optimization models (Konno and Inori, 1989), Konno and Yajima (1992) have addressed this problem for two ratios using a parametric simplex algorithm. Their method does not seem to easily extend to products of more than two ratios.

3.1. SOLUTION ALGORITHM

Following the same approach described in Section 2.1 for problem \mathcal{P}_1 , we map each ratio in the objective function of problem \mathcal{P}_2 into a single coordinate w_i and define the image of the feasible region as

$$T^2 = \{(w_1, \dots, w_n) \mid w_i = n_i(x)/d_i(x), i = 1, \dots, m; \text{ for some } x \in S\}.$$

Since the objective function of \mathcal{P}_2 is a product of functions, the isovalue contours of the problem in its image space are hyperbolas of the form $w_1 w_2 = K$ (for $m = 2$). If we assume that all of the ratios are nonnegative, then the image space T^2 is contained in the first quadrant (see Figure 6). In Section 3.2, we discuss the application of the algorithm to cases involving nonpositive ratios.

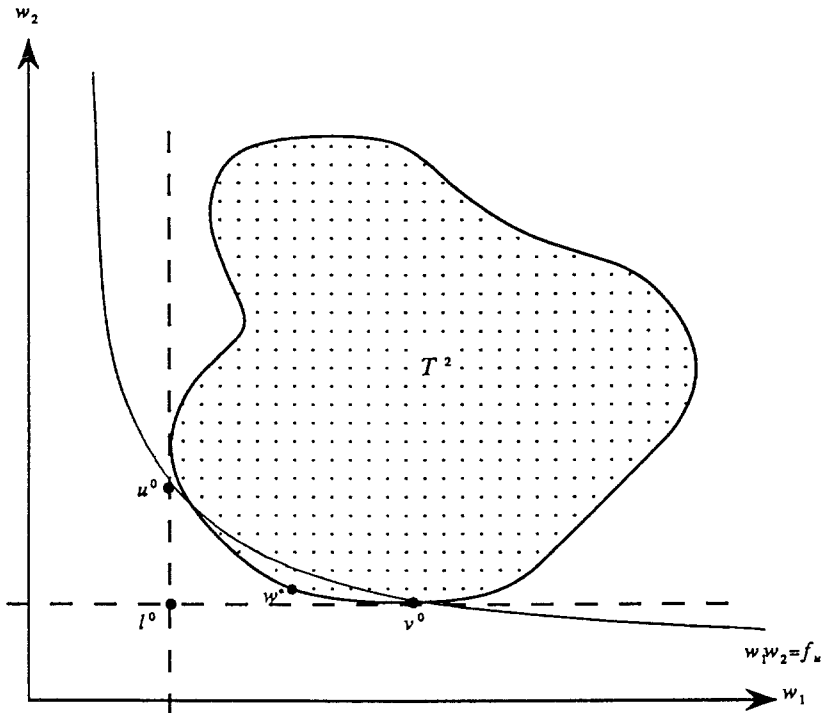


Fig. 6. Initial bounds on the optimal solution of \mathcal{P}_2 in image space.

We get an initial lower bound l^0 on the optimal solution in the image space by solving the linear fractional programs

$$l_1^0 = \text{minimum}_{x \in S} \frac{n_1(x)}{d_1(x)} \quad \text{and} \quad l_2^0 = \text{minimum}_{x \in S} \frac{n_2(x)}{d_2(x)}$$

as before. The solution points for these two problems, $x^{1,0}$ and $x^{2,0}$, define the feasible points $w(x^{1,0})$ and $w(x^{2,0})$ in image space where

$$w(x^{1,0}) = (w_1(x^{1,0}), w_2(x^{1,0})) = \left(\frac{n_1(x^{1,0})}{d_1(x^{1,0})}, \frac{n_2(x^{1,0})}{d_2(x^{1,0})} \right)$$

and

$$w(x^{2,0}) = (w_1(x^{2,0}), w_2(x^{2,0})) = \left(\frac{n_1(x^{2,0})}{d_1(x^{2,0})}, \frac{n_2(x^{2,0})}{d_2(x^{2,0})} \right).$$

An upper bound f_u on the optimal solution value in the image space is provided by one of these two feasible points as

$$f_u = \text{minimum} \{w_1(x^{1,0})w_2(x^{1,0}), w_1(x^{2,0})w_2(x^{2,0})\}.$$

Using the isovalue contour $w_1w_2 = f_u$ and the lower bound l^0 , we determine an initial subset of the image space which is guaranteed to contain the optimal solution as shown in Figure 6. The points u^0 and v^0 in Figure 6 are defined mathematically as $u^0 = (u_1^0, u_2^0) = (l_1^0, f_u/l_1^0)$ and $v^0 = (v_1^0, v_2^0) = (f_u/l_2^0, l_2^0)$.

It turns out that the following result holds (analogous to Theorem 2.1) and therefore we can use this result to determine if either u^0 or v^0 (whichever point is feasible) is an optimal solution to problem \mathcal{P}_2 before continuing on to the iterative phase of the algorithm. The proof of this theorem parallels the one given in Falk and Palocsay (1992) for Theorem 2.1. (Note that the function $G(w)$ is concave whereas the function $H(r)$ is convex.)

