A Cauchy–Khinchin integral inequality

Bao Qi Feng, Andrew Tonge

A R T I C L E   I N F O

Article history:
Received 12 October 2009
Accepted 15 April 2010
Available online 13 May 2010
Submitted by R.A. Brualdi

AMS classification:
15A39
15A48
15A69
46G12
60E15

Keywords:
Cauchy inequality
Cauchy–Schwarz inequality
Conditional expectation
Khinchin inequality
The Monotone Convergence Theorem
Probability measure
Riemann integral
Simple function
Tensor product of matrices

T A B S T R A C T

This paper discusses some Cauchy–Khinchin integral inequalities. Khinchin [2] obtained an inequality relating the row and column sums of 0-1 matrices in the course of his work on number theory. As pointed out by van Dam [6], Khinchin’s inequality can be viewed as a generalization of the classical Cauchy inequality. Van Dam went on to derive analogs of Khinchin’s inequality for arbitrary matrices. We carry this work forward, first by proving even more than general matrix results, and then by formulating them in a way that allows us to apply limiting arguments to create new integral inequalities for functions of two variables. These integral inequalities can be interpreted as giving information about conditional expectations.

The classical Cauchy inequality for real numbers \( x_1, \ldots, x_m \) can be written in the form

\[
\left( \sum_{i=1}^{m} x_i \right)^2 \leq m \sum_{i=1}^{m} x_i^2.
\]

(1)
Over the years, this inequality has been generalized in many ways, and our focus in this paper is to build on a new generalization. This new generalization is simultaneously a generalization of another apparently unrelated inequality proved by Khinchin [2,3] in the course of his work in number theory.

Khinchin’s inequality asserts that if an $m \times n$ matrix of 0’s and 1’s has row sums $r_i$ ($1 \leq i \leq m$) and column sums $c_j$ ($1 \leq j \leq n$), and if all its entries sum to $\sigma$, then

$$\sum_{i=1}^{m} r_i^2 + \sum_{j=1}^{n} c_j^2 \leq \sigma^2 + l^2 \sigma,$$

(2)

where $l = \max\{m, n\}$.

A common generalization of the Cauchy and Khinchin inequalities was given in 1998 by van Dam [6].

Theorem 1 (van Dam). Let $X = (x_{ij})$ be a real $m \times n$ matrix. Then

$$m \sum_{i=1}^{m} \left( \sum_{j=1}^{n} x_{ij} \right)^2 + n \sum_{j=1}^{n} \left( \sum_{i=1}^{m} x_{ij} \right)^2 \leq \left( \sum_{i=1}^{m} \sum_{j=1}^{n} x_{ij} \right)^2 + mn \sum_{i=1}^{m} \sum_{j=1}^{n} x_{ij}^2$$

(3)

with equality if and only if $x_{ij} = y_i + z_j$ for some real vectors $y$ and $z$, and all $i$ and $j$.

To see why van Dam’s theorem is an extension of Cauchy’s inequality, take $X$ to be the $m \times 2$ matrix

$$\begin{pmatrix}
    x_1 & -x_1 \\
    \vdots & \vdots \\
    x_m & -x_m
\end{pmatrix}.$$

Then (3) immediately reduces to (1). On the other hand, if $X$ is an $m \times n$ matrix of 0’s and 1’s with row sums $r_i$ ($1 \leq i \leq m$), column sums $c_j$ ($1 \leq j \leq n$) and if all its entries sum to $\sigma$, then van Dam’s theorem reduces to

$$m \sum_{i=1}^{m} r_i^2 + n \sum_{j=1}^{n} c_j^2 \leq \sigma^2 + mn \sigma.$$  

(4)

As observed in Matúš and Tuzar [4], this is an improvement of Khinchin’s inequality (2).

The following result shows how van Dam’s theorem can be interpreted as an integral inequality.

