Mathematical
Inequalities
& Applications
Volume 1, Number 3 (1998), 319–330

OPTIMIZATION OF SCHUR–CONVEX FUNCTIONS

XIN-MIN ZHANG

(*communicated by I. Olkin*)

Abstract. In this paper, we show that Schur-convex functions share some important properties with the ordinary convex functions. We apply special properties of Schur-convex functions to establish some inequalities for the generalized power means that include many well-known classical analytic inequalities as special cases.

1. Introduction

The Schur-convex function was introduced by I. Schur in 1923 [7] and has many important applications in analytic inequalities. Hardy, Littlewood, and Pólya were also interested in some inequalities that are related to Schur-convex functions [2]. The following definitions and examples can be found in many references such as [3,7,9,10].

DEFINITION 1.1. An $n \times n$ matrix $S = [s_{ij}]$ is called a *doubly stochastic matrix* if $s_{ij} \geqslant 0$ for $1 \leqslant i,j \leqslant n$, and

$$\sum_{j=1}^{n} s_{ij} = 1, \quad i = 1, 2, \cdots, n; \qquad \sum_{i=1}^{n} s_{ij} = 1, \quad j = 1, 2, \cdots, n.$$

EXAMPLES 1.2. (a) A permutation matrix is a doubly stochastic matrix. (A permutation matrix is a matrix obtained by permuting the rows of the identity matrix.)
(b) $S = [s_{ij}]$ with $s_{ij} = \frac{1}{n}$, $1 \leqslant i,j \leqslant n$ is a doubly stochastic matrix.

Let $I^n = I \times I \times \cdots \times I$ (n copies), where I is an interval of the real number line R. We are interested in the following special multivariable functions.

DEFINITION 1.3. $f : I^n \longrightarrow R$ $(n > 1)$ is called *Schur-convex* if for every doubly stochastic matrix S,

$$f(S\mathbf{x}) \leqslant f(\mathbf{x}) \tag{1}$$

for all $\mathbf{x} \in I^n$. It is called *strictly Schur-convex* if the inequality is strict; f is called *Schur-concave* (resp. *strictly Schur-concave*) if the inequality (1) is reversed.

Mathematics subject classification (1991): 26B25, 26D05, 26D10.

Key words and phrases: Doubly stochastic matrix, Schur-convex function, symmetric function, generalized power mean, rth-power convexity.

DEFINITION 1.4. $f : I^n \longrightarrow R$ $(n > 1)$ is called *symmetric* if for every permutation matrix P,

$$f(P\mathbf{x}) = f(\mathbf{x})$$

for all $\mathbf{x} \in I^n$.

Let us recall that a twice differentiable function $f : R^n \longrightarrow R$ is convex (resp., strictly convex) if its Hessian matrix, Hess(f), is nonnegative (resp., positive) definite on R^n ([10, p. 103]). It is well-known that a convex function is not necessarily a Schur-convex function, and a Schur-convex function does not have to be convex in the ordinary sense either. However, under the assumption of ordinary convexity, f is Schur-convex if and only if it is symmetric.([10, p. 258]). In this paper, we like to show that Schur-convex functions do share some very important properties with convex functions, especially, some basic properties of convex functions about extreme values that have been used widely in theoretical and applied mathematics. In practice, due to the symmetric nature of Schur-convex functions, it is easier to verify the Schur-convexity than to check for ordinary convexity. As an application, in section 3 we provide a simple sufficient condition for a function of two variables that guarantees an unique local minimum a global one. In sections 4 and 5, we shall construct some Schur-convex and Schur-concave functions in terms of a differentiable function of single variable, and establish some inequalities for these functions. The new inequalities generalize several well-known classical analytic inequalities such as the Jensen's inequality and the arithmetic-geometric mean inequality.

2. Some Properties of Schur-Convex Functions

In this section, we discuss some special properties of Schur-convex functions that are essential in extremum problems. For a historical development of this kind of functions and the fruitful applications to statistics, economics and other applied fields, refer to the popular book by Marshall and Olkin [7].

