Comments on and Complements to
Inequalities: Theory of Majorization and Its Applications
by Albert W. Marshall and Ingram Olkin*

James V. Bondar
Department of Mathematics and Statistics
Carleton University
1125 Colonel By Drive
Ottawa, Ontario, Canada K1S 5B6

Submitted by George P. H. Styan

Dedicated to Ingram Olkin

ABSTRACT

In the years since the appearance of Marshall and Olkin's masterful survey, the influence of this work has been so great that some comments (however incomplete and idiosyncratic) on subsequent developments will be of value (Figure 1). An appendix lists minor errata to the first printing of the book.

1. INTRODUCTION

The appearance of Marshall and Olkin's *Inequalities* in 1979 had great impact on the mathematical sciences. By showing how a single concept unified a staggering amount of material from widely diverse disciplines—probability, linear algebra, geometry, statistics, operations research, etc.—this work was a revelation to those of us who had been each trying to make sense of his own corner of this material.

The *Science Citation Index* gives the following statistics: the number of citations to this book in 1982 was 26, by 1987 it grew to 39 citations in that

*Academic Press, New York, 1979.

655 Avenue of the Americas, New York, NY 10010 0024-3795/94/$7.00

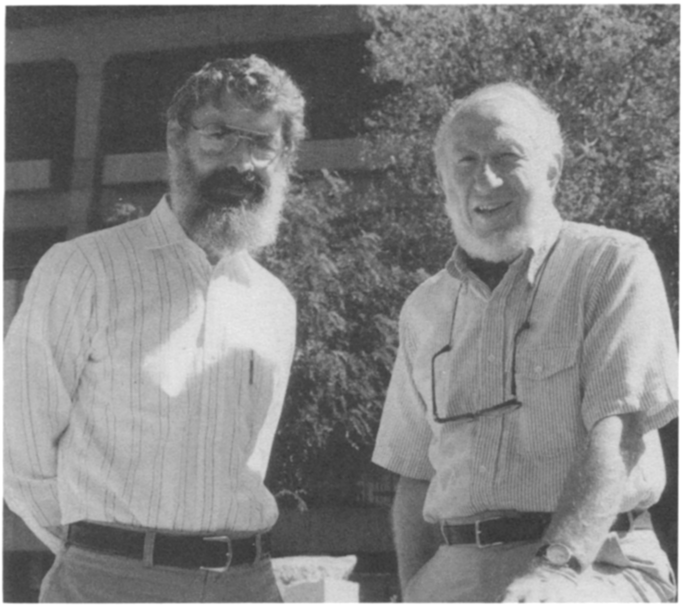

FIG. 1. Ingram Olkin (right) with coauthor A.W. Marshall (Stanford University, 1 October 1993).

year, and in 1992 it was 59. By Simpson's rule we get 570 citations; the growth rate is 8.5% per year. It is obviously impossible to do a complete survey of majorization since the publication of Marshall and Olkin. However, something of the sort is needed, because this book will remain *the* authority for the foreseeable future (for one thing, it is hard to imagine any authors with the courage and *Sitzfleisch* to attempt a project of this scale again). Hence, this paper has two goals: to give readers of Marshall and Olkin a start in pursuing some of the directions in which majorization theory has gone since publication of their book, and to pay tribute to Albert Marshall and Ingram Olkin for their dedication and insights.

2. COMMENTS

In this section, theorems and displays referring to a given chapter of Marshall and Olkin are numbered with that chapter number to the left of the decimal point.

Section 1.B. Majorization as a Partial Ordering

As remarked on p. 13 of this section, the strong and weak majorizations are partial orderings on $\mathscr{D} = \{z : z_1 \geqslant \cdots \geqslant z_n\}$, but more than this, $\mathscr{D}$ is actually a complete lattice under the weak orderings. To be precise, any $S \subset \mathscr{D}$ has a least upper bound x in $\mathscr{D}_+$ i.e., if $s \prec_w x$ for all $s \in S$, and for any y in $\mathscr{D}$ one has $s \prec_W y$ for all $s \in S$, then $x \prec_W y$. Furthermore, x is the unique least upper bound in $\mathscr{D}_+$. Similarly, S has a unique greatest lower bound in $\mathscr{D}_+$. The same result holds for the weak majorization $\prec^W$. See Bapat (1991), where this is proved for $\prec_W$ in Lemma 3 and Corollary 4; the result for $\prec^W$ follows by the relation between the two weak majorizations. For strong majorization, the same result holds, provided S lies in the hyperplane $\Sigma\, x_i = \text{const}$.

Section 1.D. Generalizations of Majorization

H. Joe has written extensively on a generalization to matrices. Whereas majorization applies to vectors with fixed totals, Joe's generalization applies to matrices with fixed row and column totals. An application is the ordering of contingency tables by degree of dependence of one variable on the other. A useful introduction is Joe (1993). A different form of multivariate majorization is found in Joe and Verducci (1993).

