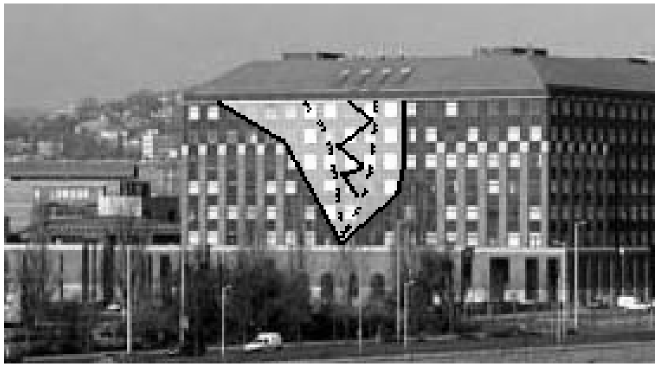
Operations Research
Report 2006-01



Simplex-type algorithm for optimizing a pseudolinear quadratic fractional function over a polytope

Miklós Ujvári

September 2006

Eötvös Loránd University of Sciences
Department of Operations Research

ISSN 1215 - 5918

Simplex-type algorithm for optimizing a pseudolinear quadratic fractional function over a polytope

Miklós Ujvári

Abstract

Recently Cambini and Carosi described a characterization of pseudolinearity of quadratic fractional functions. A reformulation of their result was given by Rapcsák. Using this reformulation, in this paper we describe an alternative proof of the Cambini–Carosi Theorem. Our proof is shorter than the proof given by Cambini–Carosi and less involved than the proof given by Rapcsák. As an application we present a simplex-type algorithm for optimizing a pseudolinear quadratic fractional function over a polytope. Our algorithm works in a more general setting than the convex simplex algorithm adapted to the above problem.

1 Introduction

Let $n \geq 2$, and let $Q \in \mathcal{R}^{n\times n}$ be a symmetric matrix, $Q \neq 0$. Let $q, p \in \mathcal{R}^n$ be vectors, $p \neq 0$, and let $q_0, p_0 \in \mathcal{R}$. Let us denote by H the set of vectors $x \in \mathcal{R}^n$ such that $p^T x + p_0 > 0$. Let us define the so called *quadratic fractional function* $f : H \to \mathcal{R}$ as follows:

$$f(x) := \frac{\frac{1}{2}x^T Qx + q^T x + q_0}{p^T x + p_0} \ (x \in H). \tag{1}$$

Then the gradient map of the function f is

$$\nabla f(x) = \frac{Qx + q - f(x)p}{p^T x + p_0} \ (x \in H). \tag{2}$$

Recently in [2] Cambini and Carosi described a characterization of pseudolinearity of quadratic fractional functions. Their result in an equivalent form was proved by Rapcsák in [8]. We start this paper with stating these equivalent theorems.

Let us recall the definition of pseudolinearity: Let $C \subseteq \mathcal{R}^n$ be an open convex set. A differentiable function $g : C \to \mathcal{R}$ is called *pseudoconvex* if

$$x_1, x_2 \in C,\ (\nabla g(x_1))^T (x_2 - x_1) \geq 0 \text{ implies } g(x_2) \geq g(x_1).$$

The function g is called *pseudoconcave* if $-g$ is pseudoconvex. The function g is called *pseudolinear* if it is pseudoconvex and pseudoconcave.

There exist several characterizations of pseudolinearity of a function in general. We collected some of them in the following well-known useful lemma. The nontrivial part of Lemma 1.1 was proved by Kortanek and Evans in [4] (see also [3]). We remark that the implication "d)⇒c)" is true also if the maps l, η are sufficiently smooth, see [7].

Lemma 1.1. *Let g be a real differentiable function defined on an open convex set $C \subseteq \mathcal{R}^n$. Then between the statements*

a) the function g is pseudolinear;
b) for every $x_1, x_2 \in C$,

$$(\nabla g(x_1))^T (x_2 - x_1) = 0 \iff g(x_1) = g(x_2);$$

c) for every $x_0 \in C$,

$$\{x \in C : g(x) = g(x_0)\} = \{x \in C : (\nabla g(x_0))^T (x - x_0) = 0\};$$

d) there exist maps $l : C \to \mathcal{R}$, $\eta : g(C) \to \mathcal{R}^n$ such that

$$\nabla g(x) = l(x) \cdot \eta(g(x)) \ (x \in C);$$

hold the following logical relations: a) is equivalent to b); b) is equivalent to c); c) implies d).

In the special case of quadratic fractional functions the following characterization of pseudolinearity was proved in [2]:

Theorem 1.1. (Cambini–Carosi) *The function f defined in (1) is pseudolinear on H if and only if f is affine or there exist constants $\alpha, \beta, \kappa \in \mathcal{R}$ such that $\alpha\kappa < 0$ and f can be rewritten in the following form:*

$$f(x) = \alpha p^T x + \beta + \frac{\kappa}{p^T x + p_0} \ (x \in H). \tag{3}$$

An equivalent form of Theorem 1.1 was given in [8]. Theorem 1.2 is constructive, it makes possible the immediate recognition of the constants $\hat{\alpha}$, $\hat{\beta}$.

Theorem 1.2. (Rapcsák) *Let us suppose that* $\nabla f(x) \neq 0$ $(x \in H)$ *where* f *is the function defined in (1). Then the function* f *is pseudolinear if and only if* f *is affine or there exist constants* $\hat{\alpha}, \hat{\beta} \in \mathcal{R}$ *such that* $Q = \hat{\alpha} pp^T$, $q = \hat{\beta} p$.

