Mathematical
Inequalities
& Applications
Volume 23, Number 2 (2020), 425–432 doi:10.7153/mia-2020-23-33

MONOTONICITY OF WEIGHTED AVERAGES OF CONVEX FUNCTIONS

G. J. O. JAMESON

Dedicated to Grahame Bennett, master of this subject,
who died in December 2016

(*Communicated by Z. Ditzian*)

Abstract. We consider weighted averages of the form $B_n(W,f) = \sum_{r=0}^n w_{n,r} f(r/n)$, where W is a summability matrix and f is convex. Conditions are given for $B_n(W,f)$ to increase or decrease with n. It decreases whenever W is a Hausdorff mean. The sequence of Bernstein polynomials for a convex function is a special case.

1. Introduction

The following result was proved in [6]:

THEOREM BJ. *For a function f on $[0,1]$, define*

$$A_n(f) = \frac{1}{n-1} \sum_{r=1}^{n-1} f\left(\frac{r}{n}\right) \qquad (n \geqslant 2),$$

$$B_n(f) = \frac{1}{n+1} \sum_{r=0}^{n} f\left(\frac{r}{n}\right) \qquad (n \geqslant 1).$$

If f is convex, then $A_n(f)$ increases with n and $B_n(f)$ decreases.

Here we present a generalisation of these results to weighted averages, proved by a refinement of the same method. A sequence of weighted averages of the type above is given by

$$B_n(W,f) = \sum_{r=0}^{n} w_{n,r} f\left(\frac{r}{n}\right), \tag{1}$$

for $n \geqslant 1$, where $W = (w_{n,r})$ is a *summability matrix*, that is:

$w_{n,r} \geqslant 0$ for all $n \geqslant 0$, $r \geqslant 0$,
$w_{n,r} = 0$ for $r > n$ (so W is lower triangular),
$\sum_{r=0}^{n} w_{n,r} = 1$ for all n.

Mathematics subject classification (2010): 26D15, 40G05, 41A10.
Keywords and phrases: Monotonicity, weighted average, convex, summability matrix, Hausdorff mean, Bernstein polynomial.

Actually, we will need to accommodate the possibility that $w_{n,r}$ is only defined for $n \geqslant n_0$ (for some n_0), while satisfying the stated conditions. We will call such W an "incomplete" summability matrix. It could be extended to a full summability matrix by defining $w_{n,0}$ to be 1 for $n < n_0$, but nothing is gained by doing so.

The $B_n(f)$ in Theorem BJ is given by the Cesàro matrix $w_{n,r} = 1/(n+1)$, while $A_n(f)$ is obtained by taking $w_{n,r}$, for $n \geqslant 2$, to be 0 for $r = 0$ and $r = n$, and $1/(n-1)$ for $1 \leqslant r \leqslant n-1$.

We will identify quite simple conditions on $w_{n,r}$ that ensure that $B_n(W,f)$ either increases or decreases with n for convex f. We show that the condition for $B_n(W,f)$ to decrease is satisfied by a wide class of summability matrices, the so-called *Hausdorff mean* matrices. A special case equates to the statement that the sequence of Bernstein polynomials for a convex function decreases with n.

We then record some further examples of our theorems, and establish necessary and sufficient conditions, albeit at the cost of greater complication.

We mention that the article [1] describes some generalisations of Theorem BJ of a different sort.

2. Sufficient conditions

We present the decreasing case first.

THEOREM 1. *Let $W = (w_{n,r})$ be a summability matrix, and define $B_n(W,f)$ by (1). Suppose that f is convex on $[0,1]$. For $n \geqslant 2$ and $0 \leqslant r \leqslant n-1$, put*

$$u_{n,r} = \frac{n-r}{n} w_{n,r} + \frac{r+1}{n} w_{n,r+1}. \tag{2}$$

If

$$u_{n,r} = w_{n-1,r} \quad \textit{for } n, r \textit{ as stated}, \tag{3}$$

then $B_n(W,f)$ decreases with n for $n \geqslant 1$.

