PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 142, Number 11, November 2014, Pages 3665–3674
S 0002-9939(2014)12119-8
Article electronically published on July 2, 2014

CAUCHY INEQUALITIES FOR THE SPECTRAL RADIUS OF PRODUCTS OF DIAGONAL AND NONNEGATIVE MATRICES

JOEL E. COHEN

(Communicated by Walter Craig)

ABSTRACT. Inequalities for convex functions on the lattice of partitions of a set partially ordered by refinement lead to multivariate generalizations of inequalities of Cauchy and Rogers-Hölder and to eigenvalue inequalities needed in the theory of population dynamics in Markovian environments: If A is an $n \times n$ nonnegative matrix, $n > 1$, D is an $n \times n$ diagonal matrix with positive diagonal elements, $r(\cdot)$ is the spectral radius of a square matrix, $r(A) > 0$, and $x \in [1, \infty)$, then $r^{x-1}(A)r(D^x A) \geq r^x(DA)$. When A is irreducible and $A^T A$ is irreducible and $x > 1$, then equality holds if and only if all elements of D are equal. Conversely, when $x > 1$ and $r^{x-1}(A)r(D^x A) = r^x(DA)$ if and only if all elements of D are equal, then A is irreducible and $A^T A$ is irreducible.

1. Introduction

The aim of this paper is to establish some inequalities for the spectral radius, dominant eigenvalue, or Perron-Frobenius root of certain nonnegative matrices. In the following sections, we first discuss inequalities for convex functions on a lattice of partitions, then inequalities for the spectral radius of nonnegative matrices. The proofs follow in a separate section. The remainder of this Introduction explains the motivation and use of these inequalities.

In modeling stochastic population growth as a Markovian multiplicative (rather than additive) random walk, we let $N(t) > 0$ represent the (real scalar) number of individuals in a population at time $t \in \mathbb{N} = \{0, 1, 2, \ldots\}$. For $t > 0$, we assume $N(t) = G(t-1)G(t-2) \cdots G(0)N(0)$, where the growth factors $G(t), t \in \mathbb{N}$ take values from a finite set $d_1, \ldots, d_n$ of positive numbers. Values of $G(t)$ are selected by a homogeneous stationary n-state Markov chain with column-to-row transition matrix A according to $\Pr\{G(t+1) = d_i | G(t) = d_j\} = a_{ij},\ t \in \mathbb{N},\ \Pr\{G(0) = d_i\} = \pi_i > 0,\ i, j = 1, \ldots, n$, and if $A = (a_{ij})_{i,j=1}^n,\ \pi = (\pi_1, \ldots, \pi_n)^T$ (π is a column n-vector), then $A\pi = \pi$, i.e., π is the stationary distribution of the Markov chain. The sum of each column of A is 1.

Let $D = diag(d_1, \ldots, d_n)$ be a diagonal matrix with $d_{ii} = d_i$. The possible values of the growth factors d_i are along the diagonal. The asymptotic long-run growth rate of the pth moment of $N(t),\ p \in \mathbb{R}$, is given by $\lim_{t\to\infty} \frac{1}{t} \log E[(N(t))^p] = \log[r(D^p A)]$ [3]. By definition, the variance of $N(t)$ is $Var(N(t)) = E(N^2(t)) - [E(N(t))]^2$. Because $Var(N(t)) \geq 0$ by Cauchy's inequality [13], we have $r(D^2 A) \geq [r(DA)]^2$ [14]. We needed a sufficient condition that $r(D^2 A) > [r(DA)]^2$ to establish

Received by the editors November 13, 2012 and, in revised form, November 14, 2012.
2010 *Mathematics Subject Classification.* Primary 15A42; Secondary 15B48, 15A16, 15A18, 26D15, 60K37.

that the asymptotic long-run growth rate of the variance satisfies

$$\lim_{t\to\infty} \frac{1}{t} \log Var(N(t)) = \log r(D^2A) > -\infty. \tag{1.1}$$

When $r(D^2A) > [r(DA)]^2$, the rate of growth of $E(N^2(t))$ dominates the rate of growth of $[E(N(t))]^2$, hence $[r(DA)]^2$ is absent from (1.1) [3]. The question of determining when $r(D^2A) > [r(DA)]^2$ was the origin of this study. The answer is in Corollary 3.5 and the discussion that follows.

2. Convex functions and a lattice of partitions

A convex cone X is defined as a subset of a vector space over $\mathbb{R}$ that is closed under linear combinations with positive coefficients. A real-valued function f on a convex cone X is defined to be convex if, for any $w \in [0,1]$ and any two distinct elements $x, y \in X$, $x \neq y$, $f(wx+(1-w)y) \leq wf(x)+(1-w)f(y)$, and f is defined to be strictly convex if the inequality is strict when $0 < w < 1$.