LEMMA 2.1. *Every Schur-convex function is a symmetric function.*

Proof. If P is a permutation matrix, so is its inverse P^{-1}. Hence if f is Schur-convex, then

$$f(\mathbf{x}) = f(P^{-1}(P\mathbf{x})) \leqslant f(P\mathbf{x}) \leqslant f(\mathbf{x}).$$

It shows that $f(P\mathbf{x}) = f(\mathbf{x})$ for every permutation matrix P. □

It is not hard to see that not every symmetric function can be a Schur-convex function [10, p. 258]. Similarly, every Schur-concave function is symmetric and the inverse is not true. However, we have the following so-called Schur's condition.

LEMMA 2.2. ([10, p. 259]) *Let* $f(\mathbf{x}) = f(x_1, x_2, \cdots, x_n)$ *be symmetric and have continuous partial derivatives on* I^n *where* I *is an open interval. Then* $f : I^n \longrightarrow R$

is Schur-convex if and only if

$$(x_i - x_j)\left(\frac{\partial f}{\partial x_i} - \frac{\partial f}{\partial x_j}\right) \geqslant 0 \tag{2}$$

on I^n. It is strictly Schur-convex if (2) is a strict inequality for $x_i \neq x_j$, $1 \leqslant i,\ j \leqslant n$.

Since $f(\mathbf{x})$ is symmetric, Schur's condition can be reduced as ([7, p. 57])

$$(x_1 - x_2)\left(\frac{\partial f}{\partial x_1} - \frac{\partial f}{\partial x_2}\right) \geqslant 0, \tag{3}$$

and f is strictly Schur-convex if (3) is a strict inequality for $x_1 \neq x_2$. The Schur's condition that guarantees a symmetric function being Schur-concave is the same as (2) or (3) except the direction of the inequality.

In Schur's condition, the domain of $f(\mathbf{x})$ does not have to be a Cartesian product I^n. Lemma 2.2 remains true if we replace I^n by a set $A \subseteq R^n$ with the following properties ([7, p. 57]):

(i) A is convex and has a nonempty interior;

(ii) A is symmetric in the sense that $\mathbf{x} \in A$ implies $P\mathbf{x} \in A$ for any $n \times n$-permutation matrix P.

In the rest of this paper, we shall be concerned mainly with Schur-convex functions and related inequalities. Most of the results have their counterparts for Schur-concave functions. In order to simplify the notations and the statements, let us set

$$I = (-a,\ a); \qquad H_k = \{\mathbf{x} = (x_1, x_2, \cdots, x_n) \in R^n \mid \sum_{i=1}^{n} x_i = k\},$$

$$J_n = I^n \cap H_k; \qquad \Omega = (\sigma, \sigma, \cdots, \sigma) \quad \text{where} \quad \sigma = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{k}{n}.$$

The following property of Schur-convex function plays a key role in the proof of Theorem 2.4 and implies other important results.

LEMMA 2.3. *If $f : I^n \longrightarrow R$ is a Schur-convex function, then $f(\Omega)$ is a global minimum in J_n. If f is strictly Schur-convex on I^n, then $f(\Omega)$ is the unique global minimum in J_n.*

Proof. Since f is Schur-convex in I^n, therefore,

$$f(S\mathbf{x}) \leqslant f(\mathbf{x}),$$

for every doubly stochastic matrix S and every $\mathbf{x}$ in I^n. If we take $S = [s_{ij}]$ with $s_{ij} = 1/n$, $1 \leqslant i,j \leqslant n$, then $\Omega = S\mathbf{x}$, and

$$f(\Omega) \leqslant f(\mathbf{x})$$

for all $\mathbf{x} \in J_n$. If f is strictly Schur-convex, then $f(\Omega) < f(\mathbf{x})$ for all $\mathbf{x} \in J_n$, $\mathbf{x} \neq \Omega$. □

Recall that one of the most important properties of a strictly convex function is that it admits at most one global minimum ([10]). For a Schur-convex function, we are able to prove the following result.

