Journal of Inequalities in Pure and Applied Mathematics

http://jipam.vu.edu.au/

Volume 7, Issue 3, Article 79, 2006

A VARIANCE ANALOG OF MAJORIZATION AND SOME ASSOCIATED INEQUALITIES

MICHAEL G. NEUBAUER AND WILLIAM WATKINS

DEPARTMENT OF MATHEMATICS
CALIFORNIA STATE UNIVERSITY NORTHRIDGE
NORTHRIDGE, CA 91330
michael.neubauer@csun.edu

bill.watkins@csun.edu
URL: www.csun.edu/~vcmth006

Received 03 March, 2006; accepted 21 March, 2006
Communicated by I. Olkin

ABSTRACT. We introduce a partial order, variance majorization, on $\mathbb{R}^n$, which is analogous to the majorization order. A new class of monotonicity inequalities, based on variance majorization and analogous to Schur convexity, is developed.

Key words and phrases: Inequality, Symmetric polynomial, Majorization, Schur convex, Variance.

2000 *Mathematics Subject Classification.* Primary: 26D05, 26D07; Secondary: 15A42.

1. INTRODUCTION

Let $(x_1, \dots, x_n)$ and $(y_1, \dots, y_n)$ be two sequences of real numbers in nonincreasing order. The sequence x *majorizes* y if

$$\sum_{k=1}^{i} x_k \geq \sum_{k=1}^{i} y_k,$$

for $i = 1, \dots, n$ with equality for $i = n$. Majorization is a partial order on the set of nonincreasing sequences having the same sum and it plays a large role in the theory of inequalities dating back to the work of I. Schur [7]. Indeed a function $F(x_1, \dots, x_n)$ of n real variables is said to be *Schur convex* if $F(x) \geq F(y)$ whenever the sequence x majorizes y. Marshall and Olkin [6] catalog many functions and results of this type with particular emphasis on statistical inequalities. As a simple example, take the product function $F(x) = \prod_{k=1}^{n} x_k$. If x, y are n-tuples of nonnegative real numbers and if x majorizes y, then $F(x) \leq F(y)$. That is, $-F$ is a Schur convex function. In particular, if $y = (\bar{x}, \dots, \bar{x})$ then x majorizes y and therefore the product of n nonnegative numbers with fixed mean is maximized when all of them are equal to the mean

ISSN (electronic): 1443-5756

064-06

$\bar{x}$. Another way to state this well-known elementary result is that the product of a sequence x of nonnegative reals with fixed mean $\bar{x}$ attains a maximum when the variance of x is zero.

Now suppose that the variance of x is also fixed. In this paper, we define a partial order (variance majorization) on the set of sequences x having a fixed mean *and* a fixed variance. We obtain a monotonicity result similar to the one above for sequences in which one is variance-majorized by the other. In particular the maximum value of the product of a sequence of nonnegative reals with fixed mean and fixed variance is attained when the sequence takes on only two values $\alpha < \beta$ and the multiplicity of β is 1. This simple consequence of the main theorem is known as Cohn's Inequality [1]: if $x_1, \ldots, x_n$ are nonnegative reals then

$$\prod_{k=1}^{n} x_k \leq \alpha^{n-1}\beta,$$

where α and β are chosen so that the sequences $x = (x_1, \ldots, x_n)$ and $(\alpha, \ldots, \alpha, \beta)$ have the same means and the same variances.

2. Majorization and Variance Majorization

Let $\mathfrak{I}$ ($\mathfrak{I}^{\text{st}}$) be the set of nondecreasing (strictly increasing) sequences in $\mathbb{R}^n$:

$$\mathfrak{I} = \{x \in \mathbb{R}^n : x_1 \leq x_2 \leq \cdots \leq x_n\}$$
$$\mathfrak{I}^{\text{st}} = \{x \in \mathbb{R}^n : x_1 < x_2 < \cdots < x_n\}.$$

The variance majorization order involves the variances of leading subsequences of $x \in \mathfrak{I}$. So let $x[i] = (x_1, \ldots, x_i)$ be the leading subsequence of x for $i = 1, \ldots, n$. Note that $x[i]$ consists of the i smallest components of x. We denote the mean of $x[i]$ by

$$\overline{x[i]} = (1/i)\sum_{k=1}^{i} x_k,$$

and the variance of $x[i]$ by

$$\text{Var}(x[i]) = (1/i)\sum_{k=1}^{i}(x_k - \overline{x[i]})^2.$$

2.1. Definitions of Variance Majorization and Majorization.

Definition 2.1 (Variance Majorization)**.** Let $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ be sequences of real numbers in $\mathfrak{I}$ such that $\bar{x} = \bar{y}$ and $\text{Var}(x) = \text{Var}(y)$. We say that x is *variance majorized* by y (or y *variance majorizes* x), if

$$\text{Var}(x[i]) \leq \text{Var}(y[i]),$$

for $i = 2, \ldots n$. We write $x \overset{\text{vm}}{\prec} y$ or $y \overset{\text{vm}}{\succ} x$.

For fixed mean m and variance $v \geq 0$, variance majorization is a partial order on the set

$$S(m, v) = \{x \in \mathfrak{I} : \bar{x} = m, \text{Var}[x] = v\},$$

which is the intersection of $\mathfrak{I}$ with the sphere in $\mathbb{R}^n$ centered at $m(1, 1, \ldots, 1)$ with radius $\sqrt{nv}$, and the hyperplane through $m(1, 1, \ldots, 1)$ orthogonal to the vector $(1, 1, \ldots, 1)$.

By contrast, majorization is a partial order of the set of nonincreasing sequences

$$\mathfrak{D} = \{z \in \mathbb{R} : z_1 \geq z_2 \geq \cdots \geq z_n\}.$$

Definition 2.2 (Majorization)**.** Let $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$ be sequences in $\mathfrak{D}$ such that $\bar{x} = \bar{y}$. We say that x is *majorized* by y (or y majorizes x), if

$$\overline{x[i]} \leq \overline{y[i]},$$

for $i = 1, \dots, n$. In this case we write $x \overset{\text{maj}}{\prec} y$.

The definition of majorization is usually given in this equivalent form:

$$\sum_{k=1}^{i} x_k \leq \sum_{k=1}^{i} y_k,$$

for $i = 1, \dots, n$ with equality for $i = n$.

2.2. **Least and Greatest Sequences with Respect to the Variance Majorization Order.** Returning now to the variance majorization order, there is a least element and a greatest element in $S(m, v)$ for each m and $v \geq 0$.

Lemma 2.1. *Let m and $v \geq 0$ be real numbers and let*

$$x_{\min} = (\alpha_1, \dots, \alpha_1, \beta_1)$$
$$x_{\max} = (\alpha_2, \beta_2, \dots, \beta_2),$$

where

$$\alpha_1 = m - \sqrt{v/(n-1)}$$
$$\beta_1 = m + \sqrt{(n-1)v}$$
$$\alpha_2 = m - \sqrt{(n-1)v}$$
$$\beta_2 = m + \sqrt{v/(n-1)}.$$

Then $x_{\min}, x_{\max} \in S(m, v)$ and

$$x_{\min} \overset{\text{vm}}{\prec} x \overset{\text{vm}}{\prec} x_{\max},$$

for all $x \in S(m, v)$.

Figure 2.1 shows the Hasse diagram for the variance majorization partial order for all integral sequences of length six with sum 0 and sum of squares equal to 30. In this case, $x_{\min} = (-1, -1, -1, -1, -1, 5)$ and $x_{\max} = (-5, 1, 1, 1, 1, 1)$.

By contrast, the least and greatest elements in $\mathfrak{D} \cap \{x : \bar{x} = m\}$ with respect to the majorization order are $(\bar{x}, \dots, \bar{x})$ and $(n\bar{x}, 0, \dots, 0)$.

3. Variance Monotone Functions and Schur Convex Functions

Let I be a closed interval in $\mathbb{R}$ and let $F(x_1, \dots, x_n)$ be a real-valued function defined on $\mathfrak{I} \cap I^n$.

Definition 3.1 (Variance Monotone)**.** The function F is *variance monotone increasing* on $\mathfrak{I} \cap I^n$ if

$$x \overset{\text{vm}}{\prec} y \implies F(x) \leq F(y),$$

for all $x, y \in \mathfrak{I} \cap I^n$. If $-F$ is variance monotone increasing, we say that F is *variance monotone decreasing*.

