View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Elsevier - Publisher Connector

# A Vector Majorization Method for Solving a Nonlinear Programming Problem

C.-T. Pan
*Department of Mathematical Sciences*
*Northern Illinois University*
*De Kalb, Illinois 60115*

Submitted by Michael Neumann

## ABSTRACT

In the rounding error analysis of a so-called hyperbolic rotation algorithm applied to downdating the Cholesky factor, a sharp upper bound is needed for a product

$$\prod_{i=1}^{n}\left[1+\frac{a_i}{\sqrt{1-a_1^2-\cdots-a_{i-1}^2}}\right],$$

where $\mathbf{a}=(a_1,\ldots,a_n)^T\in\mathbf{R}^n$, $\|\mathbf{a}\|\leqslant\theta<1$, and $\theta$ is fixed. It is shown that

$$\operatorname*{Max}_{\|a\|\leqslant\theta<1}\prod_{i=1}^{n}\left[1+\frac{a_i}{\sqrt{1-a_1^2-\cdots-a_{i-1}^2}}\right]=\left[1+\sqrt{1-(1-\theta^2)^{1/n}}\right]^n.$$

In the original proof by Pan and Sigmon, the principle of Lagrange multipliers was used. A second proof by using vector majorization is presented here, which provides insights into this rather strange-looking bound. Moreover, the new proof can be readily generalized to solve a large class of nonlinear programming problems.

## 0. INTRODUCTION

When we perform the rounding error analysis of a so-called hyperbolic rotation algorithm applied to downdating the Cholesky factor [1], the conclusion of the weak stability of the algorithm is based on the solution of the



© Elsevier Science Publishing Co., Inc., 1989
655 Avenue of the Americas, New York, NY 10010 0024-3795/89/$3.50

following problem:

$$\text{maximize} \quad \left[ \prod_{i=1}^{n} \left( 1 + \frac{a_i}{\sqrt{1 - a_1^2 - \cdots - a_{i-1}^2}} \right) \right] \tag{0.1}$$

$$\text{subject to} \quad \|\mathbf{a}\| \leqslant \theta < 1,$$

where $\mathbf{a} = (a_1, \ldots, a_n)^T \in \mathbf{R}^n$ and $\theta$ is fixed. Notice that in (0.1) the first factor is $(1 + a_1)$. For the detail of the connection between this problem and the error analysis of the hyperbolic rotation algorithm, the reader is referred to [1] and [4]. What we are interested in in this paper is the problem (0.1) itself.

The solution of (0.1) was based on an intuitive guess. We assume that the product in (0.1) reaches its maximum value when all the terms $a_i / \sqrt{1 - a_1^2 - \cdots - a_{i-1}^2}$ are equal and $\|\mathbf{a}\| = \theta$, i.e., that the optimal vector of the problem (0.1) is the vector $\mathbf{a}$ satisfying

$$a_1 = \frac{a_2}{\sqrt{1 - a_1^2}} = \cdots = \frac{a_n}{\sqrt{1 - a_1^2 - \cdots - a_{n-1}^2}} \tag{0.2}$$

and

$$\|\mathbf{a}\| = \theta. \tag{0.3}$$

The interesting consequence of (0.2) is that the squares of the components of the optimal vector $\mathbf{a}$ form a geometric sequence, i.e., that

$$a_i^2 = a_1^2 \left(1 - a_1^2\right)^{i-1}.$$

According to (0.3),

$$\|\mathbf{a}\|^2 = \sum_{i=1}^{n} a_i^2 = \frac{a_1^2 - a_1^2\left(1 - a_1^2\right)^n}{1 - \left(1 - a_1^2\right)} = 1 - \left(1 - a_1^2\right)^n = \theta^2$$

and thus

$$a_1 = \sqrt{1 - (1 - \theta^2)^{1/n}}.$$

Therefore, if the assumption is true, the optimal value of (0.1) should be

$$(1+a_1)^n = \left(1+\sqrt{1-(1-\theta^2)^{1/n}}\right)^n.$$

To prove the intuitive guess above, the vector majorization method is naturally applied in connection with the assumption (0.2). The rest of this paper is organized as follows. In Section 1 we briefly review the concept of vector majorization and a theorem of Schur's for our needs. In Section 2 we prove the main results. Finally, in Section 3 a generalization of the main results to a class of nonlinear programming problems with symmetric functions in both objectives and constraints is discussed.