THEOREM 3.1. *Let $w^1 \in T^2$ and $w^2, w^3 \in E^2$ such that w^1, w^2 , and w^3 are the extreme points of a triangular region in image space. Let $G(w) = \text{minimum}_{x \in S} \{\sum_{i=1}^2 [n_i(x) - w_i \cdot d_i(x)]\}$. If $G(w^1) = 0$, $G(w^2) > 0$, and $G(w^3) > 0$, then $w^* = w^1$ where*

$$w_i^* = \frac{n_i(x^*)}{d_i(x^*)} \text{ and } x^* \text{ is an optimal solution to problem } \mathcal{P}_2.$$

Now we proceed using the same approach developed for problem \mathcal{P}_1 where we search for the optimal solution by minimizing in the coordinate directions; that is, for the k th iteration, we alternately minimize w_1 and w_2 within the rectangular region defined by the points l^k, u^k , and v^k and the unlabeled point having

coordinates (v_1^k, u_2^k) . We accomplish this by appropriately choosing the parameters t_1 and t_2 in each iteration (initially, $t_1 = v_1^0$ and $t_2 = u_2^0$) and solving the two linear fractional programs

$$\begin{aligned} & \text{minimize } \frac{n_1(x)}{d_1(x)} \cdot t_2 \\ & \text{subject to} \\ & \quad x \in S \\ & \quad t_2 \leq \frac{n_2(x)}{d_2(x)} \quad \text{for fixed } t_2 \end{aligned}$$

and

$$\begin{aligned} & \text{minimize } t_1 \cdot \frac{n_2(x)}{d_2(x)} \\ & \text{subject to} \\ & \quad x \in S \\ & \quad t_1 \leq \frac{n_1(x)}{d_1(x)} \quad \text{for fixed } t_1. \end{aligned}$$

The result of this process is to iteratively tighten the upper and lower bounds until either an optimal solution is obtained or the procedure stalls and cannot improve the current bounds. If the latter occurs, subproblems defined at the iteration are solved to restart the algorithm and to locate a global solution as described in Section 2.1.

3.2. EXTENSIONS OF THE ALGORITHM

We can easily extend this algorithm to the case of more than two ratios by following the general approach outlined in Section 2.2 for the sum of ratios problem.

Note that the scheme which we have described also extends easily to problems of the form of problem \mathcal{P}_2 where one is interested in maximizing, by simple inverting the objective function ($\min f$ occurs at the \max of $1/f$).

However, recall that the algorithm was restricted to cases where all of the ratios are nonnegative. To use the algorithm to solve problems involving nonpositive ratios, we first perform a variable transformation of the form $z_i = -w_i$ on each of the image space variables which correspond to one of these ratios. To illustrate this approach, consider the problem of minimizing the product of two ratios, w_1 and w_2 . The graph of T^2 can lie in any of the four quadrants, depending on the range of values for each of the ratios, and we have four possible cases:

Case 1: $w_1 \geq 0, w_2 \geq 0$

Case 2: $w_1 \leq 0, w_2 \geq 0$

Case 3: $w_1 \leq 0, w_2 \leq 0$

Case 4: $w_1 \geq 0, w_2 \leq 0$.

To determine which of these four cases need to be considered, we solve problem \mathcal{P}_2 in each quadrant of the image space. We use two additional constraints of the form $n_i(x) \geq 0$ if $w_i \geq 0$ and $n_i(x) \leq 0$ if $w_i \leq 0$ to define the boundaries of the quadrant. Under the assumption that the denominators are strictly positive, these constraints restrict the numerators of the ratios, and therefore the original ratios, to be either nonnegative or nonpositive.

If any quadrant has no feasible solutions, we will identify it when we attempt to get lower bounds in that quadrant and eliminate it from further consideration. The global optimal solution for \mathcal{P}_2 will be the minimum over the solutions in the individual quadrants.

To apply the algorithm to the situation in Case 2, we define a new variable $z_1 = -w_1$ and solve the equivalent problem of maximizing the product $z_1 w_2$. The effect of the variable transformation in this case is to multiply the objective function by -1 so that the objective changes from minimize to maximize. We see the same effect in Case 4 where we define $z_2 = -w_2$ and maximize $w_1 z_2$. But in Case 3, where both ratios are nonpositive, we define $z_1 = -w_1$ and $z_2 = -w_2$ and minimize the product $z_1 z_2$.

For the more general problem \mathcal{P}_2 with m ratios, there are 2^m possible cases. Each case will require the addition of m constraints to restrict the numerators $n_i(x)$ of the ratios appropriately. Whenever we perform an odd number of variable transformations of the form $z_i = -w_i$ to solve one of these cases, then the objective is to maximize. Otherwise, the transformed problem will remain a minimization problem.

3.3. EXAMPLE PATTERNS

In this section, we demonstrate how our algorithm is used to solve both minimization and maximization forms of problem \mathcal{P}_2 . First, we consider the example problem from Section 2.3 where the objective function is now the product of two ratios of linear functions:

$$\text{minimize } \left(\frac{-x_1 + 2x_2 + 2}{3x_1 - 4x_2 + 5} \right) \left(\frac{4x_1 - 3x_2 + 4}{-2x_1 + x_2 + 3} \right)$$

subject to

$$x_1 + x_2 \leq 1.5$$

$$x_1 \leq x_2$$

$$0 \leq x_1 \leq 1$$

$$0 \leq x_2 \leq 1.$$

Minimizing in each of the coordinate directions, we find the feasible points

$w(x^{1,0}) = (0.4, 1.3333)$ and $w(x^{2,0}) = (4, 0.25)$ with the corresponding solutions $x^{1,0} = (0, 0)$ and $x^{2,0} = (0, 1)$ in X -space. Therefore the initial hyperbolic isovalue contour $w_1 w_2$ is $f_u = 0.53333$ and the initial lower bound on the optimal solution in image space is $l^0 = (0.4, 0.25)$, as illustrated in Figure 7.