Let $m \in \mathbb{N}$, $m > 1$. A partition of $S_m = \{1, \ldots, m\}$ is a set of $p \geq 1$, $p \in \mathbb{N}$ nonempty mutually exclusive subsets P_i, $i = 1, \ldots, p$ of S_m whose union is S_m. Each subset P_i in P is called a part of the partition P and p is the number of parts. We write $P = \{P_1, \ldots, P_p\}$, where $\bigcup_{i=1}^p P_i = P$ and $P_i \cap P_j = \emptyset$. If $Q = \{Q_1, \ldots, Q_q\}$ is a partition of S_m with $q \in \mathbb{N}$ parts, we say that Q is a refinement of P and we write $P \geq Q$ if and only if (using i to index the parts of P and j to index the parts of Q) for every $j = 1, \ldots, q$ there exists $i \in S_p$ such that $Q_j \subseteq P_i$. The lattice of partitions of S_m is defined as the set of all partitions of S_m together with their partial ordering by the relation of refinement.

Example (Part 1)**.** If $m = 3$, the partitions are partially ordered from most refined (at the bottom) to least refined (at the top) as:

$$\{\{1,2,3\}\}$$
$$\{\{1\}\ \{2,3\}\} \qquad \{\{2\},\{1,3\}\} \qquad \{\{3\},\{1,2\}\}$$
$$\{\{1\},\{2\},\{3\}\}$$

Each partition in this table is a refinement of every partition in any row above its row, e.g., $\{\{1,2,3\}\} \geq \{\{2\},\{1,3\}\} \geq \{\{1\}\{2\}\{3\}\}$ but partitions in the same row are not related by refinement.

Theorem 2.1. *Let $P = \{P_1, \ldots, P_p\}$ and $Q = \{Q_1, \ldots, Q_q\}$ be partitions of S_m with $P \geq Q$. Let X be a convex cone and let x_h, $h = 1, \ldots, m$ be m distinct points in X. Let f be a convex function on X. Also let $w_h > 0$, $h = 1, \ldots, m$, satisfy $\sum_{h=1}^m w_h = 1$. Define*

$$w(P_i) = \sum_{h\in P_i} w_h,\ i = 1, \ldots, p, \qquad w(Q_j) = \sum_{h\in Q_j} w_h,\ j = 1, \ldots, q. \tag{2.1}$$

By definition, no part of any partition is an empty set, hence all these weights are positive and

$$\sum_{j=1}^q w(Q_j) f\left(\sum_{h\in Q_j} \frac{w_h x_h}{w(Q_j)}\right) \geq \sum_{i=1}^q w(P_i) f\left(\sum_{h\in P_i} \frac{w_h x_h}{w(P_i)}\right). \tag{2.2}$$

If f is strictly convex, then the inequality is strict.

Example (Part 2)**.** Corresponding to the above partial ordering of partitions is a partial ordering of functionals of the convex function f, least at the top and greatest at the bottom. If f is strictly convex, the ordering increases strictly from top to bottom. We omitted the partitions $\{\{1\}\{2,3\}\}$ and $\{\{3\},\{1,2\}\}$, as the corresponding functionals may be obtained by permuting the subscripts in the second row.

$$\begin{aligned}\{\{1,2,3\}\} &\Leftrightarrow f(w_1x_1+w_2x_2+w_3x_3)\\ \{\{2\},\{1,3\}\} &\Leftrightarrow w_2f(x_2)+(w_1+w_3)f\left(\tfrac{w_1x_1+w_3x_3}{w_1+w_3}\right)\\ \{\{1\},\{2\},\{3\}\} &\Leftrightarrow w_1f(x_1)+w_2f(x_2)+w_3f(x_3)\end{aligned}$$

3. Convex functions of nonnegative matrices

Let $m,n\in\mathbb{N},\ m,n>1$. All matrices here are $n\times n$ real unless $n\times m$ is specified. A matrix is nonnegative if each element is nonnegative real. A nonnegative matrix is column-stochastic if the sum of each column is 1. A nonnegative matrix A is irreducible if for each row i and each column j with $1\le i,j\le n$, there exists an integer p such that $(A^p)_{ij}>0$. The transpose of A is A^T. A nonnegative matrix A is two-fold irreducible if A is irreducible and A^TA is irreducible [2, Definition 22]. A matrix is positive, $A>0$, if all its elements are positive.

A matrix is diagonal if all elements off the main diagonal are 0. A matrix is positive diagonal if it is diagonal and all elements on the main diagonal are positive. Let $\mathbb{D}_n$ be the set of diagonal matrices and let $\mathbb{D}_n^+$ be the set of positive diagonal matrices. A one-to-one correspondence between $\mathbb{D}_n$ and $\mathbb{D}_n^+$ is given by $\mathbb{D}_n^+=\exp(\mathbb{D}_n)$. A positive diagonal matrix is scalar if all its diagonal elements equal some positive real number.

The spectral radius $r(A)$ of a matrix A is the maximum of the magnitudes (absolute values) of the eigenvalues of A. For any two matrices $A,\ B,\ r(AB)=r(BA)$ and for any constant $c>0,\ r(cA)=cr(A)$ and $r(A^c)=r^c(A)\equiv(r(A))^c$. If A is irreducible, then $r(A)>0$ but not conversely.