## 1. VECTOR MAJORIZATION PRELIMINARIES

The vector majorization method was formally introduced in the classical work *Inequalities* by Hardy, Littlewood, and Polya (1934) [2], but the method was known to Schur as early as 1923 under the concept of doubly stochastic matrices [3].

NOTATION. Let $\mathbf{R}^n$ be the Euclidean space with dimension $n$, and $\|\cdot\|$ be the Euclidean vector norm defined by $\|\mathbf{x}\| = \sqrt{\mathbf{x}^T\mathbf{x}}$, $\mathbf{x} \in \mathbf{R}^n$ ($T$ denotes the transpose). The set of all matrices with $m$ rows and $n$ columns is denoted by $\mathbf{R}^{m\times n}$. If $A \in \mathbf{R}^{m\times n}$, then $A_{ij}$ denotes the $(i, j)$th entry of $A$. The vector space of interest in this paper is $\mathbf{R}^n$, $n \geqslant 2$.

We will follow Schur in using the doubly stochastic matrix to define vector majorization, a partial ordering of $\mathbf{R}^n$.

DEFINITION 1.1. A matrix $P \in \mathbf{R}^{n\times n}$ with all its entries nonnegative is *doubly stochastic* if

$$\sum_{i=1}^{n} P_{ij} = 1 \qquad \text{for} \quad j = 1,\ldots,n$$

and

$$\sum_{j=1}^{n} P_{ij} = 1 \qquad \text{for} \quad i = 1,\ldots,n.$$

DEFINITION 1.2. For $\mathbf{x},\mathbf{y} \in \mathbf{R}^n$, $\mathbf{x}$ *is majorized by* $\mathbf{y}$, denoted by $\mathbf{x} \underset{m}{\leqslant} \mathbf{y}$, if there exists a doubly stochastic matrix $P$ such that

$$\mathbf{x} = P\mathbf{y}. \tag{1.1}$$

EXAMPLE 1.1. In $\mathbf{R}^3$,

$$\left(\tfrac{1}{3},\tfrac{1}{3},\tfrac{1}{3}\right)^T \underset{m}{\leqslant} (1,0,0)^T,$$

since

$$\begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

Generally, in $\mathbf{R}^n$ one has

$$(\bar{x},\dots,\bar{x})^T \underset{m}{\leqslant} (x_1,\dots,x_n)^T \tag{1.2}$$

with

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n},$$

since

$$\begin{bmatrix} \bar{x} \\ \vdots \\ \bar{x} \end{bmatrix} = \begin{bmatrix} \frac{1}{n} & \cdots & \frac{1}{n} \\ \vdots & & \vdots \\ \frac{1}{n} & \cdots & \frac{1}{n} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}. \tag{1.3}$$

EXAMPLE 1.2. All the permutation matrices (square matrices which in each row and in each column has some one entry 1, all other entries zero) are doubly stochastic. Therefore, $(x_1, x_2)^T \underset{m}{\leqslant} (x_2, x_1)^T$ and $(x_2, x_1)^T \underset{m}{\leqslant} (x_1, x_2)^T$ are both true.

Notice that in $\mathbf{R}$, $a \underset{m}{\leqslant} b$ if and only if $a = b$; therefore, as mentioned before, in this paper we are concerned with $\mathbf{R}^n$, $n \geqslant 2$.

DEFINITION 1.3. Let $I$ be an open interval of $\mathbf{R}$, and $I^n$ be the corresponding open $n$-dimensional box in $\mathbf{R}^n$. A function $\phi: I^n \to \mathbf{R}$ is called *Schur-convex* [*Schur-concave*] on $I^n$ if

$$\phi(\mathbf{x}) \leqslant \phi(\mathbf{y}) \qquad [\phi(\mathbf{x}) \geqslant \phi(\mathbf{y}), \text{ correspondingly}]$$

whenever $\mathbf{x} \underset{m}{\leqslant} \mathbf{y}$ for any $\mathbf{x},\mathbf{y} \in I^n$.

Observe that a Schur-convex (Schur-concave) function is always a symmetric function [i.e., $\phi(\mathbf{x}) = \phi(A\mathbf{x})$ for any permutation matrix $A$] on certain symmetric sets, since the permutation matrices are all doubly stochastic (Example 1.2).

Finally, we introduce the following theorem for later use.

THEOREM 1.1 [Schur (1923), Ostrowski (1952)]. *Let $I \subseteq \mathbf{R}$ be an open interval and $\phi: I^n \to \mathbf{R}$ be a continuously differentiable function. Then $\phi$ is Schur-convex* [*Schur-concave*] *if and only if*

(1) *$\phi$ is symmetric on $I^n$,*

(2) $(x_1 - x_2)[\partial\phi/\partial x_1 - \partial\phi/\partial x_2] \geqslant 0$ [ $\leqslant 0$, *correspondingly*] *for all* $\mathbf{x} = (x_1, \dots, x_n)^T \in I^n$.