Proof. Let $n \geqslant 2$. The point r/n lies between $(r-1)/(n-1)$ and $r/(n-1)$. More exactly, for $1 \leqslant r \leqslant n$,

$$\frac{r}{n} = \frac{r}{n}\,\frac{r-1}{n-1} + \frac{n-r}{n}\,\frac{r}{n-1}.$$

Write $f(r/(n-1)) = f_r$. By convexity of f,

$$f\left(\frac{r}{n}\right) \leqslant \frac{r}{n} f_{r-1} + \frac{n-r}{n} f_r.$$

Also, $f(0/n) = f_0$ and $f(n/n) = f_{n-1}$. So

$$B_n(W,f) \leqslant w_{n,0} f_0 + \sum_{r=1}^{n-1} w_{n,r}\left(\frac{r}{n} f_{r-1} + \frac{n-r}{n} f_r\right) + w_{n,n} f_{n-1}.$$

Reassembling this to combine the two terms with f_r, we see that the right-hand side equates to $\sum_{r=0}^{n-1} u_{n,r} f_r$. Given (3), it follows that $B_n(W,f) \leqslant B_{n-1}(W,f)$. $\square$

Some immediate observations on Theorem 1 will be useful.

The proof actually shows that if (3) holds for a particular n, then $B_n(W,f) \leqslant B_{n-1}(W,f)$.

Of course, if f is concave, then $B_n(W,f)$ increases with n, and if f is linear, then $B_n(W,f)$ is constant. Applied to $f(x) = x$, this says that $\frac{1}{n}\sum_{r=0}^{n} r w_{n,r} = w_{1,1}$ for all n. We return to this point later.

It is easily checked that the Cesàro matrix satisfies (3), so Theorem 1 reproduces the second statement in Theorem BJ. In the next section, we will identify a wide class of matrices that satisfy (3).

Applied to the function $f(x) = x^p$, Theorem 1 states that $\frac{1}{n^p}\sum_{r=0}^{n} w_{n,r} r^p$ increases with n if $p > 1$, and decreases if $0 < p < 1$.

The reasoning in Theorem 1, applied to $f(x) = 1$, shows that $\sum_{r=0}^{n-1} u_{n,r} = 1$. A consequence of this is that the apparently weaker hypothesis $u_{n,r} \leqslant w_{n-1,r}$ $(0 \leqslant r \leqslant n-1)$ is actually equivalent to (3).

Similar reasoning establishes the criterion for $B_n(W,f)$ to increase. In this case, for reasons that will become apparent, it is essential to present the result for incomplete summability matrices.

THEOREM 2. *Let $W = (w_{n,r})$ be an incomplete summability matrix (restricted to $n \geqslant n_0$), and define $B_n(W,f)$ by (1). Suppose that f is convex on $[0,1]$. For $n \geqslant n_0$ and $1 \leqslant r \leqslant n$, put*

$$v_{n,r} = \frac{n-r}{n} w_{n,r} + \frac{r-1}{n} w_{n,r-1}, \tag{4}$$

(also $v_{n,0} = w_{n,0}$ and $v_{n,n+1} = w_{n,n}$). If

$$v_{n,r} = w_{n+1,r} \quad \textit{for } n \geqslant n_0 \textit{ and } 0 \leqslant r \leqslant n+1, \tag{5}$$

then $B_n(W,f)$ increases with n for $n \geqslant n_0$.

Proof. Let $n \geqslant 1$. This time we write, for $0 \leqslant r \leqslant n$,

$$\frac{r}{n} = \frac{n-r}{n}\,\frac{r}{n+1} + \frac{r}{n}\,\frac{r+1}{n+1}.$$

Write $f(r/(n+1)) = f_r$. By convexity of f,

$$B_n(W,f) \leqslant \sum_{r=0}^{n} w_{n,r}\left(\frac{n-r}{n} f_r + \frac{r}{n} f_{r+1}\right),$$

Reassembling, we see that the right-hand side equates to $\sum_{r=0}^{n+1} v_{n,r} f_r$. Given (5), it follows that $B_n(W,f) \leqslant B_{n+1}(W,f)$. $\square$

It is easily checked that the matrix generating $A_n(f)$ satisfies (5) (with $n_0 = 2$). In a case like this where $w_{n,0} = w_{n,n} = 0$, the values $f(0)$ and $f(1)$ do not appear in

$B_n(W,f)$, or in $\sum_{r=0}^{n+1} v_{n,r} f_r$, so we only need f to be defined and convex on $(0,1)$. For example, Theorem BJ can be applied with $f(x) = \log x$ to show that $(n!)^{1/n}/(n+1)$ decreases with n.

Despite the apparent similarity, there is an important difference between Theorems 1 and 2: if the values of $w_{n,r}$ are known for a certain n, then (5) determines the values of $w_{n+1,r}$, hence all *later* $w_{k,r}$, while (3) determines $w_{n-1,r}$, hence all *earlier* $w_{k,r}$.

