Matrix inequalities with applications to the theory of iterated kernels

William Banks Asma Harcharras Stefan Neuwirth Éric Ricard

Abstract

For an $m \times n$ matrix A with nonnegative real entries, Atkinson, Moran and Watterson proved the inequality $s(A)^3 \leqslant mns(AA^tA)$, where A^t is the transpose of A, and $s(\cdot)$ is the sum of the entries. We extend this result to finite products of the form $AA^tAA^t \ldots A$ or $AA^tAA^t \ldots A^t$ and give some applications to the theory of iterated kernels.

1 Introduction

For any matrix A, let $s(A)$ denote the sum of its entries. For any integer $k \geq 1$, we define

$$A^{(2k)} = (AA^t)^k, \qquad A^{(2k+1)} = (AA^t)^k A,$$

where A^t denotes the transpose of A. In Section 2, we prove the following sharp inequalities:

Theorem 1.1. *Let A be an $m \times n$ matrix with nonnegative real entries. Then for every integer $k \geqslant 1$, the following matrix inequalities hold:*

$$s(A)^{2k} \leqslant m^{k-1}n^k s(A^{(2k)}), \qquad s(A)^{2k+1} \leqslant m^k n^k s(A^{(2k+1)}).$$

For the special case of symmetric matrices, this theorem was proved in 1959 by Mulholland and Smith [4], thus settling an earlier conjecture of Mandel and Hughes [3] that had been based on the study of certain genetical models. For arbitrary matrices (with nonnegative entries), Theorem 1.1 also generalises the matrix inequality

$$s(A)^3 \leqslant mns(AA^tA),$$

which was first proved in 1960 by Atkinson, Moran and Watterson [1] using methods of perturbation theory.

Theorem 1.1 has a graph theoretic interpretation when applied to matrices with entries in $\{0, 1\}$. Let G be a graph with red vertices labeled $1, \ldots, m$ and blue vertices labeled $1, \ldots, n$ such that every edge connects only vertices of distinct colours: G is a bipartite graph. Its reduced incidence matrix is an $m \times n$ matrix A such that $a_{i,j} = 1$ if red vertex i is adjacent to blue vertex j, and $a_{i,j} = 0$ otherwise. Then $s(A)$ is the size of G, while $s(A^{(\ell)})$ is the number of walks on G of length ℓ starting from a red vertex, i.e., the number of sequences $(v_0, \ldots, v_\ell)$ such that v_0 is a red vertex and every pair $\{v_i, v_{i+1}\}$ is an edge in G. Theorem 1.1 then yields the optimal lower bound of the number of walks in terms of the size of G. We do not know of a corresponding lower bound for the number of trails (walks with no edge repeated) or paths (walks with no vertex repeated).

Recall that an $m \times n$ matrix A is said to be *bistochastic* if every row sum of A is equal to $s(A)/m$, and every column sum of A is equal to $s(A)/n$. In Section 3 we prove the following asymptotic form of Theorem 1.1:

Theorem 1.2. *Let A be an $m \times n$ matrix with nonnegative real entries. If A is bistochastic, then for all $k \geqslant 1$,*

$$s(A)^{2k} = m^{k-1}n^k s(A^{(2k)}), \qquad s(A)^{2k+1} = m^k n^k s(A^{(2k+1)}).$$

If A is not bistochastic, then there exist constants $c > 0$ and $\gamma > 1$ (depending only on A) such that for all $\ell \geqslant 1$,

$$s(A)^\ell < c\gamma^{-\ell}(mn)^{\ell/2}s(A^{(\ell)}).$$

As we show in Sections 2 and 3, both of the above theorems, though stated for arbitrary rectangular matrices with nonnegative entries, follow from the special case of *square* matrices.

Theorem 1.2 has an immediate application. Atkinson, Moran and Watterson [1] conjectured that for a nonnegative symmetric kernel function $K(x, y)$ that is integrable (in a suitable sense) over the square $0 \leqslant x, y \leqslant a$, the inequality

$$\int_0^a \int_0^a K_\ell(x, y)\, \mathrm{d}x\, \mathrm{d}y \geqslant \frac{1}{a^{\ell-1}} \left(\int_0^a \int_0^a K(x, y)\, \mathrm{d}x\, \mathrm{d}y \right)^\ell \tag{1}$$

holds for all $\ell \geqslant 1$. Here $K_\ell(x, y)$ denotes the *ℓ-th order iterate of* $K(x, y)$, which is defined recursively by

$$K_1(x, y) = K(x, y), \qquad K_\ell(x, y) = \int_0^a K_{\ell-1}(x, t) K(t, y)\, \mathrm{d}t.$$

Beesack [2] showed that the Atkinson-Moran-Watterson conjecture follows from the matrix identities of Mulholland and Smith described above. Using Beesack's ideas together with Theorem 1.2, we prove in Section 4 the following asymptotic form of the Atkinson-Moran-Watterson inequality (1):

Theorem 1.3. *Let $K(x, y)$ be a nonnegative symmetric kernel function that is integrable over the square $0 \leqslant x, y \leqslant a$, and consider the function $f(x) = \int_0^a K(x, y)\, \mathrm{d}y$ defined on the interval $0 \leqslant x \leqslant a$. If $f(x)$ is constant almost everywhere, then for all $\ell \geqslant 1$*

$$\int_0^a \int_0^a K_\ell(x, y)\, \mathrm{d}x\, \mathrm{d}y = \frac{1}{a^{\ell-1}} \left(\int_0^a \int_0^a K(x, y)\, \mathrm{d}x\, \mathrm{d}y \right)^\ell.$$

If not, there exist constants $c > 0$ and $\gamma > 1$ (depending only on K) such that for all $\ell \geqslant 1$

$$\int_0^a \int_0^a K_\ell(x, y)\, \mathrm{d}x\, \mathrm{d}y > \frac{c\gamma^\ell}{a^{\ell-1}} \left(\int_0^a \int_0^a K(x, y)\, \mathrm{d}x\, \mathrm{d}y \right)^\ell.$$

Remark 1.4. Using an approximation argument as in the proof of Theorem 1.3, Theorem 1.1 can be also applied to establish an analogue to inequalities (1) and Theorem 1.3 in the case of nonsymmetric kernel functions. Let $K(x, y)$ be any nonnegative kernel function that is integrable on the rectangle $0 \leqslant x \leqslant a$, $0 \leqslant y \leqslant b$ and let K_ℓ be the ℓ-th order iterate of K defined by $K_1(x, y) = K(x, y)$ and for each integer $k \geq 1$,

$$K_{2k}(x, x') = \int_0^b K_{2k-1}(x, y) K(x', y)\, \mathrm{d}y,$$

$$K_{2k+1}(x, y) = \int_0^a K_{2k}(x, x') K(x', y)\, \mathrm{d}x'.$$

In this case, inequalities (1) become

$$\int_0^a \int_0^b K_{2k+1}(x, y)\, \mathrm{d}x\, \mathrm{d}y \geq \frac{1}{a^k b^k} \left(\int_0^a \int_0^b K(x, y)\, \mathrm{d}x\, \mathrm{d}y \right)^{2k+1}$$

$$\int_0^a \int_0^a K_{2k}(x, x')\, \mathrm{d}x\, \mathrm{d}x' \geq \frac{1}{a^{k-1} b^k} \left(\int_0^a \int_0^b K(x, y)\, \mathrm{d}x\, \mathrm{d}y \right)^{2k}.$$

The analogue of Theorem 1.3 is then obvious.

2 Matrix inequality

Given a matrix $A = (a_{i,j})$ and an integer $\ell \geqslant 0$, we denote by $a^{(\ell)}_{i,j}$ the (i,j)-th entry of $A^{(\ell)}$, so that $A^{(\ell)} = (a^{(\ell)}_{i,j})$. This notation will be used often in the sequel.

Lemma 2.1. *Let $B = (b_{i,j})$ be a $d \times d$ matrix with nonnegative real entries. For any two sequences $\{\alpha_i\}$ and $\{\beta_i\}$ of nonnegative real numbers, the following inequality holds:*

$$(I'_2): \qquad \sum_{i,j=1}^{d} \alpha_i\beta_i b_{i,j} \leqslant d^{\frac{1}{2}} \Big(\sum_{i,j=1}^{d} \alpha_i^2 \beta_j^2 b^{(2)}_{i,j} \Big)^{\frac{1}{2}}.$$

Proof. To prove the lemma, we apply the Cauchy-Schwarz inequality twice as follows:

$$\sum_{i,j=1}^{d} \alpha_i\beta_i b_{i,j} = \sum_{i,k=1}^{d} \alpha_i\beta_i b_{i,k} \leqslant d^{\frac{1}{2}} \Big(\sum_{k=1}^{d} \Big(\sum_{i=1}^{d} \alpha_i\beta_i b_{i,k} \Big)^2 \Big)^{\frac{1}{2}}. \tag{2}$$

$$\begin{aligned}
\sum_{i,j=1}^{d} \alpha_i\beta_i b_{i,j} &\leqslant d^{\frac{1}{2}} \Big(\sum_{i,j,k=1}^{d} \alpha_i\alpha_j\beta_i\beta_j b_{i,k} b_{j,k} \Big)^{\frac{1}{2}} \\
&= d^{\frac{1}{2}} \Big(\sum_{i,j=1}^{d} \alpha_i\alpha_j\beta_i\beta_j b^{(2)}_{i,j} \Big)^{\frac{1}{2}} \\
&= d^{\frac{1}{2}} \Big(\sum_{i,j=1}^{d} \alpha_i\beta_j (b^{(2)}_{i,j})^{\frac{1}{2}} \cdot \alpha_j\beta_i (b^{(2)}_{j,i})^{\frac{1}{2}} \Big)^{\frac{1}{2}} \\
&\leqslant d^{\frac{1}{2}} \Big(\sum_{i,j=1}^{d} \alpha_i^2\beta_j^2 b^{(2)}_{i,j} \Big)^{\frac{1}{2}}.
\end{aligned}$$

Here we have used the fact that $B^{(2)} = BB^t$ is a symmetric matrix. ∎

Theorem 2.2. *Let $B = (b_{i,j})$ be a square $d \times d$ matrix with nonnegative real entries, and let $\{\alpha_i\}$ be any sequence of nonnegative real numbers. Then for each integer $\ell \geqslant 1$, we have*

$$(I_\ell): \qquad \sum_{i,j=1}^{d} \alpha_i b_{i,j} \leqslant d^{\frac{\ell-1}{\ell}} \Big(\sum_{i,j=1}^{d} \alpha_i^\ell b^{(\ell)}_{i,j} \Big)^{\frac{1}{\ell}}.$$

Proof of Theorem 2.2. The case $\ell = 1$ is trivial while the case $\ell = 2$ is a consequence of the lemma above. We prove the general case by induction. Suppose that $p \geqslant 2$, and the inequalities $(I_1), (I_2), \ldots, (I_p)$ hold for all square matrices with nonnegative real entries. If $p = 2k-1$ is an odd integer, then the inequality (I_{p+1}) follows immediately from (I_2) and (I_k). Indeed, since $B^{(2k)} = B^{(2)(k)}$, we have

$$\sum_{i,j=1}^{d} \alpha_i b_{i,j} \leqslant d^{\frac{1}{2}} \Big(\sum_{i,j=1}^{d} \alpha_i^2 b^{(2)}_{i,j} \Big)^{\frac{1}{2}} \leqslant d^{\frac{1}{2}} \Big(d^{\frac{k-1}{k}} \Big(\sum_{i,j=1}^{d} \alpha_i^{2k} b^{(2)(k)}_{i,j} \Big)^{\frac{1}{k}} \Big)^{\frac{1}{2}}. \tag{3}$$

Thus

$$\sum_{i,j=1}^{d} \alpha_i b_{i,j} \leqslant d^{\frac{2k-1}{2k}} \Big(\sum_{i,j=1}^{d} \alpha_i^{2k} b^{(2k)}_{i,j} \Big)^{\frac{1}{2k}}.$$

If $p = 2k$ is an even integer, then the inequality (I_{p+1}) follows from Hölder's inequality, and the inequalities (I_k) and (I'_2). Indeed, by Hölder's inequality, we have

$$\sum_{i,j=1}^{d} \alpha_i b_{i,j} \leqslant d^{\frac{1}{2k+1}} \Big(\sum_{i=1}^{d} \alpha_i^{\frac{2k+1}{2k}} \Big(\sum_{j=1}^{d} b_{i,j} \Big)^{\frac{2k+1}{2k}} \Big)^{\frac{2k}{2k+1}}. \tag{4}$$

Let $\mathcal{I}$ denote the term between parentheses, and set $\beta_i = \sum_{j=1}^{d} b_{i,j}$ for each i. Then

$$\mathcal{I} = \sum_{i=1}^{d} \alpha_i^{\frac{2k+1}{2k}} \Big(\sum_{j=1}^{d} b_{i,j} \Big)^{\frac{2k+1}{2k}} = \sum_{i,j=1}^{d} \alpha_i^{\frac{2k+1}{2k}} \beta_i^{\frac{1}{2k}} b_{i,j}.$$

Applying (I_k), it follows that

$$\mathcal{I} \leqslant d^{\frac{k-1}{k}} \Big(\sum_{i,j=1}^{d} \alpha_i^{\frac{2k+1}{2}} \beta_i^{\frac{1}{2}} b_{i,j}^{(k)} \Big)^{\frac{1}{k}}.$$

Applying the lemma to the sequences $\{\alpha_i^{\frac{2k+1}{2}}\}$ and $\{\beta_i^{\frac{1}{2}}\}$, and using the fact that $B^{(k)(2)} = B^{(2k)}$, we see that

$$\mathcal{I} \leqslant d^{\frac{k-1}{k}} \Big(d^{\frac{1}{2}} \Big(\sum_{i,j=1}^{d} \alpha_i^{2k+1} \beta_j b_{i,j}^{(k)(2)} \Big)^{\frac{1}{2}} \Big)^{\frac{1}{k}} = d^{\frac{2k-1}{2k}} \Big(\sum_{i,j=1}^{d} \alpha_i^{2k+1} \beta_j b_{i,j}^{(2k)} \Big)^{\frac{1}{2k}}.$$

Putting everything together, we have therefore shown that

$$\sum_{i,j=1}^{d} \alpha_i b_{i,j} \leqslant d^{\frac{2k}{2k+1}} \Big(\sum_{i,j=1}^{d} \alpha_i^{2k+1} \beta_j b_{i,j}^{(2k)} \Big)^{\frac{1}{2k+1}}.$$

Finally, note that

$$\sum_{j=1}^{d} \beta_j b_{i,j}^{(2k)} = \sum_{\ell=1}^{d} b_{i,\ell}^{(2k)} \beta_\ell = \sum_{j,\ell=1}^{d} b_{i,\ell}^{(2k)} b_{\ell,j} = \sum_{j=1}^{d} b_{i,j}^{(2k+1)}$$

since $B^{(2k+1)} = B^{(2k)} B$. Consequently,

$$\sum_{i,j=1}^{d} \alpha_i b_{i,j} \leqslant d^{\frac{2k}{2k+1}} \Big(\sum_{i,j=1}^{d} \alpha_i^{2k+1} b_{i,j}^{(2k+1)} \Big)^{\frac{1}{2k+1}} \tag{5}$$

and (I_{p+1}) holds for the case $p = 2k$. Theorem 2.2 now follows by induction. □

Proof of Theorem 1.1. For the case of square matrices, Theorem 1.1 follows immediately from Theorem 2.2. Indeed, taking $\alpha_i = 1$ for each i, the inequality (I_ℓ) yields the corresponding inequality in Theorem 1.1.

Now, let A be an $m \times n$ matrix with nonnegative real entries, put $d = mn$, and let B be the $d \times d$ matrix with nonnegative real entries defined as the tensor product $B = A \otimes J_{n,m}$, where $J_{n,m}$ is the $n \times m$ matrix with every entry equal to 1. For any integers $\ell, k \geqslant 0$, the relations

$$B^{(\ell)} = A^{(\ell)} \otimes J_{n,m}^{(\ell)}, \quad s\big(B^{(\ell)}\big) = s\big(A^{(\ell)}\big) s\big(J_{n,m}^{(\ell)}\big),$$

$$s\big(J_{n,m}^{(2k)}\big) = m^k n^{k+1}, \quad s\big(J_{n,m}^{(2k+1)}\big) = m^{k+1} n^{k+1}.$$

are easily checked. In particular, $s(B) = mns(A)$. Applying Theorem 1.1 to the matrix B and using these identities, the inequalities of Theorem 1.1 follow for the matrix A. □

3 Asymptotic matrix inequality

As will be shown below, Theorem 1.2 is a consequence of the following more precise theorem for square matrices:

Theorem 3.1. *Let B be a square $d \times d$ matrix with nonnegative real entries and $s(B) \neq 0$. Let λ be the largest eigenvalue of $B^{(2)} = BB^t$, and put $\gamma = \lambda d^2 / s(B)^2$. Then $\gamma \geqslant 1$, and there exists a constant $c > 0$ (depending only on B) such that for all integers $\ell \geqslant 0$,*

$$s(B)^\ell < c \gamma^{-\frac{\ell}{2}} d^{\ell-1} s(B^{(\ell)}). \tag{6}$$

Moreover, the following assertions are equivalent:

(*a*) $\gamma = 1$,

(*b*) $s(B)^\ell = d^{\ell-1} s(B^{(\ell)})$ *for every integer* $\ell \geqslant 0$,

(*c*) $s(B)^\ell = d^{\ell-1} s(B^{(\ell)})$ *for some integer* $\ell \geqslant 3$,

(*d*) B *is bistochastic.*

Proof. We express $B^{(2)} = BB^t$ in the form $B^{(2)} = U^t D U$, where $U = (u_{i,j})$ is an orthogonal matrix, and D is a diagonal matrix $\mathrm{diag}(\lambda_1, \ldots, \lambda_d)$ with $\lambda_1 \geqslant \ldots \geqslant \lambda_d \geqslant 0$. Here $\lambda = \lambda_1$. For each $\nu = 1, \ldots, d$, let E_ν be the projection matrix whose (ν,ν)-th entry is 1, and all other entries are equal to 0. Put $A_\nu = U^t E_\nu U$ for each ν. Then for all integers $k \geqslant 0$,

$$B^{(2k)} = \sum_{\nu=1}^{d} \lambda_\nu^k A_\nu, \qquad B^{(2k+1)} = \sum_{\nu=1}^{d} \lambda_\nu^k A_\nu B.$$

By a straightforward calculation, we see that for each ν

$$s(A_\nu) = \Big(\sum_{i=1}^{d} u_{\nu,i} \Big)^2, \qquad s(A_\nu B) = \Big(\sum_{i=1}^{d} u_{\nu,i} \Big) \Big(\sum_{j,k=1}^{d} u_{\nu,k} b_{k,j} \Big). \tag{7}$$

In particular, $s(A_\nu) \geqslant 0$. By Theorem 2.2, it follows that

$$\frac{s(B)^2}{d} \leqslant s(B^{(2)}) = \sum_{\nu=1}^{d} \lambda_\nu s(A_\nu) \leqslant \lambda \sum_{\nu=1}^{d} s(A_\nu) = \lambda d. \tag{8}$$

Therefore, $\gamma = \dfrac{\lambda d^2}{s(B)^2} \geqslant 1$. Now, from the definition of γ, we have

$$\frac{\gamma^{\frac{\ell}{2}} s(B)^\ell}{d^{\ell-1} s(B^{(\ell)})} = d \frac{\lambda^{\frac{\ell}{2}}}{s(B^{(\ell)})}.$$

Then, in order to show inequality (6), we will show that the $\lambda^{\frac{\ell}{2}} / s(B^{(\ell)})$ are bounded above by a constant that is independent of ℓ. Indeed, let $C_\ell = B^{(\ell)} / s(B^{(\ell)})$ for every $\ell \geqslant 0$. Since each C_ℓ has nonnegative real entries, and $s(C_\ell) = 1$, the entries of C_ℓ all lie in the closed interval $[0,1]$. Thus the entries of the matrices $UC_{2k}U^t$ and $UC_{2k+1}B^tU^t$ are bounded by a constant that depends only on B. Noting that for each nonnegative integer k, we have

$$UC_{2k}U^t = \frac{D^k}{s(B^{(2k)})}, \qquad UC_{2k+1}B^tU^t = \frac{D^{k+1}}{s(B^{(2k+1)})},$$

and on examining the $(1,1)$th entry for each of these matrices, we see that $\lambda^k / s(B^{(2k)})$ and $\lambda^{k+1} / s(B^{(2k+1)})$ are both bounded above by a constant that is independent of k. Consequently, inequality (6) holds.

(*a*)⇒(*b*): If $\gamma = 1$, then $\lambda d = s(B)^2/d$, hence from (8) we see that $s(A_\nu) = 0$ whenever $\lambda_\nu \neq \lambda$. By (7), we also have that $s(A_\nu B) = 0$ whenever $\lambda_\nu \neq \lambda$. Thus

$$\begin{aligned} s(B^{(2k)}) &= \sum_{\nu=1}^{d} \lambda_\nu^k s(A_\nu) = \lambda^k \sum_{\nu:\lambda_\nu=\lambda} s(A_\nu) \\ &= \lambda^k \sum_{\nu=1}^{d} s(A_\nu) = \lambda^k d = \frac{s(B)^{2k}}{d^{2k-1}}, \\ s(B^{(2k+1)}) &= \sum_{\nu=1}^{d} \lambda_\nu^k s(A_\nu B) = \lambda^k \sum_{\nu:\lambda_\nu=\lambda} s(A_\nu B) \\ &= \lambda^k \sum_{\nu=1}^{d} s(A_\nu B) = \lambda^k s(B) = \frac{s(B)^{2k+1}}{d^{2k}}. \end{aligned}$$

$(b){\Rightarrow}(a)$: If (b) holds, then inequality (6) implies $1 < c\gamma^{-\frac{\ell}{2}}$ for some $\gamma \geq 1$ and all integers $\ell \geq 0$. This forces $\gamma = 1$.

$(b){\Rightarrow}(c)$: Trivial.

$(c){\Rightarrow}(d)$: Suppose that $\ell = 2k+1 \geqslant 3$ is an odd integer such that $s(B)^\ell = d^{\ell-1}s(B^{(\ell)})$. Taking every $\alpha_i = 1$ in the proof of Theorem 2.2, our hypothesis means that equality holds in (5), hence (4) must also hold with equality:

$$\sum_{i,j=1}^{d} b_{i,j} = d^{\frac{1}{2k+1}} \left(\sum_{i=1}^{d} \Big(\sum_{j=1}^{d} b_{i,j} \Big)^{\frac{2k+1}{2k}} \right)^{\frac{2k}{2k+1}}.$$

By Hölder's inequality, this is only possible if all of the row sums of B are equal. Since ℓ is odd and s is transpose-invariant, we also have

$$s(B^t)^\ell = d^{\ell-1} s\big((B^{(\ell)})^t\big) = d^{\ell-1} s\big((B^t)^{(\ell)}\big).$$

Thus all of the row sums of B^t are equal as well, and B is bistochastic.

Now suppose that $\ell = 2k \geqslant 4$ is an even integer such that $s(B)^\ell = d^{\ell-1}s(B^{(\ell)})$. By taking every $\alpha_i = 1$ in (3), we see that $s(B)^2 = ds(B^{(2)})$. Then, taking every $\alpha_i = \beta_i = 1$ in the proof of the lemma, we see that equality holds in (2) which is only possible if all of the column sums of B are equal. Therefore $s(BA) = \beta s(A)$ for every $d \times d$ matrix A, where $\beta = s(B)/d$ is the sum of each column of B. In particular,

$$\begin{aligned} s(B)^\ell &= d^{\ell-1} s\big(B^{(\ell)}\big) = d^{\ell-1}\beta s\big((B^t)^{(\ell-1)}\big) \\ &= d^{\ell-1}\beta s\big((B^{(\ell-1)})^t\big) = d^{\ell-1}\beta s(B^{(\ell-1)}), \end{aligned}$$

thus $s(B)^{\ell-1} = d^{\ell-2}s(B^{(\ell-1)})$. Since $\ell - 1$ is odd, we can apply the previous result to conclude that B is bistochastic.

$(d){\Rightarrow}(b)$: Suppose B is bistochastic, with every row or column sum equal to $\beta = s(B)/d$. For any $d \times d$ matrix A, one has $s(AB) = \beta s(A)$ and $s(AB^t) = \beta s(A)$. In particular, $s(B^{(2k+1)}) = \beta s(B^{(2k)})$ and $s(B^{(2k+2)}) = \beta s(B^{(2k+1)})$ for all $k \geqslant 0$. Consequently,

$$s(B^{(\ell)}) = \beta^{\ell-1}s(B) = \frac{s(B)^\ell}{d^{\ell-1}}, \qquad \ell \geqslant 0.$$

This completes the proof. ☐

Corollary 3.2. *Let B be a square $d \times d$ matrix with nonnegative real entries and $s(B) \neq 0$. Let β_j be the j-th column sum of B for each j, and put*

$$\delta = 1 + \frac{1}{2s(B)^2} \sum_{i,j=1}^{d} (\beta_i - \beta_j)^2.$$

Then there exists a constant $c > 0$ (depending only on B) such that for all $\ell \geqslant 0$, we have

$$s(B)^\ell < c\delta^{-\frac{\ell}{2}} d^{\ell-1} s(B^{(\ell)}).$$

Proof. Note first that for *any* $d \times d$ matrix B, if β_j denotes the j-th column sum of B, then it is easily seen that

$$s(B^{(2)}) = \frac{s(B)^2}{d} + \frac{1}{2d} \sum_{i,j=1}^{d} (\beta_i - \beta_j)^2. \tag{9}$$

Using the notation of Theorem 3.1 and applying the relations (8) and (9), we have

$$\gamma = \frac{\lambda d^2}{s(B)^2} \geqslant \frac{ds(B^{(2)})}{s(B)^2} = 1 + \frac{1}{2s(B)^2} \sum_{i,j=1}^{d} (\beta_i - \beta_j)^2 = \delta.$$

The corollary therefore follows from (6). ☐

Proof of Theorem 1.2. Given an $m \times n$ matrix A with nonnegative real entries, we proceed as in the proof of Theorem 1.1: put $d = mn$, and let $B = A \otimes J_{n,m}$. Note that A is bistochastic if and only if B is bistochastic. Applying the corollary above to B, Theorem 1.2 follows immediately for the matrix A. The details are left to the reader. ☐

4 Asymptotic kernel inequality

Proof of Theorem 1.3. By changing variables if necessary, we can assume that $a = 1$. For simplicity, we will also assume that $K(x,y)$ is continuous. Consider the function $f(x)$ defined by

$$f(x) = \int_0^1 K(x,y)\,\mathrm{d}y, \qquad x \in [0,1].$$

If $f(x)$ is a constant function, then since $K(x,y)$ is symmetric, the equality

$$\int_0^1 \int_0^1 K_\ell(x,y)\,\mathrm{d}x\,\mathrm{d}y = \Big(\int_0^1 \int_0^1 K(x,y)\,\mathrm{d}x\,\mathrm{d}y\Big)^\ell$$

for all $\ell \geqslant 1$ follows from an easy inductive argument.

Now suppose that $f(x)$ is not constant, and let m and M denote respectively the minimum and maximum value of $f(x)$ on $[0,1]$. Choose $\varepsilon > 0$ such that $4\varepsilon < M - m$. For every integer $d \geqslant 1$, let $\mathscr{U}_i^{[d]}$ be the open interval

$$\mathscr{U}_i^{[d]} = \Big(\frac{i-1}{d}, \frac{i}{d}\Big), \qquad 1 \leqslant i \leqslant d,$$

and let $\mathscr{U}_{i,j}^{[d]}$ be the rectangle $\mathscr{U}_i^{[d]} \times \mathscr{U}_j^{[d]}$ for $1 \leqslant i,j \leqslant d$. Let $K^{[d]}(x,y)$ be defined on $[0,1] \times [0,1]$ as follows:

$$\begin{cases} \min\Big\{K(s,t) : (s,t) \in \overline{\mathscr{U}_{i,j}^{[d]}}\Big\} & \text{if } (x,y) \in \mathscr{U}_{i,j}^{[d]} \text{ for some } 1 \leqslant i,j \leqslant d \\ K(x,y) & \text{otherwise.} \end{cases}$$

Here $\overline{\mathscr{U}_{i,j}^{[d]}}$ denotes the closure of $\mathscr{U}_{i,j}^{[d]}$. Noting that $K^{[d]}(x,y)$ is constant on each rectangle $\mathscr{U}_{i,j}^{[d]}$, let $B_{[d]}$ be the $d \times d$ matrix whose (i,j)-th entry is equal to $K^{[d]}(\mathscr{U}_{i,j}^{[d]})$. Let $K_\ell^{[d]}(x,y)$ denote the ℓ-th order iterate of $K^{[d]}(x,y)$ for each $\ell \geqslant 1$. Then

$$K_\ell^{[d]}(x,y) = \int_0^1 K_{\ell-1}^{[d]}(x,t) K^{[d]}(t,y)\,\mathrm{d}t = \sum_{k=1}^d \int_{\mathscr{U}_k^{[d]}} K_{\ell-1}^{[d]}(x,t) K^{[d]}(t,y)\,\mathrm{d}t.$$

It follows by induction that $K_\ell^{[d]}(x,y)$ is also constant on each rectangle $\mathscr{U}_{i,j}^{[d]}$, and

$$K_\ell^{[d]}(\mathscr{U}_{i,j}^{[d]}) = \frac{1}{d} \sum_{k=1}^d K_{\ell-1}^{[d]}(\mathscr{U}_{i,k}^{[d]}) K^{[d]}(\mathscr{U}_{k,j}^{[d]});$$

by induction, this is the (i,j)-th entry of the matrix $\dfrac{1}{d^{\ell-1}} B_{[d]}^{(\ell)}$. In other words,

$$\Big(K_\ell^{[d]}(\mathscr{U}_{i,j}^{[d]})\Big) = \frac{1}{d^{\ell-1}} B_{[d]}^{(\ell)}, \qquad \text{for all } \ell, d \geqslant 1. \tag{10}$$

Now since $f(x)$ is continuous, we can choose d sufficiently large such that for some integers $1 \leqslant i_m, i_M \leqslant d$, we have

$$f(x) < m + \varepsilon, \qquad \text{for all } x \in \mathscr{U}_{i_m}^{[d]},$$

$$f(x) > M - \varepsilon, \qquad \text{for all } x \in \mathscr{U}_{i_M}^{[d]}.$$

Taking d larger if necessary, we can further assume that

$$0 \leqslant K(x,y) - K^{[d]}(x,y) < \varepsilon$$

for all $0 \leqslant x, y \leqslant 1$. Fixing this value of d, we define

$$\gamma = 1 + \frac{\varepsilon^2}{2d^2\left(\int_0^1\int_0^1 K(x,y)\,\mathrm{d}x\,\mathrm{d}y\right)^2}.$$

Finally, since $\gamma^{-\frac{1}{4}} < 1$, we can choose e sufficiently large so that $K^{[de]}(x,y) > \gamma^{-\frac{1}{4}}K(x,y)$ for all $0 \leqslant x, y \leqslant 1$. For this value of e, we therefore have

$$\int_0^1\int_0^1 K^{[de]}(x,y)\,\mathrm{d}x\,\mathrm{d}y > \gamma^{-\frac{1}{4}}\int_0^1\int_0^1 K(x,y)\,\mathrm{d}x\,\mathrm{d}y.$$

By the corollary to Theorem 3.1 applied to the matrix $B_{[de]}$, there exists a constant $c > 0$, which is independent of ℓ, such that

$$s\big(B_{[de]}\big)^{\ell} < c\delta^{-\frac{\ell}{2}}(de)^{\ell-1}s\big(B_{[de]}^{(\ell)}\big)$$

for all integers $\ell \geqslant 0$, where

$$\delta = 1 + \frac{1}{2s\big(B_{[de]}\big)^2}\sum_{i,j=1}^{de}\big(\beta_{[de],i} - \beta_{[de],j}\big)^2.$$

Here $\beta_{[de],j}$ denotes the j-th column sum of $B_{[de]}$ for each j. We now claim that $\delta > \gamma$. Granting this fact for the moment, we apply (10) to $K^{[de]}(x,y)$ and obtain:

$$\begin{aligned}
\int_0^1\int_0^1 K_\ell(x,y)\,\mathrm{d}x\,\mathrm{d}y &\geqslant \int_0^1\int_0^1 K_\ell^{[de]}(x,y)\,\mathrm{d}x\,\mathrm{d}y = \frac{1}{(de)^2}\sum_{i,j=1}^{de}K_\ell^{[de]}\Big(\mathscr{U}_{i,j}^{[de]}\Big)\\
&= \frac{1}{(de)^{\ell+1}}s\big(B_{[de]}^{(\ell)}\big) > c^{-1}\delta^{\frac{\ell}{2}}(de)^{-2\ell}s\big(B_{[de]}\big)^{\ell}\\
= c^{-1}\delta^{\frac{\ell}{2}}\Bigg(\frac{1}{(de)^2}\sum_{i,j=1}^{de}K^{[de]}\Big(\mathscr{U}_{i,j}^{[de]}\Big)\Bigg)^{\ell} &= c^{-1}\delta^{\frac{\ell}{2}}\Bigg(\int_0^1\int_0^1 K^{[de]}(x,y)\,\mathrm{d}x\,\mathrm{d}y\Bigg)^{\ell}\\
> c^{-1}\delta^{\frac{\ell}{2}}\gamma^{-\frac{\ell}{4}}\Bigg(\int_0^1\int_0^1 K(x,y)\,\mathrm{d}x\,\mathrm{d}y\Bigg)^{\ell} &> c^{-1}\gamma^{\frac{\ell}{4}}\Bigg(\int_0^1\int_0^1 K(x,y)\,\mathrm{d}x\,\mathrm{d}y\Bigg)^{\ell}.
\end{aligned}$$

This completes the proof of the theorem modulo our claim that $\delta > \gamma$. To see this, let $\mathscr{V}$ be any interval of the form $\mathscr{U}_i^{[de]}$ such that $\mathscr{V} \subset \mathscr{U}_{i_m}^{[d]}$. Note that there are e such intervals. Since $B^{[de]}$ is a symmetric matrix, the column sum $\beta_{[de],\mathscr{V}}$ of $B_{[de]}$ corresponding to the interval $\mathscr{V}$ is equal to the "$\mathscr{V}$-th" row sum, which can be bounded as follows:

$$\begin{aligned}
\beta_{[de],\mathscr{V}} = \sum_{j=1}^{de}K^{[de]}\big(\mathscr{V},\mathscr{U}_j^{[de]}\big) = (de)^2\int_{\mathscr{V}}\int_0^1 K^{[de]}(x,y)\,\mathrm{d}y\,\mathrm{d}x&\\
\leqslant (de)^2\int_{\mathscr{V}}\int_0^1 K(x,y)\,\mathrm{d}y\,\mathrm{d}x = (de)^2\int_{\mathscr{V}}f(x)\,\mathrm{d}x < de(m+\varepsilon).&
\end{aligned}$$

Similarly, let $\mathscr{W}$ be any interval of the form $\mathscr{U}_i^{[de]}$ such that $\mathscr{W} \subset \mathscr{U}_{i_M}^{[d]}$. Again, there are e such intervals, and by a similar calculation, the column sum $\beta_{[de],\mathscr{W}}$ satisfies the bound

$$\beta_{[de],\mathscr{W}} = \sum_{j=1}^{de}K^{[de]}\big(\mathscr{W},\mathscr{U}_j^{[de]}\big) > de(M - 2\varepsilon).$$

Thus

$$\sum_{i,j=1}^{de} \big(\beta_{[de],i} - \beta_{[de],j}\big)^2 \geqslant \sum_{\mathscr{V},\mathscr{W}} \big(\beta_{[de],\mathscr{W}} - \beta_{[de],\mathscr{V}}\big)^2 > d^2e^4(M-m-3\varepsilon)^2 > d^2e^4\varepsilon^2.$$

On the other hand, we have

$$s\big(B_{[de]}\big) = (de)^2 \int_0^1 \int_0^1 K^{[de]}(x,y)\,\mathrm{d}x\,\mathrm{d}y \leqslant (de)^2 \int_0^1 \int_0^1 K(x,y)\,\mathrm{d}x\,\mathrm{d}y,$$

and the claim follows. □

Bibliography

[1] F. V. Atkinson, G. A. Watterson, and P. A. P. Moran. A matrix inequality. *Quart. J. Math. Oxford Ser. (2)*, 11:137–140, 1960. (pp. 1 and 2).

[2] Paul R. Beesack. Inequalities involving iterated kernels and convolutions. *Univ. Beograd. Publ. Elektrotehn. Fak. Ser. Mat. Fiz.*, 274–301:11–16, 1969. (p. 2).

[3] S. P. H. Mandel and I. M. Hughes. Change in mean viability at a multi-allelic locus in a population under random mating. *Nature*, 182:63–64, 1958. (p. 1).

[4] H. P. Mulholland and C. A. B. Smith. An inequality arising in genetical theory. *Amer. Math. Monthly*, 66:673–683, 1959. (p. 1).

Keywords. Matrix inequality, iterated kernel.

2000 *Mathematics Subject Classification.* 15A45.