Section 2.D. Doubly Superstochastic Matrices

One of the open questions on p. 31 has been answered: $yP \prec^W y$ for all $y \in R_+^n$ does indeed imply that P is doubly superstochastic (Ando, 1989, Corollary 3.4). Ando's paper is recommended for its treatment of the relation between majorization and stochastic linear transformations, among many other things. Section 1 of this paper is also of interest, since it covers some material which supplements Marshall and Olkin's coverage of the basic theory of majorization (Chapters 1 and 2).

Section 2.F

Page 37: The van der Waerden permanent conjecture was proven to be true in 1981 by Egorychev. A self-contained treatment of Egorychev's proof is found in Section 5 of Ando (1989).

Section 3.G. Muirhead's Theorem

Here is an immediate consequence of Muirhead's theorem which appears not to have been widely noted:

COROLLARY 3.1. *For given real a, the symmetrized sum* $f(y) = \sum_\pi y_{\pi(1)}^{a_1} \cdots y_{\pi(n)}^{a_n}$ *is log-Schur-convex as a function of* $(y_1, \ldots, y_n) \in \mathscr{D}$, *i.e., if* $x_1 \geqslant \cdots \geqslant x_n \geqslant 0$, $y_1 \geqslant \cdots \geqslant y_n \geqslant 0$, *and also*

$$\prod_{i=1}^{k} x_i \leqslant \prod_{i=1}^{k} y_i \qquad \text{for} \quad k = 1, \ldots, n, \tag{3.1}$$

with equality for $k = n$, *then* $f(x) \leqslant f(y)$.

Proof. Let $u_i = \ln y_i$. Then

$$\begin{aligned} f(y) &= \sum_\pi e^{a_1 u_{\pi(1)}} \cdots e^{a_n u_{\pi(n)}} \\ &= \sum_\pi \left(e^{a_{\pi(1)}}\right)^{u_1} \cdots \left(e^{a_{\pi(n)}}\right)^{u_n}. \end{aligned}$$

Now, by Muirhead's theorem, this last expression is Schur-convex as a function of $(u_1, \ldots, u_n)$; hence $\sum_1^k \ln x_i \leqslant \sum_1^k \ln y_i$ and $\sum_1^n \ln x_i = \sum_1^n \ln y_i$ [which is equivalent to (3.1)] implies $f(x) \leqslant f(y)$.

Section 3.J. Integral Transformations Preserving Schur Convexity

See Shaked and Shanthikumar (1988) for new results and applications; also Liyanage and Shanthikumar (1993). These papers are also of interest in connection with several portions of chapters 11 and 12, in particular, section 11.E (families of distributions parameterized to preserve Schur-convexity).

Section 5.A

Page 116, Theorem A.1.e: Both majorizations in the displayed expression should be strong ("$\prec$") and not weak ("$\prec^W$"). That is to say,

$$x \prec y \quad \Rightarrow \quad (g(x_1), \ldots, g(x_n)) \prec (g(y_1), \ldots, g(y_n))$$

if and only if g is linear.

If the second displayed majorization in this theorem is "$\prec^W$" (as it is in the first printing), then this displayed condition holds if and only if g is concave [this follows from A.1(2) and A.2, Theorem (ii); note the erratum in the appendix for this part of A.2].

As a corollary of A.2, we get the useful formula

$$x \prec_W y \quad \Rightarrow \quad (e^{x_1}, \ldots, e^{x_n}) \prec_W (e^{y_1}, \ldots, e^{y_n}). \tag{5.1}$$

Section 8.B. Inequalities for the Sides of a Triangle

If a_1, a_2, a_3 are the sides of a triangle with semiperimeter $s = \frac{1}{2}(a_1 + a_2 + a_3)$ and average side length $\bar{a} = \frac{1}{3}(a_1 + a_2 + a_3) = 2s/3$, then

$$(\bar{a}, \bar{a}, \bar{a}) \prec (a_1, a_2, a_3) \prec (s, s, 0) \qquad \text{for all triangles,} \quad (8.1a)$$

$$(\bar{a}, \bar{a}, \bar{a}) \prec (a_1, a_2, a_3) \prec (s, s/2, s/2) \quad \text{for isosceles triangles,} \quad (8.1b)$$

$$\frac{s}{1+\sqrt{2}}(2, \sqrt{2}, \sqrt{2}) \prec (a_1, a_2, a_3) \prec (s, s, 0) \qquad \text{for obtuse triangles.} \quad (8.1c)$$

Thus, if φ is a continuous Schur-convex function, we have

$$\varphi(\bar{a}, \bar{a}, \bar{a}) \leqslant \varphi(a_1, a_2, a_3) < \varphi(s, s, 0) \qquad \text{for all triangles} \quad (8.2a)$$

$$\varphi(\bar{a}, \bar{a}, \bar{a}) \leqslant \varphi(a_1, a_2, a_3) < \varphi(s, s/2, s/2) \qquad \text{for isosceles triangles,} \quad (8.2b)$$

$$\varphi\left(\frac{2s}{1+\sqrt{2}}, \frac{\sqrt{2}\,s}{1+\sqrt{2}}, \frac{\sqrt{2}\,s}{1+\sqrt{2}}\right) < \varphi(a_1, a_2, a_3) < \varphi(s, s, 0) \qquad \text{for obtuse triangles,} \quad (8.2c)$$

and these inequalities are best possible. For Schur-concave functions, the inequalities are reversed. The corresponding majorization for acute triangles is identical to (8.1a), so specializing to acute triangles gives the same inequality (8.2a) as one gets for all triangles. For right-angled triangles, the majorization is identical to (8.1c), yielding inequalities identical to the ones for obtuse triangles, namely (8.2c).