In §2 we present an alternative proof for Theorem 1.2 and show the equivalence of Theorems 1.1 and 1.2. The resulting proof of Theorem 1.1 is shorter than the proof in [2] and is less involved than the proof in [8]. In §3 we describe a simplex-type algorithm for optimizing a pseudolinear quadratic fractional function over a polytope. Our algorithm works under less stringent conditions than the convex simplex algorithm adapted to this problem.

2 New proof for the Cambini–Carosi Theorem

In this section we will prove the Cambini–Carosi Theorem (Theorem 1.1). First we will prove the theorem's reformulation due to Rapcsák (Theorem 1.2), and then we will show the equivalence of Theorems 1.1 and 1.2.

We begin this section with proving two propositions that will be needed in the proof of Theorem 1.2.

Proposition 2.1. *Let us suppose that the function* f *defined in (1) is pseudolinear, and* $\nabla f(x) \neq 0$ $(x \in H)$. *Then*

a) the matrix Q *has rank 1 or 2;*

b) the matrix Q *has rank 1 if and only if there exist constants* $\hat{\alpha}, \hat{\beta} \in \mathcal{R}$ *such that* $Q = \hat{\alpha} pp^T$, $q = \hat{\beta} p$.

Proof. a) By part d) of Lemma 1.1 there exist maps $l : H \to \mathcal{R}$, $\eta : f(H) \to \mathcal{R}^n$ such that

$$\nabla f(x) = l(x)\eta(f(x)) \ (x \in H). \tag{4}$$

Let $x_0 \in H$. Then for any $x \in H$, $f(x) = f(x_0)$ implies $\eta(f(x)) = \eta(f(x_0))$, that is, by (2) and (4),

$$\frac{1}{l(x)} \cdot \frac{Qx + q - f(x)p}{p^T x + p_0} = \frac{1}{l(x_0)} \cdot \frac{Qx_0 + q - f(x_0)p}{p^T x_0 + p_0}.$$

Rewriting this equation we have

$$Q(x - x_0) = \left(\frac{l(x)(p^T x + p_0)}{l(x_0)(p^T x_0 + p_0)} - 1 \right) (Qx_0 + q - f(x_0)p) \tag{5}$$

for every $x \in H$, $f(x) = f(x_0)$. Furthermore, by part c) of Lemma 1.1,

$$\text{lin}\,\{x - x_0 : x \in H,\ f(x) = f(x_0)\} = \{\nabla f(x_0)\}^\perp \tag{6}$$

where lin denotes linear hull and $\perp$ denotes orthogonal complement. From (5) and (6) follows that

$$Q(\{\nabla f(x_0)\}^\perp) \subseteq \text{lin}\,\{\nabla f(x_0)\}, \tag{7}$$

and so

$$Q(\mathcal{R}^n) \subseteq \text{lin}\,\{\nabla f(x_0), Q\nabla f(x_0)\} \tag{8}$$

where $Q(\mathcal{R}^n)$ denotes the image space of the matrix Q. Hence $Q(\mathcal{R}^n)$ is at most two-dimensional, which means that the rank of the matrix Q is at most 2.

b) If the matrix Q has rank 1 then there exist a constant $\gamma \in \mathcal{R}$, $\gamma \neq 0$ and a vector $c \in \mathcal{R}^n$, $c^T c = 1$ such that $Q = \gamma cc^T$.

It follows easily from (7) that for any $x_0 \in H$ we have

$$Q(\{\nabla f(x_0)\}^\perp) = \{0\}, \text{ or } Q(\{\nabla f(x_0)\}^\perp) = \text{lin}\,\{\nabla f(x_0)\}.$$

We will show that the latter case can not hold if the matrix Q has rank 1. Really, otherwise there would exist a vector $z \in \mathcal{R}^n$ such that

$$\nabla f(x_0)^T z = 0 \text{ and } Qz = \nabla f(x_0).$$

But then $z^T Q z = 0$, that is $\gamma(c^T z)^2 = 0$, and we would have $c^T z = 0$. As $\nabla f(x_0) = Qz = \gamma(c^T z)c$, so $\nabla f(x_0) = 0$ would follow, contradicting the assumption that $\nabla f(x) \neq 0$ $(x \in H)$.

This contradiction shows that

$$Q(\{\nabla f(x)\}^\perp) = \{0\} \ (x \in H).$$

In other words the vectors orthogonal to the vector $\nabla f(x)$ are orthogonal to the vector c also. Hence $\{\nabla f(x)\}^\perp = \{c\}^\perp$, or taking orthogonal complements $\text{lin}\,\{\nabla f(x)\} = \text{lin}\,\{c\}$. Specially,

$$\nabla f(x) \in Q(\mathcal{R}^n) \ (x \in H). \tag{9}$$

We will show that (9) implies that $p, q \in Q(\mathcal{R}^n)$. Really, if (9) holds then by (2) $q - f(x)p \in Q(\mathcal{R}^n)$ for every $x \in H$. Thus $q' - f(x)p' = 0$ where q' and p' denote the vectors we get projecting the vectors q and p resp. to the null space of the matrix Q. As the function f is not a constant function on H, we have $q' = 0 = p'$, that is $p, q \in Q(\mathcal{R}^n)$ as was to be shown.

Thus the vectors p, q are the constant multiples of the vector c. Then there exist constants $\hat{\alpha}, \hat{\beta} \in \mathcal{R}$ such that $Q = \hat{\alpha} pp^T$, $q = \hat{\beta} p$, which proves the nontrivial implication in part b) of the proposition. □

The affine function $f_0 : (x_1, x_2) \mapsto (x_1^2 - x_2^2)/(x_1 + x_2)$ shows that in Proposition 2.1 the matrix Q can have rank 2.

