J Glob Optim
DOI 10.1007/s10898-011-9785-z

On some convexity properties of the Least Squares Method for pairwise comparisons matrices without the reciprocity condition

J. Fülöp · W. W. Koczkodaj · S. J. Szarek

Received: 20 December 2010 / Accepted: 11 September 2011

Abstract The relaxation of the reciprocity condition for pairwise comparisons is revisited from the optimization point of view. We show that some special but not extreme cases of the Least Squares Method are easy to solve as convex optimization problems after suitable nonlinear change of variables. We also give some other, less restrictive conditions under which the convexity of a modified problem can be assured, and the global optimal solution of the original problem found by using local search methods. Mathematical and psychological justifications for the relaxation of the reciprocity condition as well as numerical examples are provided.

Keywords Pairwise comparisons · Convexity properties

1 Introduction

The use of pairwise comparisons (PCs) is one of the most puzzling, intriguing, and controversial scientific methods although, technically speaking, many scientific measurements (e.g., in physics, chemistry, or engineering) reduce to it. Using PCs, we simply compare a

J. Fülöp (✉)
Research Group of Operations Research and Decision Systems, Computer and Automation Research Institute, Hungarian Academy of Sciences, P.O. Box 63, 1518 Budapest, Hungary
e-mail: fulop@sztaki.hu

W. W. Koczkodaj
Computer Science, Laurentian University, Sudbury, ON P3E 2C6, Canada
e-mail: wkoczkodaj@cs.laurentian.ca

S. J. Szarek
Department of Mathematics, Case Western Reserve University, Cleveland, OH 44106-7058, USA
e-mail: szarek@cwru.edu

S. J. Szarek
Université Pierre et Marie Curie-Paris 6, 75252 Paris, France

Journal: **10898-JOGO** Article No.: **9785** ☐ TYPESET ☑ DISK ☐ LE ☑ CP Disp.:**2011/9/21** Pages: **18** Layout: **Small**

distance to an assumed unit of measure, for example, a meter for distance or one liter for liquid capacities in the International System of Units. However, once we allow more general comparisons, the situation becomes much more involved and non-trivial methodological and mathematical issues arise.

The first published use of PCs is attributed to Fechner in 1860 (see [14]). The use of PCs for an election method was promoted by Condorcet in 1785 (see [7], four years before the French Revolution) and by Ramon Llull, or Raimundus Lullus, around 1275 in [29]. However, it was not done in terms of expressing intensity of preferences by real numbers but by a simpler version of "win or tie" only.

Mathematically, an $n \times n$ real matrix $A = [a_{ij}]$ is a pairwise comparison matrix if $a_{ij} > 0$ and $a_{ij} = 1/a_{ji}$ for all $i, j = 1, \ldots, n$. A PC matrix A is consistent if $a_{ij}a_{jk} = a_{ik}$ for all $i, j, k = 1, \ldots, n$. It is easy to see that a PC matrix A is consistent if and only if there exists a positive n-vector w such that $a_{ij} = w_i/w_j$, $i, j = 1, \ldots, n$. For a consistent PC matrix A, the values w_i serve as priorities or implicit weights of the importance of the alternatives.

Several mathematical methods have been put forward for eliciting a suitable vector w of weights of importance (or priority weights) from an inconsistent PC matrix, or for finding the "nearest" consistent PC matrix for a given inconsistent PC matrix. The Eigenvector Method (EM) was proposed in [32]. EM suggests as w the principal (right) eigenvector of A. Another class of approaches is based on optimization methods and proposes different ways for minimizing the difference between A and consistent PC matrices. If the difference to be minimized is measured in the least-squares sense, i.e., in the Frobenius norm, then we get the Least Squares Method (LSM) (analyzed in [6]); we present here the normalized version, see [16]

$$\begin{aligned} &\min \sum_{i=1}^{n} \sum_{j=1}^{n} \left(a_{ij} - \frac{w_i}{w_j}\right)^2 \\ &\text{s.t.} \ \sum_{i=1}^{n} w_i = 1, \quad w_i > 0, \quad i = 1, \ldots, n. \end{aligned} \tag{1}$$

Since the $n \times n$ matrices can be considered as n^2-dimensional vectors (for example, via lexicographic ordering), the above problem may be also thought of as an instance of LSM in the standard Euclidean space. Unfortunately, problem (1) may be a difficult nonconvex optimization problem with multiple local optima and even with multiple isolated global optimal solutions [23,24].

Some authors state that problem (1) has no special tractable form and is difficult to solve [6,18,30,5]. In order to elude the difficulties caused by the possible nonconvexity of (1), several other, more easily solvable problem forms are proposed to derive priority weights from an inconsistent pairwise comparison matrix. The Weighted Least Squares Method (WLSM) [2,6] in the form of

$$\begin{aligned} &\min \sum_{i=1}^{n} \sum_{j=1}^{n} \left(a_{ij} w_j - w_i\right)^2 \\ &\text{s.t.} \ \sum_{i=1}^{n} w_i = 1, \quad w_i > 0, \quad i = 1, \ldots, n \end{aligned} \tag{2}$$

involves a convex quadratic optimization problem whose unique optimal solution is obtainable by solving a system of linear equations. The Logarithmic Least Squares Method (LLSM) [8,9] in the form

Journal: 10898-JOGO Article No.: 9785 TYPESET DISK LE CP Disp.:2011/9/21 Pages: 18 Layout: Small

$$\begin{aligned} \min \; & \sum_{i=1}^{n} \sum_{i=1}^{n} \left[\log a_{ij} - \log \left(\frac{w_i}{w_j} \right) \right]^2 \\ \text{s.t.} \; & \prod_{i=1}^{n} w_i = 1, , \quad w_i > 0, \quad i = 1, \dots, n \end{aligned} \tag{3}$$

is based on an optimization problem whose unique optimal solution is the geometric mean of the rows of matrix A.

For further approaches, see [3,4,13,16] and the references therein. We have to emphasize, however, that the motivation behind most of the optimization-based approaches different from (1) appears to have been – at least partly – eluding the difficulties resulting from the possible nonconvexity of problem (1) by sacrificing the natural approach of the Euclidean distance minimization.

As proved in [16], a necessary condition for the multiple local optima to appear in (1) is that some elements of matrix A are large enough. In [16], by replacing the normalization in (1) to $w_n = 1$, and by using the logarithmic transformation

$$x_i = \log w_i, \quad i = 1, \dots, n, \tag{4}$$

and the univariate function

$$f_a(x) = \left(e^x - a\right)^2 + \left(e^{-x} - 1/a\right)^2 \tag{5}$$

depending on a real parameter a, problem (1) was transcribed into the equivalent form of separable programming:

$$\begin{aligned} \min \; & \sum_{i=1}^{n-1} f_{a_{in}}(x_i) + \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} f_{a_{ij}}(x_{ij}) \\ \text{s.t.} \; & x_i - x_j - x_{ij} = 0, \; i = 1, \dots, n-2, \; j = i+1, \dots, n-1. \end{aligned} \tag{6}$$

As shown in [16], there exists an $a_0 > 0$ such that the univariate function f_a of (5) is strictly convex if and only if $1/a_0 \le a \le a_0$. Consequently, if the condition $1/a_0 \le a_{ij} \le a_0$ is fulfilled for all i, j, then (1) can be transformed into the convex programming problem (6) with a strictly convex objective function to be minimized (see [16], Proposition 2). This implies that both (1) and the equivalent problem (6) have then a unique solution which can be found using standard local search methods. The above-mentioned constant is $a_0 = ((123 + 55\sqrt{5})/2)^{1/4} = \sqrt{\frac{1}{2}\left(11 + 5\sqrt{5}\right)} \approx 3.330191$, which is a reasonable bound for many real-life problems [16,17].