Let $s_i = 2(s - a_i)$, $i = 1, 2, 3$, then

$$\tfrac{2}{3}(s, s, s) \prec (s_1, s_2, s_3) \prec (2s, 0, 0) \qquad \text{for all triangles,} \tag{8.3a}$$

$$\tfrac{2}{3}(s, s, s) \prec (s_1, s_2, s_3) \prec (s, s, 0) \qquad \text{for isosceles triangles,} \tag{8.3b}$$

$$2s(\sqrt{2} - 1)(1, 1, \sqrt{2} - 1) \prec (s_1, s_2, s_3) \prec (2s, 0, 0) \qquad \text{for obtuse triangles.} \tag{8.3c}$$

Regrettably, a slip in the first printing labelled the majorizations (8.1b) and (8.3b) as being for obtuse triangles, which means that on pp. 199–201 the inequalities claimed for obtuse triangles are actually for isosceles triangles. As an example we look at Section 8.B.1, which should read as follows:

B.1. The inequalities

(i) $$\frac{1}{3} \leqslant \frac{a_1^2 + a_2^2 + a_3^2}{(a_1 + a_2 + a_3)^2} < \frac{1}{2} \qquad \text{for all triangles,}$$

(ii) $$\frac{1}{3} \leqslant \frac{a_1^2 + a_2^2 + a_3^2}{(a_1 + a_2 + a_3)^2} < \frac{3}{8} \qquad \text{for isosceles triangles,}$$

(iii)

$$0.343146 = \frac{2}{(1 + \sqrt{2})^2} \leqslant \frac{a_1^2 + a_2^2 + a_3^2}{(a_1 + a_2 + a_3)^2} \leqslant \frac{1}{2} \qquad \text{for obtuse triangles}$$

follow from the Schur-convexity of the middle term. Inequality (i) is G.I. 1.19 and is attributed to Petrović, 1916.

Similar modifications must be made to B.2, B.3, B.4, and B.5. The inequalities of B.6 and B.7 follow from the majorizations (3a) and (3b); in each case the second inequality is for isosceles rather than obtuse triangles, and the obtuse triangle inequality comes directly from (3c).

Chapter 9. Matrix Theory

A useful recent survey is Chapter 3 of Horn and Johnson (1986), on inequalities for singular values and also for norms. See also Ando (1989). Many of the majorization results for eigenvalues and singular values and their logarithms in sections 11.E to 11.H generalize to compact operators in (complex) Hilbert space. There is also a discussion of functions which are increasing in the weak Schur ordering for infinite vectors. See Gohberg and Krein (1969) chapter 2 for details.

Section 9.B

An interesting generalization of Schur's inequality, Theorem B.1, is in Andersson and Perlman (1988): if G is any group of matrices, and B is the center of gravity of the points $\{gAg^{-1} : g \in G\}$, then $\lambda(B) \prec \lambda(A)$ by Fan's Theorem 9.G.1. This is used to derive inequalities for matrices with certain symmetries. If G is the group which reverses the signs of subsets of the coordinates in R^n, then we get Schur's inequality.

Section 9.H.1.g

The following note has been added in the second printing. "The authors are grateful to Prof. L. Mirsky for the following. The result (7) holds for any Hermitian matrices U and V. They need not be positive semidefinite, for we can replace U and V by $U + \tau I$ and $V + \tau I$, respectively, with τ sufficiently large. In the new version of (7), replacement of V by $-V$ yields (8). Indeed, (7) and (8) are really identical results."

Section 9.I. Schur or Hadamard Products

A great deal of work has recently been done on eigenvalue and singular value inequalities for Hadamard products of matrices; no small amount of this work was inspired by the conjectures proposed by Marshall and Olkin in this section of their book. A good survey is Sections 5.3 to 5.6 of Horn and Johnson (1986). The reader is warned that Horn and Johnson number eigenvalues λ_i and singular values σ_i in *increasing* order, the opposite of Marshall and Olkin.

