Journal of the
Operations Research
Society of Japan

VOLUME 12 *December 1969* NUMBER 1

NONLINEAR FRACTIONAL PROGRAMMING

O.L. MANGASARIAN

The University of Wisconsin

(Received October 31, 1968)

Abstracts

The purpose of this note is to single out a class of nonlinear programming problems with linear constraints and an objective function (not necessarily convex) which is a ratio of two nonlinear functions, and to show how to solve these problems by solving a sequence of linear programs (the Frank-Wolfe algorithm). As an application, we show how to handle a class of bi-nonlinear objective functions (that is, functions which are the product of two nonlinear functions).

The Problem

We consider the problem of finding an $\bar{x}$ such that

$$(1) \qquad \theta(\bar{x}) = \min_{x \in \Gamma} \theta(x), \quad \bar{x} \in \Gamma, \quad \theta(x) = \frac{\varphi(x)}{\psi(x)},$$

* Sponsored by the Mathematics Research Center, United States Army, Madison, Wisconsin, under Contract No.: DA-31-124-ARO-D-462.

where Γ is a polytope in the n-dimensional Euclidean space E^n, $\varphi(x)$ and $\psi(x)$ are numerical functions defined on Γ such that $\psi(x)\neq 0$ on Γ.

The present result hinges upon the following observations: (i) Under suitable restrictions on $\varphi(x)$ and $\psi(x)$, the function $\theta(x)=\dfrac{\varphi(x)}{\psi(x)}$ is pseudo-convex on Γ [6]. (ii) The Frank-Wolfe algorithm [5, 1] originally proposed for a continuously differentiable convex objective function $\theta(x)$ will also converge for a continuously differentiable pseudo-convex function $x(\theta)$.

Recently [4] Dinkelbach also considered the same problem (1). By using a parametric method, he reduced (1) to the solution of a sequence of *nonlinear* convex programming problems. In the present method, no parameter is used, and a succession of *linear* programming problems are solved instead.

Pseudo-Convexity and Quasi-Convexity of Fractional Functions

A numerical function $\theta(x)$ which is differentiable on some set Γ in E^n is said to be *pseudo-convex on* Γ [6] if for each x^1, $x^2\in\Gamma$

$$(2)\qquad \nabla\theta(x^1)(x^2-x^1)\geq 0 \;\Rightarrow\; \theta(x^2)\geq\theta(x^1).^{a)}$$

A number of properties and applications of pseudo-convex functions are given in [6]. We recall here that if Γ is convex then every differentiable convex function on Γ is pseudo-convex on Γ, but not conversely [6]. Also, every local minimum of a pseudo-convex function is a global minimum [6].

We establish now the pseudo-convexity of $\theta(x)=\dfrac{\varphi(x)}{\psi(x)}$ by suitably restricting $\varphi(x)$ and $\psi(x)$.

Lemma 1. Let $\varphi(x)$ and $\psi(x)$ be differentiable numerical functions on some convex set Γ in E^n. Then $\theta(x)=\dfrac{\varphi(x)}{\psi(x)}$ is pseudo-convex on Γ

a) $\theta(x)$ is **pseudo-concave** on Γ if and only if $-\theta(x)$ is pseudo-convex on Γ.

if

$$(A)\begin{cases} 1. & \varphi(x) \text{ is convex, } \psi(x)>0, \text{ on } \Gamma, \textit{ or} \\ 2. & \varphi(x) \text{ is concave, } \psi(x)<0, \text{ on } \Gamma \end{cases}$$

and

$$(B)\begin{cases} 1. & \psi(x) \text{ is linear on } \Gamma, \textit{ or} \\ 2. & \psi(x) \text{ is convex, } \varphi(x)\leq 0, \text{ on } \Gamma, \textit{ or} \\ 3. & \psi(x) \text{ is concave, } \varphi(x)\geq 0, \text{ on } \Gamma. \end{cases}$$