Now we know that the optimal solution lies under the hyperbola $w_1 w_2 = 0.53333$ in the region defined by the lower bound l^0 , the feasible point $u^0 = (0.4, 1.3333)$, and the point $v^0 = (2.1333, 0.25)$ where the hyperbola intersects the lower bound for w_2 . We compute $G(w)$ at each of these points and find $G(u^0) = -0.733$, $G(v^0) = -5.41$, and $G(l^0) = 3.25$ so that Theorem 3.1 does not identify the optimal solution.

Since the criteria for an optimal solution are not yet satisfied, we proceed with the iterative steps of the algorithm. Although we know that the optimal solution lies under the hyperbolic isovalue contour, we must actually search in the rectangular region defined by the points v^0, l^0, u^0 , and (v_1^0, u_2^0) to maintain linear constraints on the problem. In this case, we do not need to minimize in the direction of w_1 since this problem was solved during initialization (i.e., $u^k = u^0$ for each iteration k). So our next step is to minimize w_2 subject to the additional constraint $w_1 \leq 2.1333$ (in linear form). The solution of this linear fractional program, the feasible point $(2.1333, 0.4007)$, moves the linear bound on w_2 from 0.25 and 0.4007 and determines $v^1 = (1.3311, 0.4007)$ and $u^1 = (0.4, 0.4007)$.

Now we repeat the process of minimizing in the direction of w_2 (with the appropriate constraint on w_1) and finding a tighter lower bound on w_2 until the lower bound on w_2 converges to 1.3333 and we identify the optimal solution at

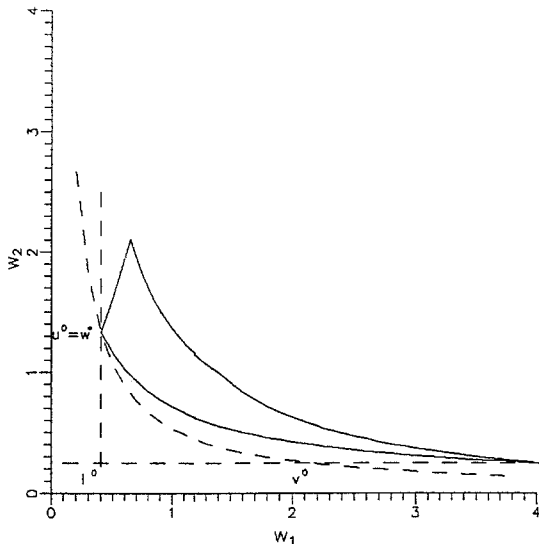


Fig. 7. Graphical representation of the first \mathcal{P}_2 example problem.

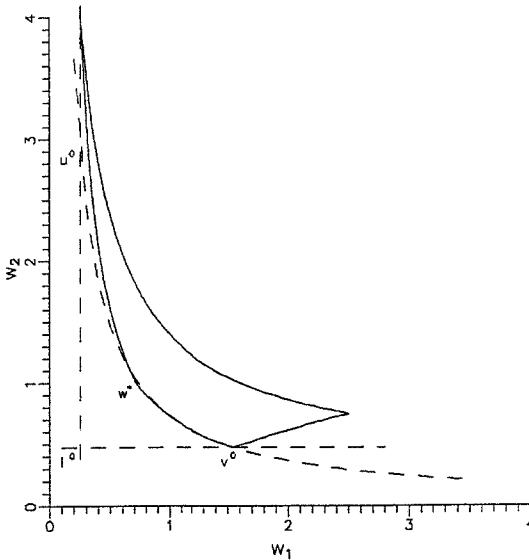


Fig. 8. Graphical representation of the second \mathcal{P}_2 example problem.

$w^* = (0.4, 1.3333)$. Thus, the corresponding point $x^* = (0, 0)$ is the global optimal solution to the problem.

If we consider the problem of maximizing (rather than minimizing) the objective function of this example problem, then we must first convert the problem to its equivalent minimization form:

$$\text{minimize} \left(\frac{3x_1 - 4x_2 + 5}{-x_1 + 2x_2 + 2} \right) \left(\frac{-2x_1 + x_2 + 3}{4x_1 - 3x_2 + 4} \right)$$

The feasible region of the image space and the initial bounds for the new problem are illustrated in Figure 8. The algorithm locates the optimal solution at $w^* = (0.7143, 1)$ so that the global solution to the max problem is $1/w^* = (1.4, 1)$ at $x^* = (0.5, 1)$.

4. Optimizing the Product of Linear Functions

The basic concept underlying the algorithms presented in Sections 2 and 3 for problems \mathcal{P}_1 and \mathcal{P}_2 is the analysis of the optimization problem in an image space. In this section, we apply this concept to problem \mathcal{P}_3 , the problem of minimizing a product of linear functions over linear polyhedra. While this is a special case of the product-of-ratios problem addressed in Section 3, the special structure here allows for a more efficient resolution.

Recently, Konno, Yajima, and Matsui (1991) have addressed this class of

problems, referred to as linear multiplicative problems. However, their parametric simplex algorithm only treats problems involving two linear functions.