Marshall and Olkin's conjecture (15) on p. 258 has now been proved [Bapat and Sunder (1985), or see Horn and Johnson (1986, 5.3.3b)]. In fact, Bapat and Sunder prove even more than the conjecture:

THEOREM 9.1. *Let A and B be nonnegative definite (hence Hermitian) $n \times n$ matrices, and β be the vector of diagonal elements of B (arranged in* *decreasing order). Then*

(a) $$\lambda(A \circ B) \prec_W \lambda(A) \circ \beta \prec_W \lambda(A) \circ \lambda(B),$$

(b) $$\prod_{i=k}^{n} \lambda_i(A \circ B) \geqslant \prod_{i=k}^{n} \lambda_i(A)\beta_i, \qquad 1 \leqslant k \leqslant n.$$

The Schur product theorem (that $A \circ B$ is also nonnegative definite) follows as a corollary of (b). If A is the identity, we get Schur's inequality (9.B.1), and Oppenheim's inequality is also a corollary of Theorem 9.1.

Horn and Johnson (5.3.3.a) conjecture that

$$\prod_{i=k}^{n} \lambda_i(AB) \leqslant \prod_{i=k}^{n} \lambda_i(A \circ B), \qquad k = 1, \dots, n \tag{9.1}$$

for positive semidefinite A and B. (For $k = 1$ and n, this follows from Theorem 9.1.)

The conjectures (17) and (18) on p. 259 of Marshall and Olkin appear to be as yet unresolved.

Corresponding to Theorem 9.1a is a result for singular values:

THEOREM 9.3. *If A and B are $n \times n$ matrices, then*

$$\sigma(A \circ B) \prec_W \sigma(A) \circ \sigma(B). \tag{9.2}$$

For a proof, see Bapat (1987, Corollary 4); the result is due to Horn and Johnson and to Okubo. If B is equal to the identity, we get a generalization of Schur's inequality: $(|a_{11}|, \dots, |a_{nn}|) \prec_W \sigma(A)$.

From Theorem 9.1a, we get:

COROLLARY 1. *If $a = [a_{ij}]$ is a nonnegative definite $n \times n$ matrix, and f is any analytic function whose Taylor series has nonnegative coefficients, then*

$$\lambda\big([f(a_{ij})]\big) \prec_W \lambda(f(A)). \tag{9.3}$$

(Proof: By Theorem 9.1a, $\lambda(A \circ \cdots \circ A) \prec_W \lambda(A) \circ \cdots \circ \lambda(A) = (\lambda(A))^k = \lambda(A^k)$. Now use the fact that weak majorization is closed under summation, positive scalar multiplication, and limits.) As an example, if f is the exponential function and A is nonnegative definite, we get

$$\lambda([e^{a_{ij}}]) \prec_W \lambda(e^A). \tag{9.4}$$

A special Hadamard product is the relative gain array $A \circ (A^T)^{-1}$ used in engineering process control. If A is positive definite, then the eigenvalues of the relative gain array are all $\geqslant 1$ (Horn and Johnson, 1986, 5.4).

Section 14.B. G-Majorization

Further theory of G-majorization and some statistical applications are given in Giovagnoli and Wynn (1985). The excellent surveys of Eaton (1982, 1987) contain generalizations of Theorem B.2. This theory has been greatly developed since 1979.

From a group-theoretic point of view, G-majorization is *the* natural generalization of majorization, and it has already been of great use in statistics. We use the definition of G-ordering on p. 422, and call a nonnegative function f on R^n G-decreasing iff $f(x) \leqslant f(y)$ whenever $y \prec_G x$. If G is the group of coordinate permutations, then these are precisely the Schur-concave positive functions. The convolution theorem 3.J.1 was strengthened for finite reflection groups by Eaton and Perlman (1977):

THEOREM 14.1. *Let G be a finite reflection group, and f_1 and f_2 be G-decreasing. Then the convolution $f_1 \bigstar f_2$ defined by*

$$f_1 \bigstar f_2(\theta) = \int_{R^n} f_1(\theta - x) f_2(x)\, dx$$

is a G-decreasing function on R^n (whenever the integral exists).

For further discussion and applications, see Eaton (1982, 1987). Suren Fernando (to be published) has recently shown the converse: if $f_1 \bigstar f_2$ is G-decreasing for every pair of G-decreasing functions, then G must be a finite reflection group (also known as a Coxeter group).

Chapter 15. Much work has recently been done on forms of multivariate majorization. See Karlin and Rinott (1983) and a series of important papers by these authors, studying the version of majorization defined by doubly stochastic matrices (Marshall and Olkin, Definition 15.A.2). An important question is the nonequivalence of the definition of multivariate majorization using T-transforms (Definition 15.A.1) and that using doubly stochastic matrices S: if every S is a *limit* of products of T-transforms (i.e. if the products of the T-transforms are a dense subset of the doubly stochastic matrices), then the class of *continuous* functions which is increasing in one majorization order will be increasing in the other. This is now known to be false (Berg, 1984, p. 251); hence the ordering using products of T-transforms is strictly stronger than the order using doubly stochastic matrices.

A new version of multivariate majorization, designed for studying resource allocation problems, is Joe and Verducci (1993). In this work, $X \prec Y$ if $x'X \prec x'Y$ for all positive vectors x. If each row is considered as a distribution of a resource over different organizations (a different row corresponds to and a different resource), then x is a vector of costs. The authors study convexity properties and an algorithm for determining when a given matrix majorizes another.