Proof: Let x^1, $x^2 \in \Gamma$. By assumption (A) we have that

$$(3)\qquad \psi(x^1)(\varphi(x^2)-\varphi(x^1))\geq(\nabla\varphi(x^1)(x^2-x^1))\psi(x^1),$$

and by assumption (B)

$$(4)\qquad \varphi(x^1)(\psi(x^2)-\psi(x^1))\leq(\nabla\psi(x^1)(x^2-x^1))\varphi(x^1).$$

Hence

$$\begin{aligned}\nabla\theta(x^1)(x^2-x^1) &= \frac{1}{(\psi(x^1))^2}[\psi(x^1)\nabla\varphi(x^1)-\varphi(x^1)\nabla\psi(x^1)](x^2-x^1)\\ &\leq \frac{1}{(\psi(x^1))^2}[\psi(x^1)(\varphi(x^2)-\varphi(x^1))-\varphi(x^1)(\psi(x^2)-\psi(x^1))]\\ &\qquad\qquad \text{(by (3) and (4))}\\ &= \frac{\psi(x^2)}{\psi(x^1)}[\theta(x^2)-\theta(x^1)].\end{aligned}$$

Hence,

$$\nabla\theta(x^1)(x^2-x^1)\geq 0 \quad\Rightarrow\quad \theta(x^2)\geq\theta(x^1),$$

and $\theta(x)$ is pseudo-convex on Γ. Q.E.D.

Since every pseudo-convex function $\theta(x)$ on a convex set Γ is also quasi-convex[a] on Γ [6] it follows from Lemma 1 that under assumptions

---

[a] A numerical function $\theta(x)$ defined on a convex set $\Gamma\subset E^n$ is said to be quasi-convex on Γ if the set $\Omega=\{x|x\in\Gamma,\ \theta(x)\leq\alpha\}$ is convex for each real number α.

(A) and (B) $\theta(x)$ is also quasi-convex on Γ. The quasi-convexity of $\theta(x)$ however can also be established ***without*** any differentiability requirements on $\theta(x)$.

Lemma 2. Let $\varphi(x)$ and $\psi(x)$ be numerical functions on some convex set $\Gamma\subset E^n$. Then $\theta(x)=\dfrac{\varphi(x)}{\psi(x)}$ is quasi-convex on Γ if assumptions (A) and (B) above hold.

Proof: Let α be any real number. Define

$$\Omega=\{x|x\in\Gamma,\ \theta(x)\leq\alpha\}$$

$$\Omega_1=\{x|x\in\Gamma,\ \varphi(x)-\alpha\psi(x)\leq 0\}$$

$$\Omega_2=\{x|x\in\Gamma,\ \varphi(x)-\alpha\psi(x)\geq 0\}.$$

(**A1**) *and* (**B1**): $\Omega=\Omega_1$; Ω_1 is convex by the convexity of φ and the linearity of ψ.

(**A2**) *and* (**B1**): $\Omega=\Omega_2$; Ω_2 is convex by the concavity of φ and the linearity of ψ.

(**A1**) *and* (**B2**): $\Omega=\Omega_1$. For $\alpha>0$, $\Omega=\Gamma$, which is convex. For $\alpha\leq 0$, Ω_1 is convex by the convexity of φ and ψ.

(**A2**) *and* (**B2**): $\Omega=\Omega_2$. For $\alpha<0$, $\Omega=\phi$, which is convex. For $\alpha\geq 0$, Ω_2. is convex by the concavity of φ and the convexity of ψ.

(**A1**) *and* (**B3**): $\Omega=\Omega_1$. For $\alpha<0$, $\Omega=\phi$, which is convex. For $\alpha\geq 0$, Ω_1 is convex by the convexity of φ and the concavity of ψ.

(**A2**) *and* (**B3**): $\Omega=\Omega_2$. For $\alpha>0$, $\Omega=\Gamma$, which is convex. For $\alpha\leq 0$, Ω_2 is convex by the concavity of φ and ψ. Q.E.D.

Remark: It is simple to show that both lemmas above hold if the words "convex" and "concave" are interchanged throughout the statements of the lemmas except that the phrase "convex set Γ" remains unchanged.

