Inequalities Involving Multivariate Convex Functions

EDWARD NEUMAN

Department of Mathematics, Southern Illinois University,
Carbondale, Illinois 62901

AND

JOSIP PEČARIĆ

Faculty of Civil Engineering, University of Beograd,
Yugoslavia

Submitted by R. P. Boas

Received September 15, 1986

INTRODUCTION

Let $x_0, \ldots, x_n$ be given real numbers, where $a \leq x_0 \leq \cdots \leq x_n \leq b$. Further let $f$ be a real-valued function defined on $[a, b]$ and let $[x_0, \ldots, x_n]f$ denote the $n$th divided difference of $f$ at the points $x_0, \ldots, x_n$.

In [4] Farwig and Zwick established the following.

**THEOREM A.** Let $f^{(n)}$ be a convex function on $(a, b)$. Then

$$f^{(n)} \left( \frac{1}{n+1} \sum_{i=0}^{n} x_i \right) \leq n! \ [x_0, \ldots, x_n]f$$

$$\leq \frac{1}{n+1} \sum_{i=0}^{n} f^{(n)}(x_i).$$

If $x_0 \neq x_n$, then strict inequalities hold if and only if $f \notin P_{n+1}$ (the space of all polynomials of degree $\leq n+1$).

The following theorem gives a generalization of a result by Pečarić [11].

**THEOREM B [12].** Let $f$ be $(n+2)$-convex on $(a, b)$. Then the function

$$g(x) = [x + h_0, \ldots, x + h_n]f$$

is

541

Copyright © 1989 by Academic Press, Inc.
All rights of reproduction in any form reserved.
is a convex function of \( x \) for all \( x \) and \( h_0, \ldots, h_n \) such that \( x + h_i \in (a, b) \) \((i = 0, 1, \ldots, n)\).

In this paper we offer, among other things, generalizations of these results to the case of multivariate functions. In Section 2 we give some notation and definitions. The main results are presented in Section 3.

2. Notation and Definitions

Let us now introduce some notation and definitions which will be used throughout the sequel. By \( x, y, \ldots \), we denote elements of Euclidean space \( \mathbb{R}^k \) \((k \geq 1)\), i.e., \( x = (x_1, \ldots, x_k) \). Superscripts are used to number vectors \( x^i \) \((i = 0, 1, \ldots, n; n \geq 0)\). The inner product of \( x, y \in \mathbb{R}^k \) is denoted by \( x \cdot y = \sum_{i=1}^{k} x_i y_i \). For a given set \( A \subset \mathbb{R}^k \), \( \chi_A(x) \), \( \text{vol}_k(A) \), and \([A]\) represent the characteristic function, the \( k \)-dimensional Lebesgue measure, and the convex hull of \( A \), respectively. We use standard multi-index notation, i.e., for \( \alpha \in \mathbb{Z}^k_+ \), \(|\alpha| = \alpha_1 + \cdots + \alpha_k\), \( \alpha! = \alpha_1! \cdots \alpha_k! \), \( x^\alpha = x_1^{\alpha_1} \cdots x_k^{\alpha_k} \). Thus

\[
P_m(\mathbb{R}^k) = \left\{ \sum_{|\alpha| \leq m} c_\alpha x^\alpha : c_\alpha \in \mathbb{R} \right\}
\]

is the space of all polynomials of (total) degree \( \leq m \) which has dimension \( (k + m)^k \). Furthermore, we set

\[
D^\alpha = \left( \frac{\partial}{\partial x_1} \right)^{\alpha_1} \cdots \left( \frac{\partial}{\partial x_k} \right)^{\alpha_k}.
\]

By

\[
S^n = \left\{ (\lambda_1, \ldots, \lambda_n) \in \mathbb{R}^n : \lambda_i \geq 0, \text{ all } i, \sum_{i=1}^{n} \lambda_i \leq 1 \right\}
\]

we denote the standard \( n \)-simplex. Let us denote as in [7]

\[
\int_{[x^0, \ldots, x^n]} f = \int_{S^n} f(\lambda_0 x^0 + \cdots + \lambda_n x^n) d\lambda,
\]

where \( \lambda_0 = 1 - \lambda_1 - \cdots - \lambda_n \), \( d\lambda = d\lambda_1 \cdots d\lambda_n \), and \( f \) is a function defined and integrable over \([x^0, \ldots, x^n]\), \( x^0, \ldots, x^n \in \mathbb{R}^k \).