Sections 16.E7 and 16.E8

Pages 468–473: Several of the comments on matrix convexity of powers in these sections have been rendered obsolete by Ando (1979); this is Ando's Corollary 4.1 on p. 215:

THEOREM 16.1. *The function $\varphi(A) = A^p$ on the positive definite symmetric matrices is convex if $1 \leqslant p \leqslant 2$ or $-1 \leqslant p \leqslant 0$, and concave if $0 \leqslant p \leqslant 1$.*

By Shorrock and Rizvi's paper (referenced in Marshall and Olkin), φ is neither concave nor convex for other values of p. A more readable treatment of this theorem than Ando's is that of R. Farrell (1985, Theorem 14.5.16 and the rest of Section 14.5). A survey of related questions is Kwong (1989).

Chapter 18. Total Positivity

A recent survey is Ando (1987).

3. COMPLEMENTS

The dominant eigenvalue of nonnegative matrices. If A is an $n \times n$ nonnegative matrix (i.e., $a_{ij} \geqslant 0$ for all i, j), then it is well known that the eigenvalue of maximum modulus is real and nonnegative; it is called the dominant or Perron eigenvalue or spectral radius $r(A)$. Now consider a function from some domain in R^m into the class of $n \times n$ nonnegative matrices, which function we shall denote $x \to A(x)$. Each entry $a_{ij}(x)$ can be considered as a positive real valued function of x. The following is more or less found in Kingman (1961):

THEOREM 1. *Let S be a class of positive real valued functions, closed under addition, multiplication, raising to any positive power, and the taking of pointwise lim sups of countable subsequences. If $x \to A(x)$ is a matrix valued function whose entries $x \to a_{ij}(x)$ are functions in S, then the spectral radius function $x \to r(A(x))$ is also a member of S.*

Kingman's proof is stunningly neat. Let $f_p(x) = [\mathrm{tr}\, A^p(x)]^{1/p}$. Then by our hypotheses, $x \to f_p(x)$ is also in S, $p = 1, 2, \ldots$. Furthermore $f_p(x)$ is the l_p norm of $\lambda(A(x))$; hence $r(A(x)) = \|\lambda(A(x))\|_\infty = \limsup_p f_p(x)$ lies in S. Three corollaries come immediately to mind:

COROLLARY 1. *If the functions $x \to a_{ij}(x)$ are nonnegative increasing functions from R^m to R, then $r(A(x))$ is an increasing function of x.*

COROLLARY 2. *If the functions $x \to a_{ij}(x)$ are log-convex or identically zero, then so is $r(A(x))$.*

That these functions satisfy the hypotheses of the theorem is proven in Kingman (1961).

COROLLARY 3. *If the functions* $x \to a_{ij}(x)$ *are nonnegative and Schur-convex, then* $r(A(x))$ *is Schur-convex.*

That the hypotheses hold for these functions is easily checked.

Corollary 1 is classical, but Corollary 3 appears to be new.

An example of Corollary 2 at work is the following (Horn and Johnson, 1986, 5.7.7): Let A_i be a nonnegative matrix, $i = 1, \dots, k$, and suppose that $a_i \geqslant 0$, $i = 1, \dots, k$, satisfy $a_1 + \cdots + a_k \geqslant 1$. Then

$$r\left(A_1^{(a_1)} \circ \cdots \circ A_k^{(a_k)}\right) \leqslant r(A_1)^{a_1} \cdots r(A_k)^{a_k}, \tag{1}$$

where $A^{(a)}$ is the matrix whose ij entry is $(a_{ij})^a$.

What is the real nature of majorization, and why is it so important? Kemperman (1981) discussed this question at length from the point of view of cone orderings: these are important objects, and majorization is the ordering defined by the cone of convex functions (which are themselves important). In a nice argument, Kemperman shows that the equivalence of the cone ordering definition of majorization and the doubly stochastic matrix definition follows from a theorem of Cartier in potential theory (in potential theory cone orderings are called sweeping orderings or balayage orderings). It is clear that a monograph setting majorization in the context of cone orderings would be of great value. A first step to such a theory is Marshall and Olkin's Section 14.C. Further material on cone orderings is in Dellacherie and Meyer (1988) chapter X, section 2.

Thompson (1983) focuses on the relation with group representation theory. A fascinating survey by Hazewinkel and Martin (1983) brings together a lot of connections with group representations, multilinear algebra, physics, etc. and is highly recommended. Hazewinkel calls majorization "the specialization order" and seems unaware of the classical literature in inequalities. More connections with group representations can be glimpsed in Eaton and Perlman (1977), Zobin and Zobina (1993), and Kostant (1973). A survey of these connections would be of great value to the profession. Anyone want to volunteer?