Theorem 3.1. *Let A be a nonnegative matrix such that $r(A)>0$. Let $D(1),D(2),\dots,D(m)\in\mathbb{D}_n^+$. Let $P=\{P_1,\dots,P_p\}$ and $Q=\{Q_1,\dots,Q_q\}$ be partitions of S_m with $P\ge Q$. Define the weights w as in Theorem* 2.1 *and* (2.1)*. Then*

$$\prod_{j=1}^{q} r^{w(Q_j)}\left(\left[\prod_{h\in Q_j}D(h)^{w_h}\right]^{\frac{1}{w(Q_j)}}A\right)\ge\prod_{i=1}^{p} r^{w(P_i)}\left(\left[\prod_{h\in P_i}D(h)^{w_h}\right]^{\frac{1}{w(P_i)}}A\right). \tag{3.1}$$

If, for each $P_i\in P$, there exists $D_i\in\mathbb{D}_n^+$ such that, for every part $Q_j\subseteq P_i$, $[\prod_{h\in Q_j}D(h)^{w_h}]^{\frac{1}{w(Q_j)}}$ is a scalar multiple of D_i, then equality holds. If A is two-fold irreducible, then equality holds only if, for each $P_i\in P$, there exists $D_i\in\mathbb{D}_n^+$ such that, for every part $Q_j\subseteq P_i$, $[\prod_{h\in Q_j}D(h)^{w_h}]^{\frac{1}{w(Q_j)}}$ is a scalar multiple of D_i. Conversely, when equality holds only if, for each $P_i\in P$, there exists $D_i\in\mathbb{D}_n^+$ such that, for every part $Q_j\subseteq P_i$, $[\prod_{h\in Q_j}D(h)^{w_h}]^{\frac{1}{w(Q_j)}}$ is a scalar multiple of D_i, then A is two-fold irreducible.

Example (Part 3)**.** Corresponding to the above partial ordering of functionals of the convex function f, the following ordering of functionals of the spectral radius $r(\cdot)$ is greatest at the bottom and least at the top:

$$\begin{array}{rcl} \{\{1,2,3\}\} & \Leftrightarrow & r\big(D(1)^{w_1}D(2)^{w_2}D(3)^{w_3}A\big) \\ \{\{2\},\{1,3\}\} & \Leftrightarrow & r^{w_2}(D(2)A)r^{w_1+w_3}\big([D(1)^{w_1}D(3)^{w_3}]^{\frac{1}{w_1+w_3}}A\big) \\ \{\{1\},\{2\},\{3\}\} & \Leftrightarrow & r^{w_1}(D(1)A)r^{w_2}(D(2)A)r^{w_3}(D(3)A) \end{array}$$

If we set $w_h = \frac{1}{3},\ h = 1,2,3$, replace each $D(h)^{1/3}$ by $D(h)$, and then cube the left, middle, and right members of the inequalities, we get $r^3(D(1)D(2)D(3)A) \leq r(D(2)^3A)r^2([D(1)^3D(3)^3]^{\frac{1}{2}}A) \leq r(D(1)^3A)r(D(2)^3A)r(D(3)^3A)$. If all the $D(h)$ are scalar multiples of some fixed $D \in \mathbb{D}_n^+$, then equality holds on the left and the right. When A is two-fold irreducible, equality holds on the left if and only if, for some $c > 0,\ D(2) = c[D(1)D(3)]^{1/2}$, and equality holds on the right if and only if, for some $c > 0,\ D(1) = cD(3)$.

Corollary 3.2. *Let $P = \{P_1,\dots,P_p\}$ and $Q = \{Q_1,\dots,Q_q\}$ be partitions of S_m with $P \geq Q$. Define the weights w as in Theorem 2.1 and (2.1). Let X be a positive $n \times m$ matrix with element $x_{gh} > 0$ in row g and column h. Then*

$$\prod_{j=1}^{q}\sum_{g=1}^{n}\Big[\prod_{h\in Q_j} x_{gh}^{w_h}\Big]^{\frac{1}{w(Q_j)}} \geq \prod_{i=1}^{p}\sum_{g=1}^{n}\Big[\prod_{h\in P_i} x_{gh}^{w_h}\Big]^{\frac{1}{w(P_i)}}. \tag{3.2}$$

Equality holds if and only if, for each $P_i \in P$, for every part $Q_j \subseteq P_i$, the vectors with n elements

$$\Big[\prod_{h\in Q_j} x_{gh}^{w_h}\Big]^{\frac{1}{w(Q_j)}},\ g = 1,\dots,n,$$

are scalar multiples of one another.

A special case of (3.2) with $P = \{\{1,2,\dots,m\}\}$ and $Q = \{\{1\},\{2\},\dots,\{m\}\}$ is [13, p. 152, Eq. (9.35)].