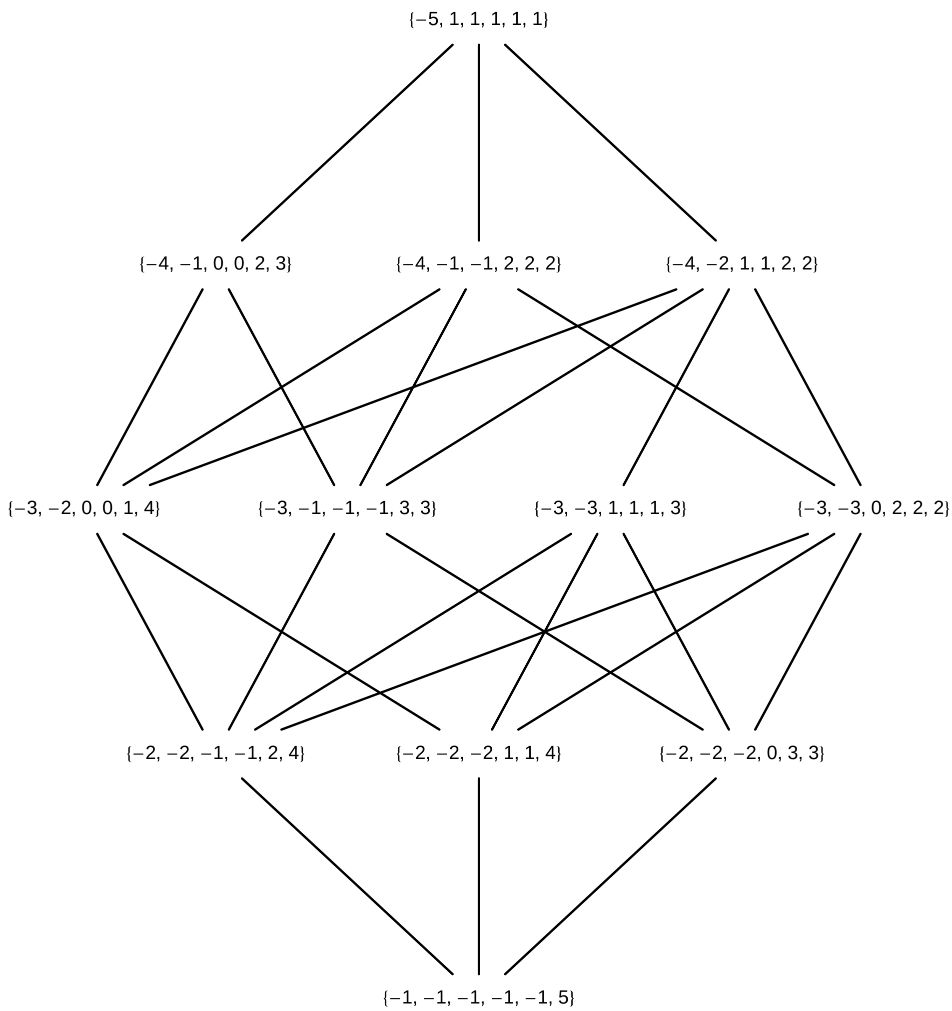

Figure 2.1: Variance majorization partial order for integral sequences in $S(0,30)$.

Definition 3.2 (Schur Convex)**.** The function F is *Schur convex* on $\mathfrak{D} \cap I^n$ if

$$x \overset{\text{maj}}{\prec} y \implies F(x) \leq F(y),$$

for all $x, y \in \mathfrak{D} \cap I^n$.

3.1. **The Main Result.** The next theorem is the main result.

Theorem 3.1. *Let I be a closed interval in $\mathbb{R}^n$, and let $F(z_1, \dots, z_n)$ be a continuous, real-valued function on $\mathfrak{I} \cap I^n$ that is differentiable on the interior of $\mathfrak{I} \cap I^n$ with gradient $\nabla F(z) = (F_1(z), \dots, F_n(z))$. Suppose that*

$$\frac{F_2(z) - F_1(z)}{z_2 - z_1} \geq \frac{F_3(z) - F_2(z)}{z_3 - z_2} \geq \cdots \geq \frac{F_n(z) - F_{n-1}(z)}{z_n - z_{n-1}}, \tag{3.1}$$

for all $z \in \mathfrak{I}^{\text{st}} \cap I^n$. Then F is variance monotone increasing on $\mathfrak{I} \cap I^n$, that is

$$x \overset{\text{vm}}{\prec} y \implies F(x) \leq F(y).$$

for all $x, y \in \mathfrak{I} \cap I^n$.

For functions of the form $F(z) = \phi(x_1) + \cdots + \phi(x_n)$, Theorem 3.1 specializes to the following corollary:

Corollary 3.2. *Let $\phi(t)$ be a continuous, real-valued function on a closed interval I such that ϕ is twice-differentiable on the interior of I and ϕ'' is nonincreasing. Then the function*

$$F(x_1, \ldots, x_n) = \phi(x_1) + \cdots + \phi(x_n)$$

is variance monotone increasing on the set of nondecreasing sequences in I^n. That is,

$$x \overset{\mathrm{vm}}{\prec} y \implies \phi(x_1) + \cdots + \phi(x_n) \leq \phi(y_1) + \cdots + \phi(y_n),$$

for all $x, y \in \mathfrak{I} \cap I^n$.

It turns out that $S(m, v) \subset I^n$ when the interval $I = \left[m - \sqrt{(n-1)v}, m + \sqrt{(n-1)v)}\right]$ (see Corollary 4.8). Thus the sequences $x_{\max}$ and $x_{\min}$ (described in Lemma 2.1) are in I^n. So, if F is variance monotone increasing on $\mathfrak{I} \cap I^n$, then the maximum and minimum values of F are attained at $x_{\max}$ and $x_{\min}$. This means we can bound $F(x)$ by expressions involving only the mean and variance of x:

Corollary 3.3. *Let m and $v \geq 0$ be real numbers and $I = \left[m - \sqrt{(n-1)v}, m + \sqrt{(n-1)v}\right]$. Let F be a variance monotone increasing function on $\mathfrak{I} \cap I^n$. Then*

$$F(\alpha_1, \ldots, \alpha_1, \beta_1) \leq F(x) \leq F(\alpha_2, \beta_2, \ldots, \beta_2),$$

for all $x \in S(m, v)$, where $\alpha_1, \beta_1, \alpha_2, \beta_2$ are defined as in Lemma 2.1.

3.2. **Schur Convex Functions.** By comparison, the following theorem by Schur is the result analogous to Theorem 3.1 for majorization. It plays the central role in the theory of majorization inequalities:

Theorem 3.4 ([7]). *Let $F(z)$ be a continuous, real-valued function on $\mathfrak{D}$ that is differentiable on in the interior of $\mathfrak{D}$. Then*

$$x \overset{\mathrm{maj}}{\prec} y \implies F(x) \leq F(y),$$

for all $x, y \in \mathfrak{D}$ if and only if

$$F_1(z) \geq F_2(z) \geq \cdots \geq F_n(z), \tag{3.2}$$

for all z in the interior of $\mathfrak{D}$.

The result analogous to Corollary 3.2 for majorization is known as Karamata's Theorem:

Corollary 3.5 ([4]). *Let ϕ be a continuous, real-valued function on a closed interval I such that ϕ is twice differentiable on the interior of I and ϕ'' is nonnegative. Then*

$$x \overset{\mathrm{maj}}{\prec} y \implies \phi(x_1) + \cdots + \phi(x_n) \leq \phi(y_1) + \cdots + \phi(y_n),$$

for all $x, y \in \mathfrak{D} \cap I^n$.

4. SOME VARIANCE MONOTONE AND SCHUR CONVEX FUNCTIONS

In this section, we give a short list of some common functions that are monotone in both the regular majorization and variance majorization orders. And we give as corollaries some samples of the kinds of inequalities one can obtain from Corollary 3.3.

4.1. **Elementary Symmetric Functions.** Let $E_k(z)$ denote the kth elementary symmetric function of the sequence $z = (z_1, \ldots, z_n) \in \mathbb{R}^n$:

$$E_k(z) = \sum z_{i_1} z_{i_2} \cdots z_{i_k},$$

where the sum is taken over all sets of k indices with $1 \le i_1 < \cdots < i_k \le n$.

Theorem 4.1. *Let E_k be the kth elementary symmetric function. Then*

$$x \overset{\text{vm}}{\prec} y \implies E_k(y) \le E_k(x),$$
$$x \overset{\text{vm}}{\prec} y \implies \frac{E_{k+1}(x)}{E_k(x)} \le \frac{E_{k+1}(y)}{E_k(y)},$$

for all $x, y \in \mathfrak{I} \cap [0, \infty)^n$.

The next corollary is obtained from Corollary 3.3 by evaluating the elementary symmetric function E_k at $x_{\min}$ and $x_{\max}$.

Corollary 4.2. *Let $x \in \mathfrak{I} \cap [0, \infty)^n$. Then $A(v, m, k) \le E_k(x) \le B(v, m, k)$, where*

$$A(v, m, k) = E_k(x_{\max}) = \binom{n}{k} \left(m + \sqrt{\frac{v}{n-1}} \right)^{k-1} \left(m - (k-1)\sqrt{\frac{v}{n-1}} \right)$$
$$B(v, m, k) = E_k(x_{\min}) = \binom{n}{k} \left(m - \sqrt{\frac{v}{n-1}} \right)^{k-1} \left(m + (k-1)\sqrt{\frac{v}{n-1}} \right)$$

The inequality analogous to Theorem 4.1 for (regular) majorization is given next:

Theorem 4.3 ([6, p. 80])**.**

$$x \overset{\text{maj}}{\prec} y \implies E_k(y) \le E_k(x)$$
$$x \overset{\text{maj}}{\prec} y \implies \frac{E_{k+1}(y)}{E_k(y)} \le \frac{E_{k+1}(x)}{E_k(x)},$$

for all $x, y \in \mathfrak{D} \cap [0, \infty)^n$.

