Mathematical Programming 52 (1991) 191-207
North-Holland

Algorithms for generalized fractional programming

Jean-Pierre Crouzeix

Département de mathématiques appliquées, Université de Clermont II, France

Jacques A. Ferland

Département d'informatique et de recherche opérationnelle, Université de Montréal, Canada

Received 15 October 1988
Revised manuscript received 7 September 1989

A generalized fractional programming problem is specified as a nonlinear program where a nonlinear function defined as the maximum over several ratios of functions is to be minimized on a feasible domain of $\mathbb{R}^n$. The purpose of this paper is to outline basic approaches and basic types of algorithms available to deal with this problem and to review their convergence analysis. The conclusion includes results and comments on the numerical efficiency of these algorithms.

1. Introduction

A *generalized fractional programming* problem is specified as a nonlinear program

$$(P) \qquad \bar{\lambda} = \inf_{x\in X}\left\{\max_{1\leqslant i\leqslant m}\{f_i(x)/g_i(x)\}\right\}$$

where X is a nonempty subset of $\mathbb{R}^n$, f_i and g_i are continuous on an open set $\tilde{X}$ in $\mathbb{R}^n$ including $\mathrm{cl}(X)$ (the closure of X), and $g_i(x)>0$ for all $x\in\tilde{X}$, $1\leqslant i\leqslant m$.

When $m=1$, then (P) reduces to a classical *fractional programming* problem which has been extensively investigated in the last two decades. Many of the results in fractional programming are reviewed in [20] and an extensive bibliography can be found in [19]. This type of problem occurs frequently in models where some kind of efficiency measure expressed as a ratio is to be optimized. In numerical analysis the eigenvalue problem is formulated as a fractional program. Fractional programs are also exhibited in stochastic programming. These applications and others are discussed in Schaible [20] where appropriate references are given.

This research was supported by NSERC (grant A8312) and Cooperation franco-québécoise (projet 20-02-13). Paper presented at "Workshop on Mathematical Programming" Catholic University of Rio de Janeiro, Brazil, October 10-14, 1988.

An early application of generalized fractional programming ($m > 1$) is found in the Von Neumann's model of an expanding economy [23]. The best rational approximation problem [1] is another manifestation of generalized fractional programming. Furthermore, goal programming and multicriteria optimization where several ratios are considered and Chebychev's norm is used give rise to generalized fractional programming. More specific applications are given in [9, 15, 18, 20].

The purpose of this paper is to outline the basic approaches and the basic types of algorithms to deal with generalized fractional programming and to review their convergence analysis. Note that other major developments related to duality theory for (P) [9, 16] are not included.

The algorithms reviewed here can be seen as generalizations of the approach proposed by Dinkelbach [10] for the case $m = 1$ to determine the root of the equation $F(\lambda) = 0$, where $F(\lambda)$ is the optimal value of the parametric program

$$(P_\lambda) \qquad F(\lambda) = \inf_{x \in X} \left\{ \max_{1 \leq i \leq m} \{f_i(x) - \lambda g_i(x)\} \right\}.$$

The general properties of F and the equivalence between the two problems studied in [8, 9, 13, 16] are reviewed in Section 2. Dinkelbach's algorithm for the case $m = 1$ is first formulated as a Newton procedure to determine the root of the equation $F(\lambda) = 0$ and then generalized for problems with arbitrary finite values of m. Section 3 is dedicated to the convergence analysis. In Section 4, referring to Robinson's method to deal with the Von Neumann problem, the Dinkelbach-type algorithm is formulated as a partial linearization procedure to deal with (P). Interval-type algorithms to determine the root of the equation $F(\lambda) = 0$ are derived in Section 5 as a consequence of the graphical interpretation of the Dinkelbach-type algorithm. In Section 6, we analyse the linear case where f_i, g_i are affine and X is a polyhedral convex set. Finally, in the conclusion, we review the numerical efficiency of these algorithms as reported in [2, 4, 11].

2. Dinkelbach-type algorithm

The equivalence between the generalized fractional programming problem (P) and the problem of finding the root of the equation $F(\lambda) = 0$ is a consequence of the following result.

Proposition 2.1 [8, Proposition 2.1].

(a) $F(\lambda) < +\infty$ *since* X *is nonempty*; F *is nonincreasing and upper semicontinuous.*

(b) $F(\lambda) < 0$ *if and only if* $\lambda > \bar{\lambda}$; *hence* $F(\bar{\lambda}) \geq 0$.

(c) *If* (P) *has an optimal solution, then* $F(\bar{\lambda}) = 0$.

(d) *If* $F(\bar{\lambda}) = 0$, *then programs* (P) *and* $(P_{\bar{\lambda}})$ *have the same set of optimal solutions (which may be empty).* □

Whenever X is compact, these results can be strengthened as follows.

Proposition 2.2 [8, Theorem 4.1]. *Assume that X is compact. Then:*
(a) *$F(\lambda) < +\infty$, F is decreasing and continuous.*
(b) *(P) and (P_λ) always have optimal solutions.*
(c) *$\bar{\lambda}$ is finite and $F(\bar{\lambda}) = 0$.*
(d) *$F(\lambda) = 0$ implies $\lambda = \bar{\lambda}$.* □

Notice that properties of F were already studied for the linear case in [9] where they are used to establish duality relations for (P).

2.1. Original Dinkelbach's procedure

The function F was first introduced in the case where $m = 1$ (see [10, 14, 21]), and Dinkelbach's procedure can be seen as a method to identify a root of the equation $F(\lambda) = 0$, where $F(\lambda)$ reduces to

$$F(\lambda) = \inf_{x \in X} \{f(x) - \lambda g(x)\}. \tag{2.1}$$

Since the function $(f(x) - \lambda g(x))$ is affine in λ, it follows that $F(\lambda) = \inf_{x \in X} \{f(x) - \lambda g(x)\}$ is concave in λ on $\mathbb{R}$. Now, let $\lambda_k \in \mathbb{R}$ and assume that x^k is an optimal solution of

$$\inf_{x \in X} \{f(x) - \lambda_k g(x)\}. \tag{2.2}$$

It is easy to verify that $g(x^k)$ is a subgradient of $-F$ at λ_k. Indeed, for any $\lambda \in \mathbb{R}$,

$$\begin{aligned} &F(\lambda) \leq f(x^k) - \lambda g(x^k), \\ &F(\lambda) \leq [f(x^k) - \lambda_k g(x^k)] - g(x^k)(\lambda - \lambda_k), \\ &-F(\lambda) \geq -F(\lambda_k) + g(x^k)(\lambda - \lambda_k). \end{aligned} \tag{2.3}$$

Relying on this result and since subgradients generalize derivatives for convex functions, a Newton method can be used to identify a root of $F(\lambda) = 0$. Hence the following sequence $\{\lambda_k\}$ is generated where

$$\lambda_{k+1} = \lambda_k - \frac{F(\lambda_k)}{-g(x_k)} = \lambda_k + \frac{f(x^k)}{g(x^k)} - \lambda_k = \frac{f(x^k)}{g(x^k)}. \tag{2.4}$$

Of course, it is implicitly assumed that, for all λ_k of the sequence, an optimal solution x^k of (2.2) exists.