In particular, (5) dictates that $w_{n,0}$ takes the same value for all n, and similarly for $w_{n,n}$. Consider the case $n_0 = 1$ in Theorem 2. Take starting values $w_{1,0} = \alpha$ and $w_{1,1} = 1-\alpha$. Then $w_{n,0} = \alpha$ and $w_{n,n} = 1-\alpha$ for all $n \geqslant 1$, so $B_n(W,f)$ equals $\alpha f(0) + (1-\alpha) f(1)$ for all n, and the theorem says nothing.

Now try $n_0 = 2$, with starting values $w_{2,0} = \alpha$, $w_{2,2} = \beta$ and $w_{2,1} = \gamma$, where $\alpha+\beta+\gamma = 1$. Then one can check that (5) implies that for all $n \geqslant 2$, $w_{n,0} = \alpha$, $w_{n,2} = \beta$ and $w_{n,r} = \gamma/(n-1)$ for $1 \leqslant r \leqslant n-1$. Hence $B_n(W,f) = \alpha f(0) + \beta f(1) + \gamma A_n(f)$, and the conclusion is still simply that $A_n(f)$ increases with n.

So Theorem 2 only begins to say anything beyond Theorem BJ when $n_0 \geqslant 3$. A later example will show that there really are non-trivial cases of this type.

3. Hausdorff means and Bernstein polynomials

Given any probability measure μ on $[0,1]$, the corresponding *Hausdorff mean* matrix H_μ is the summability matrix $(h_{n,r})$ defined by

$$h_{n,r} = \binom{n}{r} \int_0^1 \theta^r (1-\theta)^{n-r} d\mu(\theta). \tag{6}$$

This class of matrices is well known in summability theory. An introductory account of them can be seen in [9, chapter 11] and the theory has been greatly developed in a series of articles by Bennett, for example [2], [3], [4], [5]. However, none of these results are needed for present purposes.

Different choices of the measure μ deliver a rich variety of examples. When μ is the point mass at a chosen point x in $[0,1]$, H_μ becomes the *Euler* matrix $E(x)$, with entries $e_{n,r}(x) = \binom{n}{r} x^r (1-x)^{n-r}$. By evaluation of beta integrals, one sees that the choice $d\mu(\theta) = d\theta$ (i.e. Lebesgue measure) gives the ordinary Cesàro matrix. The choice $m\theta^{m-1}\,d\theta$ gives the "Gamma matrix of order m", and the dual choice $m(1-\theta)^{m-1}\,d\theta$ the "Cesàro matrix of order m"; the entries in these matrices can be written explicitly as quotients of binomial coefficients (e.g. see [5, p. 24]).

THEOREM 3. *Let H_μ be a Hausdorff mean matrix. If f is convex on $[0,1]$, then $B_n(H_\mu, f)$ decreases with n.*

Proof. We verify condition (3). By the elementary identities $\frac{r+1}{n}\binom{n}{r+1} = \binom{n-1}{r}$

and $\frac{n-r}{n}\binom{n}{r} = \binom{n-1}{r}$, the $u_{n,r}$ defined by (2) is given by

$$\begin{aligned} u_{n,r} &= \binom{n-1}{r} \int_0^1 \left(\theta^r (1-\theta)^{n-r} + \theta^{r+1}(1-\theta)^{n-r-1} \right) d\mu(\theta) \\ &= \binom{n-1}{r} \int_0^1 \theta^r (1-\theta)^{n-r-1}\, d\mu(\theta) \\ &= h_{n-1,r}. \quad \square \end{aligned}$$

The nth *Bernstein polynomial* for a function f on $[0,1]$ is the function $B_n(f)$ defined by

$$(B_n f)(x) = \sum_{r=0}^{n} \binom{n}{r} f\left(\frac{r}{n}\right) x^r (1-x)^{n-r}.$$

It is well known that for any continuous f, the sequence $B_n(f)$ converges uniformly to f, thereby giving one proof of Weierstrass's approximation theorem.

In our notation, $(B_n f)(x)$ is exactly $B_n[E(x), f]$, where $E(x)$ is the Euler matrix. So the following is simply a restatement of Theorem 3 applied to this matrix.

THEOREM 4. *If f is convex on $[0,1]$, then the Bernstein polynomials for f form a decreasing sequence of functions.*

This result is far from new. It was proved by Schoenberg [10]; see also [8, Corollary 4.2] or [7, section 4.4]. However, we have exhibited it as a special case of Theorem 3. We mention that it is quite easy to show directly that $B_n(f) \geqslant f$ for convex functions f, using positivity of the operator B_n.