4.2. **Moment Functions.** Let p be a positive real number and let $\phi(t) = t^p$. The pth moment function of $z \in [0, \infty)^n$ is given by

$$M_p(z) = z_1^p + \cdots + z_n^p.$$

The following results are applications of Corollaries 3.5 and 3.2:

Theorem 4.4. *Let M_p be the pth moment function. Then*

$$x \overset{\text{maj}}{\prec} y \implies M_p(x) \le M_p(y), \text{ for } p \in (-\infty, 0] \cup [1, \infty)$$
$$x \overset{\text{maj}}{\prec} y \implies M_p(y) \le M_p(x), \text{ for } p \in [0, 1],$$

for all $x, y \in \mathfrak{D} \cap [0, \infty)^n$ and

$$x \overset{\text{vm}}{\prec} y \implies M_p(x) \le M_p(y), \text{ for } p \in (-\infty, 0] \cup [1, 2]$$
$$x \overset{\text{vm}}{\prec} y \implies M_p(y) \le M_p(x), \text{ for } p \in [0, 1] \cup [2, \infty),$$

for all $x, y \in \mathfrak{I} \cap [0, \infty)^n$.

Again we obtain bounds on $M_p(x)$, which depend only on the mean and variance of x, from Corollary 3.3:

Corollary 4.5. *Let $x \in \mathfrak{I} \cap [0,\infty)^n$ with mean m are variance v. Let*

$$\begin{aligned} A(m,v,p) &= M_p(x_{\min}) = (n-1)\left(m - \sqrt{\tfrac{v}{n-1}}\right)^p + \left(m + \sqrt{(n-1)v}\right)^p \\ B(m,v,p) &= M_p(x_{\max}) = \left(m - \sqrt{(n-1)v}\right)^p + (n-1)\left(m + \sqrt{\tfrac{v}{n-1}}\right)^p. \end{aligned}$$

Then

$$A(m,v,p) \le M_p(x) \le B(m,v,p), \text{ for } p \in (-\infty,0) \cup [1,2]$$

and

$$B(m,v,p) \le M_p(x) \le A(m,v,p), \text{ for } p \in [0,1] \cup [2,\infty).$$

4.3. **Entropy Function.** The *entropy* function is defined for $x \in [0,\infty)^n$ by

$$H(x) = -(x_1 \log x_1 + \cdots + x_n \log x_n).$$

Letting $\phi(t) = -t \log t$, we have $\phi''(t) = -1/t$, which is nonpositive and increasing on $[0,\infty)$. Thus $-\phi$ satisfies the conditions of Corollaries 3.2 and 3.5. Thus we have the following inequalities:

Theorem 4.6. *Let H be the entropy function. Then*

$$x \overset{\text{vm}}{\prec} y \implies H(y) \le H(x),$$

for all $x, y \in \mathfrak{I} \cap [0,\infty)^n$, and

$$x \overset{\text{maj}}{\prec} y \implies H(y) \le H(x),$$

for all $x, y \in \mathfrak{D} \cap [0,\infty)^n$.

4.4. **Coordinates of x.** The smallest and the largest coordinates of a sequence are variance monotone decreasing.

Lemma 4.7. *Let $x, y \in \mathfrak{I}$. Then*

$$x \overset{\text{vm}}{\prec} y \implies x_1 \ge y_1 \text{ and } x_n \ge y_n.$$

We call this result a lemma because it is a part of the proof of Theorem 3.1 rather than a consequence of it.

When combined with Corollary 3.3, Lemma 4.7 gives bounds for the smallest and largest coordinates of a sequence in $S(m,v)$ in terms of m and v:

Corollary 4.8. *Let $x = (x_1, \ldots, x_n)$ be a sequence in $S(m,v)$. Then*

$$\begin{array}{ccccc} m - \sqrt{(n-1)v} & \le & x_1 & \le & m - \sqrt{v/(n-1)} \\ m + \sqrt{v/(n-1)} & \le & x_n & \le & m + \sqrt{(n-1)v}. \end{array}$$

Applying Corollary 4.8 to the eigenvalues of a symmetric matrix, we recover an equivalent form of an inequality of Wolkowicz and Styan [8, Theorem 2.1] that bounds the maximum and minimum eigenvalues by expressions involving only the trace and Euclidean norm of the matrix:

Corollary 4.9. *Let G be a symmetric matrix with eigenvalues $\lambda_1 \le \cdots \le \lambda_n$. Then*

$$\begin{array}{c} \frac{\text{tr}(G)}{n} + \frac{1}{n\sqrt{n-1}}\sqrt{n||G||^2 - (\text{tr}(G))^2} \le \lambda_n \le \frac{\text{tr}(G)}{n} + \frac{\sqrt{n-1}}{n}\sqrt{n||G||^2 - (\text{tr}(G))^2} \\ \frac{\text{tr}(G)}{n} - \frac{\sqrt{n-1}}{n}\sqrt{n||G||^2 - (\text{tr}(G))^2} \le \lambda_1 \le \frac{\text{tr}(G)}{n} - \frac{1}{n\sqrt{n-1}}\sqrt{n||G||^2 - (\text{tr}(G))^2}. \end{array}$$

Corollary 4.9 follows from the fact that the mean m and variance v of the eigenvalues can be expressed in terms of the trace and Euclidean norm $||G||$ of G as follows:

$$\begin{aligned} m &= \frac{\text{tr}(G)}{n} \\ v &= \frac{1}{n}\sum_i (\lambda_i - m)^2 \\ &= \frac{1}{n}\left(\sum_i \lambda_i^2 - 2m\lambda_i + m^2\right) \\ &= \frac{1}{n}(\text{tr}(G^2) - nm^2) \\ &= \frac{1}{n^2}\left(n||G||^2 - \text{tr}(G)^2\right). \end{aligned}$$

5. RESTRICTING $S(m,v)$ TO AN INTERVAL

Lemma 2.1 guarantees that there is a least and a greatest element in $S(m,v)$ with respect to the variance majorization order. Now we restrict the set $S(m,v)$ to an interval $I = [m-\delta, m+\beta]$, with $\delta, \beta \geq 0$, containing m. From Corollary 4.8, we have

$$\left[m - \sqrt{v/(n-1)}, m + \sqrt{v/(n-1)}\right]^n \subset S(m,v) \subset \left[m - \sqrt{(n-1)v}, m + \sqrt{(n-1)v}\right]^n.$$

So if either $\delta, \beta < \sqrt{v/(n-1)}$, then $S(m,v) \cap I^n = \emptyset$. However, if $S(m,v) \cap I^n$ is not empty, then it contains a least element, but it may not contain a greatest element.

Lemma 5.1. *Let m and $v \geq 0$ be real numbers. Let I be the interval $I = [m-\delta, m+\beta]$ such that $S(m,v) \cap I^n$ is not empty. Then there exist unique real numbers $m - \delta \leq \alpha \leq \gamma < m + \beta$ and an integer $1 \leq j \leq n-1$ such that the sequence*

$$x_{\min} = (\overbrace{\alpha, \dots, \alpha}^{j}, \gamma, \overbrace{m+\beta, \dots, m+\beta}^{n-j-1}) \in S(m,v) \cap I^n.$$

Moreover $x_{\min} \overset{\text{vm}}{\prec} x$ for all $x \in S(m,v) \cap I^n$.

Example 5.1. Let $n = 5$. The least element in $S(0, 1944/5) \cap [-36, 24]^5$ is $(-18, -18, -12, 24, 24)$. There is no greatest element in $S(0, 1944/5) \cap [-36, 24]^5$. See Section 6.7.

The situation for restricting the sequences to a closed interval is a little different for majorization. There is always a least element and a greatest in $\mathfrak{D} \cap [m, M]^n$. The sequence $(\bar{x}, \dots, \bar{x}) \in \mathfrak{D} \cap [m, M]^n$ is the least element. The greatest element with respect to majorization takes at most three values for its coordinates, two of which are the end points of the closed interval. That is, the greatest element in for majorization in $\mathfrak{D} \cap [m, M]^n$ is of the form

$$(M, \dots, M, \theta, m, \dots, m).$$

A discussion of restricting the majorization order to an interval is given in [5] and [6, page 132].

6. PROOFS

The main technique used in the proofs is to express the sequences in $\mathfrak{I}$ as linear combinations of a special basis for $\mathbb{R}^n$, the so-called Helmert basis.