Corollary: Let $a\in E^n$, $b\in E^n$, $\alpha\in E^1$, $\beta\in E^1$ be fixed. Then $\theta(x)=\dfrac{ax+\alpha}{bx+\beta}$

is ***both*** pseudo-convex and pseudo-concave (and hence also both quasi-convex and quasi-concave) on each convex set $\Gamma \subset E^n$ on which $bx+\beta \neq 0$.

The Frank-Wolfe Algorithm

The Frank-Wolfe algorithm [5, 1] solves the problem

$$\operatorname*{Min}_{x \in \Gamma} \theta(x)$$

under the following assumptions

(C)

1. Γ is a polytope in E^n
2. The function $\theta(x)$ is a continuously differentiable pseudo-convex function on Γ. (The original convergence proof of the Frank-Wolfe algorithm [5, 1] required the convexity of $\theta(x)$. The convergence proof of Ghouila-Houri [1, p. 91] goes through if we relax the convexity to pseudo-convexity.)
3. For each $\hat{x} \in \Gamma$, the linear function $\nabla\theta(\hat{x})x$ is bounded from below on Γ. (This assumption is satisfied if Γ is a polyhedron, that is a bounded polytope.)

The algorithm consists of the following steps:

(D)

1. Find an $x^1 \in \Gamma$ (by the simplex algorithm, say [2]).
2. Construct a sequence $x^1, x^2, \cdots, x^k, \cdots$, of points in Γ as follows: Knowing x^k use the simplex method of linear programming [2] (or any other method) to find a vertex v^k of Γ such that

$$\nabla\theta(x^k)v^k = \min_{x \in \Gamma} \nabla\theta(x^k)x,$$

(that is v^k is a solution of the linearized problem about x^k), then choose x^{k+1} such that

$$\theta(x^{k+1}) \leq (1-\rho^k)\theta(x^k) + \rho^k \min_{x \in [x^k, v^k]} \theta(x),$$

where ρ^k is some number such that $0 < \sigma \leq \rho^k \leq 1$.

Under assumptions (C) the algorithm (D) generates a sequence of points $x^1, x^2, \cdots, x^k, \cdots$ in Γ which has an accumulation point x^0 in Γ, such that

$$\theta(x^1) \geq \theta(x^2) \geq \cdots \geq \theta(x^k) \geq \cdots,$$

$$\lim_{k\to\infty} \theta(x^k) = \min_{x\in\Gamma} \theta(x) = \theta(x^0).$$

By combining these convergence criteria of the Frank-Wolfe algorithm with the results of the previous section the following convergence criterion is obtained for the nonlinear fractional programming problem (1).

Convergence Criterion for Nonlinear Fractional Programs: Let Γ be a polytope in E^n, let $\varphi(x)$ and $\psi(x)$ be continuously differentiable numerical functions on Γ such that $\psi(x) \neq 0$ on Γ, let $\theta(x) = \dfrac{\varphi(x)}{\psi(x)}$ and let $\nabla\theta(\hat{x})x$ be bounded from below on Γ for each fixed $\hat{x}$ in Γ.[a] If assumptions (A) and (B) hold, then the Frank-Wolfe algorithm (D) converges for problem (1), that is

$$\lim_{k\to\infty} \theta(x^k) = \min_{x\in\Gamma} \theta(x) = \theta(x^0),$$

where x^0 is an accumulation point of the sequence $x^1, x^2, \cdots$.

In view of the Corollary of the previous section, linear fractional programs can also be solved by the Frank-Wolfe algorithm. However, the methods proposed specifically for linear fractional problems (see bibliography of [4]) are probably more efficient. The method of [4] can be considered as an extension of some of the methods of linear fractional programming to nonlinear fractional programming. One extremely simple and little known method for solving linear fractional programs [3, pp. 22-23] does not seem to extend to nonlinear fractional programs.

[a] This last assumption may be replaced by the less general assumption that Γ is bounded.