Theorem 2. Suppose $f(x, y)$ is Riemann integrable on the square $\Delta = \{(x, y) : 0 \leq x < 1, 0 \leq y \leq 1\}$. If for each $x \in [0, 1]$, the Riemann integral $\int_0^1 f(x, y)dy$ exists and for each $y \in [0, 1]$, the Riemann integral $\int_0^1 f(x, y)dx$ exists, then

$$\int_0^1 \left( \int_0^1 f(x, y)dy \right)^2 dx + \int_0^1 \left( \int_0^1 f(x, y)dx \right)^2 dy \leq \left( \int_0^1 \int_0^1 f(x, y)dxdy \right)^2 + \int_0^1 \int_0^1 f^2(x, y)dxdy.$$  

(5)

Proof. That $f(x, y)$ is Riemann integrable on the square $\Delta = \{(x, y) : 0 \leq x < 1, 0 \leq y \leq 1\}$ ensures the four integrals in the inequality (5) exist. Thus, under any partition of $\Delta$ the Darboux sums converge to the integrals respectively when the maximum diameter of the pieces in the partition approach to zero. Hence we can choose the special partition of $\Delta$ as follows.

$$P_y = [x_i, x_{i+1}; y_j, y_{j+1}],$$

with $x_i = i/m$ ($1 \leq i \leq m$), $y_j = j/n$ ($1 \leq j \leq n$).
Denoting \( f(x_i, y_j) = f_{ij} \). By Theorem 1, we have
\[
m \sum_{i=1}^{m} \left( \sum_{j=1}^{n} f_{ij} \right)^2 + n \sum_{j=1}^{n} \left( \sum_{i=1}^{m} f_{ij} \right)^2 \leq \left( \sum_{i=1}^{m} \sum_{j=1}^{n} f_{ij} \right)^2 + mn \sum_{i=1}^{m} \sum_{j=1}^{n} f_{ij}^2.
\]

Divide both sides of it by \( m^2 n^2 \) to obtain
\[
\frac{1}{mn^2} \sum_{i=1}^{m} \left( \sum_{j=1}^{n} f_{ij} \right)^2 + \frac{1}{m^2 n} \sum_{j=1}^{n} \left( \sum_{i=1}^{m} f_{ij} \right)^2 \leq \frac{1}{m^2 n^2} \left( \sum_{i=1}^{m} \sum_{j=1}^{n} f_{ij} \right)^2 + \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} f_{ij}^2.
\]
that is
\[
\sum_{i=1}^{m} \left( \frac{n}{n} \sum_{j=1}^{n} f_{ij} \right)^2 \frac{1}{m} + \sum_{j=1}^{n} \left( \frac{m}{m} \sum_{i=1}^{m} f_{ij} \right)^2 \frac{1}{n} \leq \left( \sum_{i=1}^{m} \sum_{j=1}^{n} f_{ij} \right)^2 + \sum_{i=1}^{m} \sum_{j=1}^{n} f_{ij}^2 \frac{1}{mn}.
\]
Let \( m, n \to \infty \), we have the inequality (5) in Theorem 2. \( \Box \)

We wish to point out the relationship between the four integrals in inequality (5). From the integration form of the Cauchy inequality we immediately have
\[
\left( \int_0^1 \int_0^1 f(x, y) \, dx \, dy \right)^2 \leq \int_0^1 \left( \int_0^1 f(x, y) \, dy \right)^2 \, dx \leq \int_0^1 \int_0^1 f^2(x, y) \, dx \, dy,
\]
\[
\left( \int_0^1 \int_0^1 f(x, y) \, dx \, dy \right)^2 \leq \int_0^1 \left( \int_0^1 f(x, y) \, dx \right)^2 \, dy \leq \int_0^1 \int_0^1 f^2(x, y) \, dx \, dy.
\]
Hence
\[
2 \left( \int_0^1 \int_0^1 f(x, y) \, dx \, dy \right)^2 \leq \int_0^1 \left( \int_0^1 f(x, y) \, dx \right)^2 \, dy + \int_0^1 \left( \int_0^1 f(x, y) \, dy \right)^2 \, dx \leq 2 \int_0^1 \int_0^1 f^2(x, y) \, dx \, dy.
\]
So, although Theorem 2 might appear to be a straightforward result, it is actually quite delicate.

It is now natural to ask whether the integral inequality in Theorem 2 can be extended to an integration with respect to arbitrary probability measures? The answer is ‘Yes’, and, as before, the key is the following discrete inequality. It extends van Dam’s Theorem 1.