THEOREM 2.4. *Let $f : R^n \longrightarrow R$ be a Schur-convex function and have only one critical point $\mathbf{a} = (a_1, a_2, \cdots, a_n)$. If $f(\mathbf{a})$ is a local minimum, then it must be the global minimum.*

Proof. First, we claim that $a_1 = a_2 = \cdots = a_n$, that is, the only critical point $\mathbf{a}$ must be in the subset D of R^n where $D = \{(x, x, \cdots, x) \mid x \in R\}$, because of the symmetric property of f and the uniqueness of the critical point. Then we may split the whole space R^n into a disjoint union of the hyperplanes $H_k = \{(x_1, x_2, \cdots, x_n) \mid \sum_{i=1}^n x_i = k\}$, $k \in R$. Lemma 2.3 tells us that f attains the minimum on each subset H_k at $(k/n, k/n, \cdots, k/n)$, a point in $D \cap H_k$, $k \in R$. Finally, we can show that there is no point in D other than $\mathbf{a}$ whose function value is smaller than $f(\mathbf{a})$, thereby $f(\mathbf{a})$ is the global minimum of f on R^n. If there exists a point $\mathbf{c} = (c, c, \cdots, c) \in D$ such that $f(\mathbf{a}) > f(\mathbf{c})$, then there must be a real number b between a and c, and a point $\mathbf{b} = (b, b, \cdots, b)$ in D such that the directional derivative $f_{\vec{u}}(\mathbf{b}) = 0$ where $\vec{u} = \vec{e}_1 + \vec{e}_2 + \cdots + \vec{e}_n$, and $\vec{e}_1, \vec{e}_2, \cdots, \vec{e}_n$ are the canonical base vectors of R^n. However, on the subset D, we have $f_{x_1} = f_{x_2} = \cdots = f_{x_n}$ by the symmetric property of f. Therefore $f_{\vec{u}}(\mathbf{b}) = \bigtriangledown f(\mathbf{b}) \cdot \vec{u}/\sqrt{n} = 0$ implies that $f_{x_1}(\mathbf{b}) = f_{x_2}(\mathbf{b}) = \cdots = f_{x_n}(\mathbf{b}) = 0$, and we get a second critical point $\mathbf{b}$. This is a contradiction to the uniqueness of $\mathbf{a}$. □

REMARKS. Optimization problems could be complicated when the objective function f involves n variables ($n \geqslant 3$). Not only locating critical points requires to solve a system of n equations, the second derivative test requires the calculations of a large number of determinants with different orders. For details, refer to [6, p. 228]. On the other hand, when the domain of definition of f is closed, checking function values on the boundary could be time and energy consuming. For example, even for $n = 3$ and f is defined on the cube $[0,1]^3$. Besides solving $\bigtriangledown f = 0$ in the interior of the cube, to make sure that f does not attain any extremum on the boundary, one has to check 6 faces, 12 edges, and 8 vertices of the cube. Bear these in mind, whenever one has noticed the objective function is symmetric, it is always worth-while to check the Schur's condition (2) or (3). If it is Schur-convex, then the optimization problem could be much simpler. In a lower dimensional case, say, $n = 2$, the two statements and the two examples in the following section along with Theorems 3.5 and 3.6 are accessible to undergraduates.

3. A Calculus Problem

First, let us consider the following simple statement in a first semester calculus:

STATEMENT 1.

Let $f(x)$ be a differentiable function defined on an open interval (a,b). If x_0 is the only critical point of f in that interval, and $f(x_0)$ is a local minimum, then $f(x_0)$ must be the global minimum of f on that interval.

Second, when we consider a similar question for a function of two variables, the answer is somewhat surprising. More specifically, we have the following:

STATEMENT 2.

Let $f(x,y)$ be a differentiable function defined on an open domain $D \subseteq R^2$ having only one critical point (x_0,y_0), and $f(x_0,y_0)$ is a local minimum. But $f(x_0,y_0)$ is not necessarily the global minimum of f on D.

EXAMPLE 3.1. ([1]) Consider the following function:

$$f(x,y) = \frac{-1}{1+x^2} + (2y^2 - y^4)\left(e^x + \frac{1}{1+x^2}\right).$$

A direct calculation shows that $(0,0)$ is the only critical point, i.e., $\triangledown f(0,0) = 0$. The second derivative test indicates that f has a local minimum at $(0,0)$. However, $f(0,0) = -1$ is not the global minimum of f on the plane, say, $f(0,2) = -17$ which is smaller than -1.