4.1. SOLUTION ALGORITHM

When we map problem \mathcal{P}_3 into its image space, each linear function in the objective function corresponds to a single variable y_i . The set T^3 , defined as

$$T^3 = \{(y_1, \dots, y_m) \mid y_i = c_i^T x + \gamma_i, i = 1, \dots, m; \text{ for some } x \in S\},$$

represents the image of the feasible region. Unlike T^1 and T^2 , the set of T^3 is a convex set (actually a polytope) because the mapping $y_i = c_i^T x + \gamma_i$ is a linear mapping. And since the objective function of \mathcal{P}^3 is quasiconcave (see Section 4.2 for a proof), if \mathcal{P}_3 has a solution, it must have a solution at an extreme point. We take advantage of this special structure in developing an algorithm for problem \mathcal{P}_3 .

Note that the isovalue contours of the transformed objective function are hyperbolas of the form $y_1 y_2 = K$ (for $m = 2$). Under the assumption that all of the linear functions are nonnegative, the image space T^3 is located in the first quadrant (see Figure 9). We will consider the more general case, where T^3 may overlap another quadrant, in the next section.

In the first phase of the algorithm, we determine lower bounds l_1 and l_2 on the optimal solution by solving the two linear programs

$$l_1 = \text{minimum}_{x \in S} c_1^T x + \gamma_1 \quad \text{and} \quad l_2 = \text{minimum}_{x \in S} c_2^T x + \gamma_2.$$

The solutions of these two problems, denoted by x^1 and x^2 , determine the feasible points

$$y(x^1) = (y_1(x^1), y_2(x^1)) = (c_1^T(x^1) + \gamma_1, c_2^T(x^1) + \gamma_2)$$

and

$$y(x^2) = (y_1(x^2), y_2(x^2)) = (c_1^T(x^2) + \gamma_1, c_2^T(x^2) + \gamma_2)$$

in the image space T^3 .

We now determine an initial upper bound on the optimal solution as

$$f_u = \text{minimum} \{y_1(x^1)y_2(x^1), y_1(x^2)y_2(x^2)\}.$$

Clearly the optimal solution (in image space) satisfies $y_1 y_2 \leq f_u$ and $y_1 \leq l_1, y_2 \geq l_2$. (See Figure 9). The smallest triangle (in general, simplex) containing this region has vertices $l, u,$ and v where $u = (u_1, u_2) = (l_1, f_u/l_1)$ and $v = (v_1, v_2) = (f_u/l_2, l_2)$.

Let $a_1 y_1 + a_2 y_2 = 1$ determine the line through the points u and v . Now we

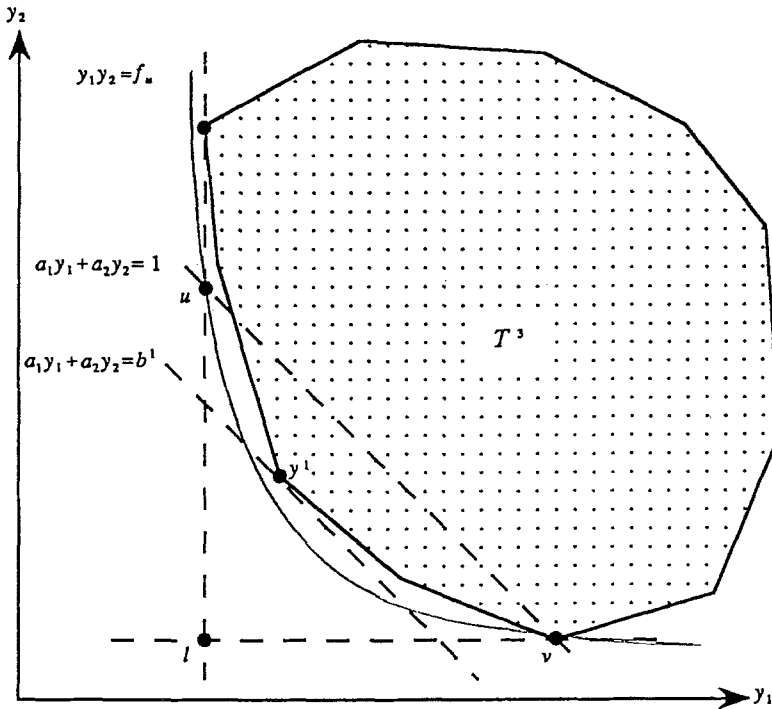


Fig. 9. Initial bounds on the optimal solution of \mathcal{P}_3 in image space.

replace the objective function in problem \mathcal{P}_3 with the linear function associated with this equation and solve the linear program

$$\text{minimize}_{x \in S} a_1 y_1 + a_2 y_2$$

where $y_1 = c_1^T x + \gamma_1$ and $y_2 = c_2^T x + \gamma_2$.

The optimal solution and objective function value of this linear program in image space are denoted y^1 and b^1 , respectively. If $b^1 = 1$, then we are done with the optimal solution to \mathcal{P}_3 , y^* , equal to either u or v , whichever point is feasible, since the entire feasible region must satisfy $a_1 y_1 + a_2 y_2 \geq b^1$. Otherwise, $b^1 < 1$ and we know that the optimal solution of \mathcal{P}_3 lies in the region between the two lines $a_1 y_1 + a_2 y_2 = 1$ and $a_1 y_1 + a_2 y_2 = b^1$, and under the hyperbola $y_1 y_2 = f_u$. (If the feasible point y^1 defines a better hyperbolic isovalue contour, we replace f_u with the new value and revise the coordinates of u and v .)

When $b^1 < 1$, the optimal solution must satisfy $a_1 y_1 + a_2 y_2 \geq b^1$ and $y_1 y_2 \leq f_u$ as well as the bounds $y_1 \geq l_1$ and $y_2 \geq l_2$. The set of such points is composed of two disjoint sets (see the shaded areas of Figure 10), one of which is enclosed in the triangle with vertices u , t^1 , and t^{12} , while the other is enclosed in the triangle v , t^2 , and t^{21} . Each of these is treated as above.