For our further purposes we need the following.
DEFINITION 2.1 [2]. Let $x^0, \ldots, x^n \in \mathbb{R}^k$. Then the $k$ dimensional divided difference of $f$ at $x^0, \ldots, x^n$, denoted by $[x^0, \ldots, x^n]f$, is a map from $C^\infty(\mathbb{R}^k)$ into $\mathbb{R}^N$, $N = (n + k + 1)$ whose $\alpha$th component is given by

$$([x^0, \ldots, x^n]f)_\alpha = \int_{[x^0, \ldots, x^n]} D^\alpha f,$$  

(2.2)

where $\alpha \in \mathbb{Z}_+^k$, $|\alpha| = n$.

Note that this definition is consistent with the fact that the divided difference (2.2) is independent of the order $x^0, \ldots, x^n$, but should be borne in mind that $[x^0, \ldots, x^n]f$ is an element of $\mathbb{R}^N$.

We also will need the following.

DEFINITION 2.2 [3]. For $\{x^0, \ldots, x^n\} \subset \mathbb{R}^k$ with $\text{vol}_k([x^n, \ldots, x^n]) > 0$ the multivariate B-spline $M(\cdot | x^0, \ldots, x^n)$ is defined for $n \geq k$ by requiring that

$$\int_{\mathbb{R}^k} f(x) M(x | x^0, \ldots, x^n) \, dx = n! \int_{S^n} f(\lambda_0 x^0 + \cdots + \lambda_n x^n) \, d\lambda$$

(2.3)

holds for all $f \in C(\mathbb{R}^k)$. When $n = k$, the B-spline $M(\cdot | x^0, \ldots, x^k)$ is given by

$$M(x | x^0, \ldots, x^k) = \frac{\chi_\sigma(x)}{\text{vol}_k(\sigma)}$$

(2.4)

with $\sigma = [x^0, \ldots, x^k]$.

If $n > k$, then (2.3) determines $M(\cdot | x^0, \ldots, x^n)$ everywhere on $\mathbb{R}^k$ (cf. [7]).

Following Micchelli [7] we will say that the points $x^0, \ldots, x^n$ are in general position if every subset of $k + 1$ points of $\{x^0, \ldots, x^n\}$ forms a simplex of dimension $k$.

For the reader's convenience we list below some well-known properties of these splines. We assume that the points $x^0, \ldots, x^n$ are in general position.

1. $M(\cdot | x^0, \ldots, x^n)$ is a polynomial of total degree $\leq n - k$ in each region bounded by, but not cut by, convex sets passing through subsets of $k$ points of $x^0, \ldots, x^n$.

2. $M(\cdot | x^0, \ldots, x^n) \in C_+^{n-k-1}(\mathbb{R}^k)$.

3. $M(x | x^0, \ldots, x^n) \{ \geq 0, \ x \in \text{interior} [x^0, \ldots, x^n]\}$.

Thus $\text{supp} M(\cdot | x^0, \ldots, x^n) = [x^0, \ldots, x^n]$.

4. $\int_{\mathbb{R}^k} M(x | x^0, \ldots, x^n) \, dx = 1$ (see, e.g., [3, 7, 8]).

For our further purposes we need the following version of Jensen's inequality for multivariate convex functions.
**Lemma A** [6]. Let $f$ be a multivariate convex function on $\mathbb{R}^k$ and let $\varphi_1, \ldots, \varphi_k \in C(\mathbb{R}^k)$. If $d\mu$ is a probability measure on $\mathbb{R}^k$, then

$$f\left(\int_{\mathbb{R}^k} \varphi_1 d\mu, \ldots, \int_{\mathbb{R}^k} \varphi_k d\mu\right) \leq \int_{\mathbb{R}^k} f(\varphi_1, \ldots, \varphi_k) d\mu. \quad (2.5)$$

The sign of equality holds in (2.5) if and only if $f \in P_1(\mathbb{R}^k)$.