6.1. **Helmert basis.** The purpose of this section is to describe the relationship between the coordinates of an n-tuple $x \in \mathbb{R}^n$ and the coordinates of x with respect to the so-called Helmert basis for $\mathbb{R}^n$ (see [6, p. 47] for a discussion of the Helmert basis). The *Helmert basis* for $\mathbb{R}^n$ is defined as follows:

$$
\begin{aligned}
w_0^T &= \frac{1}{\sqrt{n}}(1, 1, \dots, 1)\\
w_1^T &= \frac{1}{\sqrt{2}}(-1, 1, 0, \dots, 0)\\
w_2^T &= \frac{1}{\sqrt{6}}(-1, -1, 2, 0, \dots, 0)\\
&\vdots\\
w_i^T &= \frac{1}{\sqrt{i(i+1)}}(\overbrace{-1, \dots, -1}^{i}, i, 0, \dots, 0)\\
&\vdots\\
w_{n-1}^T &= \frac{1}{\sqrt{(n-1)n}}(-1, \dots, -1, n-1).
\end{aligned}
$$

It is clear that $\{w_0, w_1, \dots, w_{n-1}\}$ is an orthonormal basis for $\mathbb{R}^n$. Thus every vector $x \in \mathbb{R}^n$ is a linear combination $x = \sum_{k=0}^{n-1} a_k w_k$. The Helmert coefficient a_0 is determined by the mean $\bar{x}$ of the sequence x. Specifically, $a_0 = \sqrt{n}\bar{x}$. The other Helmert coefficients, $a_1, \dots, a_{n-1}$ are related to the variance, partial variances, and order of the coordinates of x.

Lemma 6.1. *Let x be a sequence of real numbers with $x = \sum_{k=0}^{n-1} a_k w_k$. Then*

$$
\mathit{Var}(x[i]) = \frac{1}{i}(a_1^2 + \cdots + a_{i-1}^2),
$$

for $i = 2, \dots, n$. In particular,

$$
\mathit{Var}(x) = \frac{1}{n}(a_1^2 + \cdots + a_{n-1}^2).
$$

Proof. Since $a_k = x \cdot w_k$,

$$
\begin{aligned}
\sum_{k=1}^{i-1} a_k^2 &= x\left(\sum_{k=1}^{i-1} w_k^T w_k\right) x^T\\
&= x\left((I_i - (1/i)J_i) \oplus 0_{n-i}\right) x^T\\
&= \left(\sum_{k=1}^{i} x_k^2 - (1/i)(\sum_{k=1}^{i} x_k)^2\right)\\
&= i\mathrm{Var}(x[i]).
\end{aligned}
$$

The $i \times i$ identity matrix is denoted by I_i and J_i denotes the $i \times i$ matrix all of whose entries are one. The fact that $\sum_{k=1}^{i-1} w_k^T w_k = I_i - (1/i)J_i$ follows from a simple inductive argument. ☐

Let $x = \sum a_i w_i$. In the next definition and lemma, we give necessary and sufficient conditions on the sequence $(a_1, \dots, a_{n-1})$ for the sequence x to be nondecreasing. (Clearly, the coefficient a_0 does not influence the relative ordering of the coordinates of x.)

Definition 6.1 (Admissible Sequence). Let $\alpha = (\alpha_1, \dots, \alpha_{n-1})$ be a sequence of nonnegative real numbers. Then α is *admissible* if

$$(i-1)i\alpha_{i-1} \le i(i+1)\alpha_i, \tag{6.1}$$

for $2 \le i \le n-1$.

Lemma 6.2. *Let $x = (x_1, \dots, x_n)$ be a vector in $\mathbb{R}^n$ and $a_0, \dots, a_{n-1}$ be scalars such that $x = \sum_{i=0}^{n-1} a_i w_i$. The following conditions are equivalent:*

1. *x is nondecreasing.*
2. *$a_i \ge 0$, for $i = 1, \dots, n-1$, and the sequence $a^{(2)} = (a_1^2, \dots, a_{n-1}^2)$ is admissible.*
3. *the kth component of $a_{i-1}w_i - a_i w_{i-1}$ is nonnegative for all $k \ne i$ and nonpositive for $k = i$.*

Proof. Let $1 \le i \le n-1$. Then $x_2 - x_1 = \sqrt{2}a_1$, and

$$\begin{aligned} x_{i+1} - x_i &= \frac{ia_i}{\sqrt{i(i+1)}} - \left(\frac{(i-1)a_{i-1}}{\sqrt{(i-1)i}} - \frac{a_i}{\sqrt{i(i+1)}} \right) \\ &= \frac{(i+1)a_i}{\sqrt{i(i+1)}} - \frac{(i-1)a_{i-1}}{\sqrt{(i-1)i}} \\ &= \frac{1}{i}\left(\sqrt{i(i+1)}a_i - \sqrt{(i-1)i}a_{i-1} \right), \end{aligned} \tag{6.2}$$

for $i \ge 2$. Thus x is nondecreasing if and only if $a_i \ge 0$, for $i = 1, \dots, n-1$ and $a^{(2)}$ is admissible. So Conditions 1 and 2 are equivalent.

Now let $(a_{i-1}w_i - a_i w_{i-1})_k$ be the k component of $a_{i-1}w_i - a_i w_{i-1}$. Then

$$(a_{i-1}w_i - a_i w_{i-1})_k = \begin{cases} -\frac{a_{i-1}}{\sqrt{i(i+1)}} + \frac{a_i}{\sqrt{(i-1)i}} & \text{, if } k < i, \\ -\frac{a_{i-1}}{\sqrt{i(i+1)}} - \frac{(i-1)a_{i-1}}{\sqrt{(i-1)i}} & \text{, if } k = i, \\ \frac{ia_{i-1}}{\sqrt{i(i+1)}} & \text{, if } k = i+1, \\ 0 & \text{, if } k > i+1. \end{cases} \tag{6.3}$$

To prove that Condition 2 implies Condition 3, suppose that $a_i \ge 0$ for $i = 1, \dots, n-1$ and that $a^{(2)}$ is admissible. It is clear from Equation (6.3) that $(a_{i-1}w_i - a_i w_{i-1})_k \ge 0$ for all $k \ne i$ and that $(a_{i-1}w_i - a_i w_{i-1})_i \le 0$.

Conversely, suppose that Condition 3 holds. With $k = 3, i = 2$ in Equation (6.3), we get $a_1 \ge 0$. With $k = 1 < i$ we get that $a^{(2)}$ is admissible and $a_i \ge 0$ for all i. Thus Condition 2 holds. □

6.2. **Proof of Theorem 3.1 and Lemma 4.7.** Let F be a differentiable, real-valued function on $\mathfrak{I} \cap I^n$ satisfying Inequality (3.1). Let x, y be nondecreasing sequences in I^n such that $\bar{x} = \bar{y}$, $\text{Var}(x) = \text{Var}(y)$ and $x \overset{\text{vm}}{\prec} y$. Let

$$x = \sum_{k=0}^{n-1} a_k w_k$$

$$y = \sum_{k=0}^{n-1} b_k w_k,$$

for scalars a_k, b_k, $k = 0, \dots, n-1$, and let

$$a^{(2)} = (a_1^2, \dots, a_{n-1}^2), \qquad b^{(2)} = (b_1^2, \dots, b_{n-1}^2).$$

Since $\bar{x} = \bar{y}$, we have $a_0 = b_0$. By Lemma 6.2, $a_k, b_k \geq 0$ for $k = 1, \dots, n-1$, $a^{(2)}$ and $b^{(2)}$ are admissible, and by Lemma 6.1

$$\sum_{k=1}^{i} a_k^2 \leq \sum_{k=1}^{i} b_k^2, \tag{6.4}$$

for $1 \leq i \leq n-1$ with equality in Inequality (6.4) for $i = n-1$ since Var$(x) =$ Var(y).

Next define a path $c(t) = (c_0, c(t)_1, \dots, c(t)_{n-1})$ from a to b by $c_0 = a_0 = b_0$, and

$$c(t)_k = \sqrt{(1-t)a_k^2 + tb_k^2},$$

for $t \in [0,1]$ and $k = 1, \dots, n-1$. Then $c(t)^{(2)} = (1-t)a^{(2)} + tb^{(2)}$, from which it follows that $c^{(2)}$ is admissible and that

$$\sum_{k=1}^{i} a_k^2 \leq \sum_{k=1}^{i} c(t)_k^2 \leq \sum_{k=1}^{i} b_k^2,$$

for $i = 1, \dots, n-1$.

Now define a path $z(t)$ from x to y by

$$z(t) = a_0 w_0 + \sum_{k=1}^{n-1} c(t)_k w_k.$$

Since $c(t)_k \geq 0$ and $c(t)^{(2)}$ is admissible, $z(t)$ is a nondecreasing sequence.