Even if the largest element of a PC matrix is greater than a_0, it is possible to come up with conditions assuring that (1) has a single local minimum, see Corollary 2 of [16]. In all such cases local search methods can be used to find the solution. In this paper we shall consider matrices A for which the reciprocity constraint $a_{ij} = 1/a_{ji}$ for all $i, j = 1, \dots, n$ is not assumed. From the mathematical point of view, A is simply a positive matrix, but we shall call it a nonreciprocal PC matrix in order to emphasize that it is obtained from pairwise comparisons. We will focus on the LSM problem (1) for a nonreciprocal PC matrix A, and we will show how the convexity findings of [16] can be extended to the nonreciprocal case.

In Sect. 2 we discuss the practical importance of nonreciprocal PC matrices, and advert briefly to some basic approaches for eliciting the weights of importance from a nonreciprocal PC matrix, or for approximating it by a consistent PC matrix. Section 3 focuses on a special univariate function of two parameters that plays in our analysis the role similar to that of f_a from [16], and describes the region of the parameters for which the univariate function

Journal: 10898-JOGO Article No.: 9785 TYPESET DISK LE CP Disp.:2011/9/21 Pages: 18 Layout: Small

is convex. In Sect. 4 we present other conditions which assure that a local optimal solution of problem (1) is also a global optimal solution. In Sect. 5 we show numerical examples. A relation to the issue of the small scale, conclusions and final remarks are addressed in Sect. 6.

2 Methods for approximating a nonreciprocal PC matrix by a consistent PC matrix

When we compare X to Y, it is tacitly assumed that Y to X is the reciprocal, and this is called the reciprocity condition. In some cases, however, it may not be reasonable to assume that reciprocity holds. A double blind wine tasting is a crown example for the necessity of relaxation of this condition. However, testing wine is mostly fun and, as such, it would probably not deserve our attention. A more serious reason may be hinted by an urban legend. It is said that the origin of the first freeway has taken place when two construction teams worked from opposite directions missing each other. We do not know whether such event did actually take place, but it is not uncommon for large projects to involve several teams working independently. If such work entails making pairwise comparisons, it may happen that one team compares a component X to Y, while another team may compare Y to X receiving non reciprocal values. In this situation, we have a limited choice: to discard one of the comparisons (facing a non trivial choice problem unless tossing a coin is used) or to come up with some kind of a "golden mean." Our study is exactly about this magic "middle point" solution.

The lack of reciprocity may occur even if the same person performs the comparisons, but at different times. In [10] an experiment is reported. Postgraduate students were asked to fill in the upper part of a PC matrix, where the items to be compared were in the scope of their studies. Four weeks later, they were asked to fill in the lower part of the PC matrix. It turned out that for none of the 23 PC matrices obtained in this way did the reciprocity property hold.

As mentioned in [22], nonreciprocal PC matrices may also appear when one compares financial assets denominated in different currencies. Because of transaction costs, the resulting PC matrices are not reciprocal, even if no subjectivity is involved.

As far as the authors know, the first publication concerning the approximation of a nonreciprocal PC matrix by a consistent one is [25]. They extended problem (3) of LLSM to the nonreciprocal case, and obtained the following optimal solution to (3):

$$w_i = \sqrt[2n]{\frac{\prod_{j=1}^{n} a_{ij}}{\prod_{j=1}^{n} a_{ji}}}, \quad i = 1, \ldots, n. \tag{7}$$

In the case of a reciprocal PC matrix, (7) gives the well-known formula where w_i is the geometric mean of the elements in the i-th row of A. The formula (7) was reinvented and extended by using the maximum likelihood approach in [22]. The results of [25] were also extended in [28] to the case when several decision makers express different opinions per comparison.

The other classical methods for reciprocal PC matrices can also be extended to the nonreciprocal case in more or less direct ways. EM can be interpreted without the reciprocal condition as well [33]. Problem (2) of WLSM has a unique optimal solution that can be obtained, also in the nonreciprocal case, by solving a system of linear equations. Several variants of the goal programming approach were proposed in [10,12,19–21]. Nonreciprocity also appears in fuzzy preference relations, see [11] and its references; for other aspects of the interface between fuzzy sets and pairwise comparisons see [34,35].

Journal: 10898-JOGO Article No.: 9785 TYPESET DISK LE CP Disp.:2011/9/21 Pages: 18 Layout: Small

In this paper the methodology applied in [16] to the LSM problem (1) will be extended to the nonreciprocal case. Since $a_{ij} = 1/a_{ji}$ no longer can be assumed, instead of the univariate function f_a of (5), the univariate function

$$f_{a,b}(x) = \left(e^x - a\right)^2 + \left(e^{-x} - b\right)^2 \tag{8}$$

depending on two (real positive) parameters a and b will now play an important role. Similarly to [16], changing the normalization in (1) to $w_n = 1$, we obtain the equivalent form

$$\begin{aligned} &\min \sum_{i=1}^{n} \sum_{j=1}^{n} \left(a_{ij} - \frac{w_i}{w_j}\right)^2 \\ &\text{s.t.} \quad w_n = 1, \quad w_i > 0, \quad i = 1, \ldots, n-1. \end{aligned} \tag{9}$$

Next, using again the logarithmic transformation (4), problem (3) can be transcribed into another equivalent form of separable programming:

$$\begin{aligned} &\min \sum_{i=1}^{n-1} f_{a_{in},a_{ni}}(x_i) + \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} f_{a_{ij},a_{ji}}(x_{ij}) \\ &\text{s.t.} \quad x_i - x_j - x_{ij} = 0, \; i = 1, \ldots, n-2, \; j = i+1, \ldots, n-1. \end{aligned} \tag{10}$$

Although $a_{ii} = 1, i = 1, \ldots, n$, does not have to be assumed in the nonreciprocal case, the diagonal elements of A do not appear in (10) since they give only the constant contribution $\sum_{i=1}^{n}(a_{ii} - 1)^2$ to the objective function of (9) and can thus be ignored.

Of course, problems (9) and (10) do inherit the numerical difficulties present in the reciprocal case. It would be clearly useful to characterize the matrices $A = [a_{ij}]$ for which the objective function in (10) is convex (or, better, strictly convex) as for such A the difficulties disappear. However, this appears difficult even for small dimensions. Still, an analogue of the condition "$1/a_0 \le a_{ij} \le a_0$ for all i, j," considered in [16] and assuring that the objective functions in (1) and (6) have a single local minimum, can be established also in the nonreciprocal case. Convexity properties of the univariate functions (8) will play a primary role in the investigation.

3 Some convexity properties of the univariate function depending on two parameters

In this section we focus on the question for which parameters a, b the function $f_{a,b}$ is convex, i.e., on identifying the set

$$U := \{(a, b) \in \mathbb{R}^2 \;:\; f_{a,b} \text{ is convex on } \mathbb{R}\}. \tag{11}$$

The set that is of interest is actually

$$U_+ := U \cap \mathbb{R}^2_+ = \{(a, b) \in \mathbb{R}^2_+ \;:\; f_{a,b} \text{ is convex on } \mathbb{R}\}, \tag{12}$$

but its properties will follow immediately from those of U. We will present several different descriptions of U. Propositions 1 and 2 approach U from the point of view of convex analysis, show its convexity and other geometric properties. Proposition 3 gives a tangible description of the set (intercepts, special points) and allows for its easy graphical representation (such as on Fig. 2) using standard software. Finally, Proposition 4 provides an easy test for membership in U.

Journal: 10898-JOGO Article No.: 9785 ☐ TYPESET ☑ DISK ☐ LE ☑ CP Disp.:2011/9/21 Pages: 18 Layout: Small

Since strict convexity is a useful property to assure the uniqueness of the optimal solution in convex programming, we start with the following

Remark 1 If $f_{a,b}$ is convex on $\mathbb{R}$, then it is actually strictly convex. In other words,

$$U = \{(a, b) \in \mathbb{R}^2 \ : \ f_{a,b} \ \text{ is strictly convex on } \mathbb{R}\}.$$

Indeed, assume that $f_{a,b}$ is convex but not strictly convex. Then $f'_{a,b}$ is constant and $f''_{a,b}$ is zero over a nondegenerate interval of $\mathbb{R}$. A direct calculation gives

$$f''_{a,b}(x) = 4e^{-2x} - 2be^{-x} - 2ae^{x} + 4e^{2x}. \tag{13}$$

Substituting $v = e^x > 0$, we have $f''_{a,b}(x) = 0$ exactly when $\frac{2}{v^2} - \frac{b}{v} - av + 2v^2 = 0$, or

$$2v^4 - av^3 - bv + 2 = 0. \tag{14}$$

The quartic polynomial in (14) may have at most a finite number (more precisely, by Descartes's rule of sign [27], at most 2) of positive roots; consequently, (13) can not have an infinite number of roots, either. This contradicts $f''_{a,b}$ vanishing on an interval and proves the remark.