It is interesting to note that the original Dinkelbach's procedure is precisely the Newton method outlined above. For the sake of completeness, the procedure is summarized as follows:

Step 0. Let $x^0 \in X$, $\lambda_1 = f(x^0)/g(x^0)$, and $k = 1$.
Step 1. Determine an optimal solution x^k of

$$\inf_{x \in X} \{f(x) - \lambda_k g(x)\}.$$

Step 2. If $F(\lambda_k)=0$, x^k is an optimal solution of (P) and λ_k is the optimal value, and STOP.

Step 3. Let $\lambda_{k+1}=f(x^k)/g(x^k)$. Replace k by $k+1$ and repeat Step 1.

The convergence analysis of this procedure is summarized in the following result.

Proposition 2.3. *Assume that X is compact. Denote by M the set of optimal solutions of* (P), $\bar{\lambda}^*=\sup\{g(x)\colon x\in M\}$ *and* $N=\{x\in M\colon g(x)=\bar{\lambda}^*\}$. *Then M and N are nonempty and compact. The sequences* $\{\lambda_k\}, \{g(x^k)\}$ *and* $\{f(x^k)\}$ *decrease and converge to* $\bar{\lambda}$, $\bar{\lambda}^*$ *and* $\bar{\lambda}\bar{\lambda}^*$, *respectively. Furthermore each convergent subsequence of* $\{x^k\}$ *converges to a point in N and the following relation holds*:

$$0\leqslant \lambda_{k+1}-\bar{\lambda}\leqslant(\lambda_k-\bar{\lambda})(1-\bar{\lambda}^*/g(x^k)).$$

It follows that the convergence of $\{\lambda_k\}$ *to* $\bar{\lambda}$ *is superlinear.*

Proof. The convergence of $\{\lambda_k\}$ to $\bar{\lambda}$ is a consequence of the convergence of the Newton method for convex functions. Let $\theta=-F$. The θ is convex and increasing, and by construction

$$\theta(\lambda_k)=(\lambda_k-\lambda_{k+1})g(x^k). \tag{2.5}$$

Since $g(x^k)\in\delta\theta(\lambda_k)$ (where $\delta\theta(\lambda_k)$ is the set of subgradients of θ at λ_k),

$$0=\theta(\bar{\lambda})\geqslant\theta(\lambda_k)+(\bar{\lambda}-\lambda_k)g(x^k). \tag{2.6}$$

Combining (2.5) and (2.6), we obtain that for all k,

$$\lambda_{k+1}\geqslant\bar{\lambda},\qquad \theta(\lambda_k)\geqslant 0,\qquad \lambda_k\geqslant\lambda_{k+1}. \tag{2.7}$$

Now, denote

$$\theta'_+(\bar{\lambda})=\lim_{\lambda\to\bar{\lambda}_+}\frac{\theta(\lambda)-\theta(\bar{\lambda})}{\lambda-\bar{\lambda}}.$$

Then $0<\theta'_+(\bar{\lambda})\in\delta\theta(\bar{\lambda})$, and

$$\theta(\lambda_k)\geqslant\theta(\bar{\lambda})+(\lambda_k-\bar{\lambda})\theta'_+(\bar{\lambda})=(\lambda_k-\bar{\lambda})\theta'_+(\bar{\lambda}). \tag{2.8}$$

Using (2.5) and (2.8), it follows that

$$(\lambda_{k+1}-\bar{\lambda})\leqslant(\lambda_k-\bar{\lambda})\left(1-\frac{\theta'_+(\bar{\lambda})}{g(x^k)}\right). \tag{2.9}$$

Referring to (2.7), it follows from the monotonicity of the subgradients that

$$\theta'_+(\bar{\lambda})\leqslant g(x^{k+1})\leqslant g(x^k)\leqslant\cdots\leqslant g(x^1). \tag{2.10}$$

Then clearly $\{\lambda_k\}$ converges at least linearly to $\bar{\lambda}$. Furthermore, since the subgradient is uppersemicontinuous and $\{\lambda_k\}\to\bar{\lambda}$, then $\{g(x^k)\}\to\theta'_+(\bar{\lambda})$. Hence the convergence of $\{\lambda_k\}$ is superlinear and $f(x^k)$ decreases to $\bar{\lambda}\theta'_+(\bar{\lambda})$ since $f(x^k)=\lambda_{k+1}g(x^k)$.

Now, let $\bar{x}$ be the limit point of a convergent subsequence of $\{x^k\}$. Then $f(\bar{x}) = \bar{\lambda} g(\bar{x})$, and consequently $\bar{x} \in M$. But from (2.10) it follows that $\theta'_+(\bar{\lambda}) \leq g(\bar{x})$, and hence $\bar{x} \in N$. □

When M or N reduces to a singleton $\{\bar{x}\}$, then the whole sequence $\{x^k\}$ converges to $\bar{x}$. Furthermore, if θ happens to be twice differentiable at $\bar{\lambda}$, then the sequence $\{\lambda_k\}$ converges quadratically to $\bar{\lambda}$. However, even if θ is not differentiable in general, the rate of convergence can be made equal to 1.618 or even 2 under appropriate regularity conditions. This point is discussed in Section 3.2 for generalized fractional programming.

Remark. Recently, Sniedovich [22] has derived the Dinkelbach parametric approach using the theory of classical first order necessary and sufficient optimality conditions. As such, the Dinkelbach procedure can also be regarded as a classical method of mathematical programming.

2.2. Extension with several ratios

When (P) includes several ratio the preceding procedure can be extended by modifying the selection of λ_{k+1} as follows:

$$\lambda_{k+1} = \max_{1 \leq i \leq m} \{f_i(x^k)/g_i(x^k)\}.$$

The Dinkelbach-type procedure DT-1 to deal with problem (P) (by finding a root of the equation $F(\lambda) = 0$) is summarized as follows:

Step 0. Let $x^0 \in X$, $\lambda_1 = \max_{1 \leq i \leq m}\{f_i(x^0)/g_i(x^0)\}$, and $k = 1$.
Step 1. Determine an optimal solution x^k of

$$(P_{\lambda_k}) \qquad F(\lambda_k) = \inf_{x \in X} \left\{ \max_{l \leq i \leq m} \{f_i(x) - \lambda_k g_i(x)\} \right\}.$$

Step 2. If $F(\lambda_k) = 0$, x^k is an optimal solution of (P) and λ_k is the optimal value, and STOP.

Step 3. Let

$$\lambda_{k+1} = \max_{1 \leq i \leq m} \{f_i(x^k)/g_i(x^k)\}.$$

Replace k by $k+1$ and repeat Step 1.