Let j be the smallest index for which $a_j \neq b_j$. Then $c(t)_k = a_k$ for $k < j$ and $c(t)_j > 0$ for $t > 0$. It is easy to verify that $c'(t)_k = (b_k^2 - a_k^2)/c(t)_k$ (unless $c(t)_k = 0$). Thus the tangent vector $z'(t)$ is given by

$$\begin{aligned} z'(t) &= \sum_{k=j}^{n-1} \frac{b_k^2 - a_k^2}{c(t)_k} w_k \\ &= (b_j^2 - a_j^2)\left(\frac{w_j}{c_j} - \frac{w_{j+1}}{c_{j+1}}\right) + (b_{j+1}^2 + b_{j+2}^2 - a_{j+1}^2 - a_{j+2}^2)\left(\frac{w_{j+2}}{c_{j+2}} - \frac{w_{j+3}}{c_{j+3}}\right) \\ &\quad + \cdots + \left(b_1^2 + b_2^2 + \cdots + b_{n-1}^2 - a_1^2 - a_2^2 - \cdots - a_{n-1}^2\right)\left(\frac{w_{n-2}}{c_{n-2}} - \frac{w_{n-1}}{c_{n-1}}\right), \end{aligned} \tag{6.5}$$

for $t > 0$. It follows from Inequality (6.4) that $z'(t)$ is a nonnegative linear combination of the vectors $\frac{w_{i-1}}{c_{i-1}} - \frac{w_i}{c_i}$, for $i = j+1, \dots, n-1$.

We now show that $z(t) \in I^n$ for all $t \in [0,1]$. Indeed, both the first and the last coordinates of $z(t)$ are nonincreasing functions of t. Thus

$$y_1 = z(1)_1 \leq z(t)_1 \leq z(t)_2 \leq \cdots \leq z(t)_n \leq z(0)_n = x_n.$$

Since $y_1, x_n \in I$, $z(t) \in I^n$ and thus $F(z(t))$ is defined for all $t \in [0,1]$. To see that $z(t)_1$ and $z(t)_n$ are decreasing, we examine the first and last coordinates of the vectors $\frac{w_{i-1}}{c_{i-1}} - \frac{w_i}{c_i}$. The first coordinate is

$$\frac{-1}{\sqrt{(i-1)i}c_{i-1}} + \frac{1}{\sqrt{i(i+1)}c_i},$$

which is nonpositive since $c^{(2)}$ is an admissible sequence. Thus $z'(t)_1 \leq 0$ and $z(t)_1$ is nonincreasing in t. This proves the first part of Lemma 4.7.

The last coordinate of $\frac{w_{i-1}}{c_{i-1}} - \frac{w_i}{c_i}$ is zero unless $i = n-1$ and in that case it is

$$\frac{-(n-1)}{\sqrt{(n-1)n}c_{n-1}},$$

which is also nonpositive. Thus $z(t)_n$ is nonincreasing. This proves the other part of Lemma 4.7.

Finally to prove that $F(z(t))$ is an increasing function in t, we show that

$$\frac{dF}{dt} = \nabla F \cdot z'(t) \geq 0.$$

In view of Equation (6.5), it suffices to show that

$$\nabla F \cdot \left(\frac{w_{i-1}}{c_{i-1}} - \frac{w_i}{c_i}\right) \geq 0,$$

for $i = j+1, \ldots, n-1$.

Since w_k is orthogonal to the all-ones vector e,

$$\nabla F \cdot \left(\frac{w_{i-1}}{c_{i-1}} - \frac{w_i}{c_i}\right) = (\nabla F - F_i e) \cdot \left(\frac{w_{i-1}}{c_{i-1}} - \frac{w_i}{c_i}\right).$$

For each $i = 1, \ldots, n-1$ define a function K_i on $\mathfrak{I}^{\text{st}} \cap I^n$ by

$$K_i(z) = \frac{F_{i+1}(z) - F_i(z)}{z_{i+1} - z_i}.$$

Now let $i < j$. Then $\frac{F_j(z) - F_i(z)}{z_j - z_i}$ is a convex combination of $K_k(z)$ for $k = i, \ldots, j-1$:

$$\frac{F_j(z) - F_i(z)}{z_j - z_i} = \sum_{k=i}^{j-1} \frac{z_{k+1} - z_k}{z_j - z_i} K_k(z).$$

Thus by Condition (3.1),

$$\frac{F_j(z) - F_i(z)}{z_j - z_i} \leq K_i(z)$$

and so

$$F_j(z) - F_i(z) \leq (z_j - z_i) K_i(z)$$

for all $i < j$ and $z \in \mathfrak{I}^{\text{st}} \cap I^n$.

Since $c(t)^{(2)}$ is an admissible sequence, Lemma 6.2 guarantees that all components of $\frac{w_{i-1}}{c_{i-1}} - \frac{w_i}{c_i}$ are nonpositive except the ith component. Of course the ith component of $\nabla F - F_i e$ is zero. It follows that

$$\begin{aligned}(\nabla F - F_i e) \cdot \left(\frac{w_{i-1}}{c_{i-1}} - \frac{w_i}{c_i}\right) &\geq K_i(z)(z - z_k e) \cdot \left(\frac{w_{i-1}}{c_{i-1}} - \frac{w_i}{c_i}\right) \\ &= K_k(z) z \cdot \left(\frac{w_{i-1}}{c_{i-1}} - \frac{w_i}{c_i}\right) \\ &= 0.\end{aligned}$$

The last equality holds because $z = \sum_{k=j}^{n-1} c_k w_k$ is clearly orthogonal to $\frac{w_{i-1}}{c_{i-1}} - \frac{w_i}{c_i}$.

We have shown that $\frac{dF}{dt} \geq 0$, for $z \in \mathfrak{I}^{\text{st}} \cap I^n$ and $t \in (0,1)$. Thus $F(z(t))$ is an increasing function of t. So $F(x) = F(z(0)) \leq F(z(1)) = F(y)$.

By the continuity of F, F is variance monotone increasing on $\mathfrak{I} \cap I^n$.

6.3. **Proof of Corollary 3.2.** Let $\phi(t)$ be a continuous, real-valued function on a closed interval I such that $\phi(t)$ is twice-differentiable on the interior of I and $\phi''(t)$ is nonincreasing on I. Let z be an increasing sequence in I^n. Let i be an integer satisfyng $1 \leq i \leq n-1$. By the mean-value theorem, there exists ξ_i in the interval $z_i < \xi_i < z_{i+1}$ such that

$$\frac{\phi'(z_{i+1}) - \phi'(z_i)}{z_{i+1} - z_i} = \phi''(\xi_i).$$

Inequality (3.1) follows since $\xi_1 \leq \xi_2 \leq \cdots \leq \xi_{n-1}$, ϕ'' is nonincreasing, and $F_i(z) = \phi'(z_i)$.

6.4. **Proof of Theorem 4.1.** Since $E_2(z)$ and $E_1(z)$ are constant on $S(m,v)$, we assume that $k \geq 3$. Let i be an integer satisfying $1 \leq i \leq n-1$ and let $E_k^{(i,i+1)}$ be the kth elementary symmetric polynomial of the $n-2$ variables $z_1, z_2, \ldots, z_{i-1}, z_{i+2}, \ldots, z_n$. Then

$$E_k(z) = E_k^{(i,i+1)} + (z_i + z_{i+1})E_{k-1}^{(i,i+1)} + z_i z_{i+1} E_{k-2}^{(i,i+1)}.$$

Thus

$$\frac{\partial E_k(z)}{\partial z_{i+1}} - \frac{\partial E_k(z)}{\partial z_i} = -(z_{i+1} - z_i)E_{k-2}^{(i,i+1)},$$

and so

$$\left(\frac{1}{z_{i+1} - z_i}\right)\left(\frac{\partial E_k(z)}{\partial z_{i+1}} - \frac{\partial E_k(z)}{\partial z_i}\right) = -E_{k-2}^{(i,i+1)},$$

for all $z \in \Im \cap [0,\infty]^n$. But the sequence z is nondecreasing so $E_{k-2}^{(i-1,i)} \geq E_{k-2}^{(i,i+1)}$ for $i = 1, \ldots, n-1$. It follows from Theorem 3.1 that $-E_k(z)$ is variance monotone increasing. Thus $E_k(z)$ is variance monotone decreasing. This proves the first inequality in Theorem 4.1.

