

Università degli Studi di Pisa Dipartimento di Statistica e Matematica Applicata all'Economia

Report n. 189

On the Pseudoconvexity of a Quadratic Fractional Function

Alberto Cambini - Jean-Pierre Crouzeix Laura Martein

Pisa, September 2000

- Stampato in Proprio -

On the pseudoconvexity of a quadratic fractional function

Alberto Cambini * Jean-Pierre Crouzeix †
Laura Martein *

September 2000

Abstract

In this paper we give a necessary and sufficient condition for the pseudoconvexity of a function f which is the ratio of a quadratic function over an affine function. The obtained results allow to suggest a simple algorithm to test the pseudoconvexity of f and also to characterize the pseudoconvexity of the sum of a linear and a linear fractional function.

Key Words: Generalized convexity, fractional programming. AMS: subject classifications 26B25, 90C26, 90C29, 90C32

1 Introduction

A fractional programming problem arises whenever the optimization of ratios such as performance/cost, income/investment and cost/time is required. Depending on the nature of the functions describing for instance income, cost, investment, we can obtain linear, quadratic or concave-convex fractional programs. A wide class of problems requires to optimize the ratio of a convex quadratic function over an affine function [9, 11]. This class is particularly

^{*}Department of Statistics and Applied Mathematics, University of Pisa, Via Ridolfi 10, 56124 Pisa, Italy

important since the ratio is then a pseudoconvex function and this property ensures that a local minimum is also global [1].

The aim of this paper is to point out that pseudoconvexity may be still achieved even if the quadratic function is not convex. Unlike [4], the followed approach allows to establish a necessary and sufficient condition for the pseudoconvexity of a quadratic fractional function which can be checked by means of a simple algorithm. These results, when applied to the sum of a linear and a linear fractional function, allow to give a complete characterization of this important class.

2 The fundamental theorem

Through the paper, we are concerned with the pseudoconvexity of the quadratic fractional function

$$f(x) = \frac{n(x)}{d(x)} = \frac{\frac{1}{2}x^{T}Ax - a^{T}x + \alpha}{b^{T}x + \beta}$$
 (2.1)

on the set

$$X = \{x : b^T x + \beta > 0\},$$

where A is a $n \times n$ symmetric matrix, $a, x, b \in \mathbb{R}^n$, $b \neq 0$ and $\alpha, \beta \in \mathbb{R}$. In the following, we denote by $\nu_+(A), \nu_-(A)$ and $\nu_0(A)$ the numbers of positive, negative and zero eigenvalues of A respectively.

Because it is well known (see for instance [1]) that f is pseudoconvex on X when A is positive semidefinite (i.e. when $\nu_{-}(A) = 0$), we assume in this paper that A is not positive semidefinite (i.e. $\nu_{-}(A) \geq 1$). In [4], a necessary and sufficient condition for the pseudoconvexity of f has been established, but this condition is not easily checked. In this section we give a new necessary and sufficient condition leading to a simple algorithm for testing this pseudoconvexity.

Some preliminaries are needed in order to achieve the main result. A twice differentiable function f on an open convex set X is pseudoconvex on X if and only if the two following conditions hold ([6]):

$$x \in X$$
 and $h^T \nabla f(x) = 0 \Rightarrow h^T \nabla^2 f(x) h \ge 0$ (2.2)

$$\hat{x} \in X$$
 and $\nabla f(\hat{x}) = 0 \Rightarrow f(\hat{x}) \le f(x) \ \forall x \in X$ (2.3)

where ∇f and $\nabla^2 f$ denote the gradient vector and the Hessian matrix of f, respectively.

We apply this characterization to the case where f is a quadratic fractional function.