The proof can be found, for example, in [3].

## 2. MAIN RESULTS

Precisely, we will prove the following main result.

THEOREM 2.1. *Fix $\theta$ $(0 < \theta < 1)$ and let $\mathbf{a} = (a_1, \dots, a_n)^T \in \mathbf{R}^n$ with $\|\mathbf{a}\| \leqslant \theta$. Then*

$$\operatorname*{Max}_{\|\mathbf{a}\| \leqslant \theta} \left[ \prod_{i=1}^{n} \left( 1 + \frac{a_i}{\sqrt{1 - a_1^2 - \cdots - a_{i-1}^2}} \right) \right] = \left[ 1 + \sqrt{1 - (1 - \theta^2)^{1/n}} \right]^n, \quad (2.1)$$

*where the maximum value is reached with the* optimal vector $\mathbf{a}$ *such that*

$$a_1 = \frac{a_i}{\sqrt{1 - a_1^2 - \cdots - a_{i-1}^2}} \quad \text{for } i = 2, \dots, n \quad \text{and} \quad \|\mathbf{a}\| = \theta.$$

In order to prove Theorem 2.1, we start from several lemmas.

LEMMA 2.1. *The function*

$$F(\mathbf{t}) = \sum_{i=1}^{n} t_i^2 - \sum_{i \neq j} t_i^2 t_j^2 + \cdots + (-1)^{n-1} t_1^2 t_2^2 \cdots t_n^2$$

*is a Schur-convex function on* $(-1,1)^n$.

*Proof.* According to Theorem 1.1, we only need to check that

$$(t_1 - t_2)\left[\frac{\partial F}{\partial t_1} - \frac{\partial F}{\partial t_2}\right] \geqslant 0, \tag{2.2}$$

where $\mathbf{t} \in I^n$ and $I = (-1,1)$, since the symmetry of $F(\mathbf{t})$ on $(-1,1)^n$ is obvious.

First we claim that

$$\frac{\partial F}{\partial t_1} - \frac{\partial F}{\partial t_2} = 2(t_1 - t_2)(1 + t_1 t_2)(1 - t_3^2) \cdots (1 - t_n^2) \tag{2.3}$$

for $n \geqslant 2$. For the case $n = 2$, the formula (2.3) is easy to verify. To proceed with induction on $n$, one needs an equality

$$F(t_1, \ldots, t_n) = F(t_1, \ldots, t_{n-1}) + t_n^2(1 - t_1^2) \cdots (1 - t_{n-1}^2), \tag{2.4}$$

which is obvious. Thus,

$$\begin{aligned}\frac{\partial F}{\partial t_1} - \frac{\partial F}{\partial t_2} = {} & \frac{\partial F(t_1, \ldots, t_{n-1})}{\partial t_1} - \frac{\partial F(t_1, \ldots, t_{n-1})}{\partial t_2} \\ & + 2t_2(1 - t_1^2)(1 - t_3^2) \cdots (1 - t_{n-1}^2)t_n^2 \\ & - 2t_1(1 - t_2^2) \cdots (1 - t_{n-1}^2)t_n^2 \\ = {} & 2(t_1 - t_2)(1 + t_1 t_2)(1 - t_3^2) \cdots (1 - t_n^2),\end{aligned}$$

and (2.3) is proved. Therefore, (2.2) is true and the lemma is proved. ■

LEMMA 2.2. *The function*

$$Q(\mathbf{t}) = \prod_{i=1}^{n} (1+t_i)$$

*is Schur-concave on* $(-1,\infty)^n$.

*Proof*. Again $Q(\mathbf{t})$ is symmetric about $t_1,\dots,t_n$; therefore, one only needs to check condition (2) in Theorem 1.1. By computation one knows that

$$\frac{\partial Q}{\partial t_1} - \frac{\partial Q}{\partial t_2} = (1+t_3)\cdots(1+t_n)(t_2-t_1);$$

thus,

$$(t_1-t_2)\left[\frac{\partial Q}{\partial t_1} - \frac{\partial Q}{\partial t_2}\right] = -(t_1-t_2)^2(1+t_3)\cdots(1+t_n) \leqslant 0$$

for all $t_i \in (-1,\infty)$. ■

Before introducing the next lemma, let us introduce some new notation, which will be frequently used in the following text.