Following Mitrinović [9] we denote by $m_r$, $r \in \mathbb{Z}_+$, where

$$m_r = \binom{n+r}{r}^{-1} \sum_{i_0 + \cdots + i_n = r} t_0^{i_0} \cdots t_n^{i_n} \quad (2.6)$$

$(i_0, \ldots, i_n \in \{0, 1, \ldots, r\})$, the $r$th generalized symmetric mean of $t_0, \ldots, t_n$, $t_i \in \mathbb{R}$, all $i$. The sum (2.6) involves $(n+r)$ terms. These means were studied extensively in [10].

### 3. Main Results

In this section we give generalizations of Theorems A and B to the case of multivariate convex functions. Also a multivariate analog of Hadamard's inequality (see [9])

$$g\left(\frac{a+b}{2}\right) \leq \frac{1}{b-a} \int_a^b g(t) \, dt \leq \frac{1}{2} (g(a) + g(b)),$$

where $g$ is a convex function on $(a, b)$, $a \neq b$, is included.

The following lemmas will be used in the sequel.

**Lemma 3.1.** Let $p_0, \ldots, p_n > -1$. Then

$$\int_{\mathbb{S}^n} \lambda_0^{p_0} \cdots \lambda_n^{p_n} \, d\lambda = \frac{\Gamma(p_0 + 1) \cdots \Gamma(p_n + 1)}{\Gamma(p_0 + \cdots + p_n + n + 1)}, \quad (3.1)$$

where $\Gamma(\cdot)$ stands for the gamma function.

**Proof.** Follows immediately from [1, p. 33].

**Lemma 3.2.** Let $x^0, \ldots, x^n \in \mathbb{R}^k$ ($n \geq k$) with $\text{vol}_k([x^0, \ldots, x^n]) > 0$. Then for every $l_i = 0, 1, \ldots$; $i = 1, 2, \ldots, k$,

$$\int_{\mathbb{R}^k} (x_i)^l M(x \mid x^0, \ldots, x^n) \, dx = m_i, \quad (3.2)$$

where $x_i$ denotes the $i$th component of $x \in \mathbb{R}^k$, $m_i$ stands for the $i$th generalized symmetric mean of $x_i^0, \ldots, x_i^n$. 

MULTIVARIATE CONVEX FUNCTIONS

Proof: Let $y \in \mathbb{R}^k$. Then for every $h \in L^\infty(\mathbb{R})$

$$\int_{\mathbb{R}^k} h(y \cdot x) M(x_0, \ldots, x^n) \, dx = \int_{\mathbb{R}} h(t) M(t | y \cdot x_0, \ldots, y \cdot x^n) \, dt \quad (3.3)$$

(see [3, p. 53]). Here $M(\cdot | y \cdot x_0, \ldots, y \cdot x^n)$ stands for the univariate B-spline with knots $y \cdot x_0, \ldots, y \cdot x^n$. Choosing $y = e^i$ the $i$th coordinate vector in $\mathbb{R}^k$, we obtain $y \cdot x = x_i$, $y \cdot x^s = x_i^s$ ($i = 1, 2, \ldots, k$; $s = 0, 1, \ldots, n$).

Setting $h(t) = t^i$ in (3.3) one gets

$$\int_{\mathbb{R}^k} (x_i)^i M(x_0, \ldots, x^n) \, dx = \int_{\mathbb{R}} t^i M(t | x_0^i, \ldots, x^n_i) \, dt.$$ 

Since

$$\int_{\mathbb{R}} t^i M(t | x_0^i, \ldots, x^n_i) \, dt = m_i,$$

(see [10]), the desired result follows.

We are ready to state and prove the following.

**Theorem 3.1.** Let $f$ be a convex function on $\mathbb{R}^k$ and let $vol_k([x_0, \ldots, x^n]) > 0$, $x_i \in \mathbb{R}^k$, $i = 0, 1, \ldots, n$. Then

$$f(m_1, \ldots, m_k) \leq \int_{\mathbb{R}^k} f((x_1)^i, \ldots, (x_k)^i) M(x_0, \ldots, x^n) \, dx, \quad (3.4)$$

where $l_i = 1, 2, \ldots$; $i = 1, 2, \ldots, k$. Equality holds in (3.4) if and only if $f \in P_1(\mathbb{R}^k)$.