4. A pair of examples

We cannot point to a class of matrices comparable to Hausdorff means that satisfy (5). However, the following companion pair of examples demonstrates that some matrices satisfying (3) are accompanied by analagous ones satisfying (5).

EXAMPLE 1. Let

$$w_{n,r} = \frac{2(r+1)}{(n+1)(n+2)}$$

for $n \geqslant 0$ and $0 \leqslant r \leqslant n$. This is, in fact, the Hausdorff mean given by $d\mu(\theta) = 2\theta\, d\theta$ (in other words, the Gamma matrix of order 2). However, it is just as easy, and more useful for our purposes, simply to verify condition (3) directly:

$$u_{n,r} = \frac{2(r+1)}{n(n+1)(n+2)} \Big((n-r) + (r+2) \Big) = \frac{2(r+1)}{n(n+1)} = w_{n-1,r}.$$

So, for example, if $p > 1$ and $S_n(p) = \sum_{r=0}^{n} (r+1) r^p$, then

$$\frac{S_n(p)}{(n+1)(n+2)n^p}$$

decreases with n.

EXAMPLE 2. For $n \geqslant 3$, let $w_{n,0} = w_{n,n} = 0$ and

$$w_{n,r} = \frac{2(r-1)}{(n-1)(n-2)}$$

for $1 \leqslant r \leqslant n-1$. We verify that (5) holds, so that $B_n(W,f)$ increases with n for convex f:

$$v_{n,r} = \frac{2(r-1)}{n(n-1)(n-2)}\Big((n-r)+(r-2)\Big) = \frac{2(r-1)}{n(n-1)} = w_{n+1,r}.$$

We remark that this matrix is generated by the starting values $(0,0,1,0)$ of $w_{3,r}$.

5. Necessary and sufficient conditions

First, we mention an obvious necessary condition which is enough to detect numerous matrices that do not satisfy the conclusion of Theorem 1 or 2. If $B_n(W,f)$ either decreases or increases with n (for $n \geqslant n_0$) for convex f, then it is constant for the linear function $f(x) = x$. In other words, $T_n(W)$ is constant for $n \geqslant n_0$, where

$$T_n(W) = \frac{1}{n}\sum_{r=0}^{n} r w_{n,r}, \tag{7}$$

In particular, if this occurs with $n_0 = 1$, then $T_n(W) = T_1(W) = w_{1,1}$ for all $n \geqslant 1$.

EXAMPLE 3. Slightly modifying Example 1, let $w_{n,r} = 2r/[n(n+1)]$ for $n \geqslant 1$ and $0 \leqslant r \leqslant n$. Then

$$T_n(W) = \frac{2}{n^2(n+1)}\sum_{r=1}^{n} r^2 = \frac{1}{3}\left(2+\frac{1}{n}\right).$$

This is not constant, so $B_n(w,f)$ is not monotonic for every convex f.

Conditions (3) and (5) have served very well for the applications described, but they are certainly not necessary. A rather trivial example is enough to illustrate this fact.

EXAMPLE 4. Let W be the Cesàro matrix, and let W' be obtained from W by changing row 2, setting $w'_{2,0} = w'_{2,2} = \frac{1}{2}$ and $w'_{2,1} = 0$. Then $B_2(W',f) = \frac{1}{2}f(0) + \frac{1}{2}f(1) = B_1(W,f)$, while $B_n(W',f) = B_n(W,f)$ for other n. So $B_n(W',f)$ decreases with n for convex f. However, (3) is not satisfied, since $u_{3,r} = \frac{1}{3}$ for $r = 0,1,2$.

We finish by establishing necessary and sufficient conditions. The following Lemma is easily proved by two steps of Abel summation [4, Lemma 1].

LEMMA. *Let a_r $(0 \leqslant r \leqslant n)$ be real numbers. Put $A_j = \sum_{r=0}^{j} a_r$ and $A_k^* = \sum_{j=0}^{k} A_j = \sum_{r=0}^{k}(k-r+1)a_r$. Suppose that $A_n = A_n^* = 0$ and $A_k^* \geqslant 0$ for $0 \leqslant k \leqslant n-1$. Then $\sum_{r=0}^{n} a_r x_r \geqslant 0$ for all convex sequences (x_r).*

THEOREM 5. *Let W be a summability matrix (possibly incomplete). Define $T_n(W)$ by (7). Then $B_n(W,f) \leqslant B_{n-1}(W,f)$ for all convex functions f if and only if $T_{n-1}(W) = T_n(W)$ and*

$$\sum_{r=0}^{k}\left(k-r+\frac{r}{n}\right)w_{n,r} \leqslant \sum_{r=0}^{k-1}(k-r)w_{n-1,r} \tag{8}$$

for $1 \leqslant k \leqslant n-1$. Also, $B_n(W,f) \leqslant B_{n+1}(W,f)$ for all convex functions f if and only if $T_{n+1}(W) = T_n(W)$ and

$$\sum_{r=0}^{k-1}\left(k-r-\frac{r}{n}\right)w_{n,r} \leqslant \sum_{r=0}^{k-1}(k-r)w_{n+1,r} \tag{9}$$

for $1 \leqslant k \leqslant n$.