Proposition 2.1 Assume that A is not positive semidefinite. Then f is pseudoconvex on X if and only if the two following conditions hold

$$x \in X \text{ and } h^T[Ax - a - f(x)b] = 0 \Longrightarrow h^T Ah \ge 0$$
 (2.4)

and

$$x \in X \Longrightarrow \nabla f(x) \neq 0.$$
 (2.5)

Proof Easy calculations give

$$\nabla f(x) = \frac{1}{b^T x + \beta} [Ax - a - f(x)b], \qquad (2.6)$$

and

$$(b^T x + \beta) \nabla^2 f(x) = A + \frac{1}{(b^T x + \beta)} [2f(x)bb^T - (Ax - a)b^T - b(Ax - a)^T].$$

Since

$$h^T \nabla f(x) = 0 \iff h^T (Ax - a) = f(x)h^T b,$$

then

$$(b^{T}x + \beta)h^{T}\nabla^{2}f(x)h = h^{T}Ah + \frac{2}{(b^{T}x + \beta)}[f(x)(b^{T}h)^{2} - h^{T}(Ax - a)h^{T}b]$$

= $h^{T}Ah$.

Hence, we deduce that conditions (2.2) and (2.4) are equivalent. Furthermore, when $\nabla f(x) = 0$, condition (2.3) implies $h^T A h \geq 0$ for all h, in contradiction with the assumption on A.

Condition (2.4) needs to check the positive semidefineteness of A on a subspace orthogonal to a vector v. To do that we use the following result ([6]): Assume that A is not positive semidefinite and $v \neq 0$, then

$$(v^T h = 0 \Rightarrow h^T A h \ge 0) \iff \begin{cases} \nu_-(A) = 1, \ v \in A(\Re^n) \\ \text{if } A u = v \text{ then } u^T v \le 0 \end{cases}$$
 (2.7)

It is seen, from simple arguments of linear algebra, that, when A is non singular,

 $(Au_1 = Au_2 = v) \Longrightarrow u_1^T v = u_2^T v.$

Hence there is no ambiguity in (2.7).

The following result is straightforwardly derived from condition (2.7).

Proposition 2.2 Assume that A is not positive semidefinite. Then condition (2.4) is equivalent to the following conditions:

i) $\nu_{-}(A) = 1$;

ii) for all $x \in X$ there exists $u \in \mathbb{R}^n$ (depending on x) such that :

$$Au = Ax - a - f(x)b, (2.8)$$

and

$$u^{T}(Ax - a - f(x)b) \le 0.$$

$$(2.9)$$

We look at the implications of condition ii).

Proposition 2.3 Assume that A is not positive semidefinite and ∇f does not vanish on X. Then f is pseudoconvex on X if and only if the three following conditions hold:

i) $\nu_{-}(A) = 1$;

ii) $\exists c, \ \bar{x} \in \Re^n \text{ such that } Ac = b \text{ and } A\bar{x} = a;$

iii) For all $x \in X$

$$R(x) = f^{2}(x)b^{T}c + 2f(x)(b^{T}\bar{x} + \beta) - 2n(\bar{x}) \le 0.$$
 (2.10)

Proof Assume that f is pseudoconvex. It results from the assumptions that f is not constant on X. Therefore there exist $x_1, x_2 \in X$ with $f(x_1) \neq f(x_2)$. From (2.8) there exist u_1, u_2 such that $Au_1 = Ax_1 - a - f(x_1)b$ and $Au_2 = Ax_2 - a - f(x_2)b$.

Consequently, Ac = b where

$$c = \frac{u_1 - u_2 - x_1 + x_2}{f(x_2) - f(x_1)}.$$

Substituting b = Ac in (2.8), we have $A\bar{x} = a$ where $\bar{x} = x - u - f(x)c$.

In order to prove iii), it is sufficient to note that (2.9) is equivalent to

$$(x - \bar{x} - f(x)c)^T A(x - \bar{x} - f(x)c) \le 0$$

and that

$$(x - \bar{x})^T A(x - \bar{x}) = 2n(x) - 2n(\bar{x}),$$

 $c^T A(x - \bar{x}) = b^T (x - \bar{x}) = b^T x + \beta - \beta - b^T \bar{x}$

so that (2.10) holds. Conversely, if the conditions of the proposition hold, then condition (2.4) holds in view of Propositions 2.1 and 2.2 and f is pseudoconvex on X.

Now, we establish a complete characterization of the pseudoconvexity of f on X. The proof of the theorem is obtained from inequality (2.10).