Example (Part 4)**.** Corresponding to the above ordering of functionals of the spectral radius $r(\cdot)$, the following quantities are greatest at the bottom and least at the top. If column 3 of X is proportional to column 1, but neither is proportional to column 2, then the second and third rows are equal and both exceed the first.

$$\begin{array}{rcl} \{\{1,2,3\}\} & \Leftrightarrow & \sum_{g=1}^n x_{g1}^{w_1}x_{g2}^{w_2}x_{g3}^{w_3} \\ \{\{2\},\{1,3\}\} & \Leftrightarrow & (\sum_{g=1}^n x_{g2})^{w_2}\big(\sum_{g=1}^n[x_{g1}^{w_1}x_{g3}^{w_3}]^{\frac{1}{w_1+w_3}}\big)^{w_1+w_3} \\ \{\{1\},\{2\},\{3\}\} & \Leftrightarrow & (\sum_{g=1}^n x_{g1})^{w_1}(\sum_{g=1}^n x_{g2})^{w_2}(\sum_{g=1}^n x_{g3})^{w_3} \end{array}$$

If we set $w_h = 1/3,\ h = 1,2,3$, replace each $x_{gh}^{1/3}$ by x_{gh}, and then cube all terms, we get multivariate versions of Hölder's inequality [13, p. 151]:

$$\begin{aligned} \Big(\sum_{g=1}^n x_{g1}x_{g2}x_{g3}\Big)^3 &\leq \Big(\sum_{g=1}^n x_{g2}^3\Big)\Big(\sum_{g=1}^n [x_{g1}x_{g3}]^{\frac{3}{2}}\Big)^2 \\ &\leq \Big(\sum_{g=1}^n x_{g1}^3\Big)\Big(\sum_{g=1}^n x_{g2}^3\Big)\Big(\sum_{g=1}^n x_{g3}^3\Big). \end{aligned}$$

Corollary 3.3. *Let A be a nonnegative matrix such that $r(A) > 0$. Let $D(1), D(2), \ldots, D(m) \in \mathbb{D}_n^+$. Then*

$$r(D(1)^m A) r(D(2)^m A) \cdots r(D(m)^m A) \geq r^m(D(1)D(2)\cdots D(m)A). \tag{3.3}$$

If, for some $D \in \mathbb{D}_n^+$ and m positive numbers $c_1, \ldots, c_m$, $D(h) = c_h D$, $h = 1, 2, \ldots, m$, then equality holds in (3.3). If A is two-fold irreducible, then equality holds only if, for some $D \in \mathbb{D}_n^+$ and m positive numbers $c_1, \ldots, c_m$, $D(h) = c_h D$, $h = 1, 2, \ldots, m$. Conversely, when equality holds only if, for some $D \in \mathbb{D}_n^+$ and m positive numbers $c_1, \ldots, c_m$, $D(h) = c_h D$, $h = 1, 2, \ldots, m$, then A is two-fold irreducible.

Corollary 3.4. *Let A be a nonnegative matrix such that $r(A) > 0$. Let $D \in \mathbb{D}_n^+$. Then for any real $x \in [1, \infty)$,*

$$r^{x-1}(A) r(D^x A) \geq r^x(DA). \tag{3.4}$$

If D is scalar or $x = 1$, then equality holds. Assume $x > 1$. If A is two-fold irreducible, then equality holds only if D is scalar; and conversely, when equality holds only if D is scalar, then A is two-fold irreducible.

Corollary 3.5. *If A is column-stochastic, $D \in \mathbb{D}_n^+$, then*

$$r(D^2 A) \geq r^2(DA) = r([DA]^2). \tag{3.5}$$

If D is scalar, then equality holds. If A is two-fold irreducible, then equality holds only if D is scalar; and conversely, when equality holds only if D is scalar, then A is two-fold irreducible.

Assuming A is column-stochastic and irreducible and D is not scalar does not guarantee strict inequality in (3.5). For example, let $d > 1$ and

$$D = \begin{pmatrix} d & 0 \\ 0 & 1 \end{pmatrix}, \quad A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Then A is column-stochastic and irreducible and D is not scalar and for $p \in (0, \infty)$,

$$D^p = \begin{pmatrix} d^p & 0 \\ 0 & 1 \end{pmatrix}, \quad D^p A = \begin{pmatrix} 0 & d^p \\ 1 & 0 \end{pmatrix},$$

$r(D^p A) = d^{p/2}$; hence $r(D^2 A) = d = [r(DA)]^2$. Altenberg [2, Theorem 18, Proposition 31] showed that the condition that A be two-fold irreducible cannot be weakened even to the condition that A be primitive, which is stronger than irreducibility. (A nonnegative matrix A is primitive if for some finite positive integer p, every element of A^p is positive.)