Proposition 2.2. *In addition to the assumptions of Proposition 2.1 let us suppose that* $p_0 = 0$. *Then*

a) if $q_0 \neq 0$ *then the matrix* Q *has rank* 1*;*

b) if $q_0 = 0$ *and the matrix* Q *has rank* 2 *then the function* f *is affine.*

Proof. a) Suppose indirectly that the matrix Q has rank 2. Then there exist constants $\gamma, \delta \in \mathcal{R}$, $\gamma \neq 0 \neq \delta$ and vectors $c, d \in \mathcal{R}^n$, $c^T c = 1 = d^T d$, $c^T d = 0$ such that $Q = \gamma cc^T + \delta dd^T$. We can assume that $p^T c \geq 0 \leq p^T d$.

As the matrix Q has rank 2, by (8) we have

$$Q(\mathcal{R}^n) = \operatorname{lin}\{\nabla f(x), Q\nabla f(x)\} \ (x \in H).$$

Then (9) holds, and similarly as in the proof of Proposition 2.1 we have $q, p \in Q(\mathcal{R}^n)$. On the other hand, the vectors $\nabla f(x), Q\nabla f(x)$ are necessarily linearly independent. This means that the following two vectors are linearly independent also, for every $x \in H$:

$$\begin{aligned}(p^T x) \cdot \nabla f(x) &= \left(\gamma(c^T x) + (q^T c) - f(x)(p^T c)\right) c + \\ &\quad + \left(\delta(d^T x) + (q^T d) - f(x)(p^T d)\right) d \\ (p^T x) \cdot Q\nabla f(x) &= \gamma\left(\gamma(c^T x) + (q^T c) - f(x)(p^T c)\right) c + \\ &\quad + \delta\left(\delta(d^T x) + (q^T d) - f(x)(p^T d)\right) d.\end{aligned}$$

Their linear independency is obviously equivalent with the following statement: $\gamma \neq \delta$ and

$$\gamma(c^T x) + (q^T c) - f(x)(p^T c) \neq 0 \neq \delta(d^T x) + (q^T d) - f(x)(p^T d) \ (x \in H). \quad (10)$$

These inequalities will lead to contradiction.

First, if $p^T c = 0$ then $p^T d > 0$. Let

$$x := -\frac{q^T c}{\gamma} c + d.$$

Then $x \in H$ and $\gamma(c^T x) + (q^T c) = 0$, a contradiction. The case when $p^T d = 0$ can be dealt with similarly.

Thus we can suppose that $p^T c > 0 < p^T d$. The contradiction will be shown by a vector of the form $x = \lambda c$ where $\lambda > 0$ (these vectors are all in the halfspace H). For these vectors (10) means that for $\lambda > 0$

$$\gamma\lambda + (q^T c) - \frac{\frac{1}{2}\gamma\lambda^2 + (q^T c)\lambda + q_0}{(p^T c)\lambda}(p^T c) \neq 0, \quad (11)$$

$$q^T d - \frac{\frac{1}{2}\gamma\lambda^2 + (q^T c)\lambda + q_0}{(p^T c)\lambda}(p^T d) \neq 0. \quad (12)$$

Multiplying the left hand side of both inequalities by the positive constant $(p^T c)\lambda$ we get two quadratic polynomials of the variable λ, let us denote them by $\pi_1(\lambda)$ and $\pi_2(\lambda)$ resp. A quadratic polynomial in the variable λ has a positive (and a negative) root if the coefficients of λ^2 and λ^0 have different signs. From this observation follows easily that

— $0 \in \{\pi_1(\lambda) : \lambda > 0\}$ if $\gamma q_0 > 0$;

— $0 \in \{\pi_2(\lambda) : \lambda > 0\}$ if $\gamma q_0 < 0$.

These statements contradict (11) or (12), and thus (10), showing that the matrix Q can not have rank 2 if $q_0 \neq 0$.

b) Let $a \in \{p\}^\perp$, and let us define a function $\tilde{f}$ the following way:

$$\tilde{f}(y) := f(y - a) \ (y \in \mathcal{R}^n, \ p^T y > 0).$$

The pseudolinearity of the function f implies the pseudolinearity of the function $\tilde{f}$. Furthermore,

$$\tilde{f}(y) = \frac{\frac{1}{2}y^T Q y + \tilde{q}^T y + \tilde{q}_0}{p^T y} \ (y \in \mathcal{R}^n, \ p^T y > 0)$$

where

$$\tilde{q} = q - Qa, \ \tilde{q}_0 = \frac{1}{2}a^T Q a - q^T a.$$

As the matrix Q has rank 2, by part a) $\tilde{q}_0 \neq 0$ can not hold. Thus

$$\frac{1}{2}a^T Q a - q^T a = 0 \ (a \in \{p\}^\perp)$$

or equivalently

$$a^T Q a = 0 = q^T a \ (a \in \{p\}^\perp).$$

The existence of a constant $\hat{\beta} \in \mathcal{R}$ follows such that $q = \hat{\beta} p$; moreover

$$x^T V Q V x = 0 \ (x \in H)$$

where $V \in \mathcal{R}^{n\times n}$ denotes the matrix of the projection map to the subspace $\{p\}^\perp$. For notational convenience we can suppose that $p^T p = 1$. Then $x = (p^T x)p + Vx$ for every $x \in \mathcal{R}^n$. Simple calculation shows that for $x \in H$

$$\begin{aligned} f(x) &= \frac{\frac{1}{2}x^T Qx + q^T x}{p^T x} = \\ &= \frac{\frac{1}{2}p^T Qp(p^T x)^2 + (p^T QVx)(p^T x) + \hat{\beta} p^T x}{p^T x} = \\ &= \left(\frac{1}{2}(p^T Qp)p + VQp\right)^T x + \hat{\beta} \end{aligned}$$

holds. Thus f is an affine function, and the proof of part b) is finished also. □

