isid/ms/2005/04
April 11, 2005
http://www.isid.ac.in/~statmath/eprints

# Infinitely Divisible Matrices

Rajendra Bhatia

Indian Statistical Institute, Delhi Centre
7, SJSS Marg, New Delhi–110 016, India

# Infinitely Divisible Matrices

Rajendra Bhatia
Indian Statistical Institute, New Delhi 110 016, India
rbh@isid.ac.in

## 1 Introduction

### 1.1 Positive semidefinite matrices

Let $A$ be an $n \times n$ matrix with complex entries $a_{ij}$. We write this as $A = [a_{ij}]$ . $A$ is Hermitian $\left(A = A^*\right)$ if $a_{ij} = \overline{a}_{ji}$, a condition that is readily verified by inspection. A Hermitian matrix is *positive semidefinite* (psd for short) if it satisfies any of the following equivalent conditions

(i) For every vector $x \in \mathbb{C}^n$, the inner product $\langle x, Ax\rangle \geq 0$.

(ii) All principal minors of $A$ are nonnegative.

(iii) All eigenvalues of $A$ are nonnegative.

(iv) $A = BB^*$ for some matrix $B$.

(v) $A = LL^*$ for some lower triangular matrix $L$.

(vi) There exist vectors $u_1, \ldots, u_n$ in some inner product space such that $a_{ij} = \langle u_i, u_j\rangle$. The matrix $A$ is then said to be the *Gram matrix* associated with the vectors $\{u_1, \ldots, u_n\}$.

It is not easy to verify any of the conditions (i)–(vi) and a little ingenuity is often needed in proving that a certain matrix is psd.

### 1.2 The Hadamard product

Let $A = [a_{ij}]$ and $B = [b_{ij}]$ be two $n \times n$ matrices. The *Hadamard product* (or the *entrywise product*) of $A$ and $B$ is the matrix $A \circ B = [a_{ij} b_{ij}]$.

The most interesting theorem about Hadamard products was proved by I. Schur. This says that if $A$ and $B$ are psd, then so is $A \circ B$. This theorem is so striking that the product $A \circ B$ is often called the *Schur product*. Note that the usual matrix product $AB$ (of psd matrices $A$ and $B$) is psd if and only if $AB = BA$.

For each nonnegative integer $m$ let $A^{\circ m} = \left[a_{ij}^m\right]$ be the $m$th Hadamard power of $A$, and $A^m$ the usual $m$th power. If $A$ is psd, then both $A^{\circ m}$ and $A^m$ are psd.

## 1.3 Infinitely divisible matrices

Fractional powers of psd matrices are defined via the spectral theorem. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of a psd matrix $A$ and $v_1, \ldots, v_n$ the corresponding (orthonormal) eigenvectors. Then $A = \sum \lambda_i v_i v_i^*$, and for all $r \geq 0$ the (usual) $r$th power of $A$ is the psd matrix $A = \sum \lambda_i^r v_i v_i^*$.

If the entries $a_{ij}$ are nonnegative real numbers, it is natural to define *fractional Hadamard powers* of $A$. In this case for every $r \geq 0$ we write $A^{\circ r} = \left[ a_{ij}^r \right]$.

Suppose $A$ is psd and $a_{ij} \geq 0$ for all $i, j$. We say that $A$ is *infinitely divisible* if for every $r \geq 0$ the matrix $A^{\circ r}$ is psd. Every $2 \times 2$ psd matrix with nonnegative entries is infinitely divisible. This is no longer the case when $n > 2$. The $3 \times 3$ matrix

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

is psd. But $A^{\circ r}$ is psd if and only if $r \geq 1$.

C. FitzGerald and R. Horn [11] have shown that if $A$ is an $n \times n$ psd matrix and $a_{ij} \geq 0$, then for all real numbers $r \geq n-2$, the matrix $A^{\circ r}$ is psd. The *critical exponent* $n - 2$ is best possible here; for each $r < n - 2$ one can construct a psd matrix $A$ for which $A^{\circ r}$ is not psd.

If $A^{\circ r}$ is psd for $r = \frac{1}{m}$, $m = 1, 2, \ldots$, then by Schur's Theorem $A^{\circ r}$ is psd for all positive rational numbers $r$. Taking limits, we see that $A^{\circ r}$ is psd for all nonnegative real numbers. Thus a psd matrix with nonnegative entries is infinitely divisible if and only if for every positive integer $m$ there exists a psd matrix $B$ such that $A = B^{\circ m}$. The term infinitely divisible originates from this property.

The purpose of this note is to give very simple proofs of the infinite divisibility of some interesting matrices. The ideas and methods we employ are likely to be useful in other contexts.

## 1.4 Some simple facts

Most of our proofs invoke the followig facts.

If $A_1, \ldots, A_k$ are psd, then so is any linear combination $a_1 A_1 + \cdots + a_k A_k$ with nonnegative coefficients $a_j$.

If a sequence $\{A_k\}$ of psd matrices converges to $A$, then $A$ is psd.

The matrix $E$ with all entries $e_{ij} = 1$ is called the *flat matrix.* It is a rank-one psd matrix.