Corollary 3.6. *Let A be a nonnegative matrix such that $r(A) > 0$. Let $D(1), D(2), \ldots, D(m) \in \mathbb{D}_n^+$ and let $D(1)D(2)\cdots D(m) = I$, where I is the identity matrix. Then*

$$\left[r(D(1)A) r(D(2)A) \cdots r(D(m)A)\right]^{1/m} \geq r(A). \tag{3.6}$$

If $D(h)$ is scalar for $h = 1, 2, \ldots, m$, then equality holds. If A is two-fold irreducible, then equality holds only if every $D(h)$ is scalar, $h = 1, 2, \ldots, m$. Conversely, if equality holds only if every $D(h)$ is scalar, $h = 1, 2, \ldots, m$, then A is two-fold *irreducible. In particular, if $D \in \mathbb{D}_n^+$, then*

$$\big(r(DA)r(D^{-1}A)\big)^{1/2} \geq r(A) \tag{3.7}$$

and

$$\inf\{\big(r(DA)r(D^{-1}A)\big)^{1/2} \mid D \in \mathbb{D}_n^+\} = r(A). \tag{3.8}$$

Corollary 3.6 has interesting consequences that are well known and require no detailed proof here. First [13, pp. 12-13], if $p(i) \geq 0,\ x(i) > 0,\ i = 1, \dots, n,\ p(1) + \cdots + p(n) = 1$, then $(\sum_{i=1}^n p(i)x(i))(\sum_{i=1}^n p(i)/x(i)) \geq 1$. Equality holds if and only if all elements of the set $\{x(i) \mid p(i) > 0\}$ are equal. Second, setting $x(i) = p(i)/q(i)$ gives: if $p(i) > 0,\ q(i) > 0,\ i = 1, \dots, n,\ \sum_{i=1}^n p(i) = \sum_{i=1}^n q(i) = 1$, then $\sum_{i=1}^n (p(i)^2/q(i)) \geq 1$. Equality holds if and only if all $p(i)/q(i)$ are equal. Third, if $x(i) > 0,\ y(i) > 0,\ i = 1, \dots, n$ are the elements of vectors $x,\ y$ with sums $X = \sum_{i=1}^n x(i),\ Y = \sum_{i=1}^n y(i)$, and if the corresponding normalized probability vectors are $p_x = x/X,\ p_y = y/Y$, then

$$m_x := \sum_{i=1}^n p_x(i)\frac{x(i)}{y(i)} \geq m_y := \sum_{i=1}^n p_y(i)\frac{x(i)}{y(i)} = \frac{X}{Y}. \tag{3.9}$$

Equality holds if and only if all $x(i)/y(i)$ are equal. (To prove, set $p(i) = p_x(i)$, $q(i) = p_y(i),\ i = 1, \dots, n$ in the previous inequality.) If $x(i)$ is the population size and $y(i)$ is the land area of province i of a country with n provinces, then $x(i)/y(i)$ is the population density of province i. The population-weighted mean population density is m_x, the area-weighted mean population density is m_y, $m_x \geq m_y$, and $m_x = m_y$ if and only if the population density of every province is the same. In particular, if $y(i) = 1,\ i = 1, \dots, n$, then $m_x \geq X/n$ and equality holds if and only if all $x(i)$ are equal. Inequality (3.9) is known from studies of the distribution of recurrence times [5, p. 64, Eq. (3)], the length-biased sampling of fibers of yarns [5, p. 65], the number of students in classes [7, p. 217], the numbers of friends per person [6, p. 1470], and other social scientific studies [8, pp. 143–144].

Corollary 3.7. *For any $n \times m$ positive matrix X with element $x_{ij} > 0$ in row i and column j,*

$$\prod_{j=1}^m \sum_{i=1}^n x_{ij}^m \geq \big(\sum_{i=1}^n \prod_{j=1}^m x_{ij}\big)^m. \tag{3.10}$$

Equality holds if and only if X has rank one, i.e., $X = dc^T$.

If $m = 2$, Corollary 3.7 reduces to Cauchy's inequality [13, p. 1] limited to positive numbers. The extension to all real numbers is very easy for $m = 2$.

Cohen, Friedland, Kato, and Kelly [4, p. 66, Lemma 5] proved that if A and D are nonnegative $n \times n$ matrices and D is diagonal, then $r(D^2A^2) \geq r^2(DA)$, and if A^2 and A^TA are irreducible and D is positive diagonal but not scalar, then this inequality is strict. Altenberg [2, Theorem 23] proved that A^2 and A^TA are irreducible if and only if A is two-fold irreducible. The right side of the inequality $r(D^2A^2) \geq r^2(DA)$ is the same as the right side of (3.4) with $x = 2$, which is $r(A)r(D^2A) \geq r^2(DA)$, but the left sides differ. Comparing the left sides, it is easy to find a nonscalar positive diagonal matrix D and a positive matrix A such that $r(D^2A^2) > r(A)r(D^2A)$ and another such D and A such that $r(D^2A^2) < r(A)r(D^2A)$. Thus neither upper bound on $r^2(DA)$ is always better than the other for nonscalar positive diagonal D and positive A. In an earlier version of this paper, we asked for additional conditions on D and A sufficient to guarantee one or the other ordering $r(D^2A^2) \geq r(A)r(D^2A) \geq r^2(DA)$ or $r(A)r(D^2A) \geq r(D^2A^2) \geq r^2(DA)$ and conditions for strict inequality. Lee Altenberg (personal communication, May 29, 2012) observed that [10, Theorem 5.1] implies that if A is the column-stochastic transition matrix of an ergodic reversible Markov chain with all positive eigenvalues, then $r(A) = 1$ and $r(A)r(D^2A) \geq r(D^2A^2)$ and equality holds if and only if D is scalar. Altenberg further remarked that the inequality will reverse if all the non-Perron eigenvalues of A are negative, an immediate consequence of [1, Theorem 33]. He will develop details elsewhere.