The sequence $\{\lambda_k\}$ generated has interesting properties:

(i) For all $k \geq 1$,

$$\lambda_k = \max_{1 \leq i \leq m} \{f_i(x^k)/g_i(x^k)\} \geq \bar{\lambda}$$

since $x^k \in X$. Hence referring to Proposition 2.1(b), $F(\lambda_k) \leq 0$.

(ii) The sequence $\{\lambda_k\}$ is monotone decreasing. Indeed, let $1 \leqslant \tau \leqslant m$ denote the index used to specify λ_{k+1}, i.e.

$$\lambda_{k+1} = \max_{1 \leqslant i \leqslant m} \{f_i(x^k)/g_i(x^k)\} = f_\tau(x^k)/g_\tau(x^k).$$

Then

$$F(\lambda_k) = \max_{1 \leqslant i \leqslant m} \{f_i(x^k) - \lambda_k g_i(x^k)\},$$

$$F(\lambda_k) \geqslant f_\tau(x^k) - \lambda_k g_\tau(x^k),$$

$$F(\lambda_k) \geqslant (\lambda_{k+1} - \lambda_k) g_\tau(x^k).$$

Hence

$$\lambda_{k+1} - \lambda_k \leqslant F(\lambda_k)/g_\tau(x^k) < 0$$

since

$$g_\tau(x^k) > 0 \quad \text{and} \quad F(\lambda_k) < 0.$$

These properties are important to show the convergence of the algorithm.

Remark. Furthermore, it is interesting to note that, in general, (P_{λ_k}) is easier to deal with than (P). Indeed, if, for instance, the functions f_i and g_i are linear, and if X is a polytope, then (P_{λ_k}) reduces to a linear program. Similarly, if the functions f_i are convex and g_i are concave, then (P_{λ_k}) reduces to a convex program.

3. Convergence of Dinkelbach-type algorithm

The convergence analysis of the algorithm in the case with several ratios is not as straightforward as it was when $m = 1$. Furthermore, the rate of convergence decreases.

Proposition 3.1 [8, Theorem 4.1]. *Assume that X is compact. The sequence $\{\lambda_k\}$ generated by the Dinkelbach-type procedure* DT-1, *if not finite, converges linearly to $\bar{\lambda}$, and each convergent subsequence of $\{x^k\}$ converges to an optimal solution of (P).* □

This loss of power of the procedure with respect to the case $m = 1$ can be explained by the fact that the subgradient inequality (2.3) is not verified when $m > 1$. Indeed, relation (2.3) can only be replaced by two less powerful relations derived in the following result.

Proposition 3.2 [8, Proposition 2,2]. *Let $\lambda_k \in \mathbb{R}$ and assume that x^k is an optimal solution of (P_{λ_k}). Then*

$$\begin{aligned} F(\lambda) &\leqslant F(\lambda_k) - \underline{g}(x^k)(\lambda - \lambda_k) \quad \text{if } \lambda > \lambda_k, \\ F(\lambda) &\leqslant F(\lambda_k) - \bar{g}(x^k)(\lambda - \lambda_k) \quad \text{if } \lambda < \lambda_k, \end{aligned} \tag{3.1}$$

where

$$\underline{g}(x^k) = \min_{1 \leqslant i \leqslant m} \{g_i(x^k)\} \quad \text{and} \quad \bar{g}(x^k) = \max_{1 \leqslant i \leqslant m} \{g_i(x^k)\}.$$

Proof. For any $\lambda \in \mathbb{R}$,

$$F(\lambda)=\inf_{x\in X}\left\{\max_{1\leq i\leq m}\{f_i(x)-\lambda g_i(x)\}\right\}\leq \max_{1\leq i\leq m}\{f_i(x^k)-\lambda g_i(x^k)\}. \tag{3.2}$$

If the max in (3.2) is attained at $1\leq i_k\leq m$, then

$$F(\lambda)\leq f_{i_k}(x^k)-\lambda g_{i_k}(x^k),$$

$$F(\lambda)\leq f_{i_k}(x^k)-\lambda_k g_{i_k}(x^k)-g_{i_k}(x^k)(\lambda-\lambda_k),$$

$$F(\lambda)\leq F(\lambda_k)-g_{i_k}(x^k)(\lambda-\lambda_k).$$

Hence the result in (3.2) follows directly from the definitions of $\underline{g}(x^k)$ and $\bar{g}(x^k)$. □

3.1. Superlinear rate of convergence

Now, assume that $\bar{x}$ is an optimal solution of (P). Of course, (P) can be written as

$$\bar{\lambda}=\inf_{x\in X}\left\{\max_{1\leq i\leq m}\left\{\frac{f_i(x)/g_i(\bar{x})}{g_i(x)/g_i(\bar{x})}\right\}\right\}.$$

Denote $(\bar{P}_\lambda)$ the parametric program corresponding to this formulation of (P):

$$(\bar{P}_\lambda) \qquad \bar{F}(\lambda)=\inf_{x\in X}\left\{\max_{1\leq i\leq m}\left\{\frac{f_i(x)-\lambda g_i(x)}{g_i(\bar{x})}\right\}\right\}. \tag{3.3}$$

Referring to Proposition 3.2, it follows that

$$\bar{F}(\lambda)\leq \bar{F}(\bar{\lambda})-\underline{\rho}(\bar{x})(\lambda-\bar{\lambda}) \quad \text{if } \lambda>\bar{\lambda},$$

$$\bar{F}(\lambda)\leq \bar{F}(\bar{\lambda})-\bar{\rho}(\bar{x})(\lambda-\bar{\lambda}) \quad \text{if } \lambda<\bar{\lambda},$$

where

$$\underline{\rho}(\bar{x})=\min_{1\leq i\leq m}\{g_i(\bar{x})/g_i(\bar{x})\}=1=\max_{1\leq i\leq m}\{g_i(\bar{x})/g_i(\bar{x})\}=\bar{\rho}(\bar{x}).$$

Hence at $\bar{\lambda}$ we recover the subgradient inequality (2.3). Thus, by analogy with the case $m=1$, the Dinkelbach-type procedure reduces to a Newton method in the neighborhood of $\bar{\lambda}$ and the rate of convergence should be at least superlinear.

But in general $\bar{x}$ is not known a priori. Nevertheless, this argument suggests another Dinkelbach-type procedure DT-2 obtained by replacing Step 1 in DT-1 by the following:

Step 1′. Determine x^k an optimal solution of

$$(Q_{\lambda_k}) \qquad F_k(\lambda_k)=\inf_{x\in X}\left\{\max_{1\leq i\leq m}\left\{\frac{f_i(x)-\lambda_k g_i(x)}{g_i(x^{k-1})}\right\}\right\}. \tag{3.4}$$

Note that $g_i(\bar{x})$ in (3.3) is approximated by $g_i(x^{k-1})$ in (3.4) where x^{k-1} is the optimal solution obtained by solving $(Q_{\lambda_{k-1}})$ at the preceding iteration. This modified version DT-2 has a better rate of convergence.