Proposition 1 *The sets U and $U_+ = U \cap \mathbb{R}^2_+$ are nonempty and convex. Moreover*

$$(a, b) \in U \text{ if and only if } (b, a) \in U, \tag{15}$$

and

$$U = \{(a, b) \in \mathbb{R} : b \leq g(a)\}, \tag{16}$$

where $g : \mathbb{R} \to \mathbb{R}$ is a decreasing, differentiable, strictly concave bijection.

Some of the assertions of Proposition 1 follow from known facts or from elementary convex analysis. For example, we know from [16] that if $1/a_0 \leq a \leq a_0$ and $b = 1/a$, where $a_0 = \sqrt{\frac{1}{2}\left(11 + 5\sqrt{5}\right)}$, then $(a, b) \in U_+$, thus U_+ (and *a fortiori* U) is nonempty. However, the finer properties of U require a more detailed analysis; our strategy will be to conceptualize the definition of U and then deduce Proposition 1 from the general theory.

To that end, we note first that the function $f_{a,b}(x)$ is convex on $\mathbb{R}$ exactly when $f''_{a,b}(x) \geq 0$ for all $x \in \mathbb{R}$; by (13) and the discussion that followed it, this is equivalent to $\frac{2}{v^2} - \frac{b}{v} - av + 2v^2 \geq 0$ for all $v > 0$, or

$$av + \frac{b}{v} \leq \frac{2}{v^2} + 2v^2 \ \text{ for all } v > 0. \tag{17}$$

This condition can be rewritten as

$$as + bt \leq 2(s^2 + t^2) \quad \text{for all} \quad s, t > 0 \quad \text{such that} \quad st = 1$$

or, in the homogeneous form,

$$as + bt \leq \frac{2(s^2 + t^2)}{\sqrt{st}} \quad \text{for all} \quad s, t > 0.$$

In other words, we need to figure out for which $(a, b) \in \mathbb{R}^2_+$ we have the implication

$$s, t > 0 \text{ and } \frac{2(s^2 + t^2)}{\sqrt{st}} \leq 1 \ \Rightarrow \ as + bt \leq 1 \tag{18}$$

Journal: 10898-JOGO Article No.: 9785 TYPESET DISK LE CP Disp.:2011/9/21 Pages: 18 Layout: Small

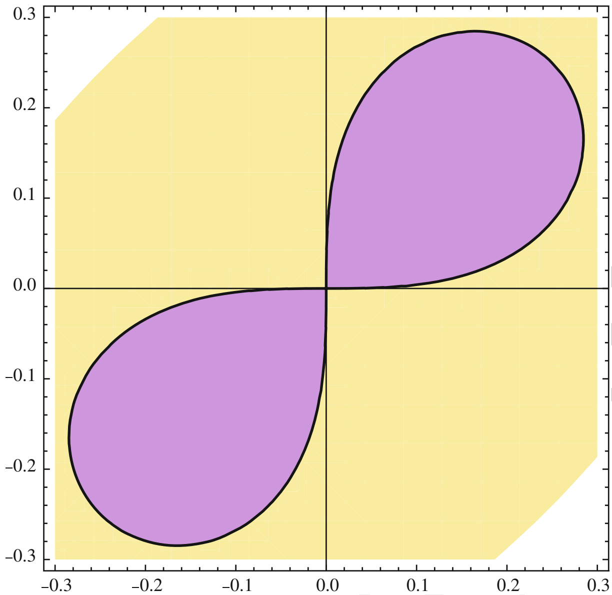

Fig. 1 The lemniscate $r^2 = \frac{1}{8}\sin(2\theta)$; the region L defined by (20) is its interior, while the region L_+ is the leaf of L lying in the first quadrant

(alternatively, we could have used the constraint $\frac{2(s^2+t^2)}{\sqrt{st}} = 1$). To interpret the above condition we recall the concept of a *polar body*. For a subset $K \subset \mathbb{R}^d$, we define K°, the polar body of K, by

$$K^\circ := \{\mathbf{u} \in \mathbb{R}^d : \forall\, \mathbf{y} \in K \quad \langle \mathbf{u}, \mathbf{y} \rangle \leq \mathbf{1}\}, \tag{19}$$

where $\langle \cdot, \cdot \rangle$ is the scalar product. This concept is extensively used in convex analysis, convex optimization, geometry and functional analysis. In the functional-analytic context, if K is the unit ball with respect to some norm, K° is the unit ball with respect to the dual norm (for example, if K is the unit ball in ℓ_p, then K° is the unit ball in ℓ_q, where $\frac{1}{p} + \frac{1}{q} = 1$). A very useful fact is "the bipolar theorem," which says that if K is closed, convex and contains the origin (the assumptions we will implicitly make in what follows, and which are satisfied for all sets of interest), then $(K^\circ)^\circ = K$. Sample references for these topics are [1,31].

In the above terminology, it is apparent that our "convexity region" U coincides with the polar body of the set $L_+ := L \cap \mathbb{R}^2_+$, where

$$L := \left\{(s, t) \in \mathbb{R}^2 \;:\; (s^2 + t^2)^2 \leq \frac{1}{4} st\right\}. \tag{20}$$

The boundary of this set, given by the equation $(s^2 + t^2)^2 = \frac{1}{4}st$, or $r^2 = \frac{1}{8}\sin(2\theta)$ in polar coordinates, is the lemniscate of Bernoulli (see Fig. 1). The more commonly quoted form of the equation is $r^2 = \frac{1}{2}\alpha^2 \cos(2\theta)$, where $\alpha > 0$ is a parameter, which produces a (rotated by $\pi/4$) horizontal "infinity" sign.

Our discussion can be summarized in the following statement, of which Proposition 1 will be a straightforward consequence.

Proposition 2 *The set U of Proposition 1 coincides with $(L_+)^\circ$, the polar body of the "upper right" leaf of the lemniscate $r^2 = \frac{1}{8}\sin(2\theta)$. Similarly, $U_+ = L^\circ \cap \mathbb{R}^2_+$.*

The second assertion follows from the fact that if we are only interested in $a, b > 0$, the points (s, t) in the "left lower" leaf of L do not affect the condition in (18) since for such a, b, s, t we have $as + bt \leq 0 < 1$.

Proof of Proposition 1 First, any polar K° is a closed convex set: for a fixed $\mathbf{y} \in K$, the constraint $\langle \mathbf{u}, \mathbf{y} \rangle \leq \mathbf{1}$ represents a closed half-plane, and so K° is the intersection of a family of such half-planes. Next, since $L_+ \subset \mathbb{R}^2_+$, $U = (L_+)^\circ$ is a hypograph of some *concave* non-increasing function g; this is because $(a, b) \in U$ implies that every (x, y) with $x \leq a, y \leq b$ also belongs to U. Further, smoothness of K corresponds to strict convexity of K° and *vice versa* (with some caveats, the only one that is relevant here being that if the origin is a boundary point, it needs to be ignored from in the hypothesis and in the assertion; recall that we consider only closed, convex sets containing the origin); this follows essentially from the definitions, see [15] for much stronger results. Since $\partial L_+ \setminus (0, 0)$ is smooth and strictly convex, so is $\partial U = \partial(L_+)^\circ$. In particular, ∂U does not contain vertical or horizontal segments and so the function g is strictly decreasing (*a priori* defined on some interval $(-\infty, c)$), smooth and *strictly* concave by the preceding observation. The point $(0, 0)$ on the boundary of L_+ plays a special role. In general, $(0, 0) \in \partial K$ iff K° is unbounded. Similarly, $K \subset \{(x, y) : x \leq c\}$ is equivalent (just by the definition (19)) to $(c^{-1}, 0) \in K^\circ$. Consequently, since no point on the x-axis other than the origin belongs to $L_+ = ((L_+)^\circ)^\circ = U^\circ$, it follows that U extends unboundedly in the positive direction of the x-axis. In other words, the domain of g is $\mathbb{R}$ (and not some interval $(-\infty, c)$ with $c < \infty$). Similarly, the range of g is $\mathbb{R}$; this can also be deduced from the symmetry property (15) of U, which in turn follows from the analogous symmetry of the leminiscate and of its leaf L_+. □