In Figure 10, the minimum in the area under the line between t^{21} and v occurs

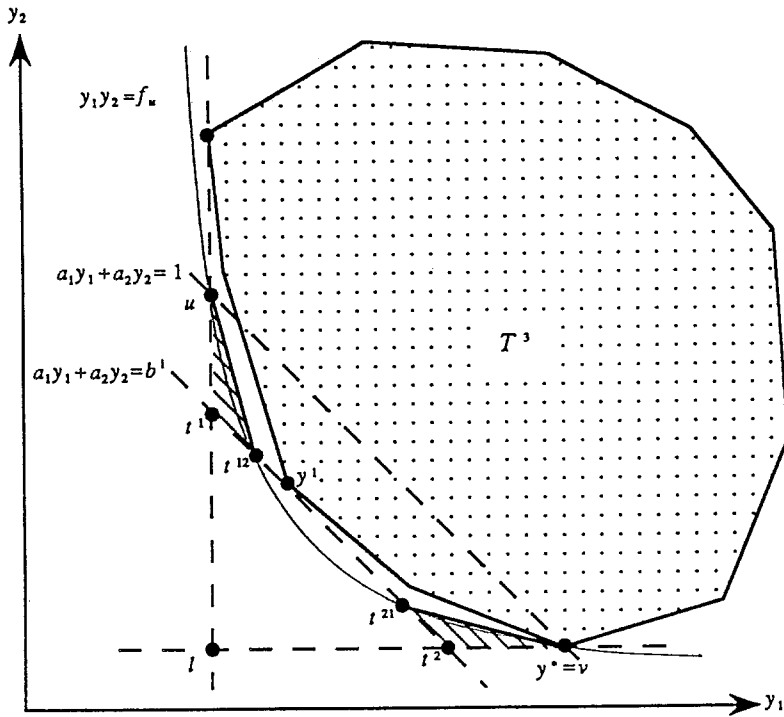


Fig. 10. Additional bounds on the optimal solution of \mathcal{P}_3 in image space.

at v so we are done in that area. But, since the area under the line between t^{12} and u does not contain any feasible points, the result of minimizing in this area will be the point y^1 . Thus, the global optimal solution must occur at $y^* = v$.

In general, we repeat the process of determining disjoint triangular subsets and minimizing their associated linear functions until we identify one of the vertices in T^3 as a global minimum. In each iteration, we will either locate a new vertex or eliminate an area from further consideration (either because we have found the optimal solution in that area or because the area does not contain any feasible points. Since the optimal solution is at an extreme point, the process is finite. Conceivably, this process could lead to a sequence of 2, 4, 8, ... linear subproblems which need to be solved. None of the example problems which we generated required the solution of more than four subproblems in total since the algorithm quickly eliminated empty subsets.

4.2. EXTENSIONS OF THE ALGORITHM

For problems involving $m > 2$ ratios, the extension is not difficult. As before, let $y_i = c_i^T x + \gamma_i$ and let $l = (l_1, \dots, l_m)$ denote a vector of lower bounds defined by $l_i = \text{minimum}_{x \in S} c_i^T x + \gamma_i$. We assume $l_i > 0$ for $i = 1, \dots, m$. Let x^i solve this problem, and set

$$f_u = \text{minimum}_{1 \leq i \leq m} \left\{ \prod_{j=1}^m c_j^T x + \gamma_j \right\}.$$

Clearly the optimal solution y^* (in image space) must satisfy the inequality $\prod_{j=1}^m y_j \leq f_u$ and also the bounds $y_j \geq l_j$. Note that $\prod_{j=1}^m l_j \leq f_u$ and if equality occurs, $y^* = f_u$.

When $\prod_{j=1}^m l_j < f_u$, set $z^i = l + \tau_i e^i$ where e^i is the i th unit vector and

$$\tau_i = \frac{f_u - \prod_{j=1}^m l_j}{\prod_{j \neq i}^m l_j}.$$

The points z^1, \dots, z^m all lie on the surface $\prod_{j=1}^m y_j = f_u$. It is easy to determine a vector a such that $\prod_{j=1}^m a_j y_j = 1$ for $y = z^j$ ($j = 1, \dots, m$). We now solve the (linear) program

$$\text{minimize}_{x \in S} \sum_{j=1}^m a_j (c_j^T x_j + \gamma_j)$$

and obtain a value b^1 . In image space, the optimal solution y^* must satisfy $\sum_{j=1}^m a_j y_j \geq b^1$ in addition to the inequality $\prod_{j=1}^m y_j \leq f_u$. If $b^1 = 1$, we are done. If the optimal solution y^1 of this problem satisfies $\prod_{j=1}^m y_j^1 < f_u$, we replace f_u by $\prod_{j=1}^m y_j^1$ and continue.

The only other possibility is that $\prod_{j=1}^m y_j^1 > f_u$ but $b^1 < 1$. In this case, the feasible region (in image space) consists of m subsets, each of which must be investigated as candidates which could contain the global optimizer. Each subset S^i contains the point t^i where the line segment $[l, z^i]$ intersects the hyperplane $\sum_{j=1}^m a_j y_j = b^1$. (See Figure 11). The point z^i is also in subset S^i .

Now consider the point t^i as the vertex of a cone where extreme rays pass from t^i through t^j ($j \neq i$). There are $n - 1$ of these, and they pierce the surface $\{y \mid \prod_{j=1}^m y_j = f_u\}$ in points t^{ij} .