4. Proofs

Proof of Theorem 2.1. First we establish an inequality for a fixed i on the right side of (2.2) and then we sum over i. Fix i. The partition Q partitions part $P_i \in P$ into $p_i \geq 1$ parts $Q_1(i), \ldots, Q_{p_i}(i) \in Q$, where

$$\sum_{i=1}^{p} p_i = q, \quad \bigcup_{i=1}^{p} \big(Q_1(i) \cup \cdots \cup Q_{p_i}(i)\big) = Q,$$
$$w(P_i) = \sum_{g=1}^{p_i} w(Q_g(i)), \qquad \bigcup_{g=1}^{p_i} Q_g(i) = P_i.$$

For this fixed i,

$$f\Big(\sum_{h\in P_i} \frac{w_h x_h}{w(P_i)}\Big) = f\Big(\sum_{g=1}^{p_i} \sum_{h\in Q_g(i)} \frac{w_h x_h}{w(P_i)}\Big) = f\Big(\sum_{g=1}^{p_i} \frac{w(Q_g)}{w(P_i)} \sum_{h\in Q_g(i)} \frac{w_h x_h}{w(Q_g)}\Big)$$

$$\leq \sum_{g=1}^{p_i} \frac{w(Q_g)}{w(P_i)} f\Big(\sum_{h\in Q_g(i)} \frac{w_h x_h}{w(Q_g)}\Big) \tag{4.1}$$

by convexity of $f(\cdot)$. Multiply by $w(P_i)$ and sum over i to get

$$\sum_{i=1}^{p} w(P_i) f\Big(\sum_{h\in P_i} \frac{w_h x_h}{w(P_i)}\Big) \leq \sum_{i=1}^{p} w(P_i) \sum_{g=1}^{p_i} \frac{w(Q_g)}{w(P_i)} f\Big(\sum_{h\in Q_g(i)} \frac{w_h x_h}{w(Q_g)}\Big) \tag{4.2}$$
$$= \sum_{j=1}^{q} w(Q_j) f\Big(\sum_{h\in Q_j} \frac{w_h x_h}{w(Q_j)}\Big).$$

If f is strictly convex, then strict inequality holds in (4.1), since all $x_h,\ h = 1, \ldots, m$ are distinct and all weights are positive, and therefore strict inequality holds in (4.2). □

The following results depend on this theorem:

Theorem 4.1 (Friedland [9, Theorem 4.2] and Altenberg [2, Theorem 21])**.** *Let A be a nonnegative matrix such that $r(A) > 0$. For any $C_1, C_2 \in \mathbb{D}_n,\ t \in (0, 1)$,*

$$\log r(e^{(1-t)C_1 + tC_2} A) \leq (1 - t) \log r(e^{C_1} A) + t \log r(e^{C_2} A). \tag{4.3}$$

If $C_2 - C_1$ is scalar, then (4.3) is an equality. Moreover, the following are equivalent:

(1) *A is two-fold irreducible (A is irreducible and $A^T A$ is irreducible);*

(2) (4.3) *is an equality only if* $C_2 - C_1$ *is scalar;*
(3) (4.3) *is a strict inequality for all* $C_1, C_2 \in \mathbb{D}_n$ *such that* $C_2 - C_1$ *is not scalar.*

The weak inequality in (4.3) follows easily from [11]. We need an obvious generalization of Theorem 4.1.

Theorem 4.2. *Let* A *be a nonnegative matrix such that* $r(A) > 0$. *For any positive integer* $m > 1$ *and any* $C_1, C_2, \ldots, C_m \in \mathbb{D}_n$ *and any* $t_1, \ldots, t_m \in (0, 1)$ *such that* $t_1 + \cdots + t_m = 1$,

$$r(e^{t_1C_1+t_2C_2+\cdots+t_mC_m}A) \leq r^{t_1}(e^{C_1}A)\cdots r^{t_m}(e^{C_m}A), \tag{4.4}$$

and in particular when all $t_i = 1/m$,

$$r^m(e^{(C_1+C_2+\cdots+C_m)/m}A) \leq r(e^{C_1}A)r(e^{C_2}A)\cdots r(e^{C_m}A). \tag{4.5}$$

If there exist $C \in \mathbb{D}_n$ *and real numbers* $c_1, c_2, \ldots, c_m$ *such that*

$$C_h = c_hI + C, \quad h = 1, 2, \ldots, m, \tag{4.6}$$

then equality holds in (4.4) *and* (4.5). *Moreover, the following are equivalent:*

(1) *A is two-fold irreducible* (*A is irreducible and* A^TA *is irreducible*)*;*
(2) (4.4) *is an equality only if* (4.6) *holds;*
(3) (4.4) *is a strict inequality for all* $C_1, C_2, \ldots, C_m \in \mathbb{D}_n$ *such that for some* C_i, C_j, $i \neq j$, $C_i - C_j$ *is not scalar.*