To prove that the function $F(z) = E_{k+1}(z)/E_k(z)$ is variance monotone increasing, we must show that Inequality (3.1) holds. It suffices to show that

$$\frac{F_2(z) - F_1(z)}{z_2 - z_1} \geq \frac{F_3(z) - F_2(z)}{z_3 - z_2}.$$

We write E_k for the kth elementary symmetric function of $z_1, \ldots, z_n$ and E'_k for the kth elementary symmetric function of $z_4, \ldots, z_n$. Then

$$E_k = E'_k + (z_1 + z_2 + z_3)E'_{k-1} + (z_1 z_2 + z_1 z_3 + z_2 z_3)E'_{k-2} + z_1 z_2 z_3 E'_{k-3}. \tag{6.6}$$

It follows that

$$\begin{aligned} F_1 &= \frac{\partial}{\partial z_1}\left(\frac{E_{k+1}}{E_k}\right) \\ &= \frac{1}{E_k^2}[E_k E'_k - E_{k+1}E'_{k-1} + (z_2 + z_3)(E_k E'_{k-1} - E_{k+1}E'_{k-2}) \\ &\qquad + z_2 z_3 (E_k E'_{k-2} - E_{k+1}E'_{k-3})]. \end{aligned}$$

Thus

$$\frac{F_2 - F_1}{z_2 - z_1} = -\frac{1}{E_k^2}\left[E_k E'_{k-1} - E_{k+1}E'_{k-2} + z_3(E'_k E'_{k-2} - E_{k+1}E'_{k-3})\right].$$

Similarly,

$$\frac{F_3 - F_2}{z_3 - z_2} = -\frac{1}{E_k^2}\left[E_k E'_{k-1} - E_{k+1}E'_{k-2} + z_1(E'_k E'_{k-2} - E_{k+1}E'_{k-3})\right],$$

so that

$$\frac{F_2 - F_1}{z_2 - z_1} - \frac{F_3 - F_2}{z_3 - z_2} = \frac{z_3 - z_1}{E_k^2}(E_k E'_{k-2} - E_{k+1}E'_{k-3}).$$

Since $z_3 - z_1$ and E_k^2 are positive, it remains to show that $E_k E'_{k-2} - E_{k+1} E'_{k-3}$ is nonnegative, which can be rewritten using Equation (6.6) as

$$E_k E'_{k-2} - E_{k+1} E'_{k-3} = (E'_k E'_{k-2} - E'_{k+1} E'_{k-3}) + (z_1 + z_2 + z_3)(E'_{k-1} E'_{k-2} - E'_k E'_{k-3}) \\ + (z_1 z_2 + z_1 z_3 + z_2 z_2)(E'_{k-2} E'_{k-2} - E'_{k-1} E'_{k-3}).$$

Each of the expressions in E'_r above are nonnegative. These weak inequalities follow from a simple counting argument. Or we can use an old result in Hardy, Littlewood and Pólya [3, p. 52]:

$$z \in [0, \infty)^n \text{ and } s > r \implies E_{s-1} E_r > E_s E_{r-1},$$

with $r = k - 2$ and $s = k + 1, k, k - 1$.

6.5. **Proof of Lemma 5.1.** We may assume that $m = 0$ so that $I = [-\delta, \beta]$.

Let $z \in \mathbb{R}^n$ with mean $\bar{z} = 0$. There exist real numbers $a = (a_1, \ldots, a_{n-1})$ such that $z = \sum_{i=1}^{n-1} a_i w_i$ Then $z \in S(0, v)$ if and only if a satisfies the following conditions:

$$\begin{aligned} &a_i \geq 0, \text{ for all } i \\ &a^{(2)} \text{ is admissible} \\ &\textstyle\sum_{i=1}^{n-1} a_i^2 = nv. \end{aligned} \tag{6.7}$$

Let $z \in S(0, v)$. The only vector among $w_1, \ldots, w_{n-1}$ having a nonzero nth coordinate is w_{n-1}. Thus

$$z_n = \sqrt{\frac{n-1}{n}} a_{n-1}.$$

To compute z_1 in terms of a_k, notice that the first coordinate of each w_k is $-1/\sqrt{k(k+1)}$. So

$$z_1 = -\sum_{k=1}^{n-1} \frac{a_k}{\sqrt{k(k+1)}}.$$

Thus $z \in I^n$ if and only if

$$(n-1) a_{n-1}^2 \leq n \beta^2, \tag{6.8}$$

and

$$-\delta \leq -\sum_{k=1}^{n-1} \frac{a_k}{\sqrt{k(k+1)}}. \tag{6.9}$$

Next, we establish another inequality for the sequences a for which $z = \sum a_k w_k \in S(0, v) \cap I^n$. Since $a^{(2)}$ is admissible and Inequality (6.8) holds, we have

$$2a_1^2 \leq 6a_2^2 \leq \cdots \leq k(k+1) a_k^2 \leq \cdots \leq (n-1) n a_{n-1}^2 \leq n^2 \beta^2.$$

Thus

$$a_k^2 \leq n^2 \beta^2 \left[\frac{1}{k} - \frac{1}{k+1} \right], \tag{6.10}$$

for $k = 1, \ldots, n - 1$.

We now define $b = (b_1, \ldots, b_{n-1})$ so that it satisfies Conditions (6.7) and so that for all other a satisfying Conditions (6.7), we have

$$\sum_{k=1}^{i} b_k^2 \leq \sum_{k=1}^{i} a_k^2, \tag{6.11}$$

for $i = 1, \ldots, n-1$. In view of the fact that $\sum_{k=1}^{n-1} a_k^2 = nv = \sum_{k=1}^{n-1} b_k^2$, the inequalities above are equivalent to

$$\sum_{k=i+1}^{n-1} a_k^2 \leq \sum_{k=i+1}^{n-1} b_k^2, \tag{6.12}$$

for $i = 0, \ldots, n-2$.

We begin by specifying the integer j. The function

$$f(j) := n\beta^2 \left[\frac{1}{j} - \frac{1}{n}\right],$$

is decreasing in j with

$$f(1) = n\beta^2 \left[1 - \frac{1}{n}\right]$$
$$f(n) = 0.$$

We also have from Inequality (6.10) that

$$v = \frac{1}{n}\sum_{k=1}^{n-1} a_k^2 \leq n\beta^2 \left[1 - \frac{1}{n}\right] = f(1).$$

Thus there exists $1 \leq j \leq n$ such that

$$n\beta^2 \left[\frac{1}{j+1} - \frac{1}{n}\right] \leq v < n\beta^2 \left[\frac{1}{j} - \frac{1}{n}\right]. \tag{6.13}$$

Define the sequence of nonnegative reals $b = (b_1, \ldots, b_{n-1})$ as follows:

$$b_i^2 = n^2\beta^2 \left[\frac{1}{i} - \frac{1}{i+1}\right], \quad \text{for } i = j+1, \ldots, n-1 \tag{6.14}$$

$$b_j^2 = nv - n^2\beta^2 \left[\frac{1}{j+1} - \frac{1}{n}\right], \tag{6.15}$$

$$b_i = 0, \quad \text{for } i = 1, \ldots, j-1.$$

It is clear that

$$\sum_{k=1}^{n-1} b_k^2 = nv.$$

To check that $b^{(2)}$ is admissible, notice that

$$i(i+1)b_i^2 = n^2\beta^2 = (i+1)(i+2)b_{i+1}^2,$$

for $i = j+1, \ldots, n-2$. Also

$$\begin{aligned}
j(j+1)b_j^2 &= j(j+1)\left[nv - n^2\beta^2\left(\frac{1}{j+1} - \frac{1}{n}\right)\right] \\
&\leq j(j+1)\left[n^2\beta^2\left(\frac{1}{j} - \frac{1}{n}\right) - n^2\beta^2\left(\frac{1}{j+1} - \frac{1}{n}\right)\right] \\
&= n^2\beta^2 \\
&= (j+1)(j+2)b_{j+1}^2.
\end{aligned}$$

The inequality follows from the choice of j in Inequality (6.13). Of course $0 = i(i+1)b_i^2 \leq (i+1)(i+2)b_{i+1}^2$ for $i = 1, \ldots, j-1$. Thus $b^{(2)}$ is admissible. It follows that b satisfies Conditions (6.7).

Next we show that Inequalities (6.12) hold. So suppose that $a = (a_1, \dots, a_{n-1})$ satisfies Conditions (6.7) and (6.8). Then from Inequality (6.10) we get

$$\sum_{k=i+1}^{n-1} a_k^2 \le n^2\beta^2 \left[\frac{1}{i+1} - \frac{1}{n}\right].$$

For $i \ge j$ we have

$$\sum_{k=i+1}^{n-1} b_k^2 = n^2\beta^2 \sum_{k=i+1}^{n-1} \left[\frac{1}{k} - \frac{1}{k+1}\right]$$
$$= n^2\beta^2 \left[\frac{1}{i+1} - \frac{1}{n}\right].$$

So Inequality (6.12) holds for $i \ge j$. For $i < j$,

$$\sum_{k=i+1}^{n-1} a_k^2 \le \sum_{k=1}^{n-1} a_k^2 = nv = \sum_{k=i+1}^{n-1} b_k^2.$$

So Inequality (6.12) holds for $i < j$ too.

Now let

$$x_{\min} = \sum_{k=1}^{n-1} b_k w_k.$$

We now show that $x = x_{\min} = (\overbrace{\alpha, \dots, \alpha}^{j}, \gamma, \overbrace{\beta, \dots, \beta}^{n-j-1})$ for some $-\delta \le \alpha \le \gamma < \beta$. From the proof of Lemma 6.2 we have $x_{i+1} = x_i$ if and only if $i(i-1)b_{i-1}^2 = i(i+1)b_i^2$. For $i = 1, \dots, j-1$ we have $b_i = 0$. So $x_1 = \cdots = x_j = \alpha$. And $i(i-1)b_{i-1}^2 = i(i+1)b_i^2$ for $i = j+2, \dots, n-1$. So $x_{j+2} = \cdots = x_n = \beta$. The last equality follows from the choice of b_{n-1}.