The same circle of ideas leads to a parametric description of U, which is more tangible than general statements such us Propositions 1 or 2. We have

Proposition 3 *The boundary of the set U of Proposition 1 is parametrized by*

$$a = \frac{2\sqrt{2}\sin(3\theta)}{\sin^{3/2}(2\theta)}, \quad b = -\frac{2\sqrt{2}\cos(3\theta)}{\sin^{3/2}(2\theta)}, \ 0 < \theta < \frac{\pi}{2}.$$

The part of the boundary that lies in the first quadrant corresponds to the range $\frac{\pi}{6} \leq \theta \leq \frac{\pi}{3}$. The x- and y-intercepts of g are $8/3^{3/4} \approx 3.50953$. The boundary intersects the line $x = y$ at $(2, 2)$.

Proof Generally, if K is smooth, strictly convex and contains the origin, then there is a canonical map $\Phi : \partial K \to\in \partial(K^\circ)$ (with the usual exclusion of the origin, if it happens to be a boundary point) uniquely defined by the condition

$$q = \Phi(p) \iff \langle q, p \rangle = 1.$$

Specifically, Φ is the Gauss map (which associates to a point p on a hypersurface S a unit vector normal to S at p) renormalized so that the above scalar product equals to 1. In particular, any parametrization of ∂K leads in a routine way to a parametrization of $\partial(K^\circ)$. In our setting we may consider the parametrization of the boundary of L_+ given by the polar coordinate θ, i.e., $s = r\cos\theta = 2\sqrt{2\sin(2\theta)}\cos\theta$ and $t = r\sin\theta = 2\sqrt{2\sin(2\theta)}\sin\theta$, $0 \leq \theta \leq \pi/2$. The tangent vector at (s, t) is $(\dot{s}, \dot{t})$ (where $\dot{}$ is differentiation with respect θ), the normal vector is $(\dot{t}, -\dot{s})$, hence $\Phi = \frac{(\dot{t}, -\dot{s})}{\langle (s,t), (\dot{t}, -\dot{s}) \rangle}$. The range in the second statement is determined by solving $b = 0$ and $a = 0$ for θ, and the intercepts by substitution. Since we are appealing to polar coordinates, the point on the line $x = y$ corresponds to $\theta = \frac{\pi}{4}$. □

Journal: **10898-JOGO** Article No.: **9785** ☐ TYPESET ☑ DISK ☐ LE ☑ CP Disp.:**2011/9/21** Pages: **18** Layout: **Small**

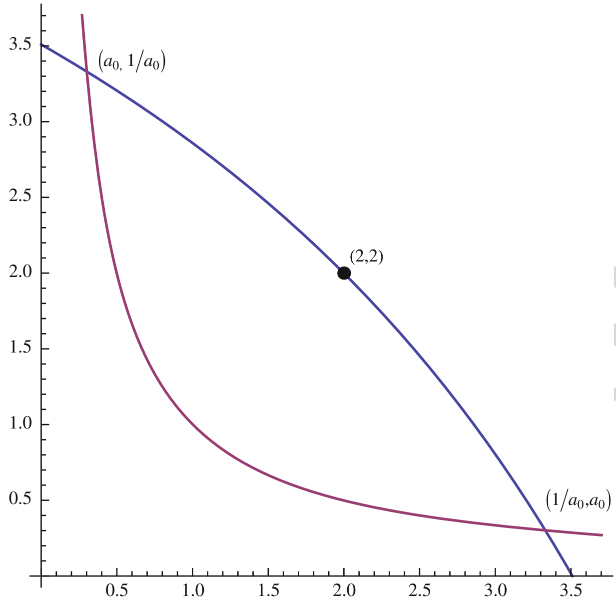

Fig. 2 The region of convexity U, which is exactly the part of the polar body of the lemniscate $r^2 = \frac{1}{8}\sin(2\theta)$ that lies in the first quadrant. Also shown is the hyperbole $b = 1/a$ representing the reciprocal restriction; the intersection points are the points $(a_0, 1/a_0)$ and $(1/a_0, a_0)$ addressed in [16]. The a- and b-intercepts are $8/3^{3/4} \approx 3.50953$

It is now very easy to represent the region U graphically using standard software. Its part lying in the first quadrant, relevant to our investigation, is shown in Fig. 2, together with the hyperbole $b = 1/a$ representing the reciprocal restriction.

The above analysis may be alternatively performed by focusing on L°, the polar of the *entire* lemniscate. This actually simplifies the argument a little since the polar of a set is the same as that of its convex hull, and $\tilde{L}$, the convex hull of L, is a compact, convex, origin-symmetric set with the origin in its interior. This class of bodies is invariant under the polar functor. The bodies $\tilde{L}$ and $(\tilde{L})^\circ = L^\circ$ are shown in Fig. 3.

While the descriptions of U given above, particularly Proposition 2, are mathematically elegant, they do not provide a simple test for membership in the set. A natural way to provide such test would be to express the function g of (16) in an explicit form. To attempt that, rewrite (17) as

$$b \le \frac{2v^4 - av^3 + 2}{v} =: G_a(v) \text{ for all } v > 0. \tag{21}$$

Since $\lim_{v\to 0+0} G_a(v) = \lim_{v\to\infty} G_a(v) = \infty$, it follows that $\inf_{v>0} G_a(v)$ is attained and hence $g(a) = \min_{v>0} G_a(v)$. On the other hand, a direct calculation shows that $G'_a(v) = 0$ iff $3v^4 - av^3 - 1 = 0$. An analysis of the coefficients of the polynomial $3v^4 - av^3 - 1$ shows (see again Descartes's rule of sign [27]) that it has always exactly one positive root v_a, which corresponds to the single global minimum of G_a over $v > 0$. In particular $g(a) = G_a(v_a)$. Ferrari's formula [27] can now be used to find an explicit expression for v_a, and substituting it into G_a we can obtain an explicit, but *very* ugly formula for $g(a)$. Hence, rather than to pursue this route, we will employ an alternative, less direct approach, namely the resultant method applied in [16].

Journal: **10898-JOGO** Article No.: **9785** ☐ TYPESET ☑ DISK ☐ LE ☑ CP Disp.:**2011/9/21** Pages: **18** Layout: **Small**

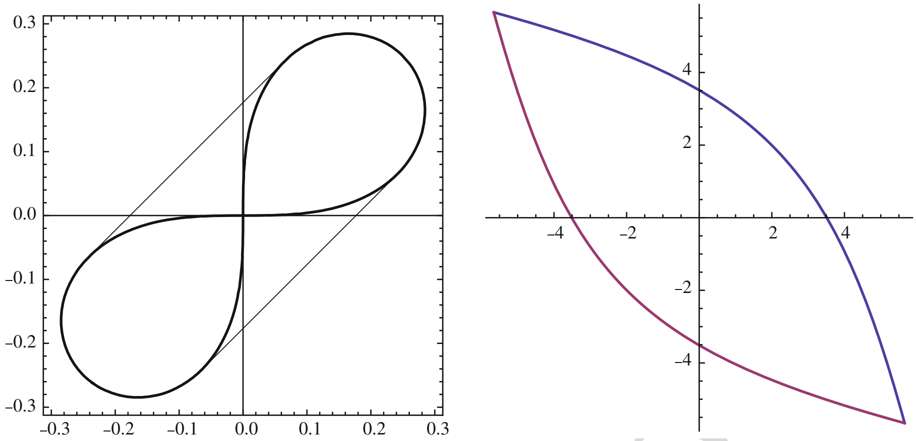

Fig. 3 $\tilde{L}$, the convex hull of the lemniscate $r^2 = \frac{1}{8}\sin(2\theta)$, and its polar $(\tilde{L})^\circ = L^\circ$, which is the intersection of U and $-U$, its symmetric image with respect to the origin. The "*sharp*" points where the boundaries of U and $-U$ meet correspond to the points where the boundaries of L and $\tilde{L}$ separate (in the picture on the *left*), and to the values $\theta = \frac{\pi}{12}$ and $\frac{5\pi}{12}$ in the parameterizations described in the text