Proposition 3.3 [7, Theorems 2.1 and 2.2]. *Assume that X is compact.*

(a) *If $F_k(\lambda_k)=0$, then $\lambda_k=\bar{\lambda}$ and x^k is an optimal solution of (P).*

(b) *The sequence $\{\lambda_k\}$ generated by* DT-2, *if not finite, converges at least linearly to $\bar{\lambda}$, and each convergent subsequence of $\{x^k\}$ converges to an optimal solution of (P).*

(c) *Furthermore, when $\{\lambda_k\}$ is not finite, if the sequence $\{x^k\}$ converges to $\bar{x}$, then $\{\lambda_k\}$ converges superlinearly to $\bar{\lambda}$.* $\square$

The differential correction algorithm due to Cheney and Loeb [6] and used by Barrodale et al. [1] to solve the rational approximation problem is the specialization of DT-2 to deal with a specific linear form of (P). Furthermore under additional assumptions, Barrodale et al. [1] show that the rate of convergence of their algorithm is at least quadratic.

3.2. Higher rate of convergence

As expected, the rate of convergence of DT-2 can be improved at the expense of more restrictive assumptions. Borde and Crouzeix [5] derived their results using sensitivity analysis based on the implicit function theorem as proposed by Fiacco in [12]. Flachs' results [13] are derived by generalizing the differential correction approach of Cheney and Loeb [6].

The analysis in [5] requires that the following specific assumptions hold.

(H1) For $1\leq i\leq m$, g_i is concave differentiable and positive on $\tilde{X}$, f_i is convex differentiable on $\tilde{X}$ and non-negative whenever g_i is not affine.

(H2) X is a compact set defined by $X=\{x\in\mathbb{R}^n\colon h_j(x)\leq 0, 1\leq j\leq q\}$ where functions h_j are convex and differentiable on $\tilde{X}$.

(H3) (Slater's condition.) There exists $\hat{x}\in\mathbb{R}^n$ such that $h_j(\hat{x})<0$, $1\leq j\leq q$.

(H4) For $1\leq i\leq m$, $1\leq j\leq q$, functions f_i, g_i, h_j are twice continuously differentiable in a neighborhood of $\bar{x}$, an optimal solution of (P).

(H5) Denote by $\bar{\mu}_i$ and $\bar{\nu}_j$ the optimal multipliers associated with the optimal solution $\bar{x}$ of the problem

$$\begin{aligned} &\inf && t\\ &\text{subject to} && f_i(x)-\bar{\lambda}g_i(x)-tg_i(\bar{x})\leq 0,\ 1\leq i\leq m,\\ &&& h_j(x)\leq 0,\ 1\leq j\leq q, \end{aligned}$$

and let the set of active constraints be

$$\bar{I}=\{1\leq i\leq m\colon f_i(\bar{x})-\bar{\lambda}g_i(\bar{x})=0\},$$

$$\bar{J}=\{1\leq j\leq q\colon h_j(\bar{x})=0\}.$$

The following conditions are satisfied:

(a) (Strict complementary slackness.)

$$\bar{\mu}_i>0 \quad \text{if } i\in\bar{I}, \qquad \bar{\nu}_j>0 \quad \text{if } i\in\bar{J}.$$

(b) (Regularity condition.) The vectors $[(\nabla f_i(\bar{x})-\bar{\lambda}\nabla g_i(\bar{x}))^{\mathrm{T}}, -g_i(\bar{x})]^{\mathrm{T}}$, $i\in\bar{I}$, and $[\nabla h_j(\bar{x})^{\mathrm{T}}, 0]^{\mathrm{T}}$, $j\in\bar{J}$, are linearly independent.

(c) (Augmentability condition.) Denote by Γ the Hessian matrix of the function $\sum_{i=1}^{m} \mu_i(f_i(x) - \lambda g_i(x)) + \sum_{j=1}^{q} \nu_j h_j(x)$ with respect to x evaluated at $\bar{x}$, $\bar{\lambda}$, $\bar{\mu}$, $\bar{\nu}$. Also denote by $\tilde{D}$ the matrix having its columns equal to $\nabla f_i(\bar{x}) - \bar{\lambda}\nabla g_i(\bar{x})$, $i \in \bar{I}$, by $\tilde{E}$ the matrix having its columns equal to $\nabla h_j(\bar{x})$, $j \in \bar{J}$, and by $\tilde{g}$ the vector having its components equal to $g_i(\bar{x})$, $i \in \bar{I}$.

Then there exists a scalar $\omega > 0$ such that

$$\Gamma + \omega(\tilde{E}\tilde{E}^{\mathrm{T}} + \tilde{D}(I - (\tilde{g}^{\mathrm{T}}\tilde{g})^{-1}\tilde{g}\tilde{g}^{\mathrm{T}})\tilde{D}^{\mathrm{T}})$$

is positive definite (where I is the identity matrix).

Combining sensitivity analysis applied to problem $(Q_{\bar{\lambda}})$ and Flach's result [13, Theorem 1], Borde and Crouzeix obtain a convergence rate at least equal to 1.618.

Theorem 3.4 [5. Theorem 4.1]. *Assume that* (H1) *to* (H5) *hold. Then* $\bar{x}$ *is the unique optimal solution of* (P) *and the rate of convergence of the sequences* $\{x^k\}$ *and* $\{\lambda^k\}$ *is at least equal to* 1.618. □

Furthermore, introducing two additional assumptions, they obtain quadratic rate of convergence.

(H6) For $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant q$, f_i, g_i, h_j are three times differentiable on $\tilde{X}$.

(H7) The system

$$\tilde{D}z_1 + \tilde{E}z_2 = \tilde{L}\tilde{\mu},$$

$$\tilde{g}^{\mathrm{T}}z_1 = 0,$$

has a solution which is unique. Here $\tilde{\mu}$ is the vector having its components equal to $\bar{\mu}_i$, $i \in \bar{I}$, and $\tilde{L}$ is the matrix having its column equal to $\nabla g_i(\bar{x})$, $i \in \bar{I}$.

Theorem 3.5 [5, Theorem 5.1]. *Assume that* (H1) *to* (H7) *hold. The rate of convergence of the sequences* $\{x^k\}$ and $\{\lambda^k\}$ is quadratic. □

4. Partial linearization procedure

In Section 2 Dinkelbach-type algorithms are derived as some kind of Newton procedures to identify a root of the equation $F(\lambda) = 0$. Extending Robinson's approach [17] to deal with the irreducible Von Neumann economic model, DT-2 can be derived as a partial linearization procedure to solve (P) [3].