**Theorem 3.** Suppose \( m, n \) are positive integers, and let the non-negative real numbers \( s_i (1 \leq i \leq m) \) and \( t_j (0 \leq j \leq n) \) satisfy \( \sum_{i=1}^{m} s_i = \sum_{j=1}^{n} t_j = 1 \). Let \( X = (x_{ij}) \) be a real \( m \times n \) matrix. Then
\[
\sum_{i=1}^{m} \left( \sum_{j=1}^{n} x_{ij} t_j \right)^2 + \sum_{j=1}^{n} \left( \sum_{i=1}^{m} x_{ij} s_i \right)^2 \leq \left( \sum_{i=1}^{m} \sum_{j=1}^{n} x_{ij} s_i t_j \right)^2 + \sum_{i=1}^{m} \sum_{j=1}^{n} x_{ij}^2 s_i t_j.
\]

**Proof.** Let \( D_m = \text{diag}[s_1, \ldots, s_m] \) and \( E_n = \text{diag}[t_1, \ldots, t_n] \) be diagonal matrices with diagonal entries \( s_1, \ldots, s_m \) and \( t_1, \ldots, t_n \) respectively. Let
\[
A_m := \begin{pmatrix}
s_1 s_1 & s_1 s_2 & \cdots & s_1 s_m \\
s_2 s_1 & s_2 s_2 & \cdots & s_2 s_m \\
\vdots & \vdots & \ddots & \vdots \\
s_m s_1 & s_m s_2 & \cdots & s_m s_m
\end{pmatrix}.
\]
\[ B_n := \begin{pmatrix} t_1 t_1 & t_1 t_2 & \cdots & t_1 t_n \\ t_2 t_1 & t_2 t_2 & \cdots & t_2 t_n \\ \vdots & \vdots & \ddots & \vdots \\ t_n t_1 & t_n t_2 & \cdots & t_n t_n \end{pmatrix}, \]

and let
\[ \tilde{X} = (x_{11}, \ldots, x_{1n}, x_{21}, \ldots, x_{2n}, \ldots, x_{m1}, \ldots, x_{mn}). \]

It is easy to verify that
\[
\tilde{X}(A_m \otimes B_n + D_m \otimes E_n - D_m \otimes B_n - A_m \otimes E_n)\tilde{X}^t \\
= \left( \sum_{i=1}^m \sum_{j=1}^n x_{ij}s_it_j \right)^2 + \sum_{i=1}^m \sum_{j=1}^n x_{ij}^2s_it_j - \sum_{i=1}^m \sum_{j=1}^n x_{ij}^2s_it_j \leq \left( \sum_{i=1}^m \sum_{j=1}^n x_{ij}s_it_j \right)^2 + \sum_{i=1}^m \sum_{j=1}^n x_{ij}^2s_it_j
\]

where \( A \otimes B \) is the tensor product of \( A \) and \( B \). Recall that if \( A = (a_{ij}) \) is an \( m \times n \) matrix and \( B = (b_{ij}) \) is an \( s \times t \) matrix, then their tensor (or Kronecker) product is the \( ms \times nt \) matrix
\[
A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{bmatrix}.
\]

Hence
\[
\sum_{i=1}^m \left( \sum_{j=1}^n x_{ij}s_it_j \right)^2 \leq \left( \sum_{i=1}^m \sum_{j=1}^n x_{ij}s_it_j \right)^2 + \sum_{i=1}^m \sum_{j=1}^n x_{ij}^2s_it_j
\]

\[ \iff \tilde{X}(A_m \otimes B_n + D_m \otimes E_n - D_m \otimes B_n - A_m \otimes E_n)\tilde{X}^t \succeq 0 \]

\[ \iff A_m \otimes B_n + D_m \otimes E_n - D_m \otimes B_n - A_m \otimes E_n \succeq 0 \]

\[ \iff (D_m - A_m) \otimes (E_n - B_n) \succeq 0. \]

We know that if two matrices are positive semidefinite, then so is their tensor product. Thus, to prove (6) we just need to prove that
\[ D_m - A_m \succeq 0 \quad \text{and} \quad E_n - B_n \succeq 0. \]