The subset S^i is contained in the simplex defined by the points $\{t^i, z^i, t^{ij}$ ($j \neq i\})$. For example, in Figure 11 one simplex is determined by t^1, z^1, t^{12} , and t^{13} . The algorithm continues by minimizing a linear function whose isovalue contours are parallel to the hyperplane passing through the points z^i, t^{ij} ($j \neq i$), all of which lie on the surface $\{y \mid \prod_{j=1}^m y_j = f_u\}$. As there are only a finite number of extreme points, the process is finite.

In general, the extension of this algorithm to the case where $m \geq 3$ is straightforward when all terms are known to be nonnegative. If problem \mathcal{P}_3 involves nonpositive terms, then we use the approach described in Section 3.2 for problem \mathcal{P}_2 to perform variable transformations on the nonpositive functions.

We note that a global optimal solution can be obtained directly for problems of

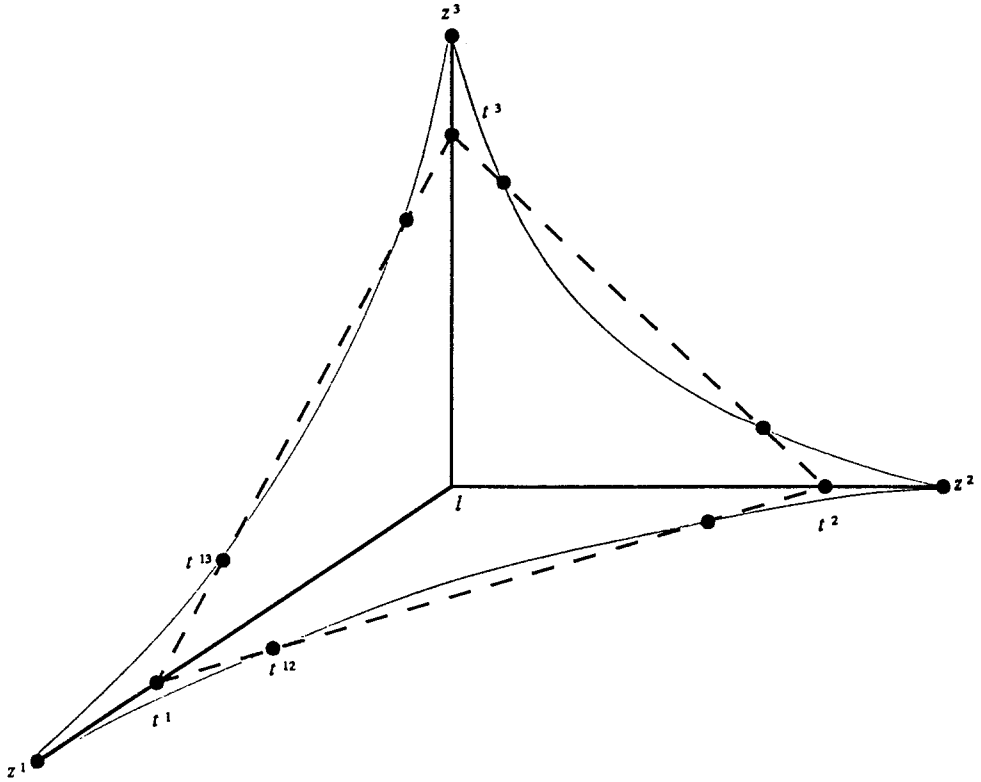


Fig. 11. Extension of the algorithm for \mathcal{P}_3 .

the form of \mathcal{P}^3 where the objective is to maximize rather than minimize by using standard nonlinear programming techniques. It is well-known that quasi-concave programs share the property with convex programs that a local optimum is a global optimum (see, e.g., Mangasarian, 1969). Since the constraints of \mathcal{P}_3 are linear, we only need to show that the objective function is quasi-concave to apply this property.

THEOREM 4.1. *The function $h(x) = \prod_{i=1}^m (c_i^T x + \gamma_i)$ where $c_i^T x + \gamma_i > 0$, c_i are n -component vectors, and γ_i are constants is quasi-concave.*

Proof. Define a level set of h as $L = \{x \in S \mid h(x) \geq \tau\}$. Since $h(x)$ is quasi-concave if and only if all of its level sets are convex, it suffices to show that L is a convex set of all τ . But the set equals

$$\left\{ x \in S \mid \sum_{i=1}^m \ln(c_i^T x + \gamma_i) \geq \ln \tau \right\}$$

and this set is convex since each function $\ln(c_i^T x + \gamma_i)$ is a concave function. \square

4.3. EXAMPLE PROBLEMS

In this section, we present several examples to demonstrate the algorithm for minimizing a product of linear functions. Consider the following problem due to Konno and Kuno (1989), where the second linear term in the objective function has been modified to make the image space variable w_2 nonnegative:

$$\text{minimize } (x_1 + x_2)(x_1 - x_2 + 7)$$

subject to

$$2x_1 + x_2 \leq 14$$

$$x_1 + x_2 \leq 10$$

$$-4x_1 + x_2 \leq 0$$

$$2x_1 + x_2 \geq 6$$

$$x_2 + 2x_2 \geq 6$$

$$x_1 - x_2 \leq 3$$

$$x_1 \leq 5.$$

To get initial bounds on the optimal solution in image space, we minimize in each of the coordinate directions and find $y(x^1) = (4, 7)$ and $y(x^2) = (10, 1)$ at $x^1 = (2, 2)$, $x^2 = (2, 8)$. The best lower bound is determined to be $l = (4, 1)$ and the initial hyperbolic isovalue contour is $y_1 y_2 = 10$, with $v = (10, 1)$ and $u = (4, 2.5)$. Figure 12 illustrates the image space for this problem.

