Appendix A
Matrix Algebra

A.1 Introduction

This appendix gives (i) a summary of basic definitions and results in matrix algebra with comments and (ii) details of those results and proofs which are used in this book but normally not treated in undergraduate Mathematics courses. It is designed as a convenient source of reference to be used in the rest of the book. A geometrical interpretation of some of the results is also given. If the reader is unfamiliar with any of the results not proved here he should consult a text such as Graybill (1969, especially pp. 4–52, 163–196, and 222–235) or Rao (1973, pp. 1–78). For the computational aspects of matrix operations see for example Wilkinson (1965).

Definition A matrix $\mathbf{A}$ *is a rectangular array of numbers. If* $\mathbf{A}$ *has n rows and p columns we say it is of order* $n \times p$. *For example, n observations on p random variables are arranged in this way.*

Notation 1 We write matrix $\mathbf{A}$ of order $n \times p$ as

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{bmatrix} = (a_{ij}), \tag{A.1.1}$$

where a_{ij} is the element in row i and column j of the matrix $\mathbf{A}$, $i = 1, \ldots, n$; $j = 1, \ldots, p$. Sometimes, we write $(\mathbf{A})_{ij}$ for a_{ij}.

We may write the matrix $\mathbf{A}$ as $\mathbf{A}(n \times p)$ to emphasize the row and column order. In general, matrices are represented by boldface upper case letters throughout this book, e.g. $\mathbf{A}, \mathbf{B}, \mathbf{X}, \mathbf{Y}, \mathbf{Z}$. Their elements are represented by small letters with subscripts.

Definition *The* transpose *of a matrix* $\mathbf{A}$ *is formed by interchanging the rows and columns:*

$$\mathbf{A}' = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ a_{1p} & a_{2p} & \cdots & a_{np} \end{pmatrix}.$$

Definition *A matrix with column-order one is called a* column vector. *Thus*

$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_n \end{pmatrix}$$

is a column vector with n components.

In general, boldface lower case letters represent column vectors. Row vectors are written as column vectors transposed, i.e.

$$\mathbf{a}' = (a_1, \ldots, a_n).$$

Notation 2 We write the columns of the matrix $\mathbf{A}$ as $\mathbf{a}_{(1)}, \mathbf{a}_{(2)}, \ldots, \mathbf{a}_{(p)}$ and the rows (if written as column vectors) as $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ so that

$$\mathbf{A} = (\mathbf{a}_{(1)}, \mathbf{a}_{(2)}, \ldots, \mathbf{a}_{(p)}) = \begin{bmatrix} \mathbf{a}_1' \\ \mathbf{a}_2' \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{a}_n' \end{bmatrix}, \tag{A.1.2}$$

where

$$\mathbf{a}_{(j)} = \begin{bmatrix} a_{1j} \\ \cdot \\ \cdot \\ a_{nj} \end{bmatrix}, \quad \mathbf{a}_i = \begin{bmatrix} a_{i1} \\ \cdot \\ \cdot \\ \cdot \\ a_{ip} \end{bmatrix}.$$

Definition *A matrix written in terms of its sub-matrices is called a partitioned matrix.*

Notation 3 Let $\mathbf{A}_{11}$, $\mathbf{A}_{12}$, $\mathbf{A}_{21}$, and $\mathbf{A}_{22}$ be submatrices such that $\mathbf{A}_{11}(r\times s)$ has elements a_{ij}, $i=1,\ldots,r$; $j=1,\ldots,s$ and so on. Then we write

$$\mathbf{A}(n\times p)=\begin{bmatrix}\mathbf{A}_{11}(r\times s) & \mathbf{A}_{12}(r\times(p-s))\\ \mathbf{A}_{21}((n-r)\times s) & \mathbf{A}_{22}((n-r)\times(p-s))\end{bmatrix}.$$

Obviously, this notation can be extended to contain further partitions of $\mathbf{A}_{11}$, $\mathbf{A}_{12}$, etc.

A list of some important types of particular matrices is given in Table A.1.1. Another list which depends on the next section appears in Table A.3.1.

Table A.1.1 Particular matrices and types of matrix (List 1). For List 2 see Table A.3.1.