EXAMPLE 3.2. The following example was provided by the referee of the paper [14]:

$$f(x,y) = x^2(1+y)^3 + y^2.$$

Once again, one may verify that f has only one critical point $(0,0)$, and $f(0,0) = 0$ is a local minimum. But it is not the global minimum of f on a large open domain containing $(0,0)$, for example, $(-5,5)\times(-5,5)$.

There are other examples that support the statement 2 such as [11]. By virtue of different softwares in mathematics, constructing examples of multivariable functions with expected properties becomes not only easier, but also very stimulating. However, the more important question here perhaps is that "what extra condition(s) should be added to a function of two variables in order its only critical point which is a local minimum to be the global minimum ?" Ash and Sexton proved the following result:

THEOREM 3.3. ([1]) *Let $f : R^2 \longrightarrow R$ be continuously differentiable and have only one critical point (a,b) which is a local minimum. If f is proper ($f^{-1}(K)$ is compact whenever K is compact), then $f(a,b)$ must be the global minimum.*

This theorem imposes a "topological condition" to the function in addition to the differentiability.

THEOREM 3.4. ([10]) *Let $f : R^2 \longrightarrow R$ be continuously differentiable and have only one critical point (a,b) which is a local minimum. If f is convex, i.e., Hess(f) is positive definite, then $f(a,b)$ must be the global minimum.*

This theorem adds "convexity" to the objective function, where Hess(f) denotes the Hessian of f. As a special case of Theorem 2.4 when $n = 2$, we present the following result as another criterion that ensures a local extremum a global one.

THEOREM 3.5. *Let $f : R^2 \longrightarrow R$ be continuously differentiable and have only one critical point (a,b) which is a local minimum. If*

$$(x-y)(f_x - f_y) > 0 \qquad for\, x \neq y, \tag{4}$$

then $f(a,b)$ must be the global minimum.

REMARKS. (i) In Theorem 3.5, the function f does not have to be symmetric. The condition (4) ensures that the only critical point must be on the line $y = x$. The rest of the proof of Theorem 3.5 follows from a similar argument as we used in the proof of Theorem 2.4. However, when the objective function f involves more than two variables, we need a family of inequalities like (4) in order to have a theorem that is similar to Theorem 2.4 without the assumption on the symmetry of f. (ii) If we are concerned with only functions of two variables, Theorem 3.5 could be generalized even further to the following one.

THEOREM 3.6. *Let $f : R^2 \longrightarrow R$ be continuously differentiable and have only one critical point (a,b) which is a local minimum. If*

$$(Ax + By + C)(Af_x + Bf_y) > 0 \tag{5}$$

for (x,y) such that $Ax + By + C \neq 0$, then $f(a,b)$ must be the global minimum.

4. Generalized Power Means

In this section, we shall discuss more symmetric functions that are either Schur-convex or Schur-concave, and establish some inequalities for these functions. Some well-known classical analytic inequalities can be included as special cases.

Symmetric functions is a very important class of functions and have been used widely in many areas of mathematics. For instance, the following symmetric function is used to define the power mean [3,9]: Let $\mathbf{x} = (x_1, x_2, \cdots, x_n) \in R^n_+$, and $r \neq 0$, define

$$F_r(x_1, x_2, \cdots, x_n) = \left(\frac{x_1^r + x_2^r + \cdots + x_n^r}{n} \right)^{1/r},$$

and define $F_0(x_1, x_2, \cdots, x_n) = \left[\prod_{i=1}^n x_i\right]^{1/n}$ if $r = 0$. It is clear that $F_1(\mathbf{x})$ is the arithmetic mean of $\mathbf{x}$, $F_0(\mathbf{x})$ is the geometric mean of $\mathbf{x}$, and $F_{-1}(\mathbf{x})$ is the harmonic mean of $\mathbf{x}$. It is well-known that $F_r \longrightarrow F_0$ as r approaches zero. Moreover, for $r < s$, we have $F_r(\mathbf{x}) \leqslant F_s(\mathbf{x})$ with equality holding if and only if $x_1 = x_2 = \cdots = x_n$ ([9, p. 76]).