While we are at it, there are connections with control theory [Rosenbock (1970, p. 190) or Dickson (1974)]. Alberti and Uhlmann (1982) apply majorization to physics and physical chemistry. In their terminology, $a \prec b$ is read "a is more chaotic than b," and the ordering is generalized to dual spaces of C^* algebras to study solutions of evolution equations (e.g. the Boltzman equation) in state spaces. Warning: the inequality which Marshall and Olkin write as $a \succ b$ is reversed to read $b \prec a$ Alberti and Uhlmann.

4. APPENDIX: ERRATA

These errata are for the first printing, which can be identified by the bottom line of the back of the title page, which reads "79 80 81 82 9 8 7 6 5 4 3 2 1." Many of these errata were corrected in the second printing. The later printings (which have a "2" or a "3" at the end of the bottom line of the back of the title page) still have errata on the following pages: 49, 79, 139 (in A.1.a), 141, 196, 216, 240, 242, 243, 246, 249, 273, 293, 305, 357, 373, 377, 485, 510, 566.

The authors have taken pains to ensure that all printings have the same pagination, so the page numbers are the same in all copies. (Negative line numbers are counted from the bottom.)

An anonymous FTP has been set up to record the latest list of errata to the book, and will be periodically updated. Contributions are welcomed by the author (Bondar). To get the errata file from this FTP by e-mail, send: "ftp alfred.ccs.carleton.ca". When asked for name, send "anonymous". When asked for password, send your e-mail address (e.g., mine is jbondar@carleton.ca), then change directories at the ftp prompt with "cd/pub/math/olkin get errata.tex". At the end, "quit".

P. 15, line 9: Read 14.B for 14.C.

P. 16, line above display (2): add ",j" before equals sign.

p. 17, top display: Multiply $\Sigma\phi(y_i)$ by $1/n$.

p. 49, Theorem I.2. Replace the ρ's by r's.

P. 73, 4 lines above section E: product in denominator is over $j \neq i$, which is $n-1$ factors and not n.

P. 74, second display: left hand side is $\phi(r)$.

P. 79, line 1: Use (10) of A.4, not (12).

P. 86, display (5): For $\binom{k}{n}$, read $\left(\dfrac{k}{n}\right)$.

P. 87, G.2.g: If $y_i \geqslant 1$, then (7) holds if $a \prec_W b$, but not in general for $a \prec^W b$. The statement given for $y_i \leqslant 1$ is correct.

P. 91, 3 lines above section H: second inequality should be "$x \succ (1/n, \ldots, 1/n)$".

P. 116, A.2, Theorem (ii): both majorizations are weak supermajorizations, i.e., they are both $\prec^W$.

P. 117, A.2.c: change $\mathscr{R}_{++}$ to $\mathscr{R}_{+}$. In last display, the last part should be changed to "$k = 1, \ldots, n$."

P. 121, last line of display in A.4.e: change $\prec$ to $\prec_W$.

P. 139, statement of Proposition A.1: The second inequality should read $\sum_{i=1}^m x^{(i)} \prec \sum_{i=1}^m x_{\downarrow}^{(i)}$.

P. 139, Proposition A.1.a: The condition "provided all integrals are finite" should be added to the statement. Both integrals in the display are with respect to $\mu(d\alpha)$, and not $\mu(dx)$. The first line of the proof should be changed to: "Let π be a permutation-valued function of α (to avoid complex formulae, we will write this as π, and not π_α) such that for each α, $x_{\pi(j)}^{(\alpha)}$ is decreasing in $j = 1, \ldots, n$." In the first integral of the display, remove the subscript $\pi(j)$ (but leave this subscript in the second integral).

P. 141, A.3.a: Display (2a) holds only under some conditions on the signs of the a_i's and b_i's. This is shown by taking $(a_1, a_2) = (-3, -2)$ and $(b_1, b_2) = (2, 1)$. It is sufficient to require that all a_i and b_i be $\geqslant 0$. A more complicated argument shows it is sufficient for the number of negative a's to equal to the number of negative b's.

P. 196, beginning of section on "The Cosine Function": The function $\log\cos kx$ is strictly concave in $(0, \pi/2k)$, not in $(0, \pi)$ nor in $(0, k\pi/2)$.

P. 197, sentence after top display: change (i) and (ii) to (i) to (iv). Next sentence: change (iii) to (v). Second last line of A.13: insert $\sin\alpha_3$ in numerator of formula.

P. 216, line 2 of A.1.a: The product at the end of the line should be "BA" not "AB."

P. 216, line -12: instead of "For an $n \times n$ matrix H," read "For an $n \times n$ Hermitian matrix H." Also, if some eigenvalues of H are negative, then the $\sigma_i(H)$ may be in a different order from the $|\lambda_i(H)|$, so the vector $\sigma(H)$ is equal to $|\lambda(H)|$ only up to rearrangement. An example is

$$H = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}.$$

P. 226, line 5: change θH_{12} to θH_{21}.

P. 232, section E.1.b: change $\prec$ to $\prec_w$.

P. 238, line -5: change the first α_i to ω_i.

P. 240, display (5) is incorrect. $|\lambda((A + A^*)/2)|$ need not be $\leqslant \sigma(A)$; a counterexample is

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}.$$

However, F.4.a is true, and the proof in G.1.f. is correct.