**Proof:** Let $d\mu(x) = M(x_0, \ldots, x^n) \, dx$. It follows from (3) and (4) that $d\mu$ is a probability measure on $\mathbb{R}^k$. In order to establish (3.4) we set $\varphi_i(x) = x_i$; $i = 1, 2, \ldots, k$, in (2.5). Hence the desired inequality follows because of (3.2). The proof is complete.

Our next result reads as follows.

**Theorem 3.2.** Under the assumptions of Theorem 3.1 the following inequalities

$$f\left(\frac{1}{n + 1} (x_0 + \cdots + x^n)\right) \leq \int_{\mathbb{R}^k} f(x) M(x_0, \ldots, x^n) \, dx$$

$$\leq \frac{1}{n + 1} \sum_{j=0}^{n} f(x_j) \quad (3.5)$$

holds. The sign of equalities holds in (3.5) if and only if $f \in P_1(\mathbb{R}^k)$. 

Proof. The left-hand side inequality of (3.5) follows immediately from (3.4) and (2.6) by letting \( l_1 = \cdots = l_k = 1 \). In order to prove the right-hand side inequality of (3.5) we apply (2.3) and next (3.1). We obtain

\[
\int_{\mathbb{R}^k} f(x) M(x | x^0, \ldots, x^n) \, dx = n! \int_{\mathbb{R}^n} f(\lambda_0 x^0 + \cdots + \lambda_n x^n) \, d\lambda
\]

\[
\leq n! \sum_{j=0}^{n} f(x^j) \int_{\mathbb{R}^n} \lambda_j \, d\lambda = \frac{1}{n+1} \sum_{j=0}^{n} f(x^j),
\]

The last assertion of our theorem is obvious. This completes the proof. \( \square \)

**Corollary 3.1 (Hadamard's Inequality for Multivariate Convex Functions).** Let \( \sigma = [x^0, \ldots, x^n] \), with vol\(_k(\sigma) > 0 \), \( k \geq 1 \). If \( f: \mathbb{R} \to \mathbb{R} \) is a convex function, then

\[
f\left( \frac{1}{k+1} (x^0 + \cdots + x^n) \right) \leq \frac{1}{\text{vol}_k(\sigma)} \int_{\sigma} f(x) \, dx \leq \frac{1}{k+1} \sum_{j=0}^{n} f(x^j).
\]

The sign of equality holds if and only if \( f \in P_1(\mathbb{R}^k) \).

*Proof.* Follows immediately from (3.5) and (2.4). \( \square \)

A generalization of Theorem A to the case of multivariate functions is contained in the following.

**Theorem 3.3.** Let \( D^af, \) with \( |a| = n \), be a convex function on \([x^0, \ldots, x^n] \); \( x^0, \ldots, x^n \in \mathbb{R}^k \), \( k \geq 1 \). Then

\[
D^af \left( \frac{1}{n+1} (x^0 \cdots x^n) \right) \leq n! \left( [x^0, \ldots, x^n] f \right)_a
\]

\[
\leq \frac{1}{n+1} \sum_{j=0}^{n} D^a f(x^j). \quad (3.6)
\]

If \( \text{vol}_k([x^0, \ldots, x^n]) > 0 \), then the sign of equality holds in (3.6) if and only if \( f \in P_{n+1}(\mathbb{R}^k) \).

*Proof.* It follows from (2.2), (2.1), and (2.3) that

\[
([x^0, \ldots, x^n] f)_a = \frac{1}{n!} \int_{\mathbb{R}^k} D^a f(x) M(x | x^0, \ldots, x^n) \, dx,
\]

where \( |a| = n \). Next making use of (3.5), we arrive at (3.6). The proof is complete. \( \square \)

An application of Theorem 3.3 is given at the end of this section.
A multivariate analog of Theorem B reads as follows.

**Theorem 3.4.** If $f$ is a $k$-variate function having convex $\alpha$th derivative on $\mathbb{R}^k$ ($|\alpha| = n$), then the function

$$G_\alpha(x) = ([x + h^0, \ldots, x + h^n]f)_x$$

is a convex function of $x$ for all $x$ and all $h^0, \ldots, h^n \in \mathbb{R}^k$.