Name	Definiton	Notation	Trivial Examples
1 Scalar	$p=n=1$	a, b	(1)
2a Column vector	$p=1$	$\mathbf{a}, \mathbf{b}, \ldots$	$\begin{pmatrix}1\\2\end{pmatrix}$
2b Unit vector	$(1,\ldots,1)'$	$\mathbf{1}$ or $\mathbf{1}_n$	$\begin{pmatrix}1\\1\end{pmatrix}$
3 Rectangular	$p\times n$	$\mathbf{A}(n\times p)$	
4 Square	$p=n$	$\mathbf{A}(p\times p)$	$\begin{pmatrix}1&3\\4&5\end{pmatrix}$
4a Diagonal	$p=n$, $a_{ij}=0$, $i\neq j$	$\operatorname{diag}(a_{ii})$	$\begin{pmatrix}2&0\\0&1\end{pmatrix}$
4b Identity	$\operatorname{diag}(\mathbf{1})$	$\mathbf{I}$ or $\mathbf{I}_p$	$\begin{pmatrix}1&0\\0&1\end{pmatrix}$
4c Symmetric	$a_{ij}=a_{ji}$		$\begin{pmatrix}3&2\\2&5\end{pmatrix}$
4d Unit matrix	$p=n$, $a_{ij}=1$	$\mathbf{J}_p=\mathbf{11}'$	$\begin{pmatrix}1&1\\1&1\end{pmatrix}$
4e Triangular matrix (upper)	$a_{ij}=0$ below the diagonal	Δ'	$\begin{pmatrix}1&0&0\\2&2&0\\3&2&5\end{pmatrix}$
Triangular matrix (lower)	$a_{ij}=0$ above the diagonal	Δ	
5 Asymmetric	$a_{ij}\neq a_{ji}$		$\begin{pmatrix}1&1\\2&3\end{pmatrix}$
6 Null	$a_{ij}=0$	$\mathbf{0}$	$\begin{pmatrix}0&0&0\\0&0&0\end{pmatrix}$

As shown in Table A.1.1 a square matrix $\mathbf{A}(p\times p)$ is *diagonal* if $a_{ij}=0$ for all $i\neq j$. There are two convenient ways to construct diagonal matrices. If $\mathbf{a}=(a_1,\ldots,a_p)'$ is any vector and $\mathbf{B}(p\times p)$ is any square matrix then

$$\operatorname{diag}(\mathbf{a})=\operatorname{diag}(a_i)=\operatorname{diag}(a_1,\ldots,a_p)=\begin{pmatrix}a_1&\cdots&0\\ \cdot&&\cdot\\ \cdot&&\cdot\\ \cdot&&\cdot\\ 0&\cdots&a_p\end{pmatrix}$$

and

$$\operatorname{Diag}(\mathbf{B})=\begin{pmatrix}b_{11}&\cdots&0\\ \cdot&&\cdot\\ \cdot&&\cdot\\ \cdot&&\cdot\\ 0&\cdots&b_{pp}\end{pmatrix}$$

each defines a diagonal matrix.

A.2 Matrix Operations

Table A.2.1. gives a summary of various important matrix operations. We deal with some of these in detail, assuming the definitions in the table.

Table A.2.1 Basic matrix operations

Operation	Restrictions	Definitions	Remarks
1 Addition	$\mathbf{A}$, $\mathbf{B}$ of the same order	$\mathbf{A}+\mathbf{B}=(a_{ij}+b_{ij})$	
2 Subtraction	$\mathbf{A}$, $\mathbf{B}$ of the same order	$\mathbf{A}-\mathbf{B}=(a_{ij}-b_{ij})$	
3a Scalar multiplication		$c\mathbf{A}=(ca_{ij})$	
3b Inner product	$\mathbf{a}$, $\mathbf{b}$ of the same order	$\mathbf{a}'\mathbf{b}=\sum a_i b_i$	
3c Multiplication	Number of columns of $\mathbf{A}$ equals number of rows of $\mathbf{B}$	$\mathbf{AB}=(\mathbf{a}_i'\mathbf{b}_{(j)})$	$\mathbf{AB}\neq\mathbf{BA}$
4 Transpose		$\mathbf{A}'=(\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_n)$	Section A.2.1
5 Trace	$\mathbf{A}$ square	$\operatorname{tr}\mathbf{A}=\sum a_{ii}$	Section A.2.2.
6 Determinant	$\mathbf{A}$ square	$\lvert\mathbf{A}\rvert$	Section A.2.3.
7 Inverse	$\mathbf{A}$ square and $\lvert\mathbf{A}\rvert\neq 0$	$\mathbf{AA}^{-1}=\mathbf{A}^{-1}\mathbf{A}=\mathbf{I}$	$(\mathbf{A}+\mathbf{B})^{-1}\neq\mathbf{A}^{-1}+\mathbf{B}^{-1}$, Section A.2.4
8 g-inverse $(\mathbf{A}^-)$	$\mathbf{A}(n\times p)$	$\mathbf{AA}^-\mathbf{A}=\mathbf{A}$	Section A.8