P. 242, statement of Theorem G.1.b: The inequality (5) must be reversed. The last two lines of the proof should be changed to: "...i.e., $\lambda(A^*A) = (\lambda(AA^*), 0)$. Using C.1,

$$\begin{aligned}(\lambda(G), \lambda(H)) &= (\lambda(XX^*), \lambda(YY^*)) \prec \lambda(A^*A) = (\lambda(AA^*), 0)\\ &= (\lambda(G + H), 0)."\end{aligned}$$

P. 243, G.1.3: This result is false; a counterexample is $A = -B =$ identity matrix.

P. 246, G.4.b: This result is false; same counterexample as for p. 243. The problem with the proof is that min R tr UBV^* can be as low as $-\Sigma\sigma_{n-i+1}(B)$. Thus, the inequality may be repaired by replacing the plus sign on the right hand of display (10) by a minus sign.

P. 249, proof of H.1.i: The last inequality is an application of (8), not (9).

P. 250, last two lines of display: change U_m to U whenever it appears.

P. 252, statement of Theorem H.4.a: For $\prec$, read $\prec_W$. The strong majorization ($\prec$) holds only if A and B commute, as is clear from H.4.b.

P. 269: The third equality in display (13) should read $(AA^+)^* = AA^+$.

P. 270, top display: the norm on the right side of the inequality has an E-norm, not an I-norm.

P. 273, line 17: Change (6) to (7).

P. 290, first display: insert "ϕ" after "E" in the right side.

P. 293, Theorem C.3 is incorrect. Rolski (1985) shows convexity of ψ fails if ϕ is the max function. The argument in the proof fails to hold in an integer lattice.

P. 297, line below display (2): The function is ϕ and not θ.

P. 300, second last line of E.5.C: change "chi-square" to "non-central chi-square".

P. 302, second display: change $E_\lambda f$ to $E_\lambda \phi$

P. 305, last display: $P(l) = P_\theta\{K \leqslant l - 1\}$.

P. 357, line -5: In the right-hand side of the display, $\sigma_1(A_i)$ and $\sigma_n(A_i)$ should be replaced by $\sigma_1(EA_i)$ and $\sigma_n(EA_i)$.

P. 373, Theorem J.1: The proof for $n = 2$ is correct; however, the proof given for $n \geqslant 3$ won't work for $t < 0$. The proof in Proschan (1965) can be used instead.

P. 377, statement of Theorem K.3: Bock et al. (1987) point out that this is stated incorrectly. The error may be repaired by interchanging "convex" and "concave," hence S-convex for *small* t, and S-concave for *large* t. The reader is referred to Bock et al. (1987), where Theorem K.3 is generalized to Gamma(α, β) densities, done for Weibull densities, and partially extended to sums of more than two gamma variables.

P. 468, Proposition E.6.a, statement (ii): Change "β" to "B" in the equation $g(\alpha) = \varphi(\alpha A + \bar{\alpha}\beta)$.

P. 485, third line of Proposition C.1: j goes from 2 to n, not $n - 1$.

P. 510, display at bottom: The subscript on the right of the inequality should be "$n - i + 1$," not "$l - i + 1$."

P. 566: The page reference for "Norm, matrix, consistent with vector norm" should be 274, not 276. The reference for "Norm, unitarily invariant" should be 263, not 236.

P. 568: The page reference for "Symmetric gauge function and unitarily invariant norms" should be 263, not 236.

I wish to thank Chandler Davis, Morris Eaton, Harry Joe, A. W. Marshall, and Y. L. Tong for assistance. Photography by George Styan. Michael Perlman has made many valuable comments.

REFERENCES

Alberti, P. M. and Uhlmann, A. 1982. *Stochasticity and Partial Order*, Deutsche Verlag Wiss., Berlin. Second edn. published by Reidel, Dordrecht.

Andersson, S. and Perlman, M. 1988. Group invariant analogues of Hadamard's inequality, *Linear Algebra Appl.* 110:91.

Ando, T. 1979. Concavity of certain maps on positive definite matrices and applications to Hadamard products, *Linear Algebra Appl.* 26:203–241.

Ando, T. 1987. Totally positive matrices, *Linear Algebra Appl.* 90:165–219.

Ando, T. 1989. Majorization, doubly stochastic matrices, and comparison of eigenvalues, *Linear Algebra Appl.* 118:163–248.

Ando, T., Horn, R., and Johnson, C. 1987. The singular values of a Hadamard product: A basic inequality, *Linear and Multilinear Algebra* 21:345–365.

Bapat, R. B. 1987. Majorization and singular values, *Linear and Multilinear Algebra* 21:211–214.

Bapat, R. B. 1991. Majorization and singular values III, *Linear Algebra Appl.* 145:59–70.

Bapat, R. and Sunder, V. 1985. On majorization and Schur products, *Linear Algebra Appl.* 72:107–117.