Bi-Nonlinear Programs

If the objective function is the product of two nonlinear functions we call the problem a bi-nonlinear program. The problem is to find an $\bar{x}$ such that

$$(5) \qquad \theta(\bar{x}) = \operatorname*{Min}_{x \in \Gamma} \theta(x),\ \bar{x} \in \Gamma,\ \theta(x) = \varphi(x)\sigma(x).$$

We use the previous results now to establish convergence of the Frank-Wolfe algorithm for this problem.

Convergence Criterion for Bi-Nonlinear Programs: Let Γ be a polytope in E^n, let $\varphi(x)$ and $\sigma(x)$ be continuously differentiable numerical functions on Γ such that $\sigma(x) \neq 0$ on Γ, let $\theta(x) = \varphi(x)\sigma(x)$, and let $\nabla\theta(\hat{x})x$ be bounded from below on Γ for each fixed $\hat{x}$ in Γ (or let Γ be bounded). If either

(E) $\quad \varphi(x)$ is convex, $\varphi(x) \leq 0$, $\sigma(x)$ is concave, $\sigma(x) > 0$, on Γ

or

(F) $\quad \varphi(x)$ is concave, $\varphi(x) \geq 0$, $\sigma(x)$ is convex, $\sigma(x) < 0$, on Γ,

then the Frank-Wolfe algorithm (D) converges for problem (5).

Proof: All we have to show is that under assumption (E) or (F), $\theta(x) = \varphi(x)\sigma(x)$ is pseudo-convex on Γ. We shall use Lemma 1 and the fact that the reciprocal of a positive concave function is a positive convex function and that the reciprocal of a negative convex function is a negative concave function.[a] If we let $\psi(x) = \dfrac{1}{\sigma(x)}$, then assumption (E) implies (A1) and (B2), while assumption (F) implies (A2) and (B3). Hence by Lemma 1, $\theta(x) = \varphi(x)\sigma(x) = \dfrac{\varphi(x)}{\psi(x)}$ is pseudo-convex. Q.E.D.

We can restate the requirements for pseudo-convexity of a bi-nonlinear function in the following schematic way

---

[a] See Appendix for a proof.

$$(\text{convex}\leq 0)\,(\text{concave}>0) \;\Rightarrow\; (\text{pseudo-convex}\leq 0)$$

and

$$(\text{convex}<0)\,(\text{concave}\geq 0) \;\Rightarrow\; (\text{pseudo-convex}\leq 0).$$

By making use of the remark following Lemma 1 we can also show that

$$(\text{convex}\leq 0)\,(\text{convex}<0) \;\Rightarrow\; (\text{pseudo-concave}\geq 0)$$

and

$$(\text{concave}\geq 0)\,(\text{concave}>0) \Rightarrow (\text{pseudo-concave}\geq 0).$$

It follows then that the function $\theta(x)=x_1x_2$ defined on E^2 is pseudo-concave on the sets

$$\{x|x_1>0,\ x_2\geq 0\},\ \{x|x_1\geq 0,\ x_2>0\},\ \{x|x_1<0,\ x_2\leq 0\},$$
$$\{x|x_1\leq 0,\ x_2<0\}\ ;$$

and it is pseudo-convex on the sets

$$\{x|x_1<0,\ x_2\geq 0\},\ \{x|x_1\leq 0,\ x_2>0\},\ \{x|x_1>0,\ x_2\leq 0\},$$
$$\{x|x_1\geq 0,\ x_2<0\}.$$

Neither pseudo-concavity nor pseudo-convexity of the function is preserved on the closure of these sets.

It follows from the above remarks that the problems

$$\underset{x\in\Gamma}{\text{Min}}\ (ax+\alpha)\,(bx+\beta)$$

and

$$\underset{x\in\Gamma}{\text{Max}}\ (ax+\alpha)\,(bx+\beta)$$

can be solved by the Frank-Wolfe algorithm provided that for the first problem the linear functions $(ax+\alpha)$ and $(bx+\beta)$ have opposite signs on Γ and one of them does not vanish on Γ, and for the second problem, both linear functions have the same sign on Γ and one of them does not vanish on Γ. The above problems were treated as parametric linear programs in [8].