NOTATION. For a vector variable function $Q(\mathbf{t}) = Q(t_1, t_2,\dots,t_n)$, we define a scalar function $\bar{Q}(t)$ of one scalar variable such that $\bar{Q}(t) = Q(t,t,\dots,t)$.

LEMMA 2.3. *For the nonlinear programming problem*

$$\text{maximize} \quad \prod_{i=1}^{n} (1+t_i),$$

*subject to*

$$0 \leqslant F(\mathbf{t}) = \sum_i t_i^2 - \sum_{i\neq j} t_i^2 t_j^2 + \cdots + (-1)^{n-1} t_1^2 t_2^2 \cdots t_n^2 \leqslant \theta^2 < 1 \quad (2.5)$$

*and*

$$0 \leqslant t_i < 1,$$

*the maximum* $(1+\tilde{t})^n$, *with* $\tilde{t} = \sqrt{1-(1-\theta^2)^{1/n}}$, *is attained when* $\mathbf{t} = \tilde{\mathbf{t}} = (\tilde{t},\dots,\tilde{t})^T$.

*Proof.* Let $\mathbf{t} = (t_1,\dots,t_n) \in \mathbf{R}^n$ be a vector satisfying (2.5) and let

$$\bar{\mathbf{t}} = (\bar{t},\dots,\bar{t})^T \quad \text{with} \quad \bar{t} = (t_1 + \cdots + t_n)/n.$$

By Lemma 2.2, $\prod_{i=1}^n (1+t_i)$ is Schur-concave on $(-1,+\infty)^n$ and therefore

$$\prod_{i=1}^{n}(1+t_i) \leqslant (1+\bar{t})^n. \tag{2.6}$$

Notice that vector $\bar{\mathbf{t}}$ need not satisfy (2.5) even though $0 \leqslant \bar{t} < 1$ is guaranteed by $\bar{t} = (t_1 + \cdots + t_n)/n$ and $0 \leqslant t_i < 1$. But Lemma 2.1 implies that

$$F(\bar{\mathbf{t}}) \leqslant F(\mathbf{t}).$$

On the other hand

$$\begin{aligned} \bar{F}(\bar{t}) &= \binom{n}{1}\bar{t}^2 - \binom{n}{2}\bar{t}^4 + \cdots + (-1)^{n-1}\binom{n}{n}\bar{t}^{2n} \\ &= 1-(1-\bar{t}^2)^n. \end{aligned}$$

We now define the function

$$\bar{F}(x) = 1-(1-x^2)^n \tag{2.7}$$

and let $\tilde{t}$ be such that $\bar{F}(\tilde{t}) = \theta^2$, i.e., $\tilde{t}$ is the solution of $\bar{F}(x) = \theta^2$; actually

$$\tilde{t} = \sqrt{1-(1-\theta^2)^{1/n}}, \tag{2.8}$$

if we restrict $0 \leqslant \tilde{t} < 1$.

Summarizing from the beginning of this proof, we have

$$\bar{F}(\bar{t}) = F(\bar{\mathbf{t}}) \leqslant F(\mathbf{t}) \leqslant \theta^2 = \bar{F}(\tilde{t}) \tag{2.9}$$

for any $\mathbf{t}$ satisfying (2.5).

By directly checking the derivative, one knows that $\bar{F}(x)$ is a monotonic increasing function on $[0,1)$, and so is the inverse of $\bar{F}(x)$. Thus, from (2.9) one has

$$\bar{t} \leqslant \tilde{t}.$$

Finally we have proved

$$\prod_{i=1}^{n} (1+t_i) \leqslant (1+\bar{t})^n \leqslant (1+\tilde{t})^n \tag{2.10}$$

with an arbitrary $\mathbf{t}$ satisfying (2.5). From (2.7) and (2.8) one knows that $\mathbf{t} = \tilde{\mathbf{t}} = (\tilde{t}, \dots, \tilde{t})^T$ satisfies (2.5) also; therefore the lemma is proved. ■

We are now ready to prove Theorem 2.1.

*Proof of Theorem 2.1.* In Lemma 2.3 let

$$t_i = \frac{a_i}{\sqrt{1 - a_1^2 - \cdots - a_{i-1}^2}}, \qquad i = 2, \dots, n, \tag{2.11}$$

and

$$t_1 = a_1.$$

Notice that as the maximum value in (2.1) is of interest, one only needs to consider the case when all $a_i \geqslant 0$.