Theorem 2.1 The function f is pseudoconvex on X if and only if one of the following conditions holds:

i) $\nu_{-}(A) = 0$ (i.e. A is positive semidefinite);

ii) $\nu_{-}(A) = 1$, \bar{x} and c exist so that $A\bar{x} = a$ and Ac = b, $b^{T}c = 0$, $b^{T}\bar{x} + \beta = 0$ and $n(\bar{x}) \geq 0$;

iii) $\nu_{-}(A) = 1$, \bar{x} and c exist so that $A\bar{x} = a$, Ac = b, $b^{T}c < 0$ and $\Delta = (b^{T}\bar{x} + \beta)^{2} + 2n(\bar{x})b^{T}c \leq 0$.

Proof Necessity: Assume that f is pseudoconvex and $\nu_{-}(A) > 0$. Taking into account Proposition 2.1, conditions i), ii) and iii) of Proposition 2.3 hold. Let us study the sign of the trinomial R(x) in (2.10) with respect to the variable f(x). We examine in succession several cases according to the sign of $b^{T}c$.

1. $b^T c = 0$. Then (2.10) is equivalent to

$$x \in X \Rightarrow f(x)(b^T \bar{x} + \beta) \le n(\bar{x})$$
 (2.11)

(a) $b^T \bar{x} + \beta > 0$. Then $\bar{x} \in X$ and (2.11) is equivalent to

$$f(x) \le f(\bar{x}) \ \forall x \in X.$$

Hence $\nabla f(\bar{x}) = 0$ in contradiction with $\bar{x} \in X$ and condition (2.5).

(b) $b^T \bar{x} + \beta < 0$. Then (2.11) is equivalent to

$$f(x) \ge f(\bar{x}) \quad \forall x \in X.$$

Since $\nu_{-}(A) = 1$, there exists v such that $v^{T}Av < 0$ and $b^{T}v \geq 0$. Take some $\hat{x} \in X$, then $x_{t} = \hat{x} + tv \in X$ for all t > 0. It is seen that $f(x_{t}) \to -\infty$ when $t \to +\infty$, in contradiction with (2.10).

- (c) $b^T \bar{x} + \beta = 0$. Then (2.11) is equivalent to $n(\bar{x}) \geq 0$.
- 2. $b^T c > 0$. Consider x_t defined as in 1.(b). Then $R(x_t) \to +\infty$ when $t \to +\infty$. A contradiction since $R(x) \leq 0$ on X.
- 3. $b^T c < 0$. Set $\Delta = (b^T \bar{x} + \beta)^2 + 2n(\bar{x})b^T c$.
 - (a) $\Delta > 0$. Define

$$\gamma_{-} = -\frac{b^T \bar{x} + \beta}{b^T c} - \sqrt{\Delta} \text{ and } \gamma_{+} = -\frac{b^T \bar{x} + \beta}{b^T c} + \sqrt{\Delta}.$$

Then (2.10) holds if and only if $f(X) \subset (-\infty, \gamma_-] \cup [\gamma_+, \infty)$. Notice that f(X) is an interval since X is convex and f is continuous.

- i. $\nu_+(A) > 0$. There is w such that $w^T A w > 0$ and $b^T w \ge 0$. Then $y_t = \hat{x} + tw \in X$ for all t > 0. Consider also x_t defined as in 1.(b). Then, $f(y_t) \to +\infty$ and $f(x_t) \to -\infty$ when $t \to +\infty$. We have a contradiction.
- ii. $\nu_+(A)=0$. Since $\nu_-(A)=1$, n is concave on X so that f is pseudoconcave on X. Let us show that f is not pseudoconvex on X. Let λ be the negative eigenvalue and u be such that $u^Tu=1$ and $Au=\lambda u$. Since $\Re^n=KerA\oplus(\Re\times\{u\})$, we have $A(\Re^n)=\Re\times\{u\}$; since $b\in A(\Re^n)$, there exists $\alpha\in\Re$ such that $b=\lambda\alpha u$. Set $c=\alpha u$, then $b^Tc=\lambda\alpha^2$. Also for any $x\in\Re^n$, y and t exist such that $x-\bar x=y+tu$ and $y\in Ker(A)$. It follows that $b^T(x-\bar x)=\lambda\alpha t$. Taking into account that $a^Ty=\bar x^TAy=0$, $b^Ty=c^TAy=0$, simple calculations give

$$f(x) = \varphi(t) = \frac{\frac{\lambda}{2}t^2 + n(\bar{x})}{t\lambda\alpha + d(\bar{x})}$$

and

$$\varphi'(t) = \frac{\lambda}{(t\lambda\alpha + d(\bar{x}))^2} [\lambda\alpha \frac{t^2}{2} + td(\bar{x}) - \alpha n(\bar{x})].$$