A.2.1 Transpose

The transpose satisfies the simple properties

$$(\mathbf{A}')'=\mathbf{A},\qquad(\mathbf{A}+\mathbf{B})'=\mathbf{A}'+\mathbf{B}',\qquad(\mathbf{AB})'=\mathbf{B}'\mathbf{A}'.\qquad\text{(A.2.1)}$$

For partitioned $\mathbf{A}$,

$$\mathbf{A}' = \begin{bmatrix} \mathbf{A}'_{11} & \mathbf{A}'_{21} \\ \mathbf{A}'_{12} & \mathbf{A}'_{22} \end{bmatrix}.$$

If $\mathbf{A}$ is a symmetric matrix, $a_{ij} = a_{ji}$, so that

$$\mathbf{A}' = \mathbf{A}.$$

A.2.2 Trace

The trace function, $\operatorname{tr} \mathbf{A} = \Sigma a_{ii}$, satisfies the following properties for $\mathbf{A}(p \times p)$, $\mathbf{B}(p \times p)$, $\mathbf{C}(p \times n)$, $\mathbf{D}(n \times p)$, and scalar α:

$$\operatorname{tr} \alpha = \alpha, \qquad \operatorname{tr} \mathbf{A} \pm \mathbf{B} = \operatorname{tr} \mathbf{A} \pm \operatorname{tr} \mathbf{B}, \qquad \operatorname{tr} \alpha\mathbf{A} = \alpha \operatorname{tr} \mathbf{A} \tag{A.2.2a}$$

$$\operatorname{tr} \mathbf{CD} = \operatorname{tr} \mathbf{DC} = \sum_{i,j} c_{ij} d_{ji}, \tag{A.2.2b}$$

$$\sum \mathbf{x}_i' \mathbf{A} \mathbf{x}_i = \operatorname{tr}(\mathbf{AT}), \quad \text{where} \quad \mathbf{T} = \sum \mathbf{x}_i \mathbf{x}_i'. \tag{A.2.2c}$$

To prove this last property, note that since $\sum \mathbf{x}_i' \mathbf{A} \mathbf{x}_i$ is a scalar, the left-hand side of (A.2.2c) is

$$\begin{aligned} \operatorname{tr} \sum \mathbf{x}_i' \mathbf{A} \mathbf{x}_i &= \sum \operatorname{tr} \mathbf{x}_i' \mathbf{A} \mathbf{x}_i \quad \text{by (A.2.2a)} \\ &= \sum \operatorname{tr} \mathbf{A} \mathbf{x}_i \mathbf{x}_i' \quad \text{by (A.2.2b)} \\ &= \operatorname{tr} \mathbf{A} \sum \mathbf{x}_i \mathbf{x}_i' \quad \text{by (A.2.2a)}. \end{aligned}$$

As a special case of (A.2.2b) note that

$$\operatorname{tr} \mathbf{CC}' = \operatorname{tr} \mathbf{C}'\mathbf{C} = \sum c_{ij}^2. \tag{A.2.2d}$$

A.2.3 Determinants and cofactors

Definition *The determinant of a square matrix* $\mathbf{A}$ *is defined as*

$$|\mathbf{A}| = \sum (-1)^{|\tau|} a_{1\tau(1)} \cdots a_{p\tau(p)}, \tag{A.2.3a}$$

where the summation is taken over all permutations τ *of* $(1, 2, \ldots, p)$, *and* $|\tau|$ *equals* $+1$ *or* -1, *depending on whether* τ *can be written as the product of an even or odd number of transpositions.*

For $p = 2$,

$$|\mathbf{A}| = a_{11}a_{22} - a_{12}a_{21}. \tag{A.2.3b}$$

Definition *The cofactor of* a_{ij} *is defined by* $(-1)^{i+j}$ *times the minor of* a_{ij}, *where the minor of* a_{ij} *is the value of the determinant obtained after deleting the ith row and the jth column of* $\mathbf{A}$.

We denote the cofactor of a_{ij} by A_{ij}. Thus for $p = 3$,

$$A_{11} = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}, \quad A_{12} = -\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}, \quad A_{13} = \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}. \tag{A.2.3c}$$

Definition *A square matrix is* non-singular *if* $|\mathbf{A}| \neq 0$; *otherwise it is* singular.