If $A$ is psd, then for every $X$, the matrix $X^* A X$ is psd. In particular, choosing $X$ to be a diagonal matrix with positive diagonal entries $\lambda_i$, we see that if $A$ is psd (infinitely divisible), then the matrix $XAX = [\lambda_i \lambda_j a_{ij}]$ is also psd (infinitely divisible). The matrices $A$ and $X^* A X$ are said to be *congruent* to each other if $X$ is invertible.

We will use Schur's Theorem stated in Section 1.2. It is easy to prove this. One of the several known proofs goes as follows. Every rank-one psd matrix $A$ has the form $A = xx^*$ for some vector $x$; or in other words, $a_{ij} = x_i \overline{x}_j$ for some $x = (x_1, \ldots, x_n)$ in $\mathbb{C}^n$. If $A = xx^*$ and $B = yy^*$ are two such matrices, then $A \circ B = zz^*$, where $z = (x_1 y_1, \ldots, x_n y_n)$. Thus the

Hadamard product of two rank-one psd matrices is psd. The general case follows from this as every psd matrix is a sum of rank-one psd matrices.

The point to note is that it is easy to show that $A^{\circ m}$ is psd if $A$ is psd. It is not so easy to decide whether $A^{\circ r}$ is psd for every $r \geq 0$.

Chapter 7 of [20] is a rich source of information on psd matrices. A very interesting and lively discussion of the nomenclature, the history, and the most important properties of the Hadamard product may be found in the survey article [19] by R. Horn.

# 2 Examples

## 2.1 The Cauchy matrix

Let $0 < \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ be any positive numbers. The matrix $C$ with entries $c_{ij} = \frac{1}{\lambda_i+\lambda_j}$ is called a *Cauchy matrix.* The *Hilbert matrix* $H$ defined as $h_{ij} = \frac{1}{i+j-1}$ is a Cauchy matrix.

In 1841 Cauchy gave a formula for the determinant of $C$

$$\det C = \frac{\Pi_{1\leq i<j\leq n}\ (\lambda_i - \lambda_j)^2}{\Pi_{1\leq i,j\leq n}\ (\lambda_i + \lambda_j)}.$$

This shows that all principal minors of $C$ are nonnegative, and hence $C$ is psd. This fact can be proved more easily as follows.

Since

$$\frac{1}{\lambda_i + \lambda_j} = \int_0^\infty e^{-t(\lambda_i+\lambda_j)} dt, \tag{1}$$

the matrix $C$ is the Gram matrix associated with the elements $u_i(t) = e^{-t\lambda_i}$ in the Hilbert space $L_2(0, \infty)$. Hence $C$ is psd.

Actually $C$ is infinitely divisible. Two proofs are given below.

Choose any number $\varepsilon$ between 0 and $\lambda_1$. For every $r > 0$ we have

$$\begin{aligned} \frac{1}{(\lambda_i + \lambda_j - \varepsilon)^r} &= \Big(\frac{\varepsilon}{\lambda_i\lambda_j}\Big)^r \frac{1}{\Big(1 - \frac{(\lambda_i-\varepsilon)(\lambda_j-\varepsilon)}{\lambda_i\lambda_j}\Big)^r} \\ &= \Big(\frac{\varepsilon}{\lambda_i\lambda_j}\Big)^r \sum_{m=0}^{\infty} a_m \Big(\frac{(\lambda_i - \varepsilon)(\lambda_j - \varepsilon)}{\lambda_i\lambda_j}\Big)^m, \end{aligned} \tag{2}$$

where $a_m$ are the coefficients in the series expansion

$$\frac{1}{(1-x)^r} = \sum_{m=0}^{\infty} a_m x^m, \quad |x| < 1.$$

All $a_m$ are positive; we have $a_0 = 1$ and $a_m = \frac{r(r+1)\cdots(r+m+1)}{m!}$ for $m > 1$.

The matrix with entries $\frac{(\lambda_i-\varepsilon)(\lambda_j-\varepsilon)}{\lambda_i\lambda_j}$ is congruent to the flat matrix. Hence this matrix is psd, and by Schur's Theorem so are its $m$th Hadamard powers for $m = 0, 1, 2, \ldots$. Thus the matrix whose entries are given by the infinite series in (2) is psd. The matrix $\frac{1}{(\lambda_i\lambda_j)^r}$ is psd (being congruent to the flat matrix). So, again by Schur's Theorem the matrix with entries given by (2) is psd. Letting $\varepsilon \downarrow 0$ we see that the $r$th Hadamard power of the Cauchy matrix is psd for every $r > 0$.