By (21) and the discussion following it, $b = g(a) = G_a(v_a) = \min_{v>0} G_a(v)$, which implies that v_a is a minimum point of the polynomial $2v^4 - av^3 - bv + 2$ over $v > 0$. Since $(2v^4 - av^3 - bv + 2)' = 8v^3 - 3av^2 - b$, we have

$$8v_a^3 - 3av_a^2 - b = 0.$$

This means that v_a is a common root of the polynomial system

$$\begin{aligned} 2v^4 - av^3 - bv + 2 &= 0, \\ 8v^3 - 3av^2 - b &= 0. \end{aligned} \qquad (22)$$

According to the resultant method [27], (22) has a common root if and only if

$$\begin{vmatrix} 2 & -a & 0 & -b & 2 & 0 & 0 \\ 0 & 2 & -a & 0 & -b & 2 & 0 \\ 0 & 0 & 2 & -a & 0 & -b & 2 \\ 8 & -3a & 0 & -b & 0 & 0 & 0 \\ 0 & 8 & -3a & 0 & -b & 0 & 0 \\ 0 & 0 & 8 & -3a & 0 & -b & 0 \\ 0 & 0 & 0 & 8 & -3a & 0 & -b \end{vmatrix} = 0. \qquad (23)$$

After some computations (the use of tools of symbolic computation is recommended), we obtain that the value of the determinant on the left-hand side of (23) is

$$-216a^4 - 8a^3b^3 - 48a^2b^2 - 6144ab - 216b^4 + 32768.$$

After simplifications, we are led to an easy to verify necessary and sufficient condition for membership in the convexity region U.

Proposition 4 *The point (a, b) belongs to U (i.e., $b \le g(a)$) if and only if*

$$27a^4 + a^3b^3 + 6a^2b^2 + 768ab + 27b^4 \le 4096. \qquad (24)$$

The point (a, b) lies on the boundary of U, or $b = g(a)$, if and only if (24) holds as an equality.

4 Sufficient conditions for the global optimality of a local optimal solution

Let the function $F : \mathbb{R}^{n-1} \to \mathbb{R}$ be defined by

$$F(x_1, \ldots, x_{n-1}) = \sum_{i=1}^{n-1} f_{a_{in}, a_{ni}}(x_i) + \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} f_{a_{ij}, a_{ji}}(x_i - x_j). \tag{25}$$

Problem (10) is equivalent to

$$\min_{(x_1, \ldots, x_{n-1}) \in \mathbb{R}^{n-1}} F(x_1, \ldots, x_{n-1}). \tag{26}$$

We start our analysis of problem (26) with the following observation, which is a multivariate analogue of Remark 1.

Proposition 5 *Assume that the function F of* (25) *is convex over a convex set $C \subset \mathbb{R}^{n-1}$. Then F is strictly convex over C.*

Proof Let $\bar{\mathbf{x}}$ and $\hat{\mathbf{x}}$ be two arbitrary points of C such that $\bar{\mathbf{x}} \neq \hat{\mathbf{x}}$. Let

$$\varphi(\lambda) = F(\lambda \bar{\mathbf{x}} + (1 - \lambda)\hat{\mathbf{x}}) \tag{27}$$

and

$$\psi(\lambda) = \lambda F(\bar{\mathbf{x}}) + (1 - \lambda) F(\hat{\mathbf{x}}), \tag{28}$$

where $\lambda \in \mathbb{R}$. We will show that

$$\varphi(\lambda) < \psi(\lambda) \ \text{ for all } \ \lambda \in (0, 1). \tag{29}$$

Assume that, on the contrary, (29) does not hold. Then, from the convexity of φ over [0, 1], we obtain

$$\varphi(\lambda) = \psi(\lambda) \ \text{ for all } \ \lambda \in (0, 1). \tag{30}$$

Since $f_{a,b}$ of (8) is always a univariate analytic function, it is easy to see that the multivariate function F of (25) and the univariate function φ of (27) are also analytic, see e.g. [26], as is the affine function ψ of (28). According to the basic properties of analytic functions, (30) implies that

$$\varphi(\lambda) = \psi(\lambda) \ \text{ for all } \ \lambda \in \mathbb{R}, \tag{31}$$

see e.g. Corollary 1.2.5 in [26]. Since $\bar{\mathbf{x}} \neq \hat{\mathbf{x}}$, there exists an index i_0 such that $\bar{x}_{i_0} \neq \hat{x}_{i_0}$. Then $f_{a_{i_0 n}, a_{n i_0}}(\lambda \bar{x}_{i_0} + (1 - \lambda)\hat{x}_{i_0}) \longrightarrow \infty$ both for $\lambda \longrightarrow -\infty$ and $\lambda \longrightarrow \infty$. Taking also into account that the univariate functions appearing in (25) are nonnegative, we obtain

$$\varphi(\lambda) \longrightarrow \infty \text{ both for } \lambda \longrightarrow -\infty \text{ and } \lambda \longrightarrow \infty. \tag{32}$$

However, (32) cannot hold for the affine function ψ of (28), contradicting thus (31). This shows (29) and completes the proof. □

We are now able to state and prove the main result of this paper.

Theorem 1 *If*

$$(a_{ij}, a_{ji}) \in U, \ \ i = 1, \ldots, n-1, \ \ j = i+1, \ldots, n, \tag{33}$$

then problem (1) *has a unique local (thus global) optimal solution, the objective function of* (10) *is strictly convex over the feasible set, and* (10) *has a unique local (thus global) minimizer point.*

Journal: 10898-JOGO Article No.: 9785 TYPESET DISK LE CP Disp.:2011/9/21 Pages: 18 Layout: Small

Proof If (33) holds, then (by the definition of U) each univariate function in (25) is convex, and so the entire sum is also convex as a function on $\mathbb{R}^{n-1}$. Proposition 5 implies now that F is strictly convex on $\mathbb{R}^{n-1}$. (Actually it is not difficult to deduce this property just from Remark 1, but we choose the present argument to reduce repetitiveness.) Since $\lim_{x\to-\infty} f_{a,b}(x) = \lim_{x\to+\infty} f_{a,b}(x) = \infty$ for any $(a, b) \in \mathbb{R}^2$, the lower level sets of F are bounded, and hence F has a (unique) local, thus global, minimizer. This argument also shows that the objective function of (10) is strictly convex over the feasible set and that problem (10) has a unique local, thus global, optimal solution. The convexity of the objective function does not necessarily hold for the original problems in form of (1) or (9), but the nonlinear coordinate transformation (4) and its inverse assure the one-to-one correspondence between the local—hence global—optimal solutions of the problems in the spaces of $(w_1, \ldots, w_n)$ and $(x_1, \ldots, x_{n-1})$, respectively. □

Remark 2 If (33) holds, then the equivalent problems (1), (9) and (10) can be solved by local search techniques starting from any feasible point.

Relation (33) is a sufficient but not necessary condition of the convexity of F. Example 1 of Sect. 5 shows a numerical example where (33) does not hold but F is strictly convex. Namely, it may happen that the strict convexity of several univariate functions in (25) compensates for "small" nonconvexities of some other univariate functions in (25), and hence F is convex.

As pointed out earlier, our main result provides only a *sufficient* condition for uniqueness of local (thus global) minimizer in (10). Accordingly, it is a tempting project to characterize all matrices $A = [a_{ij}]$ for which the objective function F of (26) is convex, as we could immediately generalize Theorem 1 and Remark 2 to such matrices. However, this appears difficult even for small dimensions greater than 2. The following observation goes in the direction of the project and appears fundamental enough to warrant inclusion.