We have the following results:

(I) $$|\mathbf{A}| = \sum_{j=1}^{p} a_{ij} A_{ij} = \sum_{i=1}^{p} a_{ij} A_{ij}, \qquad \text{any } i, j, \tag{A.2.3d}$$

but

$$\sum_{k=1}^{p} a_{ik} A_{jk} = 0, \qquad i \neq j. \tag{A.2.3e}$$

(II) If $\mathbf{A}$ is triangular or diagonal,

$$|\mathbf{A}| = \prod a_{ii}. \tag{A.2.3f}$$

(III) $|c\mathbf{A}| = c^p |\mathbf{A}|$. (A.2.3g)

(IV) $|\mathbf{AB}| = |\mathbf{A}|\,|\mathbf{B}|$. (A.2.3h)

(V) For square submatrices $\mathbf{A}(p \times p)$ and $\mathbf{B}(q \times q)$,

$$\begin{vmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{0} & \mathbf{B} \end{vmatrix} = |\mathbf{A}|\,|\mathbf{B}|. \tag{A.2.3i}$$

(VI) $$\begin{vmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{vmatrix} = |\mathbf{A}_{11}|\,|\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}| = |\mathbf{A}_{22}|\,|\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}|. \tag{A.2.3j}$$

(VII) For $\mathbf{B}(p \times n)$ and $\mathbf{C}(n \times p)$, and non-singular $\mathbf{A}(p \times p)$,

$$|\mathbf{A} + \mathbf{BC}| = |\mathbf{A}|\,|\mathbf{I}_p + \mathbf{A}^{-1}\mathbf{BC}| = |\mathbf{A}|\,|\mathbf{I}_n + \mathbf{CA}^{-1}\mathbf{B}|. \tag{A.2.3k}$$

Remarks (1) Properties (I)–(III) follow easily from the definition (A.2.3a). As an application of (I), from (A.2.3b), (A.2.3c), and (A.2.3d), we have, for $p = 3$,

$$|\mathbf{A}| = a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{31}a_{22}).$$

(2) To prove (V), note that the only permutations giving non-zero terms in the summation (A.2.3a) are those taking $\{1, \ldots, p\}$ to $\{1, \ldots, p\}$ and $\{p+1, \ldots, p+q\}$ to $\{p+1, \ldots, p+q\}$.

(3) To prove (VI), simplify $\mathbf{BAB}'$ and then take its determinant where

$$\mathbf{B} = \begin{bmatrix} \mathbf{I} & -\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}.$$

From (VI), we deduce, after putting $\mathbf{A}^{11} = \mathbf{A}$, $\mathbf{A}^{12} = \mathbf{x}'$, etc.,

$$\begin{vmatrix} \mathbf{A} & \mathbf{x} \\ \mathbf{x}' & c \end{vmatrix} = |\mathbf{A}|\{c - \mathbf{x}'\mathbf{A}^{-1}\mathbf{x}\}. \tag{A.2.3l}$$

(4) To prove the second part of (VII), simplify

$$\begin{vmatrix} \mathbf{I}_p & -\mathbf{A}^{-1}\mathbf{B} \\ \mathbf{C} & \mathbf{I}_n \end{vmatrix}$$

using (VI). As special cases of (VII) we see that, for non-singular $\mathbf{A}$,

$$|\mathbf{A} - \mathbf{bb}'| = |\mathbf{A}|(1 - \mathbf{b}'\mathbf{A}^{-1}\mathbf{b}), \tag{A.2.3m}$$

and that, for $\mathbf{B}(p \times n)$ and $\mathbf{C}(n \times p)$,

$$|\mathbf{I}_p + \mathbf{BC}| = |\mathbf{I}_n + \mathbf{CB}|. \tag{A.2.3n}$$

In practice, we can simplify determinants using the property that the value of a determinant is unaltered if a linear combination of some of the columns (rows) is added to another column (row).

(5) Determinants are usually evaluated on computers as follows. $\mathbf{A}$ is decomposed into upper and lower triangular matrices $\mathbf{A} = \mathbf{LU}$. If $\mathbf{A} > 0$, then the Cholesky decomposition is used (i.e. $\mathbf{U} = \mathbf{L}'$ so $\mathbf{A} = \mathbf{LL}'$). Otherwise the Crout decomposition is used where the diagonal elements of $\mathbf{U}$ are ones.