We now show that $-\delta \le \alpha$. Since $S(0,v) \cap I^n$ is nonempty, let $z \in S(0,v) \cap I^n$ and let $z = \sum_{i=1}^{n-1} a_i w_i$ for some nonnegative reals $a = (a_1, \dots, a_{n-1})$. Then

$$-\delta \le z_1 = -\sum \frac{a_i}{\sqrt{i(i+1)}}.$$

We will show that

$$0 \le \sum_{i=1}^{n-1} \frac{a_i - b_i}{\sqrt{i(i+1)}}, \tag{6.16}$$

from which it follows that $-\delta \le \alpha$ and hence that $x_{\min} \in I^n$.

We begin with the following identity:

$$a_i - b_i = \frac{1}{a_i + b_i}\left(\sum_{k=1}^{i}(a_k^2 - b_k^2) - \sum_{k=1}^{i-1}(a_k^2 - b_k^2)\right).$$

Then

$$\sum_{i=1}^{n-1} \frac{a_i - b_i}{\sqrt{i(i+1)}} = \sum_{i=1}^{n-2}\left(\frac{1}{c_i} - \frac{1}{c_{i+1}}\right)\sum_{k=1}^{i}(a_k^2 - b_k^2), \tag{6.17}$$

where

$$c_i = \sqrt{i(i+1)}(a_i + b_i),$$

for $i = 1, \dots, n-2$. Both sequences $a^{(2)}$ and $b^{(2)}$ are admissible. Thus

$$c_i = \sqrt{i(i+1)}(a_i + b_i) \leq \sqrt{(i+1)(i+2)}(a_{i+1} + b_{i+1}) = c_{i+1},$$

for $i = 1, \dots, n-2$. It follows that $1/c_i - 1/c_{i+1} \geq 0$. Inequalities (6.11) hold so that the expression on the right-hand side of Equation (6.17) is nonnegative. Therefore Inequality (6.16) holds. This proves that $-\delta \leq \alpha$. It follows that $x_{\min} \in I^n$.

Finally, we show that $x_{\min}$ is unique. Let

$$y = (\overbrace{\alpha_1, \dots, \alpha_1}^{l}, \gamma_1, \overbrace{\beta, \dots, \beta}^{n-l-1}) \in S(0,v) \cap I^n,$$

where $-\delta \leq \alpha_1 \leq \gamma_1 < \beta$. We begin the proof that $x = x_{\min} = y$ by showing that $l = j$.

From the definition of b we have

$$b_1 = \cdots = b_{j-1} = 0.$$

Since $y_1 = \cdots = y_l = \alpha_1$ we have

$$a_1 = \cdots = a_{l-1} = 0.$$

But

$$\sum_{k=1}^{l-1} b_k^2 \leq \sum_{k=1}^{l-1} a_k^2 = 0, \tag{6.18}$$

and $b_{j+1} > 0$, so $l - 1 \leq j$.

We now show that $j \leq l$.

Since $y_{l+1} = \cdots = y_n$, we have

$$(l+1)(l+2)a_{l+1}^2 = (l+1)(l+2)a_{l+2}^2 = \cdots = (n-1)na_{n-1}^2 = n^2\beta^2.$$

In addition,

$$(j+1)(j+2)b_{j+1}^2 = (j+2)(j+3)b_{j+2}^2 = \cdots = (n-1)nb_{n-1} = n^2\beta^2.$$

Then

$$a_k^2 = n^2\beta^2 \left[\frac{1}{k} - \frac{1}{k+1}\right],$$

for $k = l+1, \dots, n-1$, and

$$b_k^2 = n^2\beta^2 \left[\frac{1}{k} - \frac{1}{k+1}\right],$$

for $k = j+1, \dots, n-1$.

Suppose $l < j$. Then $a_k = b_k$, $k = 1, \dots, j+1$ and $a_j > b_j$. This contradicts the fact that

$$\sum_{k=j}^{n-1} a_k^2 \leq \sum_{k=j}^{n-1} b_k^2.$$

Therefore $j \leq l$ and so $j = l$ or $j = l - 1$.

Suppose $j = l - 1$.Then from Inequality (6.18) we have $b_j = 0$ and $\gamma = \alpha$. Since

$$\sum_{k=1}^{j} a_k^2 \leq \sum_{k=1}^{j} b_j^2 = 0,$$

we have $a_j = 0$ and so $\gamma_1 = \alpha_1$, which contradicts the assumption that $\alpha_1 < \gamma_1$. It follows that $j = l$.

To finish the proof of uniqueness, we show that $a = b$. This follows immediately since

$$a_k = b_k = 0,$$

for $k = 1, \ldots, j-1$ and for $k = j+1, \ldots, n-1$, and $\sum_{k=1}^{n-1} a_k^2 = \sum_{k=1}^{n-1} a_k^2$. Thus $a_j = b_j$ and so $a = b$. It follows that $x_{\min} = y$.

6.6. **Proof of Lemma 2.1.** Let $x_{\min} = (\alpha_1, \ldots, \alpha_1, \beta_1)$ be defined as in the statement of Lemma 2.1. Clearly, Var$(x_{\min}[i]) = 0$, for $i = 2, \ldots, n-1$. Thus $x_{\min} \overset{\text{vm}}{\prec} x$, for all $x \in S(m, v)$.

The argument for $x_{\max} = (\alpha_2, \beta_2, \ldots, \beta_2)$ is more involved. Let b_0 and $b = (b_1, \ldots, b_{n-1})$ be the coordinates of $x_{\max}$ in the Helmert basis:

$$x_{\max} = (\alpha_2, \beta_2, \ldots, \beta_2) = b_0 w_0 + b_1 w_1 + \cdots + b_{n-1} w_{n-1}.$$

Since the second through nth coordinates of $x_{\max}$ are the same, the admissible sequence $b^{(2)}$ satisfies:

$$2b_1^2 = 6b_2^2 = 12b_3^2 = \cdots = (n-1)nb_{n-1}^2.$$

Now let $x \in S(m, v)$ with

$$x = a_0 w_0 + a_1 w_1 + \cdots + a_{n-1} w_{n-1}.$$

We must show that

$$\sum_{k=1}^{i} a_i^2 \le \sum_{k=1}^{i} b_i^2, \tag{6.19}$$

for $i = 1, \ldots, n-1$. We need the next lemma to finish the proof.

Lemma 6.3. *Let $r = (r_1, \ldots, r_m)$ be an admissible sequence of nonnegative real numbers with $\sum r_k = t$. Let $s = (s_1, \ldots, s_m)$ be the unique sequence of nonnegative reals with*

$$2s_1 = 6s_2 = 12s_3 = \cdots = m(m+1)s_m$$

and $\sum s_k = t$. Then

$$\sum_{k=1}^{i} r_k \le \sum_{k=1}^{i} s_k,$$

for $i = 1, \ldots, m$.

Equation (6.19) follows from Lemma 6.3 with $m = n-1$, $r = a^{(2)}$, $s = b^{(2)}$, and $t = nv$.

Proof. (**Lemma 6.3**) The condition on the sum and for admissibility can be expressed in matrix form as follows: $\sum r_k = t$ and r is admissible if and only if the first $m-1$ coordinates of Tr are nonnegative and the last coordinate is t, where

$$T = \begin{bmatrix} -t_1 & t_2 & 0 & & \cdots & & 0 \\ 0 & -t_2 & t_3 & 0 & \cdots & & 0 \\ \vdots & \vdots & \ddots & \ddots & & & \vdots \\ 0 & 0 & & -t_i & t_{i+1} & & 0 \\ 0 & 0 & & & \ddots & \ddots & \vdots \\ 0 & 0 & & \cdots & & -t_{m-1} & t_m \\ 1 & 1 & & \cdots & & 1 & 1 \end{bmatrix}$$

and $t_i = i(i+1)$. We can express this condition symbolically as

$$Tr = \begin{bmatrix} +_{m-1} \\ t \end{bmatrix},$$

where $+_{m-1}$ is a vector in $\mathbb{R}^{m-1}$ with nonnegative components.

From the definition of s we have

$$Ts = \begin{bmatrix} 0_{m-1} \\ t \end{bmatrix},$$

where 0_{m-1} is the zero vector in $\mathbb{R}^{m-1}$. So $Tr \geq Ts$. (The inequalities are coordinate-wise.)

Now let P be the $m \times m$ lower triangular matrix with ones on and below the main diagonal. Then

$$Pr = \begin{bmatrix} r_1 \\ r_1 + r_2 \\ \vdots \\ r_1 + r_2 + \cdots + r_m \end{bmatrix}.$$

We must show that $Pr \leq Ps$.