Since

$$\Delta = d^2(\bar{x}) + 2n(\bar{x})\lambda\alpha^2$$

 $\varphi'(t_{-}) = \varphi'(t_{+}) = 0$ where $\lambda \alpha t_{-} = -d(\bar{x}) - \sqrt{\Delta}$ and $\lambda \alpha t_{+} = -d(\bar{x}) + \sqrt{\Delta}$. Then $\nabla f(x_{-}) = \nabla f(x_{+}) = 0$ with $x_{-} = \bar{x} + y + t_{-}u$ and $x_{+} = \bar{x} + y + t_{+}u$. Since $d(x_{+}) = \sqrt{\Delta} > 0$, it results that $x_{+} \in X$ and we have a contradiction with condition (2.5).

iii. $\Delta \leq 0$. Then $R(x) \leq 0$ for all $x \in X$.

Sufficiency: If i) holds, then f is pseudoconvex as the ratio of a convex function over a positive affine function. If ii) or iii) holds, then, in view of Proposition 2.3, it is enough to prove that $\nabla f(x) \neq 0$ for all $x \in X$. Therefore, assume for contradiction that $\nabla f(x) = 0$ and $x \in X$. Then R(x) = 0 and Ax = a + f(x)b. Hence, w exists such that Aw = 0 and $x = \bar{x} + f(x)c + w$. Then, because $b^T w = c^T Aw = 0$,

$$b^T x + \beta = b^T \bar{x} + \beta + f(x)b^T c. \tag{2.12}$$

Assume that ii) holds, then $b^T x + \beta = 0$, hence $x \notin X$. If iii) holds, then, because R(x) = 0, f(x) is a root of the equation

$$\lambda^2 b^T c + 2\lambda (b^T \bar{x} + \beta) - 2n(\bar{x}) = 0.$$

If $\Delta < 0$, there is no such a root, if $\Delta = 0$, then $f(x)b^Tc + (b^T\bar{x} + \beta) = 0$ hence $(b^Tx + \beta) = 0$ in view of condition (2.12) and $x \notin X$.

Remark 2.1 When A is singular, the quantities b^Tc , $d(\bar{x})$ and $n(\bar{x})$ in Proposition 2.3 and Theorem 2.1 do not depend on the vectors c and \bar{x} chosen such that $A\bar{x} = a$ and Ac = b.

3 An algorithm to check the pseudoconvexity of f

The results stated in Theorem 2.1 allow to describe a simple algorithm to check the pseudoconvexity of a quadratic linear fractional function.

- STEP 1: Calculate $\nu_{-}(A)$. If $\nu_{-}(A) > 1$, STOP: f is not pseudoconvex. If $\nu_{-}(A) = 0$, STOP: f is pseudoconvex; otherwise go to STEP 2.
- STEP 2: Solve the linear systems Az = a and Av = b. If one of these systems has no solution STOP: f is not pseudoconvex; otherwise go to STEP 3.
- STEP 3: Calculate $b^T c$. If $b^T c > 0$ STOP : f is not pseudoconvex. If $b^T c = 0$ go to STEP 4; otherwise go to STEP 5.
- STEP 4: Calculate $d(\bar{x})$. If $d(\bar{x}) \neq 0$ STOP: f is not pseudoconvex, otherwise calculate $n(\bar{x})$. If $n(\bar{x}) < 0$ STOP: f is not pseudoconvex otherwise STOP: f is pseudoconvex.
- STEP 5: Calculate $\Delta = d^2(\bar{x}) + 2n(\bar{x})b^Tc$. If $\Delta > 0$ STOP: f is not pseudoconvex otherwise f is pseudoconvex.

It will be noticed that the calculation of $\nu_{-}(A)$ can be obtained in a finite number of steps (unlike the calculation of the eigenvalues). Use, for instance, the Schur's complement method ([5, 6]). Next, for a better understanding of the algorithm, we present some simple numerical examples.