A.2.4 Inverse

Definition *As already defined in Table A.1.1, the inverse of* $\mathbf{A}$ *is the unique matrix* $\mathbf{A}^{-1}$ *satisfying*

$$\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}. \tag{A.2.4a}$$

The inverse exists if and only if $\mathbf{A}$ *is non-singular, that is, if and only if* $|\mathbf{A}| \neq 0$.

We write the (i, j)th element of $\mathbf{A}^{-1}$ by a^{ij}. For partitioned $\mathbf{A}$, we write

$$\mathbf{A}^{-1} = \begin{bmatrix} \mathbf{A}^{11} & \mathbf{A}^{12} \\ \mathbf{A}^{21} & \mathbf{A}^{22} \end{bmatrix}.$$

The following properties hold:

(I) $\mathbf{A}^{-1} = \dfrac{1}{|\mathbf{A}|}(A_{ij})'$. (A.2.4b)

(II) $(c\mathbf{A})^{-1} = c^{-1}\mathbf{A}^{-1}$. (A.2.4c)

(III) $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$. (A.2.4d)

(IV) The unique solution of $\mathbf{Ax} = \mathbf{b}$ is $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$. (A.2.4e)

(V) If all the necessary inverses exist, then for $\mathbf{A}(p \times p)$, $\mathbf{B}(p \times n)$, $\mathbf{C}(n \times n)$, and $\mathbf{D}(n \times p)$,

$$(\mathbf{A} + \mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{C}^{-1} + \mathbf{DA}^{-1}\mathbf{B})^{-1}\mathbf{DA}^{-1}. \tag{A.2.4f}$$

(VI) If all the necessary inverses exist, then for partitioned $\mathbf{A}$, the elements of $\mathbf{A}^{-1}$ are

$$\left.\begin{aligned} \mathbf{A}^{11} &= (\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21})^{-1}, & \mathbf{A}^{22} &= (\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})^{-1}, \\ \mathbf{A}^{12} &= -\mathbf{A}^{11}\mathbf{A}_{12}\mathbf{A}_{22}^{-1}, & \mathbf{A}^{21} &= -\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\mathbf{A}^{11}. \end{aligned}\right\} \tag{A.2.4g}$$

Alternatively, $\mathbf{A}^{12}$ and $\mathbf{A}^{21}$ can be defined by

$$\mathbf{A}^{12} = -\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{A}^{22}, \qquad \mathbf{A}^{21} = -\mathbf{A}^{22}\mathbf{A}_{21}\mathbf{A}_{11}^{-1}.$$

Remarks (1) The result (I) follows on using (A.2.3d), (A.2.3e). As a simple application, note that, for $p = 2$, we have

$$\mathbf{A}^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}.$$

(2) Formulae (II)–(VI) can be verified by checking that the product of the matrix and its inverse reduces to the identity matrix, e.g. to verify (III), we proceed

$$(\mathbf{AB})^{-1}(\mathbf{AB}) = \mathbf{B}^{-1}\mathbf{A}^{-1}(\mathbf{AB}) = \mathbf{B}^{-1}\mathbf{IB} = \mathbf{I}.$$

(3) We have assumed $\mathbf{A}$ to be a square matrix with $|\mathbf{A}| \neq 0$ in defining $\mathbf{A}^{-1}$. For $\mathbf{A}(n \times p)$, a generalized inverse is defined in Section A.8.

(4) In computer algorithms for evaluating $\mathbf{A}^{-1}$, the following methods are commonly used. If $\mathbf{A}$ is symmetric, the Cholesky method is used, namely, decomposing $\mathbf{A}$ into the form $\mathbf{LL}'$ where $\mathbf{L}$ is lower triangular and then using $\mathbf{A}^{-1} = (\mathbf{L}^{-1})'\mathbf{L}^{-1}$. For non-symmetric matrices, Crout's method is used, which is a modification of Gaussian elimination.

A.2.5 Kronecker products

Definition *Let* $\mathbf{A} = (a_{ij})$ *be an* $(m \times n)$ *matrix and* $\mathbf{B} = (b_{kl})$ *be a* $(p \times q)$ *matrix. Then the Kronecker product of* $\mathbf{A}$ *and* $\mathbf{B}$ *is defined as*

$$\begin{bmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \cdots & a_{1n}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \cdots & a_{2n}\mathbf{B} \\ \vdots & \vdots & & \vdots \\ a_{m1}\mathbf{B} & a_{m2}\mathbf{B} & \cdots & a_{mn}\mathbf{B} \end{bmatrix},$$

which is an $(mp \times nq)$ *matrix. It is denoted by* $\mathbf{A} \otimes \mathbf{B}$.