Next we determine the line through the points u and v , $0.0714y_1 + 0.2857y_2 =$

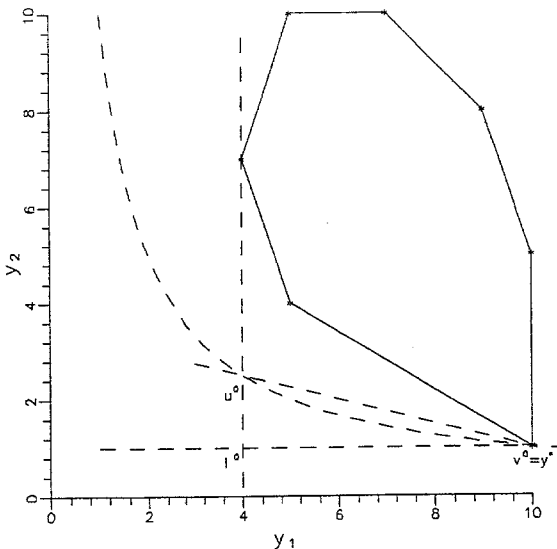


Fig. 12. Graphical representation of the first \mathcal{P}_3 example problem.

1. In this case, the minimum of $0.0741y_1 + 0.2857y_2$ occurs at v , so the global optimal solution is $y^* = (10, 1)$ at $x^* = (2, 8)$.

In the next example, we modify the feasible region to make the problem more difficult:

$$\begin{aligned} & \text{minimize } (x_1 + x_2)(x_1 - x_2 - 7) \\ & \text{subject to} \\ & \quad 2x_1 + x_2 \leq 14 \\ & \quad x_1 + x_2 \leq 10 \\ & \quad 1.44x_1 + x_2 \geq 4.89 \\ & \quad -1.58x_1 + x_2 \leq 5.65 \\ & \quad -1.03x_1 + x_2 \leq 5.93 \\ & \quad x_1 + 2x_2 \geq 6 \\ & \quad x_1 - x_2 \leq 3 \\ & \quad x_1 \leq 5. \end{aligned}$$

The initial bounds remain unchanged, but when we minimize $0.0714y_1 + 0.2857y_2$ we now get the point $(y_1, y_2) = (5, 1.5)$ as the optimal solution to this linear program (rather than v) with an objective function value of 0.7857. This new point is feasible in the image space and defines a new hyperbola, $y_1y_2 = 7.5$, which is better than the initial one. Using the new hyperbola and the lower bounds on y_1 and y_2 , we determine $u^1 = (4, 1.875)$ and $v^1 = (7.5, 1)$. At this point, we know that the optimal solution must lie above the line $0.0714y_1 + 0.2857y_2 = 0.7857$, which intersects the new hyperbola at two points, $y^1 = (5, 1.5)$ and $y^2 = (6, 1.25)$, as shown in Figure 13, and under the hyperbola $y_1y_2 = 7.5$.

Now use y^1 (with u^1) and y^2 (with v^1) to construct two lines, one for each area of uncertainty. The minimum in the first area is y^1 , so we are done in this area. But in the second area, the minimum is $(y_1, y_2) = (7, 1.05)$ so that the optimal solution in this area now lies above the line $0.0741y_1 + 0.4444y_2 = .9852$ and under a new hyperbola, $y_1y_2 = 7.35$. We continue by using the intersection points to determine the next two lines. When we solve the two associated linear programs, we find that the optimal solution in the second area, and the global optimal solution, is $y^* = (7, 1.05)$ at $x^* = (0.525, 6.475)$.

Now we return to Konno and Kuno's original problem where T^3 , the image space of the feasible region, lies in all four quadrants, as shown in Figure 14. The mathematical formulation of this problem is

$$\begin{aligned} & \text{minimize } (x_1 + x_2)(x_1 - x_2) \\ & \text{subject to} \\ & \quad 2x_1 + x_2 \leq 2 \\ & \quad x_1 + x_2 \leq 2 \end{aligned}$$

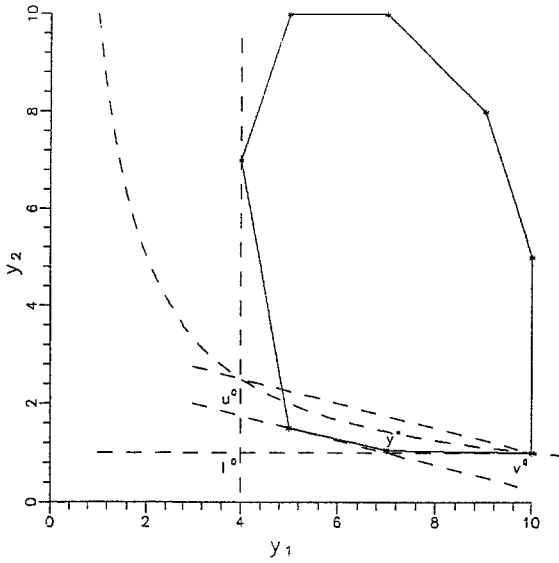


Fig. 13. Graphical representation of the second \mathcal{P}_3 example problem.