Rewrite (P) as follows:

$$\begin{aligned} &\min && t \\ &\text{subject to} && f_i(x) - t g_i(x) \leqslant 0, \quad 1 \leqslant i \leqslant m, \\ &&& x \in X. \end{aligned}$$

For $1 \leq i \leq m$, denote the ith constraint

$$H_i(x, t) = f_i(x) - tg_i(x),$$

and consider the following partial linearization of H_i with respect to t at a point (x^{k-1}, t_k):

$$H_{ik}(x, t) = H_i(x, t_k) + (t - t_k)\nabla_t H_i(x^{k-1}, t_k).$$

Using these, a partial linear approximation of (P) is specified and solved at each iteration of the procedure:

$$(L_k) \qquad \min \quad t$$
$$\text{subject to} \quad f_i(x) - t_k g_i(x) - (t - t_k) g_i(x^{k-1}) \leq 0, \ 1 \leq i \leq m,$$
$$x \in X.$$

Hence the procedure obtained is summarized as follows:

Step 0. Let $x^0 \in X$ and $t_1 = \max_{1 \leq i \leq m} \{f_i(x^0)/g_i(x^0)\}$. (Note that $t_1 = \inf\{t\colon H_i(x^0, t) \leq 0, 1 \leq i \leq m\}$ and that (x^0, t_1) is feasible for (P)). Let $k = 1$.

Step 1. Determine an optimal solution $(x^k, \bar{t}_k)$ of (L_k).

Step 2. If $\bar{t}_k = t_k$, then x^k is an optimal solution of (P) and t_k is the optimal value, and STOP.

Step 3. Let

$$t_{k+1} = \max_{1 \leq i \leq m} \{f_i(x^k)/g_i(x^k)\} = \inf\{t\colon H_i(t, x^k) \leq 0, 1 \leq i \leq m\}.$$

Replace k by $(k+1)$ and repeat Step 1.

The sequence of points $\{(x^k, t_{k+1})\}$ is generated by solving (L_k) to obtain x^k and by solving (P) with $x = x^k$ to obtain t_{k+1}. It is easy to verify that this sequence is identical to the one generated by procedure DT-2 (in Section 3) since the constraints in (L_k) can be written as

$$\frac{f_i(x) - t_k g_i(x)}{g_i(x^{k-1})} \leq (t - t_k).$$

Hence replacing $\lambda = (t - t_k)$, it follows that (L_k) and (Q_{λ_k}) are equivalent

5. Interval-type algorithms

This type of algorithms to determine the root of the equation $F(\lambda) = 0$ is a consequence of the graphical interpretation of the Dinkelbach-type algorithm. Indeed, referring to Step 3 of procedure DT-1,

$$\lambda_{k+1} = \max_{1 \leq i \leq m} \{f_i(x^k)/g_i(x^k)\} = f_\tau(x^k)/g_\tau(x^k).$$

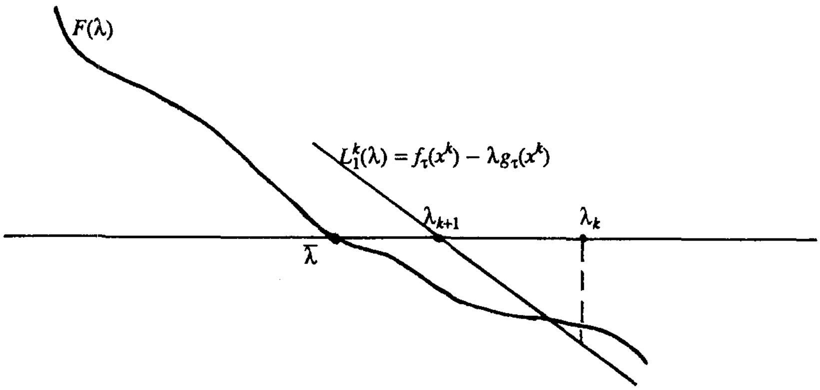

Fig. 5.1.

Hence λ_{k+1} can be interpreted as the root of

$$L_1^k(\lambda)=f_\tau(x^k)-\lambda g_\tau(x^k)$$

where L_1^k is regarded as an approximation of F near λ_k. This is illustrated in Figure 5.1.

Now, instead of L_1^k suppose that we use another approximation L_2^k of F near λ_k to determine the next iterate where

$$L_2^k(\lambda)=f_{i_k}(x^k)-\lambda g_{i_k}(x^k)$$

and $1\leqslant i_k\leqslant m$ is an index where the max is attained when $F(\lambda_k)$ is determined; i.e.

$$F(\lambda_k)=\inf_{x\in X}\left\{\max_{1\leqslant i\leqslant m}\{f_i(x)-\lambda_k g_i(x)\}\right\}=f_{i_k}(x^k)-\lambda_k g_{i_k}(x^k).$$

At first glance, L_2^k may seem more appropriate than L_1^k to determine the next iterate since

(i) $\quad L_2^k(\lambda_k)=F(\lambda_k) \quad$ while $L_1^k(\lambda_k)\leqslant F(\lambda_k)$;

(ii) $\quad f_{i_k}(x^h)/g_{i_k}(x^k)\leqslant \max_{1\leqslant i\leqslant m}\{f_i(x^k)/g_i(x^k)\}=f_\tau(x_k)/g_\tau(x^k) \qquad (5.1)$

and hence $f_{i_k}(x^k)/g_{i_k}(x^k)$ is closer to $\bar\lambda$ whenever $f_{i_k}(x^k)/g_{i_k}(x^k)\geqslant\bar\lambda$ (as illustrated in Figure 5.2).

Unfortunately nothing prevents $f_{i_k}(x^k)/g_{i_k}(x^k)$ from being smaller than $\bar\lambda$ (as illustrated in Figure 5.3). Recall that this is never the case in DT-1 algorithm since $\lambda_{k+1}=f_\tau(x^k)/g_\tau(x^k)=\max_{q\leqslant i\leqslant m}\{f_i(x^k)/g_i(x^k)\}\geqslant\bar\lambda$ because $x^k\in X$.

As mentioned in Section 2, the convergence proof of DT-1 algorithm relies heavily on the monotonicity if the sequence $\{\lambda_k\}$, i.e. $\bar\lambda\leqslant\lambda_{k+1}\leqslant\lambda_k$. Even if monotonicity of the iterates is lost when L_2^k is used, some trend toward $\bar\lambda$ is observed as follows [11]:

(i) if $\lambda_k>\bar\lambda$(i.e. $F(\lambda_k)<0$), then

$$0>F(\lambda_k)/g_{i_k}(x^k)=f_{i_k}(x^k)/g_{i_k}(x^k)-\lambda_k$$

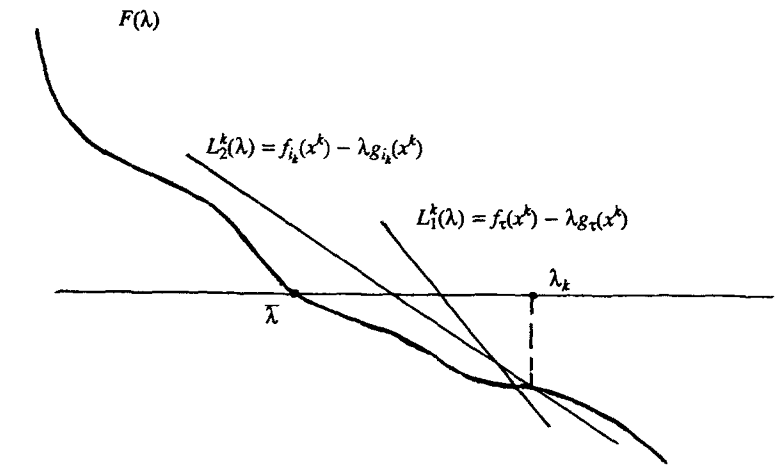

Fig. 5.2.