Example 3.1 Consider the quadratic linear fractional function $f: \mathbb{R}^3 \to \mathbb{R}$, where

$$A = \begin{pmatrix} -1 & 2 & -1 \\ 2 & -1 & -1 \\ -1 & -1 & 2 \end{pmatrix}, \quad a = \begin{pmatrix} 3 \\ -3 \\ 0 \end{pmatrix}, \quad b = \begin{pmatrix} -\sqrt{3} \\ 3 + 2\sqrt{3} \\ -3 - \sqrt{3} \end{pmatrix}$$

 $\beta = -3 - 3\sqrt{3}, \ \alpha \in \Re.$

STEP 1

The eigenvalues of A are -1, 1, 0; $\nu_{-}(A) = 1$, go to STEP 2

 $\bar{x}^T = (-1, 1, 0)$ is a solution of the system Az = a and $c^T = (2 + \sqrt{3}, 1, 0)$ is a solution of the system Av = b; go to STEP 3.

STEP 3

It results $b^T c = 0$; go to STEP 4.

STEP 4

It results $d(\bar{x}) = 0$ and $n(\bar{x}) = 3 + \alpha$. It follows that the function f is pseudoconvex if and only if $\alpha \geq -3$.

Example 3.2 Consider the quadratic linear fractional function $f: \mathbb{R}^3 \to \mathbb{R}$, where

$$A = \begin{pmatrix} -1 & 2 & -1 \\ 2 & -1 & -1 \\ -1 & -1 & 2 \end{pmatrix}, \quad a = \begin{pmatrix} 4 \\ 1 \\ -5 \end{pmatrix}, \quad b = \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}$$

 $\beta = 4, \alpha \in \Re$.

STEP 1

The eigenvalues of A are $-1, 1, 0; \nu_{-}(A) = 1$, go to STEP 2

STEP 2

 $\bar{x}^T = (1, 2, -1)$ is a solution of the system Az = a and $c^T = (1, 0, 1)$ is a solution of the system Av = b; go to STEP 3.

STEP 3

It results $b^T c = -1 < 0$; go to STEP 5.

STEP 5

It results $\Delta = 20 - 2\alpha$. It follows that the function f is pseudoconvex if and only if $\alpha \geq 10$.

4 An application

In this section, we apply the results of Section 2, in order to characterize the pseudoconvexity of a function f which is the sum between a linear and a linear fractional function, that is

$$f(x) = q^T x + \frac{d^T x + \alpha}{b^T x + \beta}$$
(4.1)

on the set $X = \{x : b^T x + \beta > 0\}$, where $q, d, b \in \Re^n, \alpha, \beta \in \Re$. We exclude the trivial cases where b = 0 or q = 0. Such a kind of function arises, for instance, in transportation problems ([8]).

This function f is of the form (2.1) with

$$A = qb^T + bq^T, \quad a = -(\beta q + d).$$

Some partial results about the quasiconvexity and/or the quasiconcavity of the function are established in [10, 7]. In this section, we give a complete characterization of the pseudoconvexity. We start with a lemma which points out some properties of the symmetric matrix A.

Lemma 4.1 Assume that q and b are linearly independent. Then $\nu_+(A) = \nu_-(A) = 1$ and $b, q \in A(\Re^n)$.

Proof Notice that $Av = (b^Tv)q + (q^Tv)b$ for all $v \in \Re^n$. Hence $\nu_0(A) \ge n-2$. Take v, w be such that $b^Tv = q^Tw = 1$ and $q^Tv = b^Tw = 0$, such v, w exist. Then $Av = q \ne 0$ and $Aw = b \ne 0$ but $v^TAv = w^TAw = 0$. Hence A cannot be neither positive semi definite nor negative semi definite and therefore $\nu_+(A) \ge 1$ and $\nu_-(A) \ge 1$.

Next, we apply Theorem 2.1 to the function defined in (4.1).

Theorem 4.1 The function is pseudoconvex on X if and only if one of the following conditions holds

i) $q = kb, k \ge 0$;

ii) there is $t \in \Re$ such that d = tb and $\alpha \ge t\beta$.