Proof. We prove the first statement; the second one is similar. Necessity is easily proved directly. We have already noted necessity of $T_{n-1}(W) = T_n(W)$. Fix $k \leqslant n-1$ and let $f(x) = \max(k/(n-1)-x, 0)$. Then

$$B_{n-1}(W,f) = \sum_{r=0}^{k-1}\frac{k-r}{n-1}w_{n-1,r}, \qquad B_n(W,f) = \sum_{r=0}^{k}\left(\frac{k}{n-1}-\frac{r}{n}\right)w_{n,r}.$$

Necessity of (8) follows.

For sufficiency, as seen in Theorem 1, we have to show that $\sum_{r=0}^{n-1} a_r f_r \geqslant 0$, where $a_r = w_{n-1,r} - u_{n,r}$. We verify the conditions of the Lemma. First, note that $A_{n-1} = 0$. We show that $A_{k-1}^* \geqslant 0$ for $k \leqslant n-1$. Note that $A_{k-1}^* = \sum_{r=0}^{k-1}(k-r)a_r$. Now

$$\begin{aligned}\sum_{r=0}^{k-1}(k-r)u_{n,r} &= \frac{1}{n}\sum_{r=0}^{k-1}(k-r)\Big((n-r)w_{n,r}+(r+1)w_{n,r+1}\Big)\\ &= \frac{1}{n}\sum_{r=0}^{k}\Big((k-r)(n-r)+(k-r+1)r\Big)w_{n,r}\\ &= \sum_{r=0}^{k}\left(k-r+\frac{r}{n}\right)w_{n,r},\end{aligned} \tag{10}$$

so (8) implies that $A_{k-1}^* \geqslant 0$. We also require $A_{n-1}^* = 0$. Applying (10) with $k = n$, and using $\sum_{r=0}^{n-1} u_{n,r} = \sum_{r=0}^{n-1} w_{n-1,r} = 1$, we find that $A_{n-1}^* = (n-1)[T_n(W) - T_{n-1}(W)]$. □

While it is satisfying to have identified necessary and sufficient conditions, it is clear that the simpler conditions (3) and (5) are more useful for applications.

Acknowledgement. The author is grateful to the referees for drawing attention to references [7], [8] and [10], and for indicating the full statement of Theorem 5.

REFERENCES

[1] S. ABRAMOVICH, G. JAMESON AND G. SINNAMON, *Inequalities for averages of convex and superquadratic functions*, J. Ineq. Pure Appl. Math., **5** (2004), issue 4, article 91 (electronic).
[2] G. BENNETT, *Inequalities complimentary to Hardy*, Quart. J. Math. **49** (1998), 395–432.
[3] G. BENNETT, *An inequality for Hausdorff means*, Houston J. Math. **25** (1999), 709–744.
[4] G. BENNETT, *Mercer's inequality and totally monotonic sequences*, Math. Ineq. Appl. **14** (2011), 747–775.
[5] G. BENNETT, *Hausdorff means and moment sequences*, Positivity **15** (2011), 17–48.
[6] G. BENNETT AND G. JAMESON, *Monotonic averages of convex functions*, J. Math. Anal. Appl. **252** (2000), 410–430.
[7] J. BUSTAMANTE, *Bernstein Operators and Their Properties*, Birkhäuser (2017).
[8] R. A. DE VORE AND G. G. LORENTZ, *Constructive Approximation*, Springer (1993).
[9] G. H. HARDY, *Divergent Series*, Oxford Univ. Press (1949).
[10] I. J. SCHOENBERG, *On variation diminishing approximation methods*, On Numerical Approximation, Univ. Wisconsin Press (1959), 249–274.

(Received February 4, 2019)

G. J. O. Jameson
Dept. of Mathematics and Statistics
Lancaster University
Lancaster LA1 4YF, UK
e-mail: g.jameson@lancaster.ac.uk