Now we can prove Theorem 1.2. The "only if" part is an easy consequence of Propositions 2.1 and 2.2. Really, if the function f is pseudolinear then the function $\hat{f}$ is also pseudolinear where $\hat{f}$ is defined the following way:

$$\hat{f}(y) := f(y - \frac{p_0}{p^T p}p) \ (y \in \mathcal{R}^n,\ p^T y > 0)$$

For this function $\hat{f}$ it holds that

$$\hat{f}(y) = \frac{\frac{1}{2}y^T Qy + \hat{q}^T y + \hat{q}_0}{p^T y} \ (y \in \mathcal{R}^n,\ p^T y > 0),$$

where $\hat{q} \in \mathcal{R}^n$, $\hat{q}_0 \in \mathcal{R}$, so we can apply Proposition 2.2. If $\hat{q}_0 \neq 0$ then by part a) of Proposition 2.2 the matrix Q has rank 1, and by part b) of Proposition 2.1 there exist constants $\hat{\alpha}, \hat{\beta} \in \mathcal{R}$ such that $Q = \hat{\alpha}pp^T$, $q = \hat{\beta}p$. In this case the proof is finished. The case when $\hat{q}_0 = 0$ and the rank of Q is 1 can be dealt with similarly. The remaining case is when $\hat{q}_0 = 0$ and the matrix Q has rank 2. In this case by part b) of Proposition 2.2 the function $\hat{f}$, and thus the function f also, is affine.

The "if" part is the consequence of Lemma 1.1. Let us suppose that f is not affine and there exist constants $\hat{\alpha}, \hat{\beta} \in \mathcal{R}$ such that $Q = \hat{\alpha}pp^T$, $q = \hat{\beta}p$. Then

$$f(x) = \alpha p^T x + \beta + \frac{\kappa}{p^T x + p_0} \ (x \in H)$$

and

$$\nabla f(x) = \left(\alpha - \frac{\kappa}{(p^T x + p_0)^2}\right) p \ (x \in H)$$

where

$$\alpha = \frac{1}{2}\hat{\alpha},\ \beta = \hat{\beta} - \frac{1}{2}\hat{\alpha}p_0,\ \kappa = q_0 - \left(\hat{\beta} - \frac{1}{2}\hat{\alpha}p_0\right)p_0.$$

Due to the assumption that $\nabla f(x) \neq 0$ $(x \in H)$, we have $\alpha\kappa < 0$.

It can be easily verified that

$$f(x_2) - f(x_1) = \left(\alpha - \frac{\kappa}{(p^T x_1 + p_0)(p^T x_2 + p_0)}\right) \cdot p^T(x_2 - x_1) \ (x_1, x_2 \in H).$$

As $\alpha\kappa < 0$, so

$$\alpha - \frac{\kappa}{(p^T x_1 + p_0)(p^T x_2 + p_0)} \neq 0 \ (x_1, x_2 \in H).$$

Thus for every $x_1, x_2 \in H$, $f(x_1) = f(x_2)$ if and only if $p^T(x_2 - x_1) = 0$. On the other hand for every $x_1, x_2 \in H$, $(\nabla f(x_1))^T(x_2 - x_1) = 0$ holds if and only if $p^T(x_2 - x_1) = 0$. The pseudolinearity of the function f now follows from part b) of Lemma 1.1, and the proof of Theorem 1.2 is complete. □

Finally, we will show that the Cambini–Carosi Theorem (Theorem 1.1) is a consequence of Rapcsák's reformulation of the theorem (Theorem 1.2).

Let us suppose first that the function f is pseudolinear. If $\nabla f(x_0) = 0$ would hold for some $x_0 \in H$ then f would be a constant function. It is easy to see that f is not a constant function (see Remark 6 in [2]). Thus $\nabla f(x) \neq 0$ $(x \in H)$, and we can apply Theorem 1.2. If f is not affine, then there exist constants $\hat{\alpha}, \hat{\beta} \in \mathcal{R}$ such that $Q = \hat{\alpha}pp^T$, $q = \hat{\beta}p$. In this case (as we have seen already in the course of the proof of Theorem 1.2) there exist constants $\alpha, \beta, \kappa \in \mathcal{R}$ such that $\alpha\kappa < 0$ and (3) holds. This proves the "only if" part of Theorem 1.1.

To prove the other direction let us suppose that there exist constants α, β, κ with the properties described in Theorem 1.1. It can be easily seen that then $\nabla f(x) \neq 0$ $(x \in H)$, so Theorem 1.2 can be applied. Furthermore,

$$\frac{1}{2}x^T Qx + q^T x + q_0 = f(x)(p^T x + p_0) = \alpha x^T pp^T x + (\alpha p_0 + \beta)(p^T x) + (\beta p_0 + \kappa).$$

Here the equality of the second, first and zeroeth derivatives implies

$$Q = 2\alpha pp^T,\ q = (\alpha p_0 + \beta)p,\ q_0 = \beta p_0 + \kappa.$$

Hence with the constants $\hat{\alpha} := 2\alpha$, $\hat{\beta} := \alpha p_0 + \beta$ it holds that $Q = \hat{\alpha}pp^T$ and $q = \hat{\beta}p$. Pseudolinearity of the function f follows from Theorem 1.2, and the Cambini–Carosi Theorem is proved. □

Similar argument shows that Theorem 1.1 implies Theorem 1.2 also, so Theorems 1.1 and 1.2 are in fact equivalent.

3 Description of the algorithm

In this section we will describe a simplex-type algorithm that can be used for optimizing a pseudolinear quadratic fractional function over a polytope.