Proof of Theorem 3.1. Let $C_h = \log D(h)$, $h = 1, \ldots, m$. Then all $C_h \in \mathbb{D}_n$ and $\mathbb{D}_n$ is a convex cone. By Theorem 4.2, for $C \in \mathbb{D}_n$, if $R(C) = \log r(e^CA)$, then $R(C)$ is a convex function of $C \in \mathbb{D}_n$. Then from (2.2), replacing f by R, and replacing x_h by C_h, we have successively

$$\sum_{j=1}^{q} w(Q_j)R\Big(\sum_{h\in Q_j}\frac{w_hC_h}{w(Q_j)}\Big) \geq \sum_{i=1}^{p} w(P_i)R\Big(\sum_{h\in P_i}\frac{w_hC_h}{w(P_i)}\Big),$$

$$\prod_{j=1}^{q} r^{w(Q_j)}\left(\exp\left[\sum_{h\in Q_j}\frac{w_hC_h}{w(Q_j)}\right]A\right) \geq \prod_{i=1}^{p} r^{w(P_i)}\left(\exp\left[\sum_{h\in P_i}\frac{w_hC_h}{w(P_i)}\right]A\right),$$

$$\prod_{j=1}^{q} r^{w(Q_j)}\left(\left[\prod_{h\in Q_j}D(h)^{w_h}\right]^{\frac{1}{w(Q_j)}}A\right) \geq \prod_{i=1}^{p} r^{w(P_i)}\left(\left[\prod_{h\in P_i}D(h)^{w_h}\right]^{\frac{1}{w(P_i)}}A\right).$$

Exponentiating both sides of (4.6) and writing $D = \exp C$ gives the equivalent condition

$$\exp C_h = D(h) = (\exp c_h)\exp C = (\exp c_h)D.$$

Conditions (i) and (ii) of Theorem 4.2 give the claimed necessary and sufficient condition for equality. □

Proof of Corollary 3.2. Let J be the $n \times n$ matrix with all elements equal to 1. Then J is two-fold irreducible. In Theorem 3.1, set $A = J$, $D(h) = diag(x_{gh},\ g = 1, \ldots, n)$, $h = 1, \ldots, m$. Since $1^TD(h)J = \big(\sum_{g=1}^n x_{gh}\big)1^T$, i.e., since all column sums of $D(h)J$ equal $\sum_{g=1}^n x_{gh}$, a theorem of Frobenius [12, p. 24] gives $r(D(h)J) = \sum_{g=1}^n x_{gh}$. The conditions for equality restate those in Theorem 3.1. □

Proof of Corollary 3.3. In Theorem 3.1, let $P = \{\{1,\ldots,m\}\}$, $Q=\{\{1\},\ldots,\{m\}\}$, $w_h = 1/m,\ h = 1,\ldots,m$. Then $P \geq Q$ and (3.1) becomes

$$\prod_{j=1}^{m} r^{1/m}(D(j)A) \geq r([\prod_{h=1}^{m} D(h)^{1/m}]A).$$

Raising both sides to the power m gives

$$\prod_{j=1}^{m} r(D(j)A) \geq r^m([\prod_{h=1}^{m} D(h)^{1/m}]A).$$

Replacing $D(h)^{1/m}$ with $D(h)$ (so that what was $D(j)$ becomes $D(j)^m$) yields (3.3).

If A is two-fold irreducible, then by Theorem 4.2 applied to $C_h = \log D(h),\ h = 1,2,\ldots,m$, equality holds in (3.3) if and only if there exist $C \in \mathbb{D}_n$ and real numbers $c_1, c_2, \ldots, c_m$ such that $C_h = \log D(h) = c_h I + C,\ h = 1,2,\ldots,m$ or equivalently $D(h) = \exp c_h \exp C = \exp c_h D$, where $D = \exp C$. □

Proof of Corollary 3.4. In (3.1), let $m = 2$, $P = \{P_1\} = \{\{1,2\}\}$, and $Q = \{Q_1, Q_2\} = \{\{1\},\{2\}\}$. Then $P \geq Q$. We are given $D \in \mathbb{D}_n^+$ and a real $x \in [1,\infty)$. If $x = 1$ or D is scalar, then both sides of (3.4) are trivially equal. Henceforth assume $x > 1$ and D is not scalar. Define $E = D^x$. Then E is scalar if and only if D is scalar, so E is not scalar. Define $D(1) = I, D(2) = E, w_1 = 1 - 1/x, w_2 = 1/x$. Because E is not a scalar multiple of I, Theorem 3.1 and (3.1) imply that $r^{1-1/x}(A)r^{1/x}(EA) > r(E^{1/x}A)$. Raising both sides of the inequality to the power x and replacing E by D^x give $r^{x-1}(A)r(D^x A) > r^x(DA)$. □

Proof of Corollary 3.5. If A is column-stochastic, then $r(A) = 1$. Apply Corollary 3.4 with $x = 2$. The condition for equality follows from that for Corollary 3.4. □