For our second proof we use the gamma function, defined for $x > 0$ by the formula

$$\Gamma(x) = \int_0^\infty e^{-t} t^{x-1} dt. \tag{3}$$

Using this one can see that for $r > 0$ we have

$$\frac{1}{(\lambda_i + \lambda_j)^r} = \frac{1}{\Gamma(r)} \int_0^\infty e^{-t(\lambda_i+\lambda_j)} t^{r-1} dt. \tag{4}$$

When $r = 1$ this formula reduces to (1). Once again we see that the matrix with entries $\frac{1}{(\lambda_i+\lambda_j)^r}$ is the Gram matrix associated with the elements $u_i(t) = e^{-t\lambda_i}$ in the space $L_2(0,\infty)$ with the measure

$$d\mu(t) = \frac{t^{r-1}}{\Gamma(r)} dt.$$

For the Hilbert matrix a proof similar to our first proof is given by M.-D. Choi [8].

## 2.2 Generalised Cauchy matrices

The ideas of the preceding section work for some other matrices.

Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be positive numbers and for each real number $t$ let $Z$ be the $n \times n$ matrix with entries

$$z_{ij} = \frac{1}{\lambda_i^2 + \lambda_j^2 + t\lambda_i\lambda_j}.$$

When is such a matrix psd, or infinitely divisible? The condition $t > -2$ is clearly necessary for $Z$ to be psd. When $n = 2$ this condition is also sufficient to ensure that $Z$ is psd, and hence infinitely divisible. For $n \geq 3$ the condition $t > -2$ is no longer sufficient. If $(\lambda_1, \lambda_2, \lambda_3) = (1, 2, 3)$ and $t = 10$, then $Z$ is not psd. However, if $-2 < t \leq 2$, then for all $n$ the matrix $Z$ is infinitely divisible. For $-2 < t < 2$ and $r > 0$ we have the expansion

$$z_{ij}^r = \frac{1}{(\lambda_i + \lambda_j)^{2r}} \sum_{m=0}^{\infty} a_m (2-t)^m \frac{\lambda_i^m \lambda_j^m}{(\lambda_i + \lambda_j)^{2m}}.$$

This shows that the matrix $Z^{\circ r}$ is a limit of sums of Hadamard products of psd matrices. So $Z$ is infinitely divisible for all $t$ in the range $-2 < t < 2$. By continuity this is true also for $t = 2$.

It is known [5] that for every $t > 2$ there exists an $n$ for which the $n \times n$ matrix $Z$ is not psd for some choice of numbers $\lambda_1, \ldots, \lambda_n$. The proof of this needs more advanced arguments (of the kind discussed in Section 3.1 of this paper). The matrix $Z$ was studied by M. K. Kwong [22] who used somewhat intricate arguments to prove that for all $n$, $Z$ is psd for $-2 < t \leq 2$.

When $t = 1$, the matrix $Z$ can be written also as

$$z_{ij} = \frac{\lambda_i - \lambda_j}{\lambda_i^3 - \lambda_j^3}.$$

The matrix $W$ with entries

$$w_{ij} = \frac{\lambda_i - \lambda_j}{\lambda_i^4 - \lambda_j^4} = \frac{\lambda_i - \lambda_j}{\lambda_i^2 - \lambda_j^2} \frac{1}{\lambda_i^2 + \lambda_j^2}$$

is infinitely divisible, being the Hadamard product of infinitely divisible matrices. This argument shows that the matrix $V$ with entries

$$v_{ij} = \frac{\lambda_i - \lambda_j}{\lambda_i^n - \lambda_j^n}$$

is infinitely divisible for positive integers $n$ that can be expressed as $2^m$ or as $3.2^m$. It is known that the matrix $V$ is infinitely divisible for all $n$. We discuss this in Section 3.3.

## 2.3 The Pascal matrix

The $n \times n$ *Pascal matrix* is the matrix $A$ with entries

$$a_{ij} = \binom{i+j}{i} \quad i, j = 0, 1, \ldots, n-1.$$

The rows of the Pascal triangle occupy the anti-diagonals of $A$. Thus the $4 \times 4$ Pascal matrix is

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 10 \\ 1 & 4 & 10 & 20 \end{bmatrix}.$$

Let $L$ be the lower triangular matrix whose rows are occupied by the rows of the Pascal triangle. Thus for $n = 4$

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 1 & 3 & 3 & 1 \end{bmatrix}.$$

It is well-known that $A = LL^*$, and hence $A$ is psd. See the recent paper [10] for several proofs of this fact, and for references to papers on Pascal matrices.

The positive definiteness of the Pascal matrix may be proved also by representing it as a Gram matrix. One such representation is

$$a_{rs} = \frac{1}{2\pi} \int_0^{2\pi} (1 + e^{i\theta})^r (1 + e^{-i\theta})^s d\theta. \tag{5}$$

Another representation is obtained using the gamma function (3). For $x, y > 0$ we have

$$\Gamma(x + y + 1) = \int_0^\infty e^{-t} t^x t^y dt. \tag{6}$$

Since $\Gamma(n + 1) = n!$ for every nonnegative integer $n$, this shows that the matrix with entries

$$(i + j)! = \Gamma(i + j + 1)$$

is a Gram matrix, and hence is psd. The Pascal matrix

$$a_{ij} = \frac{(i+j)!}{i!\, j!}$$

is congruent to it, and is therefore psd.