Proposition 6 *Let $\mathcal{C}$ be the set of matrices $A = [a_{ij}]$, for which the objective function F of* (26) *is convex on $\mathbb{R}^{n-1}$. Then*

- *$\mathcal{C}$ is convex as a subset of the space of $n \times n$ matrices*
- *$\mathcal{C}$ is monotone in the sense that if $A \in \mathcal{C}$ and $B = [b_{ij}]$ is such that $b_{ij} \leq a_{ij}$ for all i, j, then $B \in \mathcal{C}$*
- *$\mathcal{C}$ is symmetric: if $A \in \mathcal{C}$, then if $A^T \in \mathcal{C}$ and $PAP^T \in \mathcal{C}$ for every permutation matrix P.*

Since we do not need the Proposition in our analysis, we skip the proof. We only mention that, similarly as in Proposition 2, the argument depends on interpreting $\mathcal{C}$ as a polar of some set $\mathcal{L}$ of matrices with respect to the trace duality $\langle A, B\rangle := \operatorname{tr} AB^T$, with the second and the third assertion following from the "positivity" and the symmetries of $\mathcal{L}$.

From a practical aspect, it is not indispensable that F be *globally* convex on $\mathbb{R}^{n-1}$. Assume that we can somehow guarantee that all global minimizers of F are in a convex set $C \subset \mathbb{R}^{n-1}$. If we can show that F is convex over C, then (26) can be replaced by the convex programming problem

$$\min_{(x_1,\ldots,x_{n-1})\in C} F(x_1, \ldots, x_{n-1}). \tag{34}$$

A compact convex set containing all global minimizers of F can be easily constructed. To that end, let γ be the (nonnegative) value of the objective function F at some point of $\mathbb{R}^{n-1}$. We are interested only in such points for which

$$F(x_1, \ldots, x_{n-1}) \leq \gamma \tag{35}$$

Journal: 10898-JOGO Article No.: 9785 TYPESET DISK LE CP Disp.:2011/9/21 Pages: 18 Layout: Small

holds. The value of F in (25) is the sum of nonnegative values of univariate functions. For any $f_{a,b}(x)$ appearing in (25) the constraint

$$f_{a,b}(x) \leq \gamma \tag{36}$$

is a relaxation of (35). From (36) we obtain

$$(e^x - a)^2 \leq \gamma$$

and

$$\log(a - \sqrt{\gamma}) \leq x \leq \log(a + \sqrt{\gamma}). \tag{37}$$

If $a - \sqrt{\gamma} \leq 0$, the left-hand-side of (37) is considered to be $-\infty$. Similarly, from

$$(e^{-x} - b)^2 \leq \gamma$$

we obtain

$$-\log(b + \sqrt{\gamma}) \leq x \leq -\log(b - \sqrt{\gamma}). \tag{38}$$

If $b - \sqrt{\gamma} \leq 0$, the right-hand-side of (38) is considered to be $+\infty$. Taking l as the finite maximum of the left-hand-side values in (37) and (38), and u as the finite minimum of the right-hand-side values,

$$l \leq x \leq u$$

yields a lower and an upper bound for any x fulfilling (36). Performing this procedure for all univariate functions appearing in (25) leads to the constraints

$$\begin{aligned} &l_i \leq x_i \leq u_i, \ i = 1, \ldots, n-1, \\ &l_{ij} \leq x_i - x_j \leq u_{ij}, \ i = 1, \ldots, n-2, \ j = i+1, \ldots, n-1. \end{aligned} \tag{39}$$

Clearly, if (35) holds for some $(x_1, \ldots, x_{n-1})$, then also (39) holds. It is easy to see that the smaller γ is, the tighter the bounds of (39) are. Consequently, it is recommended to choose γ as the value of F at a local minimizer point.

Let $C \subset \mathbb{R}^{n-1}$ denote the convex polyhedral set of the points feasible to (39). If one can assure that F is convex over C, then adding the constraints

$$\begin{aligned} &l_i \leq x_i \leq u_i, \ i = 1, \ldots, n-1, \\ &l_{ij} \leq x_{ij} \leq u_{ij}, \ i = 1, \ldots, n-2, \ j = i+1, \ldots, n-1 \end{aligned} \tag{40}$$

to (10), we obtain a linearly constrained convex optimization problem with strictly convex objective function over the feasible set. The unique optimal solution of such problems can be easily found.

To investigate whether F is convex over a set given by (39), we again need to study the Hessian of F and the quadratic form associated with it. From (25), we obtain that

$$\begin{aligned} &(y_1, \ldots, y_{n-1}) \nabla^2 F(x_1, \ldots, x_{n-1}) (y_1, \ldots, y_{n-1})^T \\ &= \sum_{i=1}^{n-1} y_i \nabla^2 f_{a_{in}, a_{ni}}(x_i) y_i + \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} (y_i, y_j) \nabla^2 \bar{f}_{a_{ij}, a_{ji}}(x_i, x_j)(y_i, y_j)^T \end{aligned} \tag{41}$$

for any $(n-1)$-vector $(y_1, \ldots, y_{n-1})$, where

$$\bar{f}_{a_{ij}, a_{ji}}(x_i, x_j) = f_{a_{ij}, a_{ji}}(x_i - x_j). \tag{42}$$

Equation 41 means that the quadratic form with the Hessian of F can be obtained from the quadratic forms with the 1×1 and 2×2 Hessians of the functions appearing on the

Journal: 10898-JOGO Article No.: 9785 TYPESET DISK LE CP Disp.:2011/9/21 Pages: 18 Layout: Small

right-hand-side of (25). Based upon this property, we construct a quadratic matrix such that the quadratic form generated by this matrix lower-bounds the quadratic form generated by $\nabla^2 F$ at any point of $\mathbb{R}^{n-1}$.

Clearly, $\nabla^2 f_{a_{in},a_{ni}}(x_i) = f''_{a_{in},a_{ni}}(x_i)$. Let

$$\mu_i = \min\{f''_{a_{in},a_{ni}}(x_i) \mid l_i \le x_i \le u_i\}. \tag{43}$$

Then,

$$y_i \nabla^2 f_{a_{in},a_{ni}}(x_i) y_i \ge \mu_i y_i^2 \tag{44}$$

for all $x_i \in [l_i, u_i]$ and any real y_i. The computation of μ_i can be reduced to finding the positive roots of a quartic polynomial. We have,

$$f'''_{a,b}(x) = 2[4e^{2x} - ae^x - 4e^{-2x} + be^{-x}].$$

By substituting $v = e^x$, we get

$$f'''_{a,b}(x) = \frac{2}{v^2} p_{a,b}(v),$$

where

$$p_{a,b}(v) = 4v^4 - av^3 + bv - 4.$$

By determining the positive roots of the quartic polynomial $p_{a,b}(v)$, the real roots of $f'''_{a,b}(x)$ can also be obtained. Here again, finite methods can be used to solve the quartic polynomial equations [27]. Now, if the real roots of $f'''_{a_{in},a_{ni}}(x)$ are known, then taking the value $f''_{a_{in},a_{ni}}$ at the roots lying in $[l_i, u_i]$ and taking the values $f''_{a_{in},a_{ni}}(l_i)$ and $f''_{a_{in},a_{ni}}(u_i)$ also into consideration, μ_i of (43) can be obtained. According to Remark 1, if $\mu_i \ge 0$, then $f_{a_{in},a_{ni}}$ is strictly convex over $[l_i, u_i]$.