Proof of Corollary 3.6. Apply Corollary 3.3. By changing variables, $E(h)=D(h)^m$, $h = 1,\ldots,m$, in (3.3), and then replacing $E(h)$ by $D(h)$, we have

$$r(D(1)A)r(D(2)A)\cdots r(D(m)A) \geq r^m\big(D(1)^{1/m}D(2)^{1/m}\cdots D(m)^{1/m}A\big)$$
$$= r^m\big([D(1)D(2)\cdots D(m)]^{1/m}A\big) = r^m(IA) = r^m(A).$$

On the right side of (3.3), $D(1)D(2)\cdots D(m) = I$ by assumption. Inequality (3.7) is (3.6) with $m = 2$. Equality (3.8) follows because $\big(r(DA)r(D^{-1}A)\big)^{1/2}$ is a continuous function of $D \in \mathbb{D}_n^+$ and as $D \to I$, $\big(r(DA)r(D^{-1}A)\big)^{1/2} \to r(A)$. □

Proof of Corollary 3.7. Apply Corollary 3.2 with $P = \{\{1,\ldots,m\}\}$, $Q = \{\{1\},\ldots,\{m\}\}$, $w_h = 1/m,\ h = 1,\ldots,m$. □

Acknowledgements

The author thanks Lee Altenberg, Shmuel Friedland, B. J. Green, and Meng Xu for helpful comments. Lee Altenberg answered the questions in an earlier draft, gave references [6–8], and improved the exposition and notation. Shmuel Friedland directed the author to Lee Altenberg.

The author acknowledges with thanks the support of the U.S. National Science Foundation grant EF-1038337 and the Marsden Fund of the Royal Society of New Zealand (08-UOC-034), the assistance of Priscilla K. Rogerson, and the hospitality of Michael Plank and the family of William T. Golden during this work.

References

[1] L. Altenberg, *The evolution of dispersal in random environments and the principle of partial control*, Ecological Monographs **82** (3) (2012) 297–333. http://dx.doi.org/10.1890/11-1136.1.

[2] ———, *A sharpened condition for strict log-convexity of the spectral radius via the bipartite graph*, Linear Algebra Appl. **438** (2013), no. 9, 3702–3718. MR3028608

[3] Joel E. Cohen, *Stochastic population dynamics in a Markovian environment implies Taylor's power law of fluctuation scaling*, Theoret. Population Biol. **93** (2014), 30–37. DOI 10.1016/j.tpb.2014.01.001

[4] Joel E. Cohen, Shmuel Friedland, Tosio Kato, and Frank P. Kelly, *Eigenvalue inequalities for products of matrix exponentials*, Linear Algebra Appl. **45** (1982), 55–95, DOI 10.1016/0024-3795(82)90211-7. MR660979 (84h:15020)

[5] D. R. Cox, *Renewal theory*, Methuen & Co. Ltd., London, 1962. MR0153061 (27 #3030)

[6] S. L. Feld, *Why your friends have more friends than you do*, Amer. J. Sociology **96** (6) (1991), 1464–1477. http://www.jstor.org/stable/2781907

[7] S. L. Feld and B. Grofman, *Variation in class size, the class size paradox, and some consequences for students*, Research in Higher Education **6** (3) (1977), 215–222. http://www.jstor.org/stable/40195170

[8] S. L. Feld and B. Grofman, *Puzzles and paradoxes involving averages: an intuitive approach*, Collective Decision Making: Views from Social Choice and Game Theory, Ad van Deemen and Agnieszka Rusinowska, eds., Springer Verlag, Berlin, 2010, pp. 137–150. DOI 10.1007/978-3-642-02865-6_10

[9] Shmuel Friedland, *Convex spectral functions*, Linear and Multilinear Algebra **9** (1980/81), no. 4, 299–316, DOI 10.1080/03081088108817381. MR611264 (82d:15014)

[10] S. Karlin, *Classifications of selection-migration structures and conditions for a protected polymorphism*, Evolutionary Biology, 14, M. K. Hecht, B. Wallace, and G. T. Prance, eds., Plenum Publishing Corporation, New York, 1982, pp. 61–204.

[11] J. F. C. Kingman, *A convexity property of positive matrices*, Quart. J. Math. Oxford Ser. (2) **12** (1961), 283–284. MR0138632 (25 #2075)

[12] Henryk Minc, *Nonnegative matrices*, A Wiley-Interscience Publication, Wiley-Interscience Series in Discrete Mathematics and Optimization, John Wiley & Sons Inc., New York, 1988. MR932967 (89i:15001)

[13] J. Michael Steele, *The Cauchy-Schwarz master class*, An introduction to the art of mathematical inequalities. MAA Problem Books Series, Mathematical Association of America, Washington, DC, 2004. MR2062704 (2005a:26035)

[14] S. D. Tuljapurkar, *Population dynamics in variable environments. II. Correlated environments, sensitivity analysis and dynamics*, Theoret. Population Biol. **21** (1982), no. 1, 114–140, DOI 10.1016/0040-5809(82)90009-0. MR662525 (84g:92037a)

Laboratory of Populations, The Rockefeller University and Columbia University, 1230 York Avenue, Box 20, New York, New York 10065

E-mail address: cohen@rockefeller.edu