Let us suppose that f is pseudolinear, and f is not affine. Then by Theorem 1.1 there exist constants $\alpha, \beta, \kappa \in \mathcal{R}$ such that $\alpha\kappa < 0$ and (3) holds. We will consider the minimization of f over a polytope P, where $P \subseteq H$, and P is in the form

$$P = \{x \in \mathcal{R}^n : Ax = b,\, x \geq 0\},$$

for some $A \in \mathcal{R}^{m\times n}$, $b \in \mathcal{R}^m$. We will refer to this problem as Problem (P). If our task is the maximization of f over P, then we can minimize the pseudolinear function $-f$ over P.

Our algorithm works under the following assumptions:

(A1): The rank of the matrix A is m. The polytope P is nonempty. An initial feasible basis and the corresponding simplex tableau is known. (A *feasible basis* is an invertible $B \in \mathcal{R}^{m\times m}$ submatrix of the A matrix such that $B^{-1}b \geq 0$. The corresponding *simplex tableau* is

$$\begin{pmatrix} B & 0 \\ p_B^T & 1 \end{pmatrix}^{-1} \cdot \begin{pmatrix} A & b \\ p^T & -p_0 \end{pmatrix} = \begin{pmatrix} B^{-1}A & B^{-1}b \\ p^T - p_B^T B^{-1}A & -p_B^T B^{-1}b - p_0 \end{pmatrix}.$$

Here $p_B \in \mathcal{R}^m$ is the subvector of $p \in \mathcal{R}^n$ such that the matrix (B^T, p_B) is a submatrix of the matrix (A^T, p).) We can calculate such a basis and the tableau using the first phase of a finite version of the simplex method, see for example [5], [6], [10].

(A2): The set of the feasible solutions, P, is a polytope, that is the convex hull of a finite set of points. In other words P is a bounded polyhedron, that is there is no nonzero $z \in \mathcal{R}^n$ vector such that $Az = 0$, $z \geq 0$ holds. By Gordan's Theorem ([10]) we can reformulate this assumption as follows: there exists a vector $y \in \mathcal{R}^m$ such that $A^T y > 0$.

(A3): The problem is nondegenerate, that is $B^{-1}b > 0$ holds for every feasible basis B. This assumption makes possible the easy calculation of all the one-dimensional faces of the polytope P intersecting in a given extremal point, see Proposition 3.1 below. Later we will weaken this assumption.

Before describing the algorithm, we need some preliminary results concerning the geometry of the problem. These results are stated in the next three propositions.

The *basic feasible solution*, or shortly BFS, corresponding to a feasible basis $B \in \mathcal{R}^{m\times m}$ is the vector $x \in \mathcal{R}^n$ such that $x_B = B^{-1}b$ and $x_{A\setminus B} = 0$, where

the vectors x_B and $x_{A\setminus B}$ are defined similarly as the vector p_B above. The basic feasible solutions corresponding to the feasible bases are exactly the extremal points of the polytope P ([5]). (A convex set $F \subseteq \mathcal{R}^n$ is an *extremal subset*, or shortly a *face*, of P, if $F \subseteq P$ and for every $x_1, x_2 \in P$, $\varepsilon \in \mathcal{R}$, $0 < \varepsilon < 1$, $\varepsilon x_1 + (1-\varepsilon)x_2 \in F$ implies $x_1, x_2 \in F$. We denote by $F \triangleleft P$ that F is a face of P. The vector $x \in \mathcal{R}^n$ is an *extremal point* of P if $\{x\} \triangleleft P$.)

The simplex tableau corresponding to a feasible basis $B \in \mathcal{R}^{m\times m}$ will be denoted by $T(B)$. Note that due to assumption (A2) there is no nonpositive column in the matrix $B^{-1}A$. (Really, if $B^{-1}Ae_j \leq 0$ for some j, where e_j denotes the j-th column vector of the identity matrix, then let $z \in \mathcal{R}^n$ be the vector such that $z_B = -B^{-1}Ae_j$ and $z_{A\setminus B} = e_j$. Then $z \neq 0$ and $Az = 0$, $z \geq 0$, a contradiction.) So in every column of $B^{-1}A$ we can find at least one positive element.

Let K denote the following set of positions in the matrix $B^{-1} \cdot (A \setminus B)$:

$$K := \left\{ (i,j) : 0 < \frac{t_{i,m+1}}{t_{i,j}} = \min_{1\leq l\leq m,\, t_{l,j}>0} \frac{t_{l,m+1}}{t_{l,j}} \right\}, \tag{13}$$

where $t_{i,j}$ $(1 \leq i \leq m+1,\, 1 \leq j \leq n+1)$ denotes the (i,j)-th element of the simplex tableau $T(B)$. It is well-known that if $(i,j) \in K$ and $\hat{B}$ denotes the m by m matrix that we get by exchanging the i-th column vector of B for the j-th column vector of A, then $\hat{B}$ is a feasible basis, and the corresponding simplex tableau, $T(\hat{B})$ can be obtained from $T(B)$ by pivoting on the (i,j)-th position of the latter tableau ([5]).

By Theorem 1.7.7 in [6] we know that the basic feasible solutions x and $\hat{x}$ corresponding to the feasible bases B and $\hat{B}$ resp. are adjacent. (Two extremal points of P, x and $\hat{x}$ are *adjacent* if $[x, \hat{x}] \triangleleft P$, where $[x, \hat{x}]$ denotes the set of all the convex combinations of x and $\hat{x}$.) We will show here that due to the nondegeneracy assumption (A3) the converse is also true, that is all the vectors $\hat{x}$ adjacent to an extremal point x can be obtained by pivoting on positions from K.