Our argument shows that for any positive numbers $\lambda_1, \ldots, \lambda_n$ the matrix $K$ with entries

$$k_{ij} = \frac{\Gamma(\lambda_i + \lambda_j + 1)}{\Gamma(\lambda_i + 1)\Gamma(\lambda_j + 1)} \tag{7}$$

is psd. When $\lambda_j = j$, this matrix is the Pascal matrix.

In fact, the matrix $K$ is infinitely divisible. (This seems not to have been noticed before even for the Pascal matrix.)

Using Gauss's Formula

$$\Gamma(z) = \lim_{m\to\infty} \frac{m! m^z}{z(z+1)\ldots(z+m)}, \quad z \neq 0, -1, -2, \ldots \tag{8}$$

we see that

$$k_{ij} = \lim_{m\to\infty} \frac{1}{m.m!} \prod_{p=1}^{m+1} \frac{(\lambda_i + p)(\lambda_j + p)}{(\lambda_i + \lambda_j + p)}. \tag{9}$$

For each $p$ the matrix

$$\left[\frac{(\lambda_i + p)(\lambda_j + p)}{\lambda_i + \lambda_j + p}\right]$$

is congruent to the Cauchy matrix

$$\left[\frac{1}{\lambda_i + \lambda_j + p}\right],$$

and is, therefore, an infinitely divisible matrix. The expression (9) displays $K$ as a limit of products of infinitely divisible matrices. Hence by Schur's Theorem and continuity, $K$ is infinitely divisible.

A small aside may be of interest here. If $\lambda_1, \ldots, \lambda_n$ are complex numbers with positive real parts, then the matrix $C$ with entries

$$c_{ij} = \frac{1}{\lambda_i + \overline{\lambda}_j}$$

is psd. This can be proved by representing $C$ as a Gram matrix as in (1). The condition $\operatorname{Re} \lambda_i > 0$ guarantees that the integral

$$\int_0^\infty e^{-t(\lambda_i + \overline{\lambda}_j)} dt$$

is congruent. Our arguments show that with this restriction on $\lambda_i$, the matrix $K$ with entries

$$k_{ij} = \frac{\Gamma(\lambda_i + \overline{\lambda}_j)}{\Gamma(\lambda_i)\Gamma(\overline{\lambda}_j)}$$

is a psd matrix.

### 2.4 The matrix $\min(x, y)$

Consider the $n \times n$ matrix $M$ with entries $m_{ij} = \min(i, j)$. The idea behind the discussion that follows is captured by the equation

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 \\ 1 & 2 & 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix}$$
$$+ \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Each of the matrices in this sum is psd. So $M$ is psd.

This can be generalised. Let $\lambda_1, \ldots, \lambda_n$ be any positive numbers, and let $M$ be the $n \times n$ matrix with entries $m_{ij} = \min(\lambda_i, \lambda_j)$. The argument above can be modified to show that $M$ is infinitely divisible.

First, by applying a permutation similarity, we may assume that $0 < \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$. Thus

$$M = \begin{bmatrix} \lambda_1 & \lambda_1 & \ldots & \lambda_1 \\ \lambda_1 & \lambda_2 & \ldots & \lambda_2 \\ \vdots & \vdots & & \vdots \\ \lambda_1 & \lambda_2 & & \lambda_n \end{bmatrix}.$$

For $1 \leq k \leq n$, let $\tilde{E}_k$ be the $n \times n$ matrix whose bottom right corner is occupied by the $k \times k$ flat matrix and the rest of whose entries are zero. Then we may write $M$ as

$$M = \lambda_1 \tilde{E}_n + (\lambda_2 - \lambda_1)\tilde{E}_{n-1} + (\lambda_3 - \lambda_2)\tilde{E}_{n-2} + \cdots + (\lambda_n - \lambda_{n-1})\tilde{E}_1.$$

Thus $M$ is a psd matrix.

The same argument shows that if $f$ is a monotonically increasing function from $(0, \infty)$ into itself, then the matrix $[f(m_{ij})]$ is psd. The special choice $f(t) = t^r$, $r > 0$, shows that $M$ is infinitely divisible.

We could have started the discussion in this section with the factoring

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 \\ 1 & 2 & 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Building on this, alternate proofs of other results in this section may be obtained.

### 2.5 The Lehmer matrix

For positive numbers $x$ and $y$ we have

$$\frac{\min(x, y)}{xy} = \frac{1}{\max(x, y)}.$$

So from the results of Section 2.4 it follows that for positive numbers $\lambda_1, \ldots, \lambda_n$ the $n \times n$ matrix $W$ with entries

$$w_{ij} = \frac{1}{\max(\lambda_i, \lambda_j)}$$

is infinitely divisible.

Let $0 < \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ and let $L$ be the matrix with entries

$$l_{ij} = \frac{\min(\lambda_i, \lambda_j)}{\max(\lambda_i, \lambda_j)}.$$

Then $L$ is infinitely divisible. We have $l_{ij} = \lambda_i/\lambda_j$ for $1 \leq i \leq j \leq n$. For the special choice $\lambda_j = j$, $1 \leq j \leq n$, the matrix $L$ is called the *Lehmer matrix.*

Let $\lambda_1, \ldots, \lambda_n$ be positive numbers and let $K$ be the matrix with entries

$$k_{ij} = \exp(-|\lambda_i - \lambda_j|).$$

It follows from our discussion of the matrix $L$ that $K$ is an infinitely divisible matrix. This property is equivalent to the fact that the function $f(x) = e^{-|x|}$ is a positive definite function on $\mathbb{R}$. See Section 3.1.