Let $Q = P^{-1}$. Then Q is the lower triangular matrix with diagonal entries equal to 1 and subdiagonal entries equal to -1. All other entries are zero:

$$Q = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ -1 & 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ 0 & & -1 & 1 & 0 \\ 0 & \cdots & 0 & -1 & 1 \end{bmatrix}.$$

Next observe that TQ is the tridiagonal matrix:

$$TQ = \begin{bmatrix} -q_2 & t_2 & 0 & \cdots & 0 & 0 \\ t_2 & -q_3 & t_3 & \cdots & 0 & 0 \\ 0 & t_3 & -q_4 & \ddots & 0 & 0 \\ \vdots & & \ddots & \ddots & & \\ 0 & & 0 & t_{m-1} & -q_m & t_m \\ 0 & & 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} U_1 & u \\ 0_{m-1} & 1 \end{bmatrix},$$

where $t_i = i(i+1), q_i = t_{i-1} + t_i = 2i^2$, U_1 is the $(m-1) \times (m-1)$ submatrix of TQ in the upper left corner, and $u = (0, \dots, 0, t_m)^T$. Thus $(TQ)^{-1} = PT^{-1}$ has the following form:

$$PT^{-1} = \begin{bmatrix} U_1^{-1} & w \\ 0_{m-1} & 1 \end{bmatrix},$$

for some $w \in \mathbb{R}^{m-1}$.

Next we show that $-U_1$ is a symmetric M-matrix. Let D be the $n \times n$ diagonal matrix with diagonal entries $1/2, 1/3, \dots, 1/(n+1)$. Then

$$-DU_1D = \begin{bmatrix} 2 & -1 & 0 & \cdots & 0 & 0 \\ -1 & 2 & -1 & \cdots & 0 & 0 \\ 0 & -1 & 2 & \cdots & 0 & 0 \\ \vdots & & \ddots & \ddots & & \\ 0 & 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & -1 & 2 \end{bmatrix},$$

which is positive definite. It follows that $-U_1$ is also positive definite and hence an M-matrix. Thus all of the entries of U_1^{-1} are negative. (See [2], Theorem 2.5.3.)

$$Pr = PT^{-1}Tr = \begin{bmatrix} U_1^{-1} & w \\ 0_{m-1} & 1 \end{bmatrix} \begin{bmatrix} +_{m-1} \\ t \end{bmatrix} \leq t \begin{bmatrix} w \\ 1 \end{bmatrix}.$$

But

$$Ps = PT^{-1}Ts = \begin{bmatrix} U_1^{-1} & w \\ 0_{m-1} & 1 \end{bmatrix} \begin{bmatrix} 0_{m-1} \\ t \end{bmatrix} = t \begin{bmatrix} w \\ 1 \end{bmatrix}.$$

Thus, $Pr \le Ps$. □

6.7. **Example 5.1.** First we show that $z = (-18, -18, -12, 24, 24)$ is the least element in $S(0, 1944/5) \cap [-36, 24]^5$ with respect to the variance majorization order. We will construct the coordinates of $z = b_0 w_0 + b_1 w_1 + b_2 w_2 + b_3 w_3 + b_4 w_4$ in the Helmert basis, w_0, w_1, w_2, w_3, w_4.

The mean of z is 0 so $b_0 = 0$. To compute the other coordinates we first determine the index j, which is defined in Equation (6.13):

$$5 \cdot 24^2 \left[\frac{1}{j+1} - \frac{1}{5}\right] \le \frac{1944}{5} \le 5 \cdot 24^2 \left[\frac{1}{j} - \frac{1}{5}\right],$$

so $j = 2$. Now from Equations (6.14) we have:

$$\begin{aligned} b_1 &= 0, \\ b_2^2 &= 1944 - 5^2 \cdot 24^2 \left[\frac{1}{3} - \frac{1}{5}\right] = 24, \\ b_3^2 &= 5^2 \cdot 24^2 \left[\frac{1}{3} - \frac{1}{5}\right] = 1200, \\ b_4^2 &= 5^2 \cdot 24^2 \left[\frac{1}{3} - \frac{1}{4}\right] = 720. \end{aligned}$$

Then

$$\begin{aligned} z &= \sqrt{24} w_2 + \sqrt{1200} w_3 + \sqrt{720} w_4 \\ &= 2(-1, -1, 2, 0, 0) + 10(-1, -1, -1, 3, 0) + 6(-1, -1, -1, -1, 4) \\ &= (-18, -18, -12, 24, 24). \end{aligned}$$

Next we show that $S(0, 1944/5) \cap [-36, 24]^5$ has no greatest element. Suppose there is an element $w \in S(0, 1944/5)$ such that $x \overset{\text{vm}}{\prec} w$ for all $x \in S(0, 1944/5) \cap [-36, 24]^5$. We will prove that $w \notin [-36, 24]^5$.

The vectors $x = (-36, 0, 0, 18, 18)$ and $y = (-36, 0, 6, 6, 24)$ are both in $S(0, 1944/5) \cap [-36, 24]^5$. So $x, y \overset{\text{vm}}{\prec} w$. To express x, y in terms of their coordinates in the Helmert basis for $\mathbb{R}^5$, let H be the 5×5 matrix whose columns are the Helmert basis for $\mathbb{R}^5$:

$$H = \begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{12}} & -\frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{12}} & -\frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & 0 & \frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{12}} & -\frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & 0 & 0 & \frac{3}{\sqrt{12}} & -\frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & 0 & 0 & 0 & \frac{4}{\sqrt{20}} \end{bmatrix}.$$

Then the coordinate vectors of x, y with respect to the Helmert basis are

$$\begin{aligned} a &= xH = (0, 18\sqrt{2}, 6\sqrt{6}, 15\sqrt{3}, 9\sqrt{5}) \\ b &= yH = (0, 18\sqrt{2}, 8\sqrt{6}, 8\sqrt{3}, 12\sqrt{5}). \end{aligned}$$

Thus

$$a^{(2)} = (0, 648, 216, 675, 405)$$
$$b^{(2)} = (0, 648, 384, 192, 720).$$

The partial sums of $a^{(2)}$ and $b^{(2)}$ are

$$\begin{array}{l} a^{(2)} : (0,\ 648,\ \ 864,\ 1539,\ 1944) \\ b^{(2)} : (0,\ 648,\ 1032,\ 1224,\ 1944) \end{array}.$$

The maximum of each of the five partial sums is

$$(0, 648, 1032, 1539, 1944)$$

and the sequence with these partial sums is

$$c^{(2)} = (0, 648, 384, 507, 405).$$

Thus the least upper bound of x, y is

$$cH^T = (-37, -1, 5, 15, 18).$$

Now let w be a sequence in $S(0, 1944/5)$ such that $v \overset{\text{vm}}{\prec} w$ for all $v \in S(0, 1944/5)$. Express w in terms of the Helmert basis, $w = \sum d_i w_i$, for nonnegative d_i. Then

$$\sum_{k=1}^{i} a_k^2 \leq \sum_{k=1}^{i} d_k^2 \text{ and } \sum_{k=1}^{i} b_k^2 \leq \sum_{k=1}^{i} d_k^2,$$

for $i = 1, \ldots, n-1$ so that

$$\sum_{k=1}^{i} c_k^2 \leq \sum_{k=1}^{i} d_k^2,$$

for $i = 1, \ldots, n-1$. Thus $(-37, -1, 5, 15, 18) \overset{\text{vm}}{\prec} w$ and by Lemma 4.7, $w_1 \leq -37$. So $w \notin [-36, 24]^5$.

REFERENCES

[1] J.H.E. COHN, Determinants with elements ± 1, *J. London Math. Soc.*, **42** (1967), 436–442.

[2] R.A. HORN AND C.R. JOHNSON, *Topics in Matrix Analysis*. Cambridge, 1991.

[3] G.H. HARDY, J.E. LITTLEWOOD AND G. PÓLYA, *Inequalities*, Cambridge, 1967.

[4] J. KARAMATA, Sur une inégalité rélative aux fonctions convexes, *Publ. Math. Univ. Belgrade,* **1** (1932) 145–148.

[5] J.H.B. KEMPERMAN, Momentproblems for sampling without replacement, I,II,III. *Nederl. Akad. Wetensch. Proc. Ser. A,* **75** (1973), 149–188.

[6] A.W. MARSHALL AND I. OLKIN, *Inequalities: Theory of Majorization and Its Applications*, Academic Press, 1979.

[7] I. SCHUR, Über eine Klasse von Mittelbildungen mit Anwendungen auf die Determinantentheorie, *Sitzungsber. Berlin Math. Gesellschaft,* **22** (1923), 9–20. (*Issai Schur Collected Works* (A. Brauer and H. Rohrbach, eds.) Vol. II. Springer-Verlag, Berlin, 1973, 416–427.)

[8] H. WOLKOWICZ AND G. STYAN, Bounds on eigenvalues using traces, *Linear Algebra Appl.,* **29** (1980) 471–506.