It is easy to check that the conditions $a_i \geqslant 0$ and $\|\mathbf{a}\| < 1$ are equivalent to $0 \leqslant t_i < 1$ for $i = 1, \dots, n$, if one notices that

$$\begin{aligned} a_i^2 &\leqslant a_i^2 + \left(a_1^2 + \cdots + a_{i-1}^2\right)\left(1 - a_1^2 - \cdots - a_i^2\right) \\ &= \left(a_1^2 + \cdots + a_i^2\right)\left(1 - a_1^2 - \cdots - a_{i-1}^2\right). \end{aligned}$$

One also needs to check that $0 \leqslant \|\mathbf{a}\| \leqslant \theta$ is equivalent to $0 \leqslant F(\mathbf{t}) \leqslant \theta^2$. By using (2.11), it comes from direct computing that

$$a_1^2 + \cdots + a_n^2 = t_1^2 + t_2^2(1 - t_1^2) + \cdots + t_n^2(1 - t_1^2)(1 - t_2^2) \cdots (1 - t_{n-1}^2)$$
$$= F(\mathbf{t}).$$

Therefore, the two nonlinear programming problems (0.1) and (2.5) are equivalent and Theorem 2.1 is proved. ■

## 3. GENERALIZATION

It is seen in the literature that many inequalities can be proved by using the vector majorization method [3]; however, its use in solving certain optimization problems appears to be new (at least to the author). One of the possible generalizations of Lemma 2.3 is the following.

THEOREM 3.1. *Let $Q$ be a Schur-concave function on $(-q, K)^n$ with $q > 0$, $F$ be a Schur-convex function on $(-f, G)^n$ with $f > 0$, and both $\bar{Q}$ and $\bar{F}$ have the same direction of monotonicity on $[0, +\infty)$ with $\bar{F}$ being strictly montonic. Then the nonlinear programming problem*

$$\begin{aligned} &\textit{maximize} \quad Q(x_1, \dots, x_n), \\ &\textit{subject to} \quad \alpha \leqslant F(x_1, \dots, x_n) \leqslant \beta, \quad 0 \leqslant x_i < +\infty, \end{aligned}$$

*has its optimal value $\bar{Q}(\tilde{x})$ such that $\bar{F}(\tilde{x}) = \beta$, provided that $\bar{F}(x) = \beta$ has a solution $\tilde{x}$ in $[0, K)$.*

The proof is exactly the same as that of Lemma 2.3.

Table 1 shows the possible alternatives to Theorem 3.1, where the other conditions are modified correspondingly:

$$\begin{aligned} &\text{optimize} \quad Q(x) \\ &\text{subject to} \quad \alpha \leqslant F(x) \leqslant \beta, \quad 0 \leqslant x_i < +\infty. \end{aligned}$$

*The author is grateful to the referee, whose careful reading and valuable suggestion definitely improved this presentation.*

TABLE 1

| $Q$ | $F$ | Monotonicity $\bar{Q}$ | $\bar{F}$ | Optimal value $\bar{Q}(\tilde{x})$ | $\tilde{x}$ satisfies |
|---|---|---|---|---|---|
| Schur-concave | Schur-convex | $\searrow$ $\nearrow$ | $\searrow$ $\nearrow$ | Max Q | $\bar{F}(\tilde{x}) = \beta$ |
| | Schur-concave | $\searrow$ $\nearrow$ | $\nearrow$ $\searrow$ | Max Q | $\bar{F}(\tilde{x}) = \alpha$ |
| Schur-convex | Schur-concave | $\searrow$ $\nearrow$ | $\searrow$ $\nearrow$ | Min Q | $\bar{F}(\tilde{x}) = \alpha$ |
| | Schur-convex | $\searrow$ $\nearrow$ | $\nearrow$ $\searrow$ | Min Q | $\bar{F}(\tilde{x}) = \beta$ |

## REFERENCES


1 S. T. Alexander, C.-T. Pan and R. J. Plemmons, Analysis of recursive least squares hyperbolic rotation algorithm for signal processing, *Linear Algebra Appl.* 98:3–40 (1988).

2 G. H. Hardy, J. E. Littlewood, and G. Polya, *Inequalities*, 2nd ed. Cambridge U.P., New York, 1952.

3 A. W. Marshal and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*, Academic, London, 1979.

4 C.-T. Pan and K. Sigmon, A sharp bound for products of hyperbolic plane rotations, *SIAM J. Matrix Anal.* 9:4 (1988).


*Received 8 December 1987; final manuscript accepted 20 September 1988*