Remarks

The case handled by Dinkelbach [4] corresponds to the case here where assumptions (A2) and (B2) hold. However Dinkelbach allows $\varphi(x)$ to go positive contrary to our requirement $\varphi(x)\leq 0$ of (B2), but on the other hand he assumes that (i) $\min_{x\in\Gamma}\varphi(x)\leq 0$, and (ii) that his starting point x^1 satisfies the condition $\varphi(x^1)\leq 0$. Under these two assumptions, it is easy to show that the present method will generate a sequence $x^1, x^2, \cdots$, no element of which will satisfy $\varphi(x)>0$ and such that any accumulation point x^0 solves the problem.

We finally remark that the problem of determining the capacity of a discrete, constant, communication channel considered by Meister and Oettli [7] can also be solved directly by the Frank-Wolfe method proposed here for solving nonlinear fractional programs. The negative of the objective function of Meister and Oettli satisfies assumptions (A1) and (B1).[a]

Appendix

(I) Let $\sigma(x)$ be a positive concave function on the convex set $\Gamma\subset E^n$. Then its reciprocal, $\psi(x)=\dfrac{1}{\sigma(x)}$, is a positive convex function on Γ.

(II) Let $\sigma(x)$ be a negative convex function on the convex set $\Gamma\subset E^n$. Then its reciprocal, $\psi(x)=\dfrac{1}{\sigma(x)}$ is a negative concave function Γ.

Proof: (I) Let $x^1, x^2\in\Gamma$, and let $1\geq\lambda\geq 0$. Then

[a] Oettli has informed the author that the direct use of the Frank-Wolfe method for determining the channel capacity, as proposed above, was also proposed in an unpublished earlier version of [7].

$$
\begin{aligned}
\phi((1-\lambda)x^1+\lambda x^2) &= \frac{1}{\sigma((1-\lambda)x^1+\lambda x^2)} \\
&\leq \frac{1}{(1-\lambda)\sigma(x^1)+\lambda\sigma(x^2)} \\
&\qquad\qquad \text{(by concavity of } \sigma(x)\text{)} \\
&\leq \frac{1}{(\sigma(x^1))^{1-\lambda}(\sigma(x^2))^{\lambda}} \\
&\qquad \text{(arithmetic mean} \geq \text{geometric mean)} \\
&= (\phi(x^1))^{1-\lambda}(\phi(x^2))^{\lambda} \\
&\leq (1-\lambda)\phi(x^1)+\lambda\phi(x^2) \\
&\qquad \text{(arithmetic mean} \geq \text{geometric mean)}
\end{aligned}
$$

Hence, $\phi(x)$ positive and convex on Γ.

(II) This part follows from (I) if we apply (I) to the function $\bar{\sigma}(x)=-\sigma(x)$. Then $\frac{1}{\bar{\sigma}(x)}$ is positive and convex on Γ and hence $\frac{1}{\sigma(x)}$ is negative and concave on Γ. Q.E.D.

REFERENCES

[1] Berge, C., and A. Ghouila-Houri, "Programming, games, and transportation networks," John Wiley and Sons, New York, 1965.

[2] Dantzig, G. B., "Linear programming and extensions," Princeton University Press, Princeton, N.J., 1963.

[3] Derman, C., "On sequential decisions and Markov chains," *Management Science*, **9**, 1962, 16–24.

[4] Dinkelbach, W., "On nonlinear fractional programming," *Management Science*, **13**, 1967, 492–498.

[5] Frank, M., and P. Wolfe, "An algorithm for quadratic programming," *Nav. Res. Log. Quart.*, **3**, 1956, 95–110.

[6] Mangasarian, O. L., "Pseudo-convex functions," *J. SIAM Control*, **3**, 1965, 281–290.

[7] Meister, B., and W. Oettli, "On the capacity of a discrete, constant channel," *IBM Research Paper* RZ-228, Zurich, March 23, 1967.

[8] Swarup, K., "Indefinite quadratic programming," Cahiers du Centre d'Etudes de Recherche Operationelle, **8**, 1966, 217–222.