# 3 Other proofs and connections

## 3.1 Positive definite functions

The concept of infinite divisibility is important in the theory of characteristic functions of probability distributions. Infinitely divisible distributions are exactly the class of limit distributions for sums of independent random variables. See [7, pp.190–196] and [12, Chapter 10]. General techniques from this subject may be used to prove special results on matrices.

A complex-valued function $f$ on $\mathbb{R}$ is said to be a *positive definite function* if for every positive integer $n$ and for every choice of points $x_1, \ldots, x_n$ in $\mathbb{R}$, the $n \times n$ matrix $[f(x_i - x_j)]$ is positive semidefinite. A theorem of Bochner says that a function $f$, continuous at 0, is positive definite if and only if there exists a finite positive Borel measure $\mu$ such that

$$f(t) = \int_{-\infty}^{\infty} e^{-itx} d\mu(x). \tag{10}$$

We say that $f$ is the *Fourier transform* of $\mu$, and write this as $f = \hat{\mu}$.

We write $\nu_1 * \nu_2$ for the convolution of two measures $\nu_1$ and $\nu_2$. If $f_1$ and $f_2$ are the Fourier transforms of $\nu_1$ and $\nu_2$, then $f_1 f_2 = (\nu_1 * \nu_2)^\wedge$. A measure $\mu$ is said to be *infinitely divisible* if for each positive integer $m$ there exists a measure $\nu$ such that $\mu = \nu * \nu * \ldots * \nu$ (an $m$-fold convolution).

The integral representation

$$e^{-|x|} = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{e^{-itx}}{1+t^2} dt$$

shows that for every $r > 0$ the function $e^{-r|x|}$ is positive definite. This gives another proof of the infinite divisibility of the matrix $K$ of Section 2.5.

One proof of the infinite divisibility of the Cauchy matrix goes as follows. Let $\lambda_j = e^{x_j}$, $j = 1, 2, \ldots, n$. Then

$$\frac{1}{\lambda_i + \lambda_j} = \frac{1}{e^{x_i/2}\; 2\cosh\left(\frac{x_i - x_j}{2}\right) e^{x_j/2}}.$$

Thus the Cauchy matrix is congruent to the matrix whose $ij$ entry is

$$\frac{1}{\cosh\left(\frac{x_i - x_j}{2}\right)}.$$

Since congruence is an equivalence relation that preserves positive semidefiniteness, the following two statements are equivalent:

(i) For every $n$ and for every choice of positive numbers $\lambda_1, \ldots, \lambda_n$ the $n \times n$ Cauchy matrix $\left[\frac{1}{\lambda_i + \lambda_j}\right]$ is psd.

(ii) The function $f(x) = \frac{1}{\cosh\ x}$ on the real line is positive definite.

The second of these statements can be proved by calculating its Fourier transform. This turns out to be

$$\hat{f}(t) = \frac{1}{\cosh\ \frac{t\pi}{2}}.$$

Since $\hat{f}(t) > 0$ for all $t$, the function $f$ is positive definite by Bochner's Theorem.

For $r > 0$, let $g(t) = \frac{1}{(\cosh\ t)^r}$. A calculation involving contour integrals shows that the Fourier transform of $g$ is

$$\hat{g}(t) = 2^{r-2} \frac{1}{\Gamma(r)} \left| \Gamma\left(\frac{r + it}{2}\right) \right|^2.$$

This shows that $g$ is a positive definite function. Hence the Cauchy matrix is infinitely divisible.

A proof along these lines is given in [5] and in [13]. Positive definiteness of several other matrices is proved there. This method is especially useful in showing that certain functions are *not* positive definite, and hence certain matrices are not always psd.

The kernel $M(x, y) = \min(x, y)$ for $x, y > 0$, is known to be the covariance kernel of the standard Brownian motion [7]. Hence it is infinitely divisible by what we said at the begining of this section. This gives another way of looking at the matrix in Section 2.4.

## 3.2 Conditionally positive definite matrices

A Hermitian matrix $A$ is said to be *conditionally positive semidefinite* (cpd for short) if $x^*Ax \geq 0$ for all vectors $x \in \mathbb{C}^n$ for which $\sum_{j=1}^{n} x_j = 0$.

C. Loewner showed that if $A$ is a symmetric matrix with positive entries, then $A$ is infinitely divisible if and only if its Hadamard logarithm $\log^\circ(A) = [\log\ a_{ij}]$ is a cpd matrix. (See Theorem 6.3.13 in [21].) Good necessary and sufficient conditions for a matrix to be cpd are also known. One of them says that an $n \times n$ Hermitian matrix $B = [b_{ij}]$ is cpd if and only if the $(n-1) \times (n-1)$ matrix $D = [d_{ij}]$ with entries

$$d_{ij} = b_{ij} + b_{i+1,j+1} - b_{i,j+1} - b_{i+1,j}$$

is positive semidefinite. See [21, pp. 457-458] where these criteria are used to prove the infinite divisibility of some matrices.