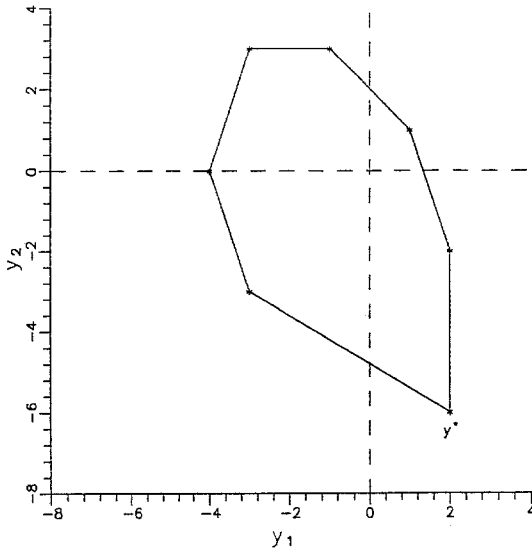


Fig. 14. Graphical representation of the third \mathcal{P}_3 example problem.

$$\begin{aligned}
 -4x_1 + x_2 &\leq 12 \\
 -2x_1 - x_2 &\leq 6 \\
 -x_1 - 2x_2 &\leq 6 \\
 x_1 - x_2 &\leq 3 \\
 x_1 &\leq 1.
 \end{aligned}$$

Following the approach described in Section 4.2, we define four cases, one for each quadrant. In Case 1, we quickly identify the optimal solution as $(y_1, y_2) = (0, 0)$. Similarly, in Case 3 (the third quadrant), we find the optimal solution is also $(0, 0)$, after performing variable transformations on both y_1 and y_2 .

Recall that when we perform variable transformations in Cases 2 and 4, the problems become maximization problems that can be solved using a conventional nonlinear programming method. We find the optimal solution in image space for Case 2 is $(-3, 3)$ compared to $(2, -6)$ for Case 4. Thus, the global optimal solution is $y^* = (2, -6)$ with $x^* = (-2, 4)$.

5. Summary

We have introduced a new general approach which facilitates the solution of a class of nonconvex programs, and we have outlined the details for three subclasses. The principal idea is that the problems can be analyzed in a transformed space where optimization is easy (at least) in the coordinate directions. The specific details are dependent on the problem structure, but are similar in application. Sequential optimizations in the easy directions can be set up so as to construct a series of nested regions which shrink towards the global optimizer.

We have not exhausted all potential applications of our approach in this paper. In particular, the ideas should extend to problems wherein one is interested in optimizing sums of products of ratios, or sums of products of sums of products, etc.

Acknowledgement

Dr Falk's research was sponsored by the Office of Naval Research, contract #N00014-89-J-1537.

References

- Almogy, Y. and O. Levin (1969), Parametric Analysis of a Multi-Stage Stochastic Shipping Problem, *Proc. of the Fifth IFORS Conference*, Venice, 359–370.
- Almogy, Y. and O. Levin (1971), A Class of Fractional Programming Problems, *Operations Research* **19**, 57–67.
- Cambini, A., L. Martein, and S. Schaible (1989), On Maximizing a Sum of Ratios, *Journal of Information and Optimization Sciences* **10**, 65–79.
- Charnes, A. and W. W. Cooper (1962), Programming with Linear Fractional Functionals, *Naval Research Logistics Quarterly* **9**, 181–186.
- Dinkelbach, W. (1967), On Nonlinear Fractional Programming, *Management Science* **13**, 492–498.
- Falk, J. E. and S. W. Palocsay (1992), Optimizing the Sum of Linear Fractional Functions, *Recent Advances in Global Optimization*, eds. C. Floudas and M. Pardalos, Princeton University Press: Princeton, New Jersey, 221–258.
- Falk, J. E., S. W. Palocsay, W. J. Sacco, W. S. Copes, and H. R. Champion (1992), Bounds on a Trauma Outcome Function via Optimization, *Operations Research* **40**, S86–S95.

- Horst, R. and H. Tuy (1990), *Global Optimization: Deterministic Approaches*, Springer-Verlag: Berlin.
- Jagannathan, R. (1966), On Some Properties of Programming Problems in Parametric Form Pertaining to Fractional Programming, *Management Science* **12**, 609–615.
- Konno, H. and M. Inori (1989), Bond Portfolio Optimization by Bilinear Fractional Programming, *Journal of the Operations Research Society of Japan* **32**, 143–158.
- Konno, H. and T. Kuno (1989), Linear Multiplicative Programming, Tokyo Institute of Technology Institute of Human and Social Sciences Technical Report IISS 89-13.
- Konno, H., Y. Yajima, and T. Matsui (1991), Parametric Simplex Algorithms for Solving a Special Class of Nonconvex Minimization Problems, *Journal of Global Optimization* **1**, 65–81.
- Konno, H. and Y. Yajima (1992), Minimizing and Maximizing the Product of Linear Fractional Functions, *Recent Advances in Global Optimization*, eds. C. Floudas and M. Pardalos, Princeton University Press: Princeton, New Jersey, 259–273.
- Mangasarian, O. L. (1969), *Nonlinear Programming*, McGraw-Hill: New York, New York.
- Ritter, K. (1967), A Parametric Method for Solving Certain Nonconcave Maximization Problems, *Journal of Computer and System Sciences* **1**, 44–54.
- Schaible, S. (1977), On the Sum of a Linear and Linear-Fractional Function, *Naval Research Logistics Quarterly* **24**, 691–693.
- Schaible, S. (1981), A Survey of Fractional Programming, *Generalized Concavity in Optimization and Economics*, Academic Press, Inc.: New York, 417–440.