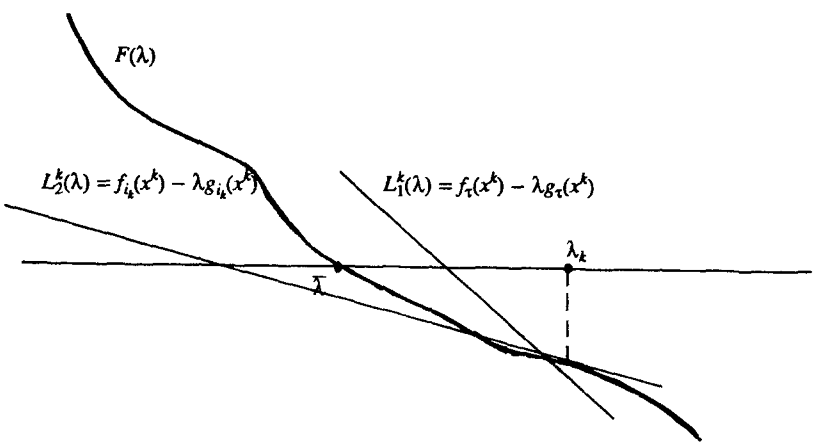

Fig. 5.3.

and

$$f_{i_k}(x^k)/g_{i_k}(x^k) < \lambda_k;$$

(ii) if $\lambda_k < \bar{\lambda}$(i.e. $F(\lambda_k) > 0$), then

$$0 < F(\lambda_k)/g_{i_k}(x^k) = f_{i_k}(x^k)/g_{i_k}(x^k) - \lambda_k$$

and

$$f_{i_k}(x^k)/g_{i_k}(x^k) > \lambda_k.$$

Nevertheless, $f_{i_k}(x^k)/g_{i_k}(x^k)$ may be very far to the left or to the right of $\bar{\lambda}$, and an interval $[\mathrm{BI}_k, \mathrm{BS}_k]$ including $\bar{\lambda}$ has to be used to limit the distance between $\bar{\lambda}$ and the next iterate. The procedure is to determine λ_{k+1} as follows:

$$\lambda_{k+1} = \begin{cases} f_{i_k}(x^k)/g_{i_k}(x^k) & \text{if } f_{i_k}(x^k)/g_{i_k}(x^k) \in [\mathrm{BI}_k, \mathrm{BS}_k], \\ \text{a point in } [\mathrm{BI}_k, \mathrm{BS}_k] & \text{otherwise.} \end{cases}$$

Several different algorithms are derived according to the way the point is selected in the interval $(\mathrm{BI}_k, \mathrm{BS}_k]$ (see [11]).

At each iteration BI_k and BS_k are updated in such a way that the length of the interval $(\mathrm{BS}_k - \mathrm{BI}_k)$, is a non-increasing function of k. The bounds BS_k and BI_k can be derived from the upper envelope $G_k(\lambda)$ and the lower envelope $T_k(\lambda)$ illustrated in Figure 5.4.

The upper envelope G_k of F at λ_k is defined by

$$G_k(\lambda) = \max_{1 \leq i \leq m} \{f_i(x^k) - \lambda g_i(x^k)\}.$$

Hence

$$G_k(\lambda_k) = \max_{1 \leq i \leq m} \{f_i(x^k) - \lambda_k g_i(x^k)\} = F(\lambda_k)$$

and for all λ,

$$G_k(\lambda) = \max_{1 \leq i \leq m} \{f_i(x^k) - \lambda g_i(x^k)\} \geq \inf_{x \in X} \left\{ \max_{1 \leq i \leq m} \{f_i(x) - \lambda g_i(x)\} \right\} = F(\lambda).$$

It follows that the root s_k of the equation $G_k(\lambda) = 0$ is an upper bound on $\bar{\lambda}$. Furthermore, it is easy to verify that

$$s_k = \max_{1 \leq i \leq m} \left\{ \frac{f_i(x^k)}{g_i(x^k)} \right\}.$$

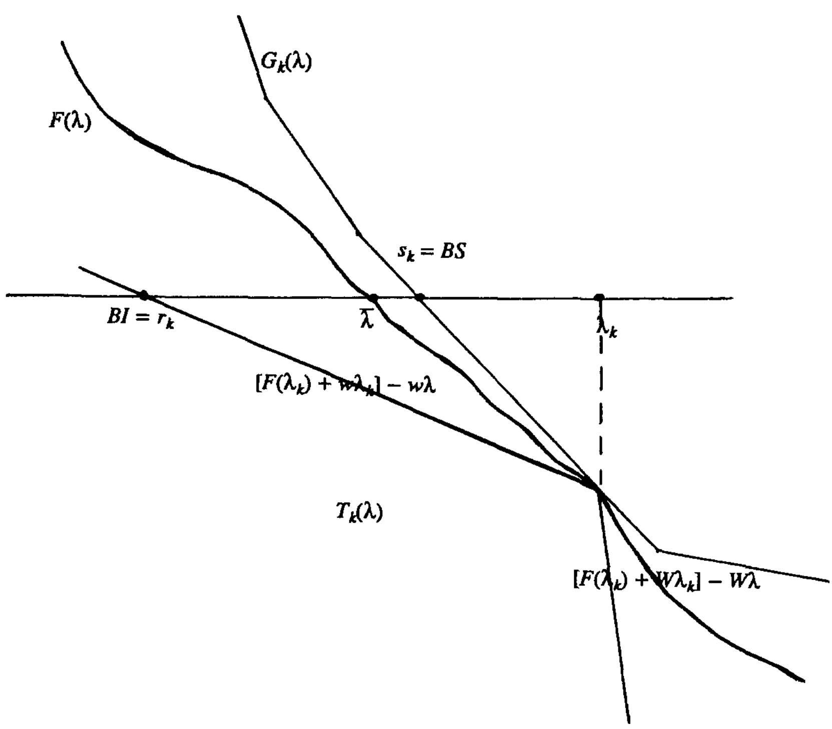

Fig. 5.4.