The computation of the Hessian of $\bar{f}_{a_{ij},a_{ji}}(x_i, x_j)$ of (42) is also simple. Let $\bar{x} = x_i - x_j$. Then

$$\begin{aligned} \frac{\partial^2 \bar{f}_{a_{ij},a_{ji}}(x_i,x_j)}{\partial x_i \partial x_i} &= \frac{\partial^2 \bar{f}_{a_{ij},a_{ji}}(x_i,x_j)}{\partial x_j \partial x_j} = f''_{a_{ij},a_{ji}}(\bar{x}), \\ \frac{\partial^2 \bar{f}_{a_{ij},a_{ji}}(x_i,x_j)}{\partial x_i \partial x_j} &= \frac{\partial^2 \bar{f}_{a_{ij},a_{ji}}(x_i,x_j)}{\partial x_j \partial x_i} = -f''_{a_{ij},,a_{ji}}(\bar{x}) \end{aligned}$$

and

$$\nabla^2 \bar{f}_{a_{ij},a_{ji}}(x_i, x_j) = \begin{bmatrix} f''_{a_{ij},a_{ji}}(\bar{x}) & -f''_{a_{ij},a_{ji}}(\bar{x}) \\ -f''_{a_{ij},a_{ji}}(\bar{x}) & f''_{a_{ij},a_{ji}}(\bar{x}) \end{bmatrix}.$$

Let

$$\mu_{ij} = \min\{f''_{a_{ij}}(\bar{x}) \mid l_{ij} \le \bar{x} \le u_{ij}\}. \tag{45}$$

Again, μ_{ij} can be obtained by solving a quartic polynomial equation, and taking into account the values of $f''_{a_{ij}}$ at l_{ij} and u_{ij}. Then

$$\begin{aligned} (y_i, y_j)\nabla^2 \bar{f}_{a_{ij},a_{ji}}(x_i, x_j)\begin{pmatrix} y_i \\ y_j \end{pmatrix} &= f''_{a_{ij},a_{ji}}(\bar{x})(y_i - y_j)^2 \\ \ge \mu_{ij}(y_i - y_j)^2 = (y_i, y_j)&\begin{bmatrix} \mu_{ij} & -\mu_{ij} \\ -\mu_{ij} & \mu_{ij} \end{bmatrix}\begin{pmatrix} y_i \\ y_j \end{pmatrix} \end{aligned} \tag{46}$$

Journal: 10898-JOGO Article No.: 9785 TYPESET DISK LE CP Disp.:2011/9/21 Pages: 18 Layout: Small

for all $l_{ij} \leq t_i - t_j \leq u_{ij}$ and for any reals y_i and y_j. According to Remark 1, if $\mu_{ij} \geq 0$, then $f_{a_{ij},a_{ji}}$ is strictly convex over $[l_{ij}, u_{ij}]$.

With this preparation, we can now state the following

Theorem 2 *If*

$$\begin{aligned} &\mu_i \geq 0,\ i = 1, \ldots, n-1, \\ &\mu_{ij} \geq 0,\ i = 1, \ldots, n-2,\ j = i+1, \ldots, n-1, \end{aligned} \tag{47}$$

then F is strictly convex over C. By adding (40) to (10) we obtain a convex programming problem whose objective function is strictly convex over the feasible set. Consequently, problems (1), (9) and (10) have a unique global optimal solution.

Proof We obtain from (47) that all $f_{a_{in},a_{ni}}$ and $f_{a_{ij},a_{ji}}$ are convex over $[l_i, u_i]$ and $[l_{ij}, u_{ij}]$, respectively. Then, according to Proposition 5, F is strictly convex over C, and hence it has a unique local (thus global) minimizer over C. The other statements follow immediately. □

Even if (47) does not hold, the convexity of F over C can sometimes be established as shown below.

Theorem 3 *Let H_{ij} be the $(n-1) \times (n-1)$ matrix having 1 in position (i, j) and zeros in all other positions. Let*

$$H = \sum_{i=1}^{n-1} \mu_i H_{ii} + \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \mu_{ij} (H_{ii} + H_{jj} - H_{ij} - H_{ji}). \tag{48}$$

If H is positive semidefinite, then F is strictly convex over C and by adding (40) to (10) we obtain a convex programming problem whose objective function is strictly convex over the feasible set, in addition, problems (1), (9) and (10) have a unique global optimal solution.

Proof It follows from (44) and (46) that for any $(x_1, \ldots, x_{n-1}) \in C$ and any $(n-1)$-vector $(y_1, \ldots, y_{n-1})$, we have

$$(y_1, \ldots, y_{n-1}) \nabla^2 F(x_1, \ldots, x_{n-1}) (y_1, \ldots, y_{n-1})^T \geq (y_1, \ldots, y_{n-1}) H (y_1, \ldots, y_{n-1})^T.$$

If H is positive semidefinite, then

$$(y_1, \ldots, y_{n-1}) H (y_1, \ldots, y_{n-1})^T \geq 0,$$

hence

$$(y_1, \ldots, y_{n-1}) \nabla^2 F(x_1, \ldots, x_{n-1}) (y_1, \ldots, y_{n-1})^T \geq 0.$$

This means that $\nabla^2 F(x_1, \ldots, x_{n-1})$ is positive semidefinite for any $(x_1, \ldots, x_{n-1}) \in C$, implying that F is convex over C. The remaining statements follow evidently from the previous results. □

Theorems 2 and 3 yield only sufficient conditions of the the convexity of F over C. It can however happen that F is convex over C, i.e., (34) is a convex programming problem, but we cannot prove it with the tools at our disposal. Another case is when F is indeed nonconvex over C. This is the territory of global optimization, and special methods need to be applied.

In this paper we primarily focus on the convexity properties of LSM for PC matrices without the reciprocity condition. The methodological and computational issues are beyond the scope of the paper but are planned to be presented in a follow-up publication.

Journal: 10898-JOGO Article No.: 9785 TYPESET DISK LE CP Disp.:2011/9/21 Pages: 18 Layout: Small

5 Numerical examples

We illustrate the convexity issues related to problem (10) and the approach of Theorem 3 with two numerical examples.

Example 1 Let

$$A = \begin{bmatrix} 1 & 2 & 4 \\ 0.3 & 1 & 1.5 \\ 0.2 & 0.6 & 1 \end{bmatrix}.$$

Although $(2, 0.3) \in U$ and $(1.5, 0.6) \in U$, we have $(4, 0.2) \notin U$ so condition (33) does not hold. We show however that F of (25) is strictly convex on $\mathbb{R}^2$. Let $C = \mathbb{R}^2$, i.e. let $l_1 = l_2 = l_{12} = -\infty$ and $u_1 = u_2 = u_{12} = \infty$, and determine the minimum of $f''_{4,0.2}$, $f''_{1.5,0.6}$ and $f''_{2,0.3}$ over the real line. These values are $\mu_1 = -1.61333$, $\mu_2 = 3.74174$ and $\mu_{1,2} = 3.18919$, respectively. According to (48), the matrix H is

$$H = \begin{bmatrix} -1.61333 + 3.18919 & -3.18919 \\ -3.18919 & 3.74174 + 3.18919 \end{bmatrix} = \begin{bmatrix} 1.57586 & -3.18919 \\ -3.18919 & 6.93093 \end{bmatrix},$$

and it is positive definite since the determinant of every leading principal submatrix of H is positive. Consequently, F is strictly convex on $\mathbb{R}^2$. □

Example 2 The second numerical example with the nonreciprocal matrix

$$A = \begin{bmatrix} 1.2 & 2 & 0.5 & 3 \\ 0.4 & 0.9 & 0.25 & 1.5 \\ 1.5 & 3 & 1 & 5 \\ 0.25 & 0.5 & 0.33 & 1.1 \end{bmatrix}$$

comes from [25].

Matrix A is not reciprocal. It is easy to see that $(a_{ij}, a_{ji}) \in U$ holds for all (i, j) except the single $(a_{3,4}, a_{4,3}) = (5, 0.33) \notin U$. Consequently, the convexity of F depends on the influence of the univariate function $f_{5,0.33}$.