Proposition 3.1. *Let $B \in \mathcal{R}^{m\times m}$ be a feasible basis with corresponding basic feasible solution x. Let $\hat{x}$ be an extremal point of P such that x and $\hat{x}$ are adjacent. Then there exists a sequence of feasible bases $B_1, \ldots, B_t$ such that the following statements hold:*

a) the first element of the sequence is $B_1 = B$;

b) the BFS corresponding to $B_1, \ldots, B_{t-1}$ is x;

c) the BFS corresponding to B_t is $\hat{x}$;

d) each two consecutive bases from the sequence $B_1, \ldots, B_t$ have $m-1$ common column vectors.

Furthermore, if $x_B > 0$, then $t = 2$.

Proof. The proof is by the simplex method. Let S denote the (finite) set of the extremal points of P. The extremal point $\hat{x}$ is also an exposed point of P, that is there exists a vector $a_1 \in \mathcal{R}^n$ such that

$$a_1^T \hat{x} < a_1^T \tilde{x} \ (\tilde{x} \in P \setminus \{\hat{x}\}).$$

Similarly there exist a vector $a_2 \in \mathcal{R}^n$ and a constant $\beta \in \mathcal{R}$ such that

$$\beta = a_2^T \tilde{x}_1 < a_2^T \tilde{x}_2 \ (\tilde{x}_1 \in [x, \hat{x}]; \ \tilde{x}_2 \in P \setminus [x, \hat{x}]).$$

Then for every $\varepsilon > 0$,

$$(a_2 + \varepsilon a_1)^T (x - \hat{x}) > 0, \tag{14}$$

and for an appropriately chosen $\varepsilon > 0$,

$$(a_2 + \varepsilon a_1)^T (\tilde{x} - x) > 0 \ (\tilde{x} \in S; \ \tilde{x} \neq x, \hat{x}). \tag{15}$$

Let $a := a_2 + \varepsilon a_1$, and let us minimize the corresponding linear (cost) function over P, using the simplex method, with initial feasible basis B. For a while the value of the cost function does not decrease, the BFS generated by the algorithm remains the same extremal point x, only the feasible bases change. Let these feasible bases be denoted by $B_1, \ldots, B_{t-1}$. After finite steps the cost function decreases; due to (14) and (15) the current BFS will be $\hat{x}$. Let us denote the current feasible basis by B_t. The feasible bases $B_1, \ldots, B_t$ defined this way obviously satisfy statements a), b), c) and d).

If $x_B > 0$, then B is the only feasible basis such that the corresponding BFS is x. Thus necessarily the cost function decreases in the first step, which proves the remaining statement. $\square$

We remark also that pivoting on the positions $(i_1, j) \in K$ and $(i_2, j) \in K$ resp. (when there is a tie in (13)) results in the same $\hat{x}$ BFS (corresponding to different feasible bases). Thus in the $x_B > 0$ case, to calculate all the $\hat{x}$ BFSs adjacent to x, it is enough to chose only one pivot position from each column.

The geometrical meaning of the following proposition is that if a convex cone with apex at an extremal point of P contains the neighbours of this extremal point then the cone contains the whole polytope. (The vector $\hat{x} \in \mathcal{R}^n$ is a *neighbour* of a vector $x \in \mathcal{R}^n$ if x and $\hat{x}$ are adjacent extremal points of the polytope P. The set of neighbours of x will be denoted by $N(x)$.)

Proposition 3.2. *For every extremal point x of P,*

$$\text{cone}\,\{\tilde{x} - x : \tilde{x} \in N(x)\} = \text{cone}\,\{\tilde{x} - x : \tilde{x} \in P\}.$$

(Here cone denotes convex conical hull.)

Proof. For notational convenience we suppose that $x = 0$. We have to prove only that $\text{cone}\,P \subseteq \text{cone}\,N(x)$, as the other inclusion is trivial. By Minkowski's Theorem ([9]) the polytope P is the convex hull of its extremal points, so it is enough to prove that $\hat{x} \in \text{cone}\,N(x)$ for every extremal point $\hat{x}$ of P. Let us denote by S the (finite) set of extremal points of P, and let $\hat{x} \in S$.

The extremal point x is also an exposed point of P, that is there exists a vector $a \in \mathcal{R}^n$ such that

$$a^T x < a^T \tilde{x} \ (\tilde{x} \in P \setminus \{x\}).$$

Let $\beta \in \mathcal{R}$ be a constant between the values $a^T x$ and $\min\{a^T \tilde{x} : \tilde{x} \in S \setminus \{x\}\}$. Let us denote by M the hyperplane $\{\tilde{x} : a^T \tilde{x} = \beta\}$. Then the set $M \cap P$ is a polytope, P', with extremal points in S'. Again by Minkowski's Theorem, P' is the convex hull of S'. There exists a constant $\varepsilon \in \mathcal{R}$, $0 < \varepsilon < 1$ such that $\varepsilon \hat{x} \in P'$. Then the vector $\varepsilon \hat{x}$ is the convex combination of points from S'. So $\hat{x} \in \text{cone}\,S'$. We will show that $\text{cone}\,S' \subseteq \text{cone}\,N(x)$. Let $x' \in S'$, we have to prove that $x' \in \text{cone}\,N(x)$.