Berg, C., Christensen, J. P. R., and Ressel, R. 1984. *Harmonic Analysis on Semigroups, Theory of Positive Definite and Related Functions*, Springer-Verlag, New York.

Bock, M. E., Diaconis, P., Huffer, F., and Perlman, M. 1987. Inequalities for linear combinations of gamma random variables, *Canad. J. Statist.* 15:387–395.

Dellacherie, C. and Meyer, P.-A. 1988. Probabilities and Potential C, North Holland Mathematics Studies No. 151. Elsevier, Amsterdam.

Dickson, B. W. 1974. On the fundamental theorem of linear state variable feedback. *IEEE Trans. Automat. Control* AC-19:577–579.

Eaton, M. L. and Perlman, M. 1977. Reflection groups, generalized Schur functions, and the geometry of majorization, *Ann. Probab.* 5:829–860.

Eaton, M. L. 1982. A review of selected topics in multivariate probability inequalities, *Ann. Statist.* 10:11–43.

Eaton, M. L. 1987. *Lectures on Topics in Probability Inequalities*, C.W.I. Tract 35, Centrum voor Wiskunde en Informatica, Amsterdam, p. 197.

Farrell, R. H. 1985. *Multivariate Calculation, Use of the Continuous Groups*, 2nd ed., Springer-Verlag, New York.

Giovagnoli, A. and Wynn, H. 1985. G-majorization with applications to matrix orderings, *Linear Algebra Appl.* 67:111–135.

Gohberg, I. C. and Krein, M. G. 1969. *Introduction to the Theory of Linear Nonselfadjoint Operators* (transl.), Amer. Math. Soc., Providence.

Hazewinkel, M. and Martin, C. F. 1983. Representations of the symmetric group, the specialization order, systems and Grassmann manifolds, *Enseign. Math.* 29:53–87.

Horn, R. and Johnson, C. 1986. *Topics in Matrix Analysis*, Cambridge U.P., Cambridge.

Joe, H. 1993. Generalized majorization orderings and applications, in *Stochastic Inequalities* (M. Shaked and Y. L. Tong, Eds.), IMS Monograph Ser. 22, Inst. of Mathematical Statistics, pp. 145–158.

Joe, H. and Verducci, J. 1993. Multivariate majorization by positive combinations, in *Stochastic Inequalities* (M. Shaked and Y. L. Tong, Eds.), IMS Monograph Ser. 22, Inst. of Mathematical Statistics, Hayward, Calif.

Karlin, S. and Rinott, Y. 1981. Entropy inequalities for classes of probability distributions II: Multivariate case, *Adv. in Appl. Probab.* 13:325–351.

Karlin, S. and Rinott, Y. 1983. Comparison of measures, multivariate majorization and applications to statistics, in *Studies in Econometrics, Time-Series and Multivariate Statistics* (S. Karlin, T. Amemiya, and L. Goodman, Eds.), Academic, New York, pp. 465–489 [MR 85f:62004].

Kemperman, J. H. B. 1981. Review of inequalities..., *Bull. Amer. Math. Soc. (N. S.)* 5:319–324.

Kingman, J. F. C. 1961. A convexity property of positive matrices. *Quart. J. Math. Oxford Ser. (2)* 12:283–284.

Kostant, B. 1973. On convexity, the Weyl group and the Iwasawa decomposition, *Ann. Sci. École Norm. Sup. (4th series)* 6:413–445.

Kwong, M. K. 1989. Some results on matrix monotone functions, *Linear Algebra Appl.* 18:129–153.

Liyanage, L. and Shanthikumar, J. G. 1993. Allocation through stochastic Schur convexity and stochastic transposition increasingness, in *Stochastic Inequalities* (M. Shaked and Y. L. Tong, Eds.), IMS Monograph Ser. 22, Inst. of Mathematical Statistics, Hayward, Calif.

Marshall, A. W. and Olkin, I. 1979. Inequalities: Theory of Majorization and its applications. Academic Pr., New York.

Pečarić, J. E., Proschan, F., and Tong, Y. L. 1992. *Convex Functions, Partial Orderings and Statistical Applications*, Academic, San Diego.

Proschan, F. 1965. Peakedness of distributions of convex combinations. *Ann. Math. Statist.* 36:1703–1706.

Rolski, T. 1985. On nonconvexity of some functions of probabilistic interest, *Bull. Polish Acad. Sci. Math.* 33:441–445 [MR 87f:60023].

Rosenbock, H. 1970. *State Space and Multivariable Theory*, Wiley, New York.

Shaked, M. and Shanthikumar, J. G. 1988. Stochastic convexity and its applications, *Adv. Appl. Probab.* 20:427–446.

Thompson, R. C. 1983. Review of inequalities..., *Linear Algebra Appl.* 54:173–175.

Zobin, N. and Zobina, V. 1993. Coxeter Groups and Interpolation of Operators. Technical report, Max Planck Inst. für Mathematik, Univ. Bonn.

Received 23 September 1993; final manuscript accepted 5 November 1993