Now
\[ D_m - A_m \succeq 0 \iff Y(D_m - A_m)Y^t \succeq 0 \]

for all real vectors \( Y = (y_1, \ldots, y_m) \). This is equivalent to
\[
\sum_{i=1}^m s_i y_i^2 \geq \sum_{i=1}^m s_i y_i y_i \geq 0,
\]

for all \( (y_1, \ldots, y_m) \in \mathbb{R}^m \). This, in turn, reduces to
\[
\left( \sum_{i=1}^m s_i y_i^2 \right)^{1/2} \geq \left( \sum_{i=1}^m s_i y_i \right).
\]

for all \( (y_1, \ldots, y_m) \in \mathbb{R}^m \). The last inequality holds, since, by the Cauchy–Schwarz inequality,
\[
\left( \sum_{i=1}^m s_i y_i \right)^2 = \left( \sum_{i=1}^m \sqrt{s_i} \sqrt{s_i y_i} \right)^2 \leq \left( \sum_{i=1}^m (\sqrt{s_i})^2 \right)^{1/2} \left( \sum_{i=1}^m (\sqrt{s_i} y_i)^2 \right)^{1/2} = \left( \sum_{i=1}^m s_i y_i^2 \right)^{1/2}.
\]
Similarly, \( E_n - B_n \geq 0 \), and the proof of (5) is complete. \( \Box \)

Before we generalize the inequality (5) into the probability measure spaces from the Riemann integrals, we need the following lemma, which is found in [5, Proposition 8, p. 373].

**Lemma 1.** Let \((X, \Sigma_X, \mu)\) and \((Y, \Sigma_Y, \nu)\) be measure spaces, and \((X \times Y, \Sigma_{X \times Y}, \mu \times \nu)\) be the product space. If \( f(x,y) \) is an \( \mu \times \nu \) measurable function on \( X \times Y \), and \( \int_{X \times Y} f(x,y) d(\mu \times \nu) < \infty \), then for each \( \varepsilon > 0 \) there exists a simple function \( h_\varepsilon \in \Sigma_X \times \Sigma_Y \) such that

\[
\int_{X \times Y} |f(x,y) - h_\varepsilon| \, d(\mu \times \nu) < \varepsilon.
\]

Where in Lemma 1 the definition of the simple function is reviewed as follows.

**Definition 1.** If \((X, \Sigma, \mu)\) is a measurable space, then a mapping \( f : X \to \mathbb{R} \) (the set of all real numbers) is called a \( \mu \)-simple (measurable) function if it can be expressed as

\[
f = \sum_{i=1}^{n} a_i \chi_{E_i}, \quad a_i \in \mathbb{R}, \ A_i \in \Sigma, \ 1 \leq i \leq n.
\]

**Theorem 4.** Let \((X, \Sigma_X, \mu)\) and \((Y, \Sigma_Y, \nu)\) be probability measure spaces with the product space \((X \times Y, \Sigma_{X \times Y}, \mu \times \nu)\). Suppose \( f(x,y) \) is an \( \mu \times \nu \) measurable function on \( X \times Y \). If \( f(x,y) \in L^2(X \times Y, \mu \times \nu) \), \( \int_{X \times Y} f(x,y) \, d\mu \in L^2(Y, \nu) \) and \( \int_{X \times Y} f(x,y) \, d\nu \in L^2(X, \mu) \), then

\[
\int_X \left( \int_Y f(x,y) \, d\nu \right)^2 \, d\mu + \int_Y \left( \int_X f(x,y) \, d\mu \right)^2 \, d\nu \\
\leq \left( \int_{X \times Y} f(x,y) \, d(\mu \times \nu) \right)^2 + \int_{X \times Y} f^2(x,y) \, d(\mu \times \nu).
\]

**Proof.** Suppose \( f(x,y) \) is a simple function on the product space \((X \times Y, \Sigma_{X \times Y}, \mu \times \nu)\) defined by

\[
f(x,y) = f_{ij}, \quad \text{if } (x,y) \in A_i \times B_j, \quad (1 \leq i \leq m \text{ and } 1 \leq j \leq n),
\]

where \( A_i \in \Sigma_X \) with \( \mu(A_i) = s_i \) (\( 1 \leq i \leq m \)), \( B_j \in \Sigma_Y \) with \( \nu(B_j) = t_j \) (\( 1 \leq j \leq n \)) and \( \sum_{i=1}^{m} s_i = \sum_{j=1}^{n} t_j = 1 \). Since \( f(x,y) \) is a step function, the four integrals in (6) exist, and for this function the inequality (6) can be rewritten in the form

\[
\sum_{i=1}^{m} \left( \sum_{j=1}^{n} f_{ij} t_j \right)^2 s_i + \sum_{j=1}^{n} \left( \sum_{i=1}^{m} f_{ij} s_i \right)^2 t_j \\
\leq \left( \sum_{i=1}^{m} \sum_{j=1}^{n} f_{ij} s_i t_j \right)^2 + \sum_{i=1}^{m} \sum_{j=1}^{n} f_{ij}^2 s_i t_j.
\]

By Theorem 3, then, inequality (6) holds for this type of simple function on the product space.