Let F denote the minimal face of P that contains x'. Then $x' \in \text{ri}\,F$, where ri denotes relative interior ([9]). So $x' \in M$ is in the relative interior of $F \cap M$. Here $F \cap M$ is a face of $P \cap M$. As $x' \in S'$ also, $x' \in \text{ri}\,(F \cap M)$ is possible only if $F \cap M = \{x'\}$. Let $\hat{F}$ denote the affine hull of F. Then

$$\underbrace{\dim(\hat{F} \cap M)}_{=0} + \underbrace{\dim(\hat{F} \cup M)}_{\leq n} = \dim(\hat{F}) + \underbrace{\dim(M)}_{=n-1},$$

so $\dim \hat{F} \leq 1$. On the other hand $F = \{x'\}$ is impossible, as otherwise x' would be an element of S and M also, but $S \cap M = \emptyset$. Thus F is the convex hull of two extremal points of P taken from the two sides of the hyperplane M. One of these two extremal points has to be x, the other point is an element of $N(x)$. So $x' \in \text{cone}\,N(x)$, and we see that $\text{cone}\,S' \subseteq \text{cone}\,N(x)$. As the inclusions $\text{cone}\,S \subseteq \text{cone}\,S'$ and $\text{cone}\,P \subseteq \text{cone}\,S$ hold also, we have $\text{cone}\,P \subseteq \text{cone}\,N(x)$, and the proof is finished. $\square$

The next proposition states that if an extremal point is optimal over the set of its neighbours then it is optimal over the whole polytope P. We will need the following result: a pseudolinear function is also quasiconcave and strictly quasiconvex (see [1]). (A real function g defined on a convex set C is *quasiconcave* if

$$g(\varepsilon_1 x_1 + \varepsilon_2 x_2) \geq \min\{g(x_1), g(x_2)\}$$

for every $\varepsilon_1, \varepsilon_2 \in \mathcal{R}$, $\varepsilon_1, \varepsilon_2 \geq 0$, $\varepsilon_1 + \varepsilon_2 = 1$ and $x_1, x_2 \in C$. By induction then this latter inequality holds for every finite number of vectors $x_1, x_2, \ldots$ also. A real function g defined on a convex set C is *strictly quasiconvex* if

$$g(\varepsilon x_1 + (1-\varepsilon)x_2) < \max\{g(x_1), g(x_2)\}$$

for every $\varepsilon \in \mathcal{R}$, $0 < \varepsilon < 1$ and $x_1, x_2 \in C$, $g(x_1) \neq g(x_2)$.)

Proposition 3.3. *Let x be an extremal point of the polytope P. If $f(\hat{x}) \geq f(x)$ holds for every $\hat{x} \in N(x)$, then $f(\tilde{x}) \geq f(x)$ holds for every $\tilde{x} \in P$.*

Proof. Again, as in the proof of Proposition 3.2, we can suppose that $x = 0$. Now suppose indirectly that $f(\tilde{x}) < f(x)$ holds for some $\tilde{x} \in P$. Then by Proposition 3.2 $\tilde{x} \in \operatorname{cone} N(x)$, and we have

$$\varepsilon\tilde{x} \in \operatorname{conv}(N(x) \cup \{x\})$$

for some $\varepsilon \in \mathcal{R}$, $0 < \varepsilon < 1$ (conv denotes convex hull). By the strict quasiconvexity and the quasiconcavity of the function f we have

$$f(x) > f(\varepsilon\tilde{x}) \geq \min\{f(\hat{x}) : \hat{x} \in N(x) \cup \{x\}\}.$$

But here the minimum on the right hand side is at least $f(x)$ by assumption. This contradiction shows the validity of the proposition. □

Now we can describe the algorithm, and prove its correctness. Let us suppose first that assumptions (A1), (A2) and (A3) hold.

Algorithm: Let us suppose that a feasible basis $B \in \mathcal{R}^{m\times m}$ is given (initially this basis is the one whose existence is guaranteed by assumption (A1)). The corresponding simplex tableau and basic feasible solution is denoted by $T(B)$ and x resp. Let τ denote the element of $T(B)$ in its lower right corner. Examining one by one the elements of K we decide if there exists an $(i,j) \in K$ such that after pivoting on position (i,j), and denoting by $\hat{\tau}$ the lower right element of the new simplex tableau,

$$(\tau - \hat{\tau}) \cdot \left(\alpha - \frac{\kappa}{\tau\hat{\tau}}\right) < 0$$

holds. If this inequality is satisfied for some $(i,j) \in K$ then let B be the new basis obtained after pivoting on the position (i,j), and start again the algorithm. If

$$(\tau - \hat{\tau}) \cdot \left(\alpha - \frac{\kappa}{\tau\hat{\tau}}\right) \geq 0 \tag{16}$$

holds for every $(i,j) \in K$ then the basic feasible solution x is optimal.

Theorem 3.1. *After finite number of steps the algorithm finds an optimal solution of Problem (P).*

Proof. The correctness of the algorithm above follows easily from Propositions 3.1 and 3.3, and the fact that

$$f(\hat{x}) - f(x) = (\tau - \hat{\tau}) \cdot \left(\alpha - \frac{\kappa}{\tau\hat{\tau}}\right),$$

where $\hat{x}$ denotes the basic feasible solution corresponding to the feasible basis obtained from B after pivoting on the (i,j)-th position. The finiteness of the algorithm follows from the fact that during the algorithm the value $f(x)$ decreases, so there can be no repetition in the sequence of the feasible bases B, and there are only a finite number of such bases. □

Similarly as in the case of the fractional linear programming problem ([1]), the convex simplex algorithm can be adapted to solve Problem (P) also.

The **convex simplex algorithm** in every step verifies if x is a Kuhn–Tucker point, that is (as $x_B > 0$ and $x_{A\setminus B} = 0$) if

$$\nabla f(x)^T - (\nabla f(x))_B^T B^{-1} A \geq 0 \tag{17}$$

holds. Simple calculation shows that (17) is equivalent with the following inequality:

$$(p^T - p_B^T B^{-1} A)\left(\alpha - \frac{\kappa}{\tau^2}\right) \geq 0. \tag{18}$$

If (18) holds, then x is a Kuhn–Tucker point, so due to the pseudolinearity of function f, x is optimal (see Theorem 4.3.7 in [1]).