Let us show how these considerations may be applied to the Pascal matrix. Using the two conditions stated above one sees that the $n \times n$ Pascal matrix is infinitely divisible if and only if the $(n-1) \times (n-1)$ matrix $D$ with entries

$$d_{ij} = \log \frac{i+j+2}{i+j+1} = \log \left(1 + \frac{1}{i+j+1}\right) \tag{11}$$

is psd.

For $x \geq 0$ we have

$$\log(1+x) = \int_1^\infty \frac{tx}{t+x} d\mu(t),$$

where $\mu$ is the probability measure on $[1, \infty)$ defined as $d\mu(t) = dt/t^2$. (See [4, p. 145].) Using this we can write the $d_{ij}$ in (11) as

$$d_{ij} = \int_1^\infty \frac{1}{i+j+1+\frac{1}{t}} d\mu(t).$$

Thus $D$ is a limit of positive linear combinations of matrices $C(t) = [c_{ij}(t)]$ where

$$C_{ij}(t) = \frac{1}{i+j+1+\frac{1}{t}}, \quad t \geq 1.$$

If we put $\lambda_i = i + \frac{1}{2}\left(1 + \frac{1}{t}\right)$, then

$$C_{ij}(t) = \frac{1}{\lambda_i + \lambda_j}.$$

Thus for each $t \geq 1$, the matrix $C(t)$ is a Cauchy matrix. Hence $D$ is psd.

Several applications of cpd matrices may be found in the book [2]. Continuous analogues and their applications are discussed in the monograph [24]. Two of the early papers on infinitely divisible matrices are [17, 18].

## 3.3 Operator monotone functions

If $A$ and $B$ are Hermitian matrices and $A - B$ is psd, then we say that $A \geq B$. Let $f$ be any map of the positive half line $[0, \infty)$ into itself. We say that $f$ is *matrix monotone of order $n$* if $f(A) \geq f(B)$ whenever $A$ and $B$ are $n \times n$ psd matrices with $A \geq B$. If $f$ is matrix monotone of order $n$ for all $n = 1, 2, \ldots$, then we say that $f$ is *operator monotone.*

Two famous theorems of C. Loewner characterise operator monotone functions. The first says that (a differentiable function) $f$ is matrix monotone of order $n$ if and only if for all positive numbers $\lambda_1, \ldots, \lambda_n$ the $n \times n$ matrix

$$\left[\frac{f(\lambda_i) - f(\lambda_j)}{\lambda_i - \lambda_j}\right] \tag{12}$$

is psd. (If $\lambda_i = \lambda_j$, the difference quotient in (12) is understood to mean $f'(\lambda_i)$. ) The matrix (12) is called the *Loewner matrix* associated with $f$.

The second theorem of Loewner says that $f$ is operator monotone if and only it has an analytic continuation mapping the upper half plane into itself. It was shown by R. Horn [16] that this analytic continuation is a one-to-one (also called *univalent* or *schlicht*) map if and only if all the Loewner matrices associated with $f$ are infinitely divisible.

From Loewner's Theorem it is clear that the function $f(t) = t^\nu$ is operator monotone for $0 \le \nu \le 1$. It follows from Horn's Theorem that for every $n$, and for all positive numbers $\lambda_1, \ldots, \lambda_n$ the $n \times n$ matrix

$$\left[\frac{\lambda_i^\nu - \lambda_j^\nu}{\lambda_i - \lambda_j}\right]$$

is infinitely divisible for $0 \le \nu \le 1$. In Section 2.2 we proved this for some special values of $\nu$.

The reader may see [9] and [4, Chapter V] for the theory of operator monotone functions.

We should point out that some of the special matrices studied in this paper are frequently used for testing the stability of numerical algorithms [14], and form a part of the matrix gallery in MATLAB [15].

# 4 More Examples

## 4.1 GCD matrices

The matrix of Section 2.4 is a cousin, in spirit, of another matrix. Given a set $S = \{x_1, \ldots, x_n\}$ of distinct positive integers the *GCD matrix associated with* $S$ is the matrix $A$ with entries $a_{ij} = (x_i, x_j)$, the greatest common divisor of $x_i$ and $x_j$. This matrix is infinitely divisible. We outline a proof of this. To make it easier reading we prove first that $A$ is psd. The proofs are borrowed from papers by Beslin and Ligh [3] and Bourque and Ligh [6]. The elementary concepts of number theory that we need may be found in a basic text such as [1].

Let $\varphi(n)$ be the *Euler $\varphi$-function.* For each positive integer $n$ this function counts integers $m$ less than or equal to $n$ such that $(m, n) = 1$. One has the equality

$$\sum_{d|n} \varphi(d) = n. \tag{13}$$

Here, as usual $d|n$ means that $1 \le d \le n$ and $n$ is divisible by $d$.