Now, let f be a positive twice differentiable function defined on an open interval (a,b) in R, and define the *generalized power mean* as follows:

$$F_{r,f}(x_1,x_2,\cdots,x_n)=\left\{\frac{f^r(x_1)+f^r(x_2)+\cdots+f^r(x_n)}{n}\right\}^{1/r},\qquad r\neq 0. \tag{6}$$

If $r=0$, we define

$$F_{0,f}(x_1,x_2,\cdots,x_n)=\left[\prod_{i=1}^{n}f(x_i)\right]^{1/n}. \tag{7}$$

Obviously, when $f(x)=x$, $F_{r,f}(\mathbf{x})$ becomes the ordinary power mean. It is also clear that $F_{r,f}\longrightarrow F_{0,f}$ as r approaches zero. In what follows, for simplicity of the statements, we shall call f *strictly rth power-convex* if $[f^r(x)]''>0$; and call f *strictly logarithmic convex* if $(\ln[f(x)])''>0$. The rth power-concavity and logarithmic concavity can be defined in the same way if the direction of inequalities are reversed. Schur-convexity of the generalized power mean can be summarized as follows.

THEOREM 4.1. *For the generalized power mean*

$$F_{r,f}(x_1,x_2,\cdots,x_n)=\begin{cases}\left[\frac{1}{n}\sum_{i=1}^{n}f^r(x_i)\right]^{1/r} & \text{if } r\neq 0\\ \left[\prod_{i=1}^{n}f(x_i)\right]^{1/n} & \text{if } r=0,\end{cases}$$

we have

(i) *If $r>0$, (resp., $r<0$,) then $F_{r,f}$ is Schur-convex if and only if f is r-th power-convex (resp., r-th power-concave);*

(ii) *If $r=0$, then $F_{r,f}$ is Schur-convex if and only if f is logarithmic convex.*

Proof. It is clear that $F_{r,f}(\mathbf{x})$ is symmetric. By Schur's condition (3), when $r\neq 0$, we have

$$\frac{\partial F_{r,f}}{\partial x_j}=\frac{1}{r}\left[\frac{1}{n}\sum_{i=1}^{n}f^r(x_i)\right]^{\frac{1-r}{r}}\cdot[f^r(x_j)]',\qquad j=1,2.$$

Therefore, if $x_1\neq x_2$ and $r>0$, we have

$$\begin{aligned}(x_1-x_2)\left(\frac{\partial F_{r,f}}{\partial x_1}-\frac{\partial F_{r,f}}{\partial x_2}\right)&\geqslant 0\\ \Longleftrightarrow (x_1-x_2)\left\{[f^r(x_1)]'-[f^r(x_2)]'\right\}&\geqslant 0\\ \Longleftrightarrow (x_1-x_2)^2[f^r(x^*)]''&\geqslant 0\\ \Longleftrightarrow [f^r(x)]''&\geqslant 0,\end{aligned}$$

where x^* is a real number between x_1 and x_2. Similarly, if $x_1\neq x_2$ and $r<0$, we have

$$(x_1-x_2)\left(\frac{\partial F_{r,f}}{\partial x_1}-\frac{\partial F_{r,f}}{\partial x_2}\right)\geqslant 0\Longleftrightarrow[f^r(x)]''\leqslant 0.$$

This proves (i). For $r=0$, observe that

$$\frac{\partial F_{0,f}}{\partial x_j}=\frac{1}{n}\left[\prod_{i=1}^{n}f(x_i)\right]^{(1-n)/n}\cdot f'(x_j)\cdot\prod_{i=1}^{n}f(x_i)/f(x_j),\qquad j=1,2.$$

Hence, if $x_1\neq x_2$, we have

$$\begin{aligned}(x_1-x_2)\left(\frac{\partial F_{0,f}}{\partial x_1}-\frac{\partial F_{0,f}}{\partial x_2}\right)&\geqslant 0\\ \Longleftrightarrow (x_1-x_2)\left\{\frac{f'(x_1)}{f(x_1)}-\frac{f'(x_2)}{f(x_2)}\right\}&\geqslant 0\\ \Longleftrightarrow (x_1-x_2)\left\{[\ln f(x_1)]'-[\ln f(x_2)]'\right\}&\geqslant 0\\ \Longleftrightarrow (x_1-x_2)^2\,[\ln f(x^*)]''&\geqslant 0\\ \Longleftrightarrow [\ln f(x)]''&\geqslant 0,\end{aligned}$$

where x^* is a point between x_1 and x_2. □

NOTE. Since $[f^r(x)]''=rf^{r-2}(x)[(r-1)(f'(x))^2+f(x)f''(x)]$, therefore when $r<0$, $[f^r(x)]''\geqslant 0$ if and only if $(r-1)[f'(x)]^2+f(x)f''(x)\leqslant 0$. On the other hand, $\{\ln[f(x)]\}''\leqslant 0$ if and only if $f(x)f''(x)-[f'(x)]^2\leqslant 0$. Hence we may include logarithmic concavity as the special case of rth power-convexity for $r=0$.

We now collect some immediate consequences of Lemma 2.3 and Theorem 4.1 as corollaries. Let us set

$$\mathbf{x}=(x_1,x_2,\cdots,x_n),\quad \mathbf{y}=(y_1,y_2,\cdots,y_n),\quad H_+^k=\{\mathbf{x}\in R_+^n\mid\sum_{i=1}^{n}x_i=k\},\ \ k\in R;$$

$$\sigma=\frac{1}{n}\sum_{i=1}^{n}x_i=\frac{k}{n}\quad\text{where}\quad \mathbf{x}\in H_+^k;\qquad \Omega_k=(\sigma,\sigma,\cdots,\sigma).$$

COROLLARY 4.2. *Let* $\mathbf{x}=S\mathbf{y}$ *for some doubly stochastic matrix* S. *If* $r>0$ *(resp.,* <0*), and* f *is strictly rth power-convex (resp., strictly rth power-concave), then*

$$F_{r,f}(\mathbf{x})\leqslant F_{r,f}(\mathbf{y}),$$

with equality holding if and only if S *is a permutation matrix.*

COROLLARY 4.3. *(a) If* $r>0$ *(resp.,* <0*) and* f *is strictly rth power-convex (resp., strictly rth power-concave), then*

$$f(\sigma)=F_{r,f}(\Omega_k)\leqslant F_{r,f}(\mathbf{x})=\left[\frac{1}{n}\sum_{i=1}^{n}f^r(x_i)\right]^{1/r}\quad on\ H_+^k;$$

(b) If $r \leqslant 0$ and f is strictly rth power convex, then

$$f(\sigma) = F_{r,f}(\Omega_k) \geqslant F_{r,f}(\mathbf{x}) = \left[\frac{1}{n}\sum_{i=1}^{n} f^r(x_i)\right]^{1/r} \quad \text{on } H_+^k,$$

with equality holding in either (a) or (b) if and only if $\mathbf{x} = \Omega_k$.

COROLLARY 4.4. *If $F_{r,f}$ is Schur-convex and $F_{s,f}$ is Schur-concave, then $F_{r,f}(\mathbf{x}) - F_{s,f}(\mathbf{x})$ is Schur-convex, and*

$$F_{r,f}(\mathbf{x}) - F_{s,f}(\mathbf{x}) \geqslant 0.$$

When $F_{r,f} - F_{s,f}$ is strictly Shcur-convex, the equality holds if and only if $x_1 = x_2 = \cdots = x_n$.

REMARKS 4.5. (i) In Corollary 4.4, if $r = 1$, $s = 0$ and $f(x) = x$, then $F_{1,f}(\mathbf{x})$ is the arithmetic mean of $\mathbf{x}$ which is Schur-convex (not strictly), and $F_{0,f}(\mathbf{x})$ is the geometric mean of $\mathbf{x}$ which is strictly Schur-concave. It is clear that $F_{1,f} - F_{0,f}$ is strictly Schur-convex, and the inequality in Corollary 4.4 reduces to the arithmetic-geometric mean inequality. (ii) In Corollary 4.3 (a), if $r = 1$, then $F_{1,f}(\mathbf{x}) = (1/n)\sum_{i=1}^{n} f(x_i)$, the rth power-convexity becomes the ordinary convexity, and the inequality is simply the well-known Jensen's inequality for positive convex functions

$$f\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) \leqslant \frac{1}{n}\sum_{i=1}^{n} f(x_i) = F_{1,f}(\mathbf{x}).$$

Notice that when $r > 1$, Corollary 4.3 (a) follows directly from the basic property of the ordinary power mean,

$$F_{1,f}(\mathbf{x}) = \frac{1}{n}\sum_{i=1}^{n} f(x_i) \leqslant F_{r,f}(\mathbf{x}) = \left[\frac{1}{n}\sum_{i=1}^{n} f^r(x_i)\right]^{1/r}.$$

When $r \leqslant 0$, the part (b) follows directly by the same reason. Because

$$F_{0,f}(\mathbf{x}) = \left[\prod_{i=1}^{n} f(x_i)\right]^{1/n}$$

which attains the maximum at $\mathbf{x} = \Omega_k$, and for $r < 0$,

$$F_{r,f}(\mathbf{x}) \leqslant F_{0,f}(\mathbf{x})$$

by the basic property of the power mean. Therefore, it seems that Corollary 4.3 has significance only for $0 < r < 1$. To see which kind of functions satisfy the inequalities in the corollary, we need to solve the following differential inequality

$$(r-1)[f']^2 + ff'' > 0, \tag{8}$$

where $0 < r < 1$, for positive differentiable functions. The following functions are some examples of solutions to inequality (8): (A) $r = 3/4$, $f(x) = x^2$ defined on $(0, +\infty)$; (B) $r = 1/2$, and $f(x) = e^x$. For the second example, the classical Jensen's inequality asserts that

$$f\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) = e^{\frac{1}{n}\sum_{i=1}^{n} x_i} \leqslant \frac{1}{n}\sum_{i=1}^{n} e^{x_i} = \frac{1}{n}\sum_{i=1}^{n} f(x_i),$$

and the power mean inequality implies that

$$F_{1/2,f}(\mathbf{x}) = \left[\frac{1}{n}\sum_{i=1}^{n} e^{x_i/2}\right]^2 \leqslant \frac{1}{n}\sum_{i=1}^{n} e^{x_i} = F_{1,f}(\mathbf{x}).$$

The equality in either of the above inequalities holds if and only if $x_1 = x_2 = \cdots = x_n$. Now, the Corollary 4.3 (a) claims that

$$f\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) \leqslant F_{1/2,f}(\mathbf{x}),$$

and with equality holding if and only if $x_1 = x_2 = \cdots = x_n$.

5. More Schur-Convex Functions

There are many different ways to construct a symmetric function F from a differentiable function f of single variable. F could be Schur-convex or Schur-concave if f satisfies certain conditions such as differential equations or differential inequalities. To conclude this paper, we shall briefly mention a few examples of this kind of functions. Some simple symmetric functions include $G(\mathbf{x}) = \sum_{i=1}^{n} g(x_i)$ where $g(x)$ is a differentiable function. As a matter of fact, the Schur-convexity of $G(\mathbf{x})$ had drawn attention of Hardy, Littlewood, and Pólya while they were investigating convex functions [2,7]. It is known that [10, p. 258],

> G is Schur-convex (resp., strictly Schur-convex) if and only if g is convex (resp., strictly convex), that is, $g'' \geqslant 0$, (resp., $g'' > 0$).

If we consider $G(\mathbf{x}) = \prod_{i=1}^{n} g(x_i)$ where $g(x)$ is a positive differentiable function, then a direct calculation shows that

> G is Schur-convex (resp., strictly Schur-convex) if and only if g is logarithmic convex (resp., strictly logarithmic convex).

In [14], we introduced the following Schur-convex functions that are made of positive solutions of a second order nonlinear differential equation. This new class of Schur-convex functions and related inequalities have been used to prove some geometric inequalities.

Let us consider the symmetric function

$$F(x_1,x_2,\cdots,x_n)=\left(\sum_{i=1}^{n} f(x_i)\right)^2-c_n\sum_{i=1}^{n} f(x_i)f'(x_i)-\left[nf(\sigma)-\sum_{i=1}^{n} f(x_i)\right]^2 \quad (9)$$

where $x_i\in(0,l)$, $i=1,2,\cdots,n$, $\sum_{i=1}^{n} x_i=ml$, $(0<m<n)$, $\sigma=(1/n)\sum_{i=1}^{n} x_i=ml/n$; and $f(x)$ is a positive solution of the equation

$$\left[f'(x)\right]^2-f(x)f''(x)=\mu, \qquad \text{where } \mu \text{ is a constant}, \quad (10)$$

such that $f'(x)f''(x)\neq 0$ on $(0,l)$; and $c_n=n[f(\sigma)/f'(\sigma)]$. In [14] we proved that

$$F(x_1,x_2,\cdots,x_n)\geqslant F(\sigma,\sigma,\cdots,\sigma)=0$$

if $f''(x)<0$ on $(0,l)$, and with equality holding if and only if $x_1=x_2=\cdots=x_n=\sigma$. Later on, we find that under the given conditions, the function $F(x_1,\cdots x_n)$ is actually a Schur-convex function if $f''(x)<0$ on $(0,l)$, and it is Schur-concave if $f''(x)>0$ on $(0,l)$. The nonlinear differential equation (10) can be generalized further so that more positive functions can be included to construct different Schur-convex or Schur-concave functions. This was done in two recent papers [13, 15] along with some applications. Although optimization problems of symmetric functions could be conducted by different techniques. Once the objective function $F(\mathbf{x})$ is confirmed to be Schur-convex or Schur-concave, it is not only easier to find the extremum, another advantage is to allow us to compare the function values at two different points where F may not attain the extremum. That is, if $\mathbf{x}$ and $\mathbf{y}$ are two points in the domain of definition of F and $\mathbf{y}=S\mathbf{x}$ for some doubly stochastic matrix S, then $F(\mathbf{y})\leqslant F(\mathbf{x})$ (resp., $\geqslant$) if F is Schur-convex (resp., Schur-concave). In this circumstance, we say " $\mathbf{y}$ is majorized by $\mathbf{x}$". Schur-convex functions preserve that majorization ([7]).

REFERENCES

[1] J. MARSHALL ASH AND HARLAN SEXTON, *A Surface with One Local Minimum,* Mathematics Magazine, Vol. **58**, No. 3, (1985) 147–149.

[2] G. H. HARDY, J. E. LITTLEWOOD, AND G. PÓLYA, *Some simple inequalities satisfied by convex functions*, Messenger Math. **58** (1929) 145–152.

[3] ———, *Inequalities*, 2nd Ed., Cambridge University Press, London, New York, 1952.

[4] MURRAYS KLAMKIN, *On Two Classes of Extremum Problems Without Calculus,* Mathematics Magazine, Vol. **65**, No. 2, (1992) 113–117.

[5] H. T. KU, M. C. KU AND X. M. ZHANG, *Generalized Power Means and Interpolating Inequalities*, Proc. AMS, to appear.

[6] J. E. MARSDEN AND A. J. TROMBA, *Vector Calculus*, 2nd Ed., W. H. Freeman and Company, 1981.

[7] ALBERT W. MARSHALL AND INGRAM OLKIN, *Inequalities: Theory of Majorization and Its Applications*, Academic Press, 1979.

[8] N. METROPOLIS AND GIAN-CARLO ROTA, *Symmetric Classes: Functions of Three Variables,* Amer. Math. Monthly, April, (1991) 328–332.

[9] D. S. MITRINOVI ć, *Analytic Inequalities*, Springer-Verlag, New York, 1970.

[10] A. WAYNE ROBERTS AND DALE E. VARBERG, *Convex Functions*, Academic Press, New York, San Francisco, London, 1973.

[11] IRA ROSENHOLTZ AND LOWELL SMYLE, *"The Only Critical Point in Town" Test,* Mathematics Magazine, Vol. **58**, No. 3, (1985) 149–150.
[12] TIBERIU TRIF, *Multiple Integrals of Symmetric Functions,* Amer. Math. Monthly, Aug/Sept (1997) 605–608.
[13] X. M. ZHANG, *Schur-Convex Functions and Isoperimetric Inequalities*, Proc. AMS, **126** (1998) 461–470.
[14] ———, *A Refinement of the Discrete Wirtinger Inequality*, J. Math. Anal. Appl., **200** (1996) 687–697.
[15] X. M. ZHANG AND DULAL K. BHAUMIK, *A New Criterion to Measure the Efficiency of a Design*, Sankhy ā –The Indian Journal of Statistics, to appear.

(Received March 3, 1998)

Xin-Min Zhang
Department of Mathematics and Statistics
University of South Alabama
Mobile, AL 36688
U. S. A.
e-mail: zhang@mathstat.usouthal.edu