Hence, if $\lambda_k > \bar{\lambda}$, s_k is equal to the next iterate in DT-1 algorithm. Finally,

$$\mathrm{BS}_k = \min\{\mathrm{BS}_{k-1}, s_k\}.$$

To define the lower envelope T_k of F at λ_k, we have to introduce the scalar w and W where

$$0 < w \leqslant \min_{x \in X}\left\{\min_{1 \leqslant i \leqslant m}(g_i(x))\right\},$$

$$0 < \max_{x \in X}\left\{\max_{1 \leqslant i \leqslant m}(g_i(x))\right\} \leqslant W.$$

Now, since

$$f_i(x) - \lambda g_i(x) = f_i(x) - \lambda_k g_i(x) + (\lambda_k - \lambda) g_i(x),$$

it follows that

$$F(\lambda) \geqslant F(\lambda_k) + (\lambda_k - \lambda) w \quad \text{if } \lambda \leqslant \lambda_k$$

and

$$F(\lambda) \geqslant F(\lambda_k) + (\lambda_k - \lambda) W \quad \text{if } \lambda \geqslant \lambda_k.$$

Hence

$$T_k(\lambda) = \begin{cases} [F(\lambda_k) + \lambda_k w] - \lambda w & \text{if } \lambda \leqslant \lambda_k, \\ [F(\lambda_k) + \lambda_k W] - \lambda W & \text{if } \lambda \geqslant \lambda_k. \end{cases}$$

The root r_k of the equation $T_k(\lambda) = 0$ is a lower bound on $\bar{\lambda}$ and

$$\mathrm{BI}_k = \max\{\mathrm{BI}_{k-1}, r_k\}.$$

Referring to the definitions of w and W, we may expect the lower bound to be less tight than the upper bound. This observation is confirmed in the numerical results, and hence we take the next iterate λ_{k+1} closer to BS_k whenever $f_{i_k}(x^k)/g_{i_k}(x^k) \in [\mathrm{BI}_k, \mathrm{BS}_k]$ (see [4]).

The detailed steps of the algorithm referred to as IT-1 are given in [4] where convergence is also analyzed. Since the sequence $\{\lambda_k\}$ includes elements on both sides of $\bar{\lambda}$, we use an approach like Ibaraki's in [14] for studying variants of Dinkelbach's algorithm for fractional programming ($m = 1$). Hence consider the following subsequences:

$$\{\lambda_j^{\mathrm{d}}\} \cup \{\lambda_j^{\mathrm{g}}\} \cup \{\lambda_j^{0}\} = \{\lambda_k\}$$

where

(i) $\lambda_k \in \{\lambda_j^{\mathrm{d}}\}$ if and only if $\lambda_k > \bar{\lambda}$ and $\lambda_{k+1} > \bar{\lambda}$;
(ii) $\lambda_k \in \{\lambda_j^{\mathrm{g}}\}$ if and only if $\lambda_k < \bar{\lambda}$;
(iii) $\{\lambda_j^{0}\} = \{\lambda_k\} - \{\lambda_j^{\mathrm{d}}\} \cup \{\lambda_j^{\mathrm{g}}\}$.

In [4, Lemma 4.1] it is shown that $\{\lambda_j^{0}\}$ does not include too many elements. Furthermore, if not finite, $\{\lambda_j^{\mathrm{d}}\}$ and $\{\lambda_j^{\mathrm{g}}\}$ converge at least linearly to $\bar{\lambda}$ (see [4, Theorems 4.2 and 4.3]).

As for the Dinkelbach-type algorithms, the rates of convergence for subsequence $\{\lambda_j^{\mathrm{d}}\}$ can be improved if subproblems (Q_{λ_k}) are used instead of (P_{λ_k}) to obtain IT-2 algorithm. In [4] it is shown that $\{\lambda_j^d\}$ converges as fast as the sequence $\{\lambda_k\}$ generated by DT-2 under the same hypothesis.

6. Linear case

Assume that f_i, g_i are affine and X is a polyhedral convex set (possibly unbounded):

$$f_i(x) = a_{i\cdot}x + \alpha_i, \qquad g_i(x) = b_{i\cdot}x + \beta_i,$$
$$X = \{x \in \mathbb{R}^n\colon Cx \leqslant \gamma, x > 0\},$$

where $a_{i\cdot}(b_{i\cdot})$ denote the ith row of $m \times n$ matrix $A(B)$, $\alpha = [\alpha_1, \alpha_2, \ldots, \alpha_m]^{\mathrm{T}}$, $\beta = [\beta_1, \beta_2, \ldots, \beta_m]^{\mathrm{T}}$, C a $q \times n$ matrix, and $\gamma \in \mathbb{R}^q$. In [5, 8] the authors make the following assumptions:

(A1) (Feasibility assumption.) There exists $\hat{x} \geqslant 0$ such that $C\hat{x} \leqslant \gamma$.

(A2) (Positivity assumption.) $B > 0$, $\beta > 0$.

They also introduce the following dual program of (P):

$$(D) \qquad \bar{\theta} = \sup_{(u,\, v) \in S} \left\{ \min \left\{ \frac{\alpha^{\mathrm{T}} u - \gamma^{\mathrm{T}} v}{\beta^{\mathrm{T}} u}, \min_{1 \leqslant j \leqslant n} \left\{ \frac{a_{\cdot j}^{\mathrm{T}} u + c_{\cdot j}^{\mathrm{T}} v}{b_{\cdot j}^{\mathrm{T}} u} \right\} \right\} \right\},$$

where $S = \{u \in \mathbb{R}^m, v \in \mathbb{R}^q\colon \sum_{i=1}^m u_i = 1, u \geqslant 0, v \geqslant 0\}$ and $a_{\cdot j}$ $(b_{\cdot j})$ denotes the jth column of A (B). Finally they verify that $\bar{\lambda} = \bar{\theta}$ whenever (A1) and (A2) hold.

Theorem 6.1 [5, Remark 5.3]. *If X is a polytope, and* (A1), (A2) *and* (H5) *hold, then* DT-2 *algorithm applied to (P) where f_i, g_i are affine generate sequences $\{x^k\}$, $\{\lambda_k\}$ converging quadratically to $\bar{x}$ and $\bar{\lambda}$.* □

Even if X is not bounded, the feasible domain S of (D) is at least bounded in u. Hence it makes sense to apply DT-2 to deal with (D) since the sequences $\{u^k, v^k\}$ and $\{\theta_k\}$ converges.

Theorem 6.2 [8, Theorem 5.1] and [5, Theorem 6.1]. *Assume that* (A1) *and* (A2) *hold.*

(i) *If not finite, the sequence $\{\theta_k\}$ converges linearly to $\bar{\theta}$ and each convergent subsequence of $\{u^k, v^k\}$ converges to an optimal solution of (D).*

(ii) *If (D) has a unique solution $\bar{u}$, $\bar{v}$, then $\{u^k, v^k\}$ converges to $\bar{u}$, $\bar{v}$ and $\{\theta_k\}$ converges superlinearly to $\bar{\theta}$.* □

Furthermore, referring to the sensitivity analysis in Section 3, Borde and Crouzeix [5, Theorem 6.2] show quadratic rate of convergence for sequence $\{u^k, v^k\}$ and $\{\theta_k\}$ if the following additional assumptions are verified:

(A3) The parametric subproblem

$$(D_{\bar{\theta}}) \qquad \sup_{(u,v) \in S} \left\{ \min \left\{ \frac{(\alpha + \bar{\theta}\beta)^{\mathrm{T}} u - \gamma^{\mathrm{T}} v}{\beta^{\mathrm{T}} \bar{u}}, \min_{1 \leqslant j \leqslant n} \left\{ \frac{(a_{\cdot j} + \bar{\theta} b_{\cdot j})^{\mathrm{T}} u + c_{\cdot j}^{\mathrm{T}} v}{b_{\cdot j}^{\mathrm{T}} \bar{u}} \right\} \right\} \right\}$$

has an unique optimal solution.