We say that a set $F$ of positive integers is *factor-closed* if whenever $x \in F$ and $d|x$, then $d \in F$. The smallest factor-closed set $F$ containing a set $S$ is called the *factor-closure* of $S$. Thus, for example, the set $\{2, 3, 5, 6, 10\}$ is the factor-closure of the set $\{2, 6, 10\}$.

Let $S = \{x_1, \ldots, x_n\}$ be any set of distinct positive integers and let $F = \{d_1, \ldots, d_t\}$ be its factor-closure. Define $n \times t$ matrices $E$ and $B$ as follows

$$\begin{aligned} e_{ij} &= \begin{cases} 1 & \text{if } d_j | x_i \\ 0 & \text{otherwise} \end{cases}, \\ b_{ij} &= e_{ij}\sqrt{\varphi(d_j)}. \end{aligned} \tag{14}$$

Then the $ij$ entry of the matrix $BB^*$ is

$$\begin{aligned} \sum_{k=1}^{t} b_{ik} b_{jk} &= \sum_{\substack{d_k | x_i \\ d_k | x_j}} \sqrt{\varphi(d_k)}\sqrt{\varphi(d_k)} \\ &= \sum_{d_k | (x_i, x_j)} \varphi(d_k) \\ &= (x_i, x_j). \end{aligned}$$

This shows that the GCD matrix $a_{ij} = (x_i, x_j)$ is psd.

A (complex) function $f$ on $\mathbb{N}$ is said to be *multiplicative* if $f(mn) = f(m)f(n)$ whenever $(m, n) = 1$. The Euler $\varphi$ function is multiplicative. The *Dirichlet convolution* of two multiplicative functions $f$ and $g$ is defined as

$$(f * g)(n) = \sum_{d|n} f(d) g\left(\frac{n}{d}\right).$$

With this binary operation, the collection of multiplicative functions is an abelian group. The identity element of this group is the function $\varepsilon(n)$ that takes the value 1 at $n = 1$ and zero at $n \neq 1$.

The *Möbius function* $\mu(n)$ is defined as follows: $\mu(1) = 1$. If $n > 1$ let

$$n = p_1^{k_1} p_2^{k_2} \ldots p_m^{k_m}$$

be a factoring of $n$ with $p_1, \ldots, p_m$ distinct primes. If $k_1 = k_2 = \cdots = k_m = 1$ (i.e., $n$ is square-free), then $\mu(n) = (-1)^m$. Otherwise $\mu(n) = 0$.

The Möbius function is multiplicative. Its inverse in the group of multiplicative functions is the function $u(n) \equiv 1$; i.e., $\mu * u = \varepsilon$. Hence, $(f * \mu) * u = f$; i.e.

$$\sum_{d|n} (f * \mu)(d) = f(n) \tag{15}$$

for every multiplicative function $f$.

We prove that for every multiplicative function $f$ such that $(f * \mu)(n) > 0$ the matrix $\left[f\Big((x_i, x_j)\Big)\right]$ is positive semidefinite. Instead of $B$ defined by (14) now consider the matrix $B$ with entries

$$b_{ij} = e_{ij}\sqrt{(f * \mu)(d_j)}. \tag{16}$$

The same calculation as before, with the equality (15) replacing (13) shows that

$$BB^* = \left[ f\Big( (x_i, x_j) \Big) \right].$$

Choosing $f(n) = n^r$, $r > 0$ we see that the GCD matrix $[(x_i, x_j)]$ is infinitely divisible.

If $l_{ij}$ is the LCM of $x_i$ and $x_j$, then the argument given in Section 2.5 shows that the matrix $A$ with entries $a_{ij} = 1/l_{ij}$ is infinitely divisible.

## 4.2 Characteristic matrices

Let $x_1, \ldots, x_n$ be vectors in the space $\mathbb{R}^k$. Associate with them an $n \times n$ matrix $A$ as follows. If exactly $m$ coordinates of the vector $x_i$ are equal to the corresponding coordinates of the vector $x_j$, then $a_{ij} = m$. Note that $0 \le m \le k$. The matrix $A$ is psd. One proof of this goes as follows.

First consider the case $k = 1$. Arrange the numbers $x_1, \ldots, x_n$ in such a way that they are grouped into disjoint classes $S_1, S_2, \ldots, S_l$ where $x_i$ and $x_j$ belong to the same class if and only if they are equal. The matrix $A$ is then a direct sum of flat matrices, and is, therefore, psd.

Now consider the case $k > 1$. For $i \le p \le k$, let $A_p$ be the $n \times n$ matrix whose $ij$ entry is 1 if the $p$th coordinate of $x_i$ is equal to the $p$th coordinate of $x_j$, and 0 otherwise. Our matrix $A$ is equal to $A_1 + \cdots + A_k$ and each matrix in this sum is psd.

This matrix, called the *characteristic matrix* associated with $\{x_1, \ldots, x_n\}$, is not always infinitely divisible. For example, let $x_1 = (1,1)$, $x_2 = (2,1)$, $x_3 = (1,2)$. Then the characteristic matrix

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 0 \\ 1 & 0 & 2 \end{bmatrix},$$

and this is not infinitely divisible.