By using local search method for minimizing F over $\mathbb{R}^3$, we found the local minimal solution

$$\left(x_1^*, x_2^*, x_3^*\right) = (1.122538, 0.459287, 1.592514) \tag{49}$$

with the objective function value $\gamma = 0.13008429$. Applying formulas (37) and (38) for $f_{5,0.33}$ and γ, the following lower and upper bounds can be obtained for x_3:

$$l_3 \le x_3 \le u_3,$$

where $l_3 = 1.534569529$ and $u_3 = 1.679089340$. By minimizing $f''_{5,0.33}$ over $[l_3, u_3]$ we get $\mu_3 = 39.74376$. Clearly, all the other values of μ_i and μ_{ij} are nonnegative since the corresponding univariate functions are convex. It follows directly from Proposition 5 that F is strictly convex over $C = \{(x_1, x_2, x_3) \mid l_3 \le x_3 \le u_3\}$. Since the local optimal solution (x_1^*, x_2^*, x_3^*) lies in C, it is the unique global optimal solution. According to $w_i = e^{x_i}$, we obtain that

$$(w_1^*, w_2^*, w_3^*, w_4^*) = (3.072644, 1.582944, 4.916091, 1)$$

is the unique global optimal solution of problem (9).

Journal: 10898-JOGO Article No.: 9785 □ TYPESET ☑ DISK □ LE ☑ CP Disp.:2011/9/21 Pages: 18 Layout: Small

6 Conclusions and final remarks

In [17], mathematical basis have been provided for using a small rather than bigger scale to express pairwise preferences. The results have been based on a convexity property of PC matrices recently found in [16]. The use of distance minimization approach in [17] has been also explored for solving the nonreciprocal case.

In this study, the Least Squares Method (LSM) was proposed for solving the relaxation of the reciprocity condition for pairwise comparisons. It has been demonstrated that, after suitable nonlinear coordinate transformations, some special cases of LSM are easy to solve convex optimization problems. It has been also shown that the global optimal solution of the original problem can be found by using local search methods under some other, less restrictive conditions when the convexity of a modified problem can be assured.

We did not detail methodological and computational issues in this paper. However, without assuring some convexity condition, the LSM may lead to a difficult global optimization problem with possible multiple local and even global minimal solutions. An adaptation of the branch-and-bound method in [16] for the nonreciprocal case is a possible approach. The methodological and computational details are planned to be presented in a follow-up publication.

Acknowledgments This research has been supported in part by NSERC grant in Canada and by OTKA grants K 60480, K 77420 in Hungary.

References

1. Barvinok, A.: A course in convexity, Graduate Studies in Mathematics, vol. 54. American Mathematical Society, Providence (2002)
2. Blankmeyer, E.: Approaches to consistency adjustments. J. Optim. Theory Appl. **54**, 479–488 (1987)
3. Bozóki, S.: Solution of the least squares method problem of pairwise comparisons matrices. Central Eur. J. Oper. Res. **16**, 345–358 (2008)
4. Carrizosa, E., Messine, F.: An exact global optimization method for deriving weights from pairwise comparison matrices. J. Global Optim. **38**, 237–247 (2007)
5. Choo, E.U., Wedley, W.C.: A common framework for deriving preference values from pairwise comparison matrices. Comput. Oper. Res. **31**, 893–908 (2004)
6. Chu, A.T.W., Kalaba, R.E., Spingarn, K.: A comparison of two methods for determining the weight belonging to fuzzy sets. J. Optim. Theory Appl. **4**, 531–538 (1979)
7. Condorcet. M.: Essai sur l'Application de l'Analyse à la Probabilité des Décisions Rendues à la Pluralité des Voix, Paris (1785)
8. Crawford, G., Williams, C.: A note on the analysis of subjective judgment matrices. J. Math. Psychol. **29**, 387–405 (1985)
9. De Jong, P.: A statistical approach to Saaty's scaling method for priorities. J. Math. Psychol. **28**, 467–478 (1984)
10. Diaz-Balteiro, L., González-Pachón, J., Romero, C.: Forest management with multiple criteria and multiple stakeholders: an application to two public forests in Spain. Scand. J. For. Res. **24**(1), 87–93 (2009)
11. Dong, Y., Li, H., Xu, Y.: On reciprocity indexes in the aggregation of fuzzy preference relations using the OWA operator. Fuzzy Sets Syst. **159**, 185–192 (2008)
12. Dopazo, E., González-Pachón, J.: Consistency-driven approximation of a pairwise comparison matrix. Kybernetika **39**(5), 561–568 (2003)
13. Farkas, A., Lancaster, P., Rózsa, P.: Consistency adjustment for pairwise comparison matrices. Numer. Linear Algebra Appl. **10**, 689–700 (2003)
14. Fechner, G.T.: Elemente der Psychophysik. Breitkopf & Härtel, Leipzig (1860)
15. Figiel, T.: On the moduli of convexity and smoothness. Studia Math. **56**, 121–155 (1976)
16. Fülöp, J.: A method for approximating pairwise comparison matrices by consistent matrices. J. Global Optim. **42**, 423–442 (2008)

17. Fülöp, J., Koczkodaj, W.W., Szarek, S.J.: A different perspective on a scale for pairwise comparisons. Transactions on Computational Collective Intelligence, Lecture Notes in Computer Science, vol. 6220, pp. 71–84 (2010)
18. Golany, B., Kress, M.: A multicriteria evaluation method for obtaining weights from ratio-scale matrices. Eur. J. Oper. Res. **69**, 210–220 (1993)
19. González-Pachón, J., Rodríguez-Galiano, M.I., Romero, C.: Transitive approximation to pairwise comparison matrices by using interval goal programming. J. Oper. Res. Soc. **54**, 532–538 (2003)
20. González-Pachón, J., Romero, C.: A method for dealing with inconsistencies in pairwise comparisons. Eur. J. Oper. Res. **158**(2), 351–361 (2004)
21. González-Pachón, J., Romero, C.: Inferring consensus weights from pairwise comparison matrices without suitable properties. Ann. Oper. Res. **154**, 123–132 (2007)
22. Hovanov, N.V., Kolari, J.W., Sokolov, M.V.: Deriving weights from general pairwise comparison matrices. Math. Soc. Sci. **55**, 205–220 (2008)
23. Jensen, R.E.: Comparison of eigenvector, least squares, chi squares and logarithmic least squares methods of scaling a reciprocal matrix. Working paper 153, Trinity University (1983)
24. Jensen, R.E.: Alternative scaling method for priorities in hierarchical structures. J. Math. Psychol. **28**, 317–332 (1984)
25. Koczkodaj, W.W., Orlowski, M.: Computing a consistent approximation to a generalized pairwise comparisons matrix. Comput. Math. Appl. **37**(2), 79–85 (1999)
26. Krantz, S.G., Parks, S.G.: A Primer of Real Analytic Functions. Birkhäuser Verlag, Basel (1992)
27. Kurosh, A.G.: Higher Algebra. Mir Publishers, Moscow (1972)
28. Limayem, F., Yannou, B.: Generalization of the RCGM and LSLR pairwise comparison methods. Comput. Math. Appl. **48**(3–4), 539–548 (2004)
29. Llull, R.: Artifitium electionis personarum (before 1283)
30. Mikhailov, L.: A fuzzy programming method for deriving priorities in the analytic hiarerchy process. J. Oper. Res. Soc. **51**, 341–349 (2000)
31. Rockafellar, R.T.: Convex Analysis. Princeton University Press, Princeton, NJ (1970)
32. Saaty, T.L.: A scaling method for priorities in hierarchical structures. J. Math. Psychol. **15**, 234–281 (1977)
33. Sekitani, K., Yamaki, N.: A logical interpretation for the eigenvalue method in AHP. J. Oper. Res. Soc. Jpn. **42**, 219–232 (1999)
34. Triantaphyllou, E., Pardalos, P.M., Mann, S.H.: The problem of determining membership values in fuzzy sets in real world situations. In: Brown, D.E., White, C.C. III. (eds.) Operations Research and Artificial Intelligence: The integration of problem solving strategies, pp. 197–214. Kluwer Academic Publishers, Dordrecht (1990)
35. Triantaphyllou, E., Lootsma, F., Pardalos, P.M., Mann, S.H.: On the evaluation and application of different scales for quantifying pairwise comparisons in fuzzy sets. J. Multi-Criteria Decis. Anal. **3**, 133–155 (1994)

Journal: **10898-JOGO** Article No.: **9785** ☐ TYPESET ☑ DISK ☐ LE ☑ CP Disp.:**2011/9/21** Pages: **18** Layout: **Small**