Proof First of all, let us note that i) is equivalent to i) of Theorem 2.1. We consider in succession the following two cases.

• q = kb with k < 0. We will prove that ii) is equivalent to iii) of Theorem 2.1. Obviously we have $\nu_{-}(A) = 1$. Since $Ab = 2k||b||^2b$, choosing $c = \frac{b}{2k||b||^2}$, we have Ac = b. On the

other hand $d = tb \Leftrightarrow a = -(\beta q + d) \in A(\Re^n) \Leftrightarrow there \ exists \ \bar{x} \ such \ that \ A\bar{x} = 0$

 \bar{x} can be chosen as

$$\bar{x} = \frac{-(\beta k + t)}{2k||b||^2}b.$$

It follows that

$$b^{T}c = \frac{1}{2k} < 0$$
, $(b^{T}\bar{x} + \beta) = \frac{k\beta - \beta t}{2k}$ and $n(\bar{x}) = \alpha - (\frac{\beta k + t}{4k})^{2}$

so that $\Delta = \frac{\alpha - t\beta}{k}$. It follows that $\alpha - t\beta \ge 0$ is equivalent to $\Delta \le 0$.

• b and q are linearly independent. We will prove that ii) is equivalent to ii) of Theorem 2.1. We choose c such that $b^Tc = 0$ and $q^Tc = 1$, such a c exists and verifies condition $Ac = (q^Tc)b + (b^Tc)q = b$. As in the previous case

 $d = tb \Leftrightarrow there \ exists \ \bar{x} \ such \ that \ A\bar{x} = a.$

Since $A\bar{x} = (q^T\bar{x})b + (b^T\bar{x})q = -tb - \beta q$, \bar{x} can be chosen such that $b^T\bar{x} = -\beta$, $q^T\bar{x} = -t$. Furthermore $n(\bar{x}) = -\beta t + \alpha$, so that $n(\bar{x}) \leq 0$ if and only if $\alpha \geq t\beta$.

Remark 4.1 In case ii) of Theorem 4.1 f is of the form

$$f(x) = q^T x + t + \frac{\gamma}{b^T x + \beta}$$
 with $\gamma = \alpha - t\beta \ge 0$.

References

- [1] M. Avriel, W. E. Diewert, S. Schaible and I. Zang, Generalized Concavity, Plenum Press, 1988.
- [2] A. I. Barros, Discrete and Fractional Programming Techniques for Location Models, Kluwer Academic Publisher, 1998.
- [3] A. Cambini, L. Martein and S. Schaible, On Maximizing a Sum of Ratios, in "Journal of Information and Optimization Sciences", Vol. 10, pp. 65-79, 1989.
- [4] A. Cambini, L. Carosi and L. Martein, On the Supremum in Fractional Programming, Technical report N. 153, Dept. Statistics and Applied Mathematics, University of Pisa, 2000.
- [5] Cottle R.W., "Manifestations of the Shur complement", Linear Algebra and its Applications 8, 1974, 189–211.
- [6] J. P. Crouzeix, Characterizations of Generalized Convexity and Monotoniticy, a Survey, in "Generalized Convexity, Generalized Monotoniticy, eds. Crouzeix et al., Kluwer Academic Publisher, pp. 237-256, 1998.

- [7] J. Hirche, On Programming Problems with a Linear-plus-linear-fractional Objective Function, "Cahiers du CERO", Vol. 26, pp. 59-64, 1984.
- [8] P. K. Kanchan, A. S. B. Holland and B. N. Sahney, Transportation Techniques in Linear-plus-fractional Programming, "Cahiers du CERO", Vol. 23, pp. 153-157, 1981.
- [9] S. Schaible, Fractional Programming, in "Handboook of Global Optimization", Horst R., Pardalos P.M. editors, Kluwer Academic Publisher, pp. 495-608, 1995.
- [10] S. Schaible, A Note on the Sum of a Linear and Linear-fractional Function, in "Naval Research Logistics Quarterly", Vol. 24, N. 24, pp. 691-693, 1977.
- [11] I.M. Stancu-Minasian, Applications of the Fractional Programming, in "Economic Computation and Economic Cybernetics Studies and Research, vol. 1, pp. 69-86, 1980.