Two comments are in order here. Our discussion suggests that there might be something special about the pattern of zero entries in an infinitely divisible matrix. The *incidence matrix* of $A$ is the matrix $G(A) = [g_{ij}]$, where $g_{ij} = 1$ if $a_{ij} \ne 0$, and $g_{ij} = 0$ if $a_{ij} = 0$. If $A$ is infinitely divisible, then $G(A)$ is psd. It is not difficult to see that $G(A)$ is psd if and only if there is a permutation matrix $X$ such that $X^*AX$ is a direct sum of flat matrices and a zero matrix. See [18], [21, p.457]. This gives a good necessary condition for infinite divisibility.

Our second remark points to a connection between characteristic matrices and positive definite functions.

Let $G$ be any additive subgroup of $\mathbb{R}$. Then the characteristic function $\chi_G$ is a positive definite function. Using this one can see that if $A$ is the matrix with $a_{ij} = m$ if exactly $m$ coordinates of the vector $x_i - x_j$ are in $G$, then $A$ is psd. The special case $G = \{0\}$ corresponds to the characteristic matrix.

Characteristic matrices arise in diverse contexts. See [25] for their use in the study of distance matrices and interpolation problems.

*This work was begun at the Manchester Centre for Computational Mathematics and continued at the Centro de Estruturas Lineares e Combinatorias of the University of Lisbon. The author is thankful to these centres for their hospitality. This work has been stimulated and enhanced by conversations with R. B. Bapat, J. A. Dias da Silva, D. Drissi, L. Elsner, N. J. Higham, R. Horn and K. R. Parthasarathy.*

## References

[1] T. M. Apostol, *Introduction to Analytic Number Theory,* Springer, 1976.

[2] R. B. Bapat and T. E. S. Raghavan, *Nonnegative Matrices and Applications,* Cambridge University Press, 1997.

[3] S. Beslin and S. Ligh, *Greatest common divisor matrices,* Linear Algebra Appl., 118 (1989) 69-76.

[4] R. Bhatia, *Matrix Analysis,* Springer, 1997.

[5] R. Bhatia and K. R. Parthasarathy, *Positive definite functions and operator inequalities,* Bull London Math. Soc., 32 (2000) 214-228.

[6] K. Bourque and S. Ligh, *Matrices associated with classes of arithmetical function,* J. Number Theory, 45 (1993) 367-376.

[7] L. Breiman, *Probability,* Addison–Wesley, 1968.

[8] M.-D. Choi, *Tricks or treats with the Hilbert matrix,* American Math. Monthly, 90 (1983) 301-312.

[9] W. F. Donoghue, *Monotone Matrix Functions and Analytic Continuation,* Springer, 1974.

[10] A. Edelman and G. Strang, *Pascal matrices,* American Math. Monthly, 111 (2004) 189-197.

[11] C. H. FitzGerald and R. A. Horn, *On fractional Hadamard powers of positive definite matrices,* J. Math. Anal. Appl., 61 (1977) 633-642.

[12] B. V. Gnedenko, *Theory of Probability,* 6th ed., translated from Russian, Gordon and Breach, 1997.

[13] F. Hiai and H. Kosaki, *Means of Hilbert Space Operators, Lecture Notes in Mathematics 1820,* Springer, 2003.

[14] N. J. Higham, *Accuracy and Stability of Numerical Algorithms,* 2nd ed., SIAM, 2002.

[15] N. J. Higham and D. J. Higham, *MATLAB Guide,* SIAM, 2000.

[16] R. A. Horn, *On boundary values of a schlicht mapping,* Proc. Amer. Math. Soc., 18(1967) 782-787.

[17] R. A. Horn, *On infinitely divisible matrices, kernels, and functions,* Z. Wahrscheinlichkeitstheorie und Verw. Gabiete, 8 (1967) 219-230.

[18] R. A. Horn, *The theory of infinitely divisible matrices and kernels,* Trans. Amer. Math. Soc., 136 (1969) 269-286.

[19] R. A. Horn, *The Hadamard product, in Matrix Theory and Applications,* C. R. Johnson, ed., Proceedings of Applied Mathematics, Vol.40, American Mathematical Society, 1990.

[20] R. A. Horn and C. R. Johnson, *Matrix Analysis,* Cambridge University Press, 1985.

[21] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis,* Cambridge University Press, 1991.

[22] M. K. Kwong, *On the definiteness of the solutions of certain matrix equations,* Linear Algebra Appl., 108 (1988) 177-197.

[23] C. Loewner, *On schlicht-monotonic functions of higher order,* J. Math. Anal. Appl., 14 (1966) 320-325.

[24] K. R. Parthasarathy and K. Schmidt, *Positive Definite Kernels, Continuous Tensor Products and Central Limit Theorems of Probability Theory, Lecture Notes in Mathematics 272,* Springer 1972.

[25] L. Reid and X. Sun, *Distance matrices and ridge function interpolation,* Canadian J. Math., 45 (1993) 1313-1323.