(A4) The gradient of the active constraints of $(D_{\bar{\theta}})$ at the optimal solution are linearly independent.

(A5) The strict complementary slackness holds at the optimal solution of $(D_{\bar{\theta}})$.

Finally, when X is a polytope, Interval-type algorithms can be applied to (P) where f_i and g_i are affine. Now, referring to (D) and weak duality theory, Ferland and Potvin [11] generate a lower bound

$$\mathrm{BI}_0 = \min\left\{ \sum_{i=1}^{m} \alpha_i \Big/ \sum_{i=1}^{m} \beta_i, \min_{1 \leq j \leq n} \left\{ \sum_{i=1}^{m} a_{ij} \Big/ \sum_{i=1}^{m} b_{ij} \right\} \right\}$$

using $u = [1/m, 1/m, \ldots, 1/m]^{\mathrm{T}}$ and $\nu = [0, 0, \ldots, 0]^{\mathrm{T}}$.

Other lower bounds are easily obtained (see [3]) by taking $u = [0, 0, \ldots, 1, \ldots, 0]^{\mathrm{T}}$ and $\nu = [0, 0, \ldots, 0]^{\mathrm{T}}$ in (D):

$$\min\left\{ \alpha_i/\beta_i, \min_{1 \leq j \leq n} \{a_{ij}/b_{ij}\} \right\}.$$

7. Conclusion

Numerical results are reported in [4, 11] for the linear case. They confirm the advantage of using subproblem (Q_{λ_k}) instead of (P_{λ_k}) in both Dinkelbach-type and Interval-type algorithms to increase the convergence rate. Indeed, the execution time of DT-1 (IT-1) is roughly equal to 1.8 time the execution time of DT-2 (IT-2) on the average.

The Dinkelbach-type algorithm DT-2 is almost as efficient as the Interval-type algorithm IT-2. Indeed the execution time of DT-2 is roughly equal to 1.07 time the execution time of IT-2. This result indicates that even if some elements of $\{\lambda_k\}$ generated by IT-2 are on the left of $\bar{\lambda}$, this is compensated by the fact that the elements of $\{\lambda_k\}$ on the right of $\bar{\lambda}$ are closer to $\bar{\lambda}$ than those generated by DT-2 (see (5.1)).

Remark. More recently, Benadada [2] has tested similar procedures on problems having quadratic functions f_i. The numerical results indicate similar relative efficiency among the different algorithms.

References

[1] I. Barrodale, M.J.D. Powell and F.D.K. Roberts, "The differential correction algorithm for rational l_∞ approximation," *SIAM Journal Numerical analysis* 9 (1972) 493–504.

[2] Y. Benadada, "Approches de résolution du problème de programmation fractionnaire généralisée," Ph.D. Thesis, Département d'informatique et de recherche opérationnelle, Université de Montréal (Montréal, Canada, 1989).

[3] Y. Benadada and J.A. Ferland, "Partial linearization for generalized fractional programming," *Zeitschrift für Operations Research* 32 (1988) 101–106.

[4] J.C. Bernard and J.A. Ferland, "Convergence of interval-type algorithms for generalized fractional programming," *Mathematical Programming* 43 (1989) 349-364.

[5] J. Borde and J.P. Crouzeix, "Convergence of a Dinkelbach-type algorithm in generalized fractional programming," *Zeitschrift für Operations Research* 31 (1987) A31-A54.

[6] E.W. Cheney and H.L. Loeb, "Two new algorithms for rational approximation," *Numerische Mathematik* 3 (1961) 72-75.

[7] J.P. Crouzeix, J.A. Ferland and S. Schaible, "A note on an algorithm for generalized fractional programs," *Journal of Optimization Theory and Applications* 50 (1986) 183-187.

[8] J.P. Crouzeix, J.A. Ferland and S. Schaible, "An algorithm for generalized fractional programs," *Journal of Optimization Theory and Applications* 47 (1985) 35-49.

[9] J.P. Crouzeix, J.A. Ferland and S. Schaible, "Duality in generalized linear fractional programming," *Mathematical Programming* 27 (1983) 342-354.

[10] W. Dinkelbach, "On non-linear fractional programming," *Management Science* 13 (1967) 492-498.

[11] J.A. Ferland and J.Y. Potvin, "Generalized fractional programming: algorithms, and numerical experimentation," *European Journal of Operational Research* 20 (1985) 92-101.

[12] A.V. Fiacco, "Sensitivity analysis for nonlinear programming using penalty methods," *Mathematical Programming* 10 (1976) 287-311.

[13] J. Flachs, "Generalized Cheney-Loeb-Dinkelbach-type algorithms," *Mathematics of Operations Research* 10 (1985) 674-687.

[14] T. Ibaraki, "Parametric approaches to fractional programs," *Mathematical Programming* 26 (1983) 345-362.

[15] R. Jagannathan, "An algorithm for a class of nonconvex programming problems with nonlinear fractional objectives," *Management Science* 31 (1985) 847-851.

[16] R. Jagannathan and S. Schaible, "Duality in generalized fractional programming via Farkas' Lemma," *Journal of Optimization Theory and Applications* 41 (1983) 417-424.

[17] S.M. Robinson, "A linearization technique for solving the irreducible Von Neumann economic model," in: J. Los, ed., *Proceedings of the conference on Von Neumann model* (Polish Scientific Publishers (PWN), Warsaw, 1972) pp. 139-150. Also appeared as Technical Summary Report No 1290, Mathematics Research Center, Wisconsin University (Madison, WI, 1972).

[18] S. Schaible, "Multiratio fractional programming — a survey," in: A. Kurzhanski, K. Neumann and D. Pallaschke, eds., *Optimization, Parallel Processing and Applications*, Lecture Notes in Economics and Mathematical Systems No. 304 (Springer, Berlin-Heidelberg, 1988) pp. 57-66.

[19] S. Schaible, "Bibliography in fractional programming," *Zeitschrift für Operations Research* 26 (1982) 211-241.

[20] S. Schaible, "A survey of fractional programming," in: S. Schaible and W.T. Ziemba, eds., *Generalized Concavity in Optimization and Economics* (Academic Press, New York, 1981) pp. 417-440.

[21] S. Schaible, "Fractional programming, II: on Dinkelbach's algorithm," *Management Science* 22 (1976) 868-873.

[22] M. Sniedovich, "Fractional programming revisited," *European Journal of Operational Research* 33 (1988) 334-341.

[23] J. Von Neumann, "A model of general economic equilibrium," *Review of Economic Studies* 13 (1945) 1-9.

[24] H. Wolf, "A parametric method for solving the linear fractional programming problem," *Operations Research* 33 (1985) 835-841.