THE BLOCK VEC OPERATOR AND THE BLOCK TENSOR UNFOLDING

Tshepang More

A dissertation submitted to the Faculty of Science, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Master of Science.

Johannesburg, 2021

SCHOOL OF MATHEMATICS

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

Declaration

I declare that this dissertation is my own, unaided work. It is being submitted for the Degree of Master of Science at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other University.

____
(Signature of candidate)

3rd day of March 20 21 at Pretoria

Abstract

The Kronecker product is the unique bilinear product of matrices which arises from vectorization via the vec operator and matrix multiplication. This approach may be generalized to the block vec operator and the block Kronecker product which are defined for arbitrary partitioned block matrices. This dissertation brings to light a novel approach of describing balanced and unbalanced partitioned block matrices and discusses properties of the block vec operator and the block Kronecker product. This novel approach is extended to block tensors and block tensor unfoldings following the work in [Lin. Alg. Appl. 149,165-184(1991)] and [SIAM J. Matrix Anal. Appl. 33(1),149-169 (2012)].

Acknowledgements

I would like to express my sincere gratitude to my supervisor Professor Yorick Hardy. Your constant belief in my ability, motivation, patience, positive critiques and kindness were all that was needed to make this journey a success it has been. I am forever indebted to you. I would also like to thank the National Research Foundation for their financial assistance, I will forever be grateful. To my family, friends and classmates, thank you for your continuous love, support and guidance, you are appreciated!

Contents

List of Symbols

$M_{m\times n}$	$m \times n$ matrix
(a_{ij})	element of an $m \times n$ matrix
$[A]_{ij}$	Block matrix
$\otimes$	Kronecker product
$\otimes_s$	Symmetric Kronecker product
$\oplus_K$	Kronecker sum
$\oplus$	Direct sum
$\odot$	Hadamard product
$\circ$	Tracey-Singh product
$*$	Khatri-Rao product
$A \boxtimes B$	Block Kronecker product of partitioned (balanced and unbalanced) block matrices A and B
$A \boxminus B$	Block Kronecker product where only matrix B is partitioned (balanced and unbalanced)
$\boxplus$	Block Kronecker sum
A^T	Transpose of the matrix A
A^*	Complex conjugate of the matrix A
$\det(A)$	Determinant of an $m \times m$ matrix A
K_{mn}	Commutation matrix
vec	The vec operator
$\mathrm{vecb_c}$	The column-wise block vec operator
$\mathrm{vecb_r}$	The row-wise block vec operator
tr	Trace
I_m	$m \times m$ identity matrix
$\mathcal{A}$	Order n tensor
$A_{(n)}$	Mode-n tensor unfolding
$[\mathcal{A}]$	Order N block tensor
$[A]_{(n)}$	Mode-n block tensor unfolding
$\|\|x\|\|$	Frobenius / Euclidean norm
$\approx$	Equivalent under some isomorphism
e_i	$m \times 1$ column vector with 1 in i-th position and 0 elsewhere
E_{ij}	Matrix $e_i e_j^T$ with 1 in its i, j-th position and 0 elsewhere

Chapter 1

Introduction

In linear algebra and multilinear algebra, we study different matrix products each with its own structure and properties. The products that arise naturally are the usual matrix product, