Internat. J. Math. & Math. Sci.
Vol. 8 No. 1 (1985) 49-55

TRANSFORMATIONS WHICH PRESERVE CONVEXITY

ROBERT A. FONTENOT

Department of Mathematics
Whitman College
Walla Walla, Washington 99362

FRANK PROSCHAN

Department of Statistics
Florida State University
Tallahassee, Florida 32306

(Received January 7, 1983)

ABSTRACT. Let C be the class of convex nondecreasing functions $f: [0,\infty) \to [0,\infty)$ which satisfy $f(0) = 0$. Marshall and Proschan [1] determine the one-to-one and onto functions $\psi: [0,\infty) \to [0,\infty)$ such that $g = \psi \circ f \circ \psi^{-1}$ belongs to C whenever f belongs to C. We study several natural models for multivariate extension of the Marshall-Proschan result. We show that these result in essentially a restatement of the original Marshall-Proschan characterization.

KEY WORDS AND PHRASES. Convexity, preservation of convexity, transformations multivariate, permutation matrices.
1980 MATHEMATICS SUBJECT CLASSIFICATION CODE. AMS Subject classification: 26B25

1. INTRODUCTION.

Let C denote the class of convex, nondecreasing functions $f: [0,\infty) \to [0,\infty)$ which satisfy $f(0) = 0$. A known and useful result is that if $p \geq 1$, then

$$g(x) = f^{p}(x^{\frac{1}{p}})$$

belongs to C whenever f belongs to C. There is an interesting geometrical interpretation of the relationship between f and g. Beginning with the equation of the graph of g,

$$y = f^{p}(x^{\frac{1}{p}}),$$

one obtains that

$$y^{\frac{1}{p}} = f(x^{\frac{1}{p}})$$

and, thus, that the graph of f is obtained from the graph of g (and vice versa) by applying the <u>same</u> transformation to both of the coordinate axes.

In [1], A. W. Marshall and F. Proschan, motivated by the special case $\psi(x) = x^{p}$, pose and solve the following problem: Determine those one-to-one and onto functions

ψ: $[0, \infty) \to [0, \infty)$ such that $g = \psi \circ f \circ \psi^{-1}$ belongs to C whenever f belongs to C. The answer to this question is of interest in several applications. Karlin ([2], pp. 368, 369) uses the answer to the question in obtaining bounds on the survival function $\bar{F}(x)$ in terms of an exponential survival function. Barlow and Proschan ([3] pp. 110, 111) also use the result in obtaining bounds on the survival function $\bar{F}$ in terms of $\bar{G}$, where $\bar{G}^{-1}\bar{F}$ is convex. Marshall and Proschan show that if ψ is continuous at some point, then g belongs to C whenever f belongs to C if and only if $\psi(x) = cx^p$ for some $c > 0$ and $p \geq 1$.

In this note we study several natural models for multivariate extension of the Marshall-Proschan result. We show that these result in essentially a restatement of the original Marshall-Proschan characterization.

2. EXTENSIONS OF MARSHALL-PROSCHAN RESULT

We begin by establishing some notation. Let R_n^+ denote the nonnegative orthant of n-dimensional Euclidean space equipped with coordinatewise ordering, multiplication, and addition. If $\underline{x}$ and $\underline{y} \in R_n^+$, let $\underline{x} + \underline{y}$ and $\underline{xy}$ denote the sum and the product, computed coordinatewise, of $\underline{x}$ and $\underline{y}$. Let $\underline{0}$ and $\underline{e}$ denote the vectors all of whose entries are 0 and 1, respectively. If $x > \underline{0}$, let $\underline{x}^{-1}$ denote the multiplicative inverse of $\underline{x}$. Finally, let $\underline{x}^n$ denote the nth power of $\underline{x}$.

We now examine several possible extensions of the Marshall-Proschan result. Let C denote the set of those convex functions f: $R_n^+ \to [0, \infty)$ such that $f(\underline{0}) = 0$ and f is nondecreasing in each argument. The following question is natural: If $p \geq 1$, f belongs to C, and

$$g(\underline{x}) = f^p(x_1^{\frac{1}{p}}, \ldots, x_n^{\frac{1}{p}})$$

for $\underline{x} = (x_1,\ldots,x_n)$ in R_n^+, does g belong to C? The following example shows that the answer is "No".

EXAMPLE 1. Let $n \geq 2$ and $p > 1$. For $\underline{x}$ in R_n^+, let $f(\underline{x}) = x_1 + x_2$. Then

$$g(\underline{x}) = (x_1^{\frac{1}{p}} + x_2^{\frac{1}{p}})^p$$

is not convex since the partial derivative of g with respect to x_2 is decreasing as a function of x_2, for fixed x_1. Thus, for fixed x_1, g is not a convex function of x_2([4],Thm.12B). Hence g does not belong to C.

Example 1, the theorem of Marshall and Proschan, and the fact that the present class C contains a copy of the class C of interest to Marshall and Proschan show that, given the present choice of C, there is no point in considering real-valued functions $\psi(x)$ which operate on each coordinate separately. If we wish to look at such functions $\psi(x)$, we must choose C differently. For example, we might let C be the class of functions f: $R_n^+ \to [0, \infty)$ which are convex and nondecreasing in each variable separately and which satisfy $f(\underline{0}) = 0$; however, Example 1 shows that this choice is not suitable. Alternatively, we might let C consist of those functions f: $R_n^+ \to [0, \infty)$ which satisfy $f(\underline{0}) = 0$, are nondecreasing in each argument, and whose restrictions to rays through the origin are convex functions of one variable. In this case, the C-preserving functions are the same as those in the one-dimensional case; the proof is a trivial and, hence, uninteresting application of the theorem of Marshall and Proschan.

In one way, the result of Example 1 is not surprising: there is an obvious difference between the dimension of the transforming function x^p and the argument $\underline{x}$ of the functions to be transformed. Thus we now consider the case where both the convex functions and the transformation function ψ map R_n^+ into R_n^+. We say that $\underline{f}: R_n^+ \to R_n^+$ is <u>convex</u> if the inequality

$$\underline{f}(\lambda x + (1 - \lambda)\underline{y}) \le \lambda \underline{f}(\underline{x}) + (1 - \lambda)\underline{f}(\underline{y})$$

holds for all $\underline{x}, \underline{y} \in R_n^+$ and all $\lambda \in [0, 1]$. It is immediate that $\underline{f}$ is convex if and only if each of the n coordinate functions of $\underline{f}$ is convex in the usual sense. Let C denote the set of convex functions $\underline{f}: R_n^+ \to R_n^+$ which are nondecreasing in each coordinate and satisfy $\underline{f}(\underline{0}) = \underline{0}$. We now pose the following problem: Determine those one-to-one and onto functions $\underline{\psi}: R_n^+ \to R_n^+$ which are continuous at some point (or bounded in a neighborhood of some point; the answer is the same) and have the property that

$$\underline{f} \in C \text{ implies that } \underline{\psi} \circ \underline{f} \circ \underline{\psi}^{-1} \in C. \tag{2.1}$$

In Example 2 (below) we shall find a necessary condition that (2.1) is satisfied by a function $\underline{\psi}$ of a certain type. To aid us, we will use the following remark:

REMARK 1. Let $a_1, a_2, \ldots, a_n$ be real numbers. Let

$$h(\underline{x}) = \prod_{j=1}^{n} x_j^{a_j}, \ \underline{x} > \underline{0}.$$

Suppose that $a_j > 0$ and $a_k > 0$ for some integers j and k such that $j \neq k$. We claim that h is <u>not</u> convex. To see this, suppose, without loss of generality, that $j = 1$, $k = 2$. The first two principal minors of the Hessian matrix, the matrix of second order partial derivatives of h, must be nonnegative if h is convex ([4],Thm. 42F). Thus the inequalities

$$a_1(1 - a_1) \le 0$$

and

$$a_1 a_2 (1 - a_1 - a_2) \ge 0$$

must simultaneously hold if h is convex. These inequalities cannot both hold; thus h is not convex.

EXAMPLE 2. Let A be a real $n \times n$ matrix. Let $\underline{f}$ be the R_n^+-valued function with domain $\underline{x} > \underline{0}$, each of whose component functions is of the type given in Remark 1 such that the exponents in the kth component function are, in order, the elements of the kth row of A, $k = 1, \ldots, n$. Represent $\underline{f}$ as follows:

$$\underline{f}(\underline{x}) = \underline{x}^A \quad \text{for} \quad \underline{x} > 0.$$

It is easy and interesting to see that if

$$\underline{g}(\underline{x}) = \underline{x}^B$$

for $\underline{x} > \underline{0}$ and for some real $n \times n$ matrix B, then

$$\underline{g} \circ \underline{f}(\underline{x}) = \underline{x}^{BA},$$

where BA is the usual matrix product of B and A. Also $\underline{f}$ is invertible if and only if A is invertible and, in this case,

$$\underline{f}^{-1}(\underline{x}) = \underline{x}^{A^{-1}}.$$

Let us call a matrix simple if it is invertible and each row contains exactly one nonzero entry. A permutation matrix is a simple matrix such that the nonzero entry in each row is 1.

We will now present a result about non-preservation of convexity. Let A be an $n \times n$ non-simple invertible matrix and let $\underline{\psi}: R_n^+ \to R_n^+$ be one-to-one and onto and also satisfy

$$\underline{\psi}(\underline{x}) = \underline{x}^A, \ \underline{x} > \underline{0}. \tag{2.2}$$

An easy argument, which we omit, shows that there exists an $n \times n$ diagonal matrix P, all of whose diagonal entries are greater than or equal to one, such that the matrix $Q = APA^{-1}$ has a row with two (strictly) positive entries. Choose such a P and let

$$\underline{f}(\underline{x}) = \underline{x}^P, \ \underline{x} \in R_n^+.$$

It is clear that $\underline{f}$ belongs to C. On the other hand, by Remark 1, the function $\underline{\psi} \circ \underline{f} \circ \underline{\psi}^{-1}$ does not belong to C. Thus $\underline{\psi}$ does not satisfy (2.1).

Suppose that we now consider an arbitrary simple matrix A. Using "test" functions in C of the form

$$\underline{f}(\underline{x}) = (g(x_1), g(x_2), \ldots, g(x_n)),$$

where $g: [0, \infty] \to [0, \infty]$ is convex, nondecreasing, and satisfies $g(0) = 0$ and using the Marshall-Proschan result, it is easy to see that if $\underline{\psi}$ satisfies (2.1) and (2.2), then A must be a permutation matrix.

3. MAIN RESULTS.

THEOREM. **Suppose that $n \geq 2$ and $\underline{\psi}: R_n^+ \to R_n^+$ is one-to-one, onto, and continuous at some point.** Then $\underline{\psi}$ satisfies (2.1), if and only if

$$\underline{\psi}(x) = \underline{c}\underline{x}^B \tag{3.1}$$

for $\underline{x}$ in R_n^+, some vector $\underline{c} > \underline{0}$, and some permutation matrix B.

PROOF. If $\underline{\psi}$ satisfies (3.1) for some vector $\underline{c} > \underline{0}$ and some permutation matrix A, then $\underline{\psi}$ clearly satisfies (2.1).

Suppose that $\underline{\psi}$ satisfies (2.1). We shall derive a functional equation which is satisfied by $\underline{\psi}$. Motivated by the proof in [1] and consideration of invertible linear functions in C, we ask the following question: If $\underline{g}$ is one-to-one, onto, and $\underline{g}$ and $\underline{g}^{-1}$ both belong to C, what must be true of $\underline{g}$? It is easy to see that the equations

$$\underline{g}(\lambda\underline{x} + \underline{y}) = \lambda\underline{g}(\underline{x}) + \underline{g}(\underline{y})$$

$$\underline{g}^{-1}(\lambda\underline{x} + \underline{y}) = \lambda\underline{g}^{-1}(\underline{x}) + \underline{g}(\underline{y})$$

must hold for all $\lambda > 0$ and all $\underline{x}, \underline{y} \in R_n^+$. It then follows that $\underline{g}(\underline{x}) = \underline{x}A$, $\underline{x}$ in R_n^+, for some nonnegative simple matrix A.

For any $\underline{a} > \underline{0}$, let $\underline{f}(\underline{x}) = \underline{ax}$ and let $\underline{g} = \underline{\psi} \circ \underline{f} \circ \underline{\psi}^{-1}$. Since $\underline{\psi}$ satisfies (2.1) and both $\underline{f}$ and $\underline{f}^{-1}$ belong to C, both $\underline{g}$ and $\underline{g}^{-1}$ also belong to C. Thus there is a nonnegative simple matrix A such that

$$\underline{\psi} \circ \underline{f} \circ \underline{\psi}^{-1}(\underline{z}) = \underline{z}A, \ \underline{z} \in R_n^+.$$

Substituting $\underline{z} = \underline{\psi}(\underline{x})$ and $\underline{f}(\underline{x}) = \underline{ax}$, we obtain

$$\underline{\psi}(\underline{ax}) = \underline{\psi}(\underline{x})A, \ \underline{x} \in R_n^+,$$

where A depends on a.

We will now show that A is a diagonal matrix. First, note that since A is simple, there is some vector $\underline{b} > \underline{0}$ and a linear transformation Π, depending on $\underline{a}$, which permutes coordinates, such that

$$\underline{\psi}(\underline{ax}) = \underline{b}\Pi(\underline{\psi}(\underline{x})), \ \underline{x} \in R_n^+. \tag{3.2}$$

Note that Π is multiplicative, that is,

$$\Pi(\underline{xy}) = \Pi(\underline{x})\Pi(\underline{y}), \ \underline{x}, \underline{y} \in R_n^+. \tag{3.3}$$

Using (3.2) and (3.3), for any positive integer m, we obtain that

$$\underline{\psi}(\underline{a}^m\underline{x}) = \underline{c}\Pi^m(\underline{\psi}(\underline{x})), \ \underline{x} \in R_n^+,$$

for some $\underline{c} > \underline{0}$ which depends on $\underline{a}$. Take $m = n!$. Using elementary group theory and the fact that the linear transformations on R_n of permutation type form a group of order m, we have that Π^m is the identity transformation. Thus

$$\underline{\psi}(\underline{a}^m\underline{x}) = \underline{c}\underline{\psi}(\underline{x}), \ \underline{x} \in R_n^+.$$

Replacing $\underline{a}^m$ by $\underline{a}$, we may write

$$\underline{\psi}(\underline{ax}) = \underline{c}\underline{\psi}(\underline{x}), \ \underline{x} \in R_n^+, \tag{3.4}$$

for all $\underline{a} > \underline{0}$ and some $\underline{c} > \underline{0}$ which depends on $\underline{a}$.

Our next step is to express $\underline{c}$ in (3.4) in terms of $\underline{\psi}$. We claim that if $\underline{z} \not> \underline{0}$, then $\underline{\psi}(\underline{z}) \not> \underline{0}$. Suppose that $z_i = 0$ for some i, $1 \le i \le n$. Let $\underline{a} > \underline{0}$ and $b > 0$ be such that $a_i \neq b_i$ and $a_j = b_j$, $j \neq i$. By (3.4) there exist $\underline{c} > \underline{0}$ and $\underline{d} > \underline{0}$ such that

$$\underline{\psi}(\underline{ax}) = \underline{c}\underline{\psi}(\underline{x})$$

and

$$\underline{\psi}(\underline{bx}) = \underline{d}\underline{\psi}(\underline{x}) \tag{3.5}$$

for all $\underline{x}$ in R_n^+. Suppose that $\underline{\psi}(\underline{z}) > \underline{0}$. Since $\underline{az} = \underline{bz}$ and $\underline{\psi}(\underline{z})$ has a multiplicative inverse, it follows from (3.5) that $\underline{c} = \underline{d}$. Using (3.5) again with $\underline{x} = \underline{e}$, we obtain $\underline{\psi}(\underline{a}) = \underline{\psi}(\underline{b})$, which contradicts the fact that $\underline{\psi}$ is one-to-one. Thus our claim is established. Furthermore, since $\underline{\psi}^{-1}$ also satisfies (3.4) with

$\underline{a}$ and $\underline{c}$ interchanged, it follows that if $\underline{z} > \underline{0}$, then $\underline{\psi}(\underline{z}) > \underline{0}$. In particular, $\underline{\psi}(\underline{e}) > 0$ and $(\underline{\psi}(\underline{e}))^{-1}$ exists.

Let $\underline{\phi}(\underline{x}) = \underline{\psi}(\underline{x})(\underline{\psi}(\underline{e}))^{-1}$. It is clear that (3.4) is equivalent to the functional equation

$$\underline{\phi}(\underline{ax}) = \underline{\phi}(\underline{a})\underline{\phi}(\underline{x}), \ \underline{a} > \underline{0}, \ \underline{x} \in R_n^+. \qquad (3.6)$$

Using (3.6) and the result in the previous paragraph, we obtain $\underline{\phi}(\underline{0}) = \underline{0}$. Considering (3.6) for $\underline{a} > \underline{0}$ and $\underline{x} > \underline{0}$ and using exponential and logarithmic functions coordinatewise as appropriate, we transform (3.6) into a functional equation of the type

$$\underline{\beta}(\underline{y} + \underline{z}) = \underline{\beta}(\underline{y}) + \underline{\beta}(\underline{z}), \ \underline{y}, \ \underline{z} \in R_n.$$

Note that $\underline{\beta}$ is bounded on some open set in R_n. The solution of this equation [5] is

$$\underline{\beta}(\underline{z}) = \underline{z}C, \ \underline{z} \in R_n,$$

for some real matrix C. Transforming and letting A denote the transpose of C, we get

$$\underline{\phi}(\underline{x}) = \underline{x}^A, \ \underline{x} > 0. \qquad (3.7)$$

Using Example 2 and the fact that $\underline{\phi}$ satisfies (2.1), we have that A is a permutation matrix.

To finish the proof, we must show that (3.7) holds for all $\underline{x}$ in R_n^+. Let

$$\Pi(\underline{x}) = \underline{x}^{A^{-1}}, \ \underline{x} \in R_n$$

Note that Π is a linear transformation, that $\underline{\alpha} = \Pi \circ \underline{\phi}$ satisfies (2.1), that $\alpha(\underline{0}) = \underline{0}$, and that

$$\underline{\alpha}(\underline{x}) = \underline{x}, \ \underline{x} > \underline{0}. \qquad (3.8)$$

To complete the argument, we require the following result, whose proof is left to the reader: If $\underline{g}\colon R_n^+ \to R_n^+$ is convex and nondecreasing then $\underline{g}$ is <u>continuous from above</u> at every point $\underline{x}$, that is, for every sequence $(\underline{x}_n)$ of points such that $\underline{x}_n \geq \underline{x}$, for all n, and $\underline{x}_n \to \underline{x}$, $g(\underline{x}_n) \to \underline{g}(\underline{x})$. Choose $\underline{f}$ in C such that $\underline{f}$ is one-to-one and $\underline{x} \neq \underline{0}$ implies $f(x) > \underline{0}$; for example, take $f(\underline{x}) = \underline{x}B$ where B is an invertible $n \times n$ matrix all of whose entries are positive.

Using (3.8) and the result about continuity from above, we obtain that

$$\underline{\alpha} \circ \underline{f} \circ \underline{\alpha}^{-1}(\underline{x}) = \underline{x} \ \text{ for all } \ \underline{x} \in R_n^+ .$$

By the choice of $\underline{f}$, $\underline{f}(\underline{\alpha}^{-1}(\underline{x})) = \underline{f}(\underline{x})$ holds for $\underline{x} \neq \underline{0}$. Thus $\underline{\alpha}^{-1}(\underline{x}) = \underline{x}$ and hence $\underline{\alpha}(\underline{x}) = \underline{x}$ holds for all $\underline{x}$ in R_n^+. This completes the proof of the theorem.

ACKNOWLEDGEMENT. Research was performed while the first author was visiting the Department of Statistics, Florida State University during the academic year 1981-82.

Research was performed under the support of the Air Force Office of Scientific Research under contract AFOSR-82-K-0007.

REFERENCES

1. MARSHALL, A. W., and PROSCHAN, F., Convexity preserving scale transformations, Inequalities, Vol. III, Academic Press, New York, 1972, pp. 225-234.
2. KARLIN, S., Total Positivity, Vol. I., Stanford University Press, Stanford, Calif., 1968.
3. BARLOW, R. E., and PROSCHAN, F., Statistical Theory of Reliability and Life Testing: Probability Problems, Second Printing, To Begin With, Silver Spring, MD, 1981.
4. ROBERTS, A. W., and VARBERG, D. E., Convex Functions, Academic Press, New York, 1973.
5. ACZEL, J., Lectures on Functional Equations and their Applications, Academic Press, New York, 1966.