If (18) does not hold, we have for some j index

$$(p^T - p_B^T B^{-1} A)e_j\left(\alpha - \frac{\kappa}{\tau^2}\right) < 0.$$

Choose an i index such that $(i,j) \in K$ ($x_B > 0$ so there exists such an i), and make a pivot on the position (i,j). For the BFS $\hat{x}$ corresponding to the new feasible basis $\hat{B}$, $f(\hat{x}) < f(x)$ holds. This can be seen as follows: Simple calculation shows that

$$\nabla f(x)^T(\hat{x} - x) < 0.$$

But then (as $\nabla f(\tilde{x}) \neq 0$ for every $\tilde{x} \in P$)

$$\nabla f(x + \varepsilon(\hat{x} - x))^T(\hat{x} - x) < 0 \ (0 < \varepsilon < 1).$$

So the function

$$\varphi : \varepsilon \mapsto f(x + \varepsilon(\hat{x} - x)) \ (0 \le \varepsilon \le 1)$$

strictly decreases on the intervallum $[0, 1]$. Consequently, $f(\hat{x}) < f(x)$. Let $x := \hat{x}$, $B := \hat{B}$, and repeat the above step until optimality is reached.

Theorem 3.2. *The convex simplex algorithm adapted to Problem (P) calculates an optimal solution in finite number of steps.* □

So far our algorithm is nothing else than the convex simplex algorithm in another disguise. Really, the extremal point x is a Kuhn–Tucker point if and only if x is optimal, and by Proposition 3.3 x is optimal if and only if x is optimal over its neighbours. Thus the optimality criteria of the two algorithms are equivalent.

What makes the difference between the two algorithms is that while the convex simplex algorithm relies on the nondegeneracy assumption (A3), our algorithm can be easily generalized to the case when only assumptions (A1) and (A2) hold. In the **generalized algorithm** instead of just *one* (the unique) feasible basis B, we consider *all* the feasible bases B corresponding to the current extremal point x (these bases can be determined in a finite number of steps by pivoting on positions (i, j) such that $x_j = 0$), and we decide whether the optimality criterion (16) holds for all these bases. The resulting algorithm in finite number of steps finds an optimal solution of Problem (P). The proof is analogous to the proof of Theorem 3.1, and is left to the reader.

Conclusion. In this paper we described an alternative proof of a recent result of Cambini–Carosi, a characterization of pseudolinearity of quadratic fractional functions. We proved the Cambini–Carosi Theorem via a less involved reformulation of the theorem due to Rapcsák. Also we used a different characterization of pseudolinearity of functions as a starting-point of our proof than the one used by Cambini–Carosi, and these facts made possible simplifications of the arguments. As an application we presented a simplex-type algorithm for optimizing a pseudolinear quadratic fractional function over a polytope. We compared our algorithm with the convex simplex algorithm adapted to the problem described above, and found that our algorithm can be more easily generalized to the case when the problem is degenerate.

Acknowledgements. I am indebted to Tamás Rapcsák, András Frank and Tibor Illés for the several consultations and for calling my attention to relevant references.

References

1. Bazaraa, M. S. and Shetty, C. M., *Nonlinear programming, theory and algorithms*, John Wiley & Sons, New York, 1979.

2. Cambini, R. and Carosi, L., On generalized linearity of quadratic fractional functions, *Journal of Global Optimization* 30 (2004) 235-251.

3. Chew, K. L. and Choo, E. U., Pseudolinearity and efficiency, *Mathematical Programming* 28 (1984) 226-239.

4. Kortanek, K. O. and Evans, J. P., Pseudo-concave programming and Lagrange regularity, *Operations Research* 15 (1967) 882-892.

5. Murty, K. G., *Linear and combinatorial programming*, John Wiley & Sons, New York, 1976.

6. Prékopa, A., *Linear programming I.*, Bolyai János Matematikai Társulat, Budapest, 1968 (in Hungarian).

7. Rapcsák, T., On pseudolinear functions, *European Journal of Operational Research* 50 (1991) 353-360.

8. Rapcsák, T., On pseudolinearity of quadratic fractional functions, to appear in *Optimization Letters* (2006).

9. Rockafellar, R. T., *Convex analysis*, Princeton University Press, Princeton, 1970.

10. Schrijver, A., *Theory of linear and integer programming*, John Wiley & Sons, New York, 1986.

11. Strang, G., *Linear algebra and its applications*, Academic Press, New York, 1980.

Miklós Ujvári
H-2600 Vác, Szent János utca 1. HUNGARY

Recent Operations Research Reports

2003-01 Zsolt Csizmadia and Tibor Illés: New criss-cross type algorithms for linear complementarity problems with sufficient matrices

2003-02 Tibor Illés and Ádám B. Nagy: A sufficient optimality criteria for linearly constrained, separable concave minimization problems

2004-01 Tibor Illés and Marianna Nagy: The Mizuno–Todd–Ye predictor–corrector algorithm for sufficient matrix linear complementarity problem

2005-01 Miklós Ujvári: On a closedness theorem

2005-02 Faluközy Tamás és Vizvári Béla: Az árutőzsde gabona szekciójának árvárakozásai a kukorica kereskedésének tükrében

2005-03 Bilen Filiz, Zsolt Csizmadia and Tibor Illés : Anstreicher-Terlaky type monotonic simplex algorithms for linear feasibility problems

2005-04 Tibor Illés, Márton Makai, Zsuzsanna Vaik: Railway Engine Assignment Models Based on Combinatorial and Integer Programming

2006-01 Újvári Miklós: Simplex-type algorithm for optimizing a pseudolinear quadratic fractional function over a polytope