**Proof.** Let $\beta_j > 0$, $j = 0, 1, \ldots, n$, where $\sum_{j=0}^n \beta_j = 1$. Also let $x = \sum_{j=0}^n \beta_j x^j$. According to (2.2) we obtain

$$G_\alpha(x) = \int_{S^n} D^\alpha f \left( \sum_{i=0}^n \lambda_i \sum_{j=0}^n \beta_j (x^j + h^j) \right) d\lambda,$$

Since $D^\alpha f$ is convex on $\mathbb{R}^k$,

$$G_\alpha(x) \leq \sum_{j=0}^n \beta_j \int_{S^n} D^\alpha f(\lambda_0(x^j + h^0) + \cdots + \lambda_n(x^j + h^n)) d\lambda,$$

which in conjunction with (2.2) yields the assertion. This completes the proof. □

We close this section with an application of Theorem 3.3 to the classical Taylor expansion for the multivariate functions. Let $c, d \in \mathbb{R}^k$ and let $\overline{cd}$ denote a line segment joining $c$ and $d$. Further, let $\Omega$ be a bounded open subset of $\mathbb{R}^k$ such that $\overline{cd} \subset \Omega$. If $f \in C^n(\Omega)$, then

$$f(d) = \sum_{|\alpha| < n} \frac{(d-c)^\alpha}{\alpha!} D^\alpha f(c) + R,$$

where $\alpha \in \mathbb{Z}^k_+$,

$$(d-c)^\alpha = (d_1 - c_1)^{\alpha_1} \cdots (d_k - c_k)^{\alpha_k}$$

and

$$R = \sum_{|\alpha| = n} \frac{(d-c)^\alpha}{\alpha!} D^\alpha f(\xi),$$

$\xi = \eta c + (1 - \eta) d$, $0 \leq \eta \leq 1$. Under some additional assumptions about the function $f$, we can get better information on the localization of the intermediate point $\xi$. To this aim, we will show that the remainder $R$ can be expressed in the form

$$R = n! \sum_{|\alpha| = n} \frac{(d-c)^\alpha}{\alpha!} ([c, \ldots, c, d]f)_n.$$

(3.8)
It follows from [8] that
\[
R = \int_{[c, ..., c, d]} \underbrace{D^n_{d - c} f}_{n \text{ times}},
\]
(3.9)

where \( D_z f \) denotes the directional derivative of \( f \) in the direction \( z \in \mathbb{R}^k \). Let \( f(x) = g(y \cdot x), \ y \in \mathbb{R}^k, \ g \in C^n(\mathbb{R}), \) be a ridge function (for this method see [5]). Then

\[
D^n_{d - c} f(x) = (y \cdot (d - c))^n g^{(n)}(y \cdot x).
\]

Setting \( q_n(y) = (y \cdot (d - c))^n \), it is easy to check that

\[
q_n(y) = n! \sum_{|\alpha| = n} \frac{(d - c)^\alpha}{\alpha!} y^\alpha.
\]

Since

\[
D^\alpha f(x) = y^\alpha g^{(n)}(y \cdot x),
\]

\[
D^n_{d - c} f = n! \sum_{|\alpha| = n} \frac{(d - c)^\alpha}{\alpha!} D^\alpha f,
\]

where the last identity holds true for the ridge functions. Since the ridge functions form a dense subset of \( C^n(\mathbb{R}^k) \), we conclude that (3.10) is valid for all \( f \in C^n(\mathbb{R}^k) \). Combining (3.10) with (3.9) and (2.2), one gets the desired result (3.8).

It follows from (3.8) and (3.7) that

\[
n! \left( [c, ..., c, d] f \right)_n = D^\alpha f(\xi).
\]

(3.11)

Let us assume now that \( D^\alpha f \), with \( |\alpha| = n \), is a convex function on \( \Omega \). Then the left-hand side inequality of (3.6) in conjunction with (3.11) yields

\[
D^\alpha f \left( \frac{1}{n + 1} (nc + d) \right) \leq D^\alpha f(\xi).
\]

Moreover, if \( D^\alpha f \) is an increasing function along \( \overline{cd} \), then \( c_n \leq \xi \leq d \), where \( c_n = (nc + d)/(n + 1) \). Otherwise, if \( D^\alpha f \) is decreasing along \( \overline{cd} \), then \( c \leq \xi \leq c_n \).

A univariate analog of our last result is due to Farwig and Zwick [4].
REFERENCES