In the general case, Lemma 1 shows that for each integer \( n \), there exists a simple function \( f_n \) (of the type considered above) on the product space such that

\[
\int_{X \times Y} |f(x,y) - f_n(x,y)| \, d(\mu \times \nu) < \frac{1}{n}.
\]

This implies there is a sequence of simple functions \( f_n \) on \( X \times Y \) such that

\[
\lim_{n \to \infty} \int_{X \times Y} f_n(x,y) \, d(\mu \times \nu) = \int_{X \times Y} f(x,y) \, d(\mu \times \nu).
\]

Without loss of generality, assume that \( f_n \leq f_{n+1} \leq f \). Then by The Monotone Convergence Theorem [1, 2.14, p. 50], we know that for almost all \( x \in X \),

\[
\lim_{n \to \infty} \int_Y f_n(x,y) \, d\nu = \int_Y f(x,y) \, d\nu.
\]
and, hence
\[ \lim_{n \to \infty} \left( \int_Y f_n(x, y) \, d\nu \right)^2 = \left( \int_Y f(x, y) \, d\nu \right)^2. \]

Since \( f_n \leq f_{n+1} \leq f \),
\[ \left( \int_Y f_n(x, y) \, d\nu \right)^2 \leq \left( \int_Y f(x, y) \, d\nu \right)^2 \]
and for almost all \( x \in X \), \( f(x, y) \in L^2(X, \mu) \), it follows from The Monotone Convergence Theorem that
\[ \lim_{n \to \infty} \int_X \left( \int_Y f_n(x, y) \, d\nu \right)^2 \, d\mu = \int_X \left( \int_Y f(x, y) \, d\nu \right)^2 \, d\mu. \]

Similarly,
\[ \lim_{n \to \infty} \int_Y \left( \int_X f_n(x, y) \, d\mu \right)^2 \, d\nu = \int_Y \left( \int_X f(x, y) \, d\mu \right)^2 \, d\nu. \]

The following equality also follows from The Monotone Convergence Theorem:
\[ \lim_{n \to \infty} \int_{X \times Y} f_n^2(x, y) \, d(\mu \times \nu) = \int_{X \times Y} f^2(x, y) \, d(\mu \times \nu). \]

The full result follows at once. \( \square \)

By scaling we can get a result for finite positive measure spaces.

**Corollary 1.** Let \((X, \Sigma_X, \mu)\) and \((Y, \Sigma_Y, \nu)\) be finite positive measure spaces with the product space \((X \times Y, \Sigma_{X \times Y}, \mu \times \nu)\). Suppose \( f(x, y) \) is an \( \mu \times \nu \) measurable function on \( X \times Y \). If \( f(x, y) \in L^2(X \times Y, \mu \times \nu) \), \( \int_X f(x, y) \, d\mu \in L^2(Y, \nu) \) and \( \int_Y f(x, y) \, d\nu \in L^2(X, \mu) \), then
\[ \mu(X) \int_Y \left( \int_X f(x, y) \, d\nu \right)^2 \, d\mu + \nu(Y) \int_X \left( \int_Y f(x, y) \, d\mu \right)^2 \, d\nu \leq \left( \int \int_{X \times Y} f(x, y) \, d(\mu \times \nu) \right)^2 + \mu(X) \nu(Y) \int \int_{X \times Y} f^2(x, y) \, d(\mu \times \nu). \]

**Proof.** The proof is immediately from Theorem 4. \( \square \)

Theorem 4 can be rephrased in terms of conditional expectation with respect to independent random variables.

**Corollary 2.** If \( X \) and \( Y \) are independent random variables, then
\[ E((E(f|X))^2) + E((E(f|Y))^2) \leq E^2(f) + E(f^2), \]
where \( E \) denotes expectation and \( E(\cdot|Z) \) denotes conditional expectation with respect to \( Z \).

**Acknowledgement**

Many thanks to the referee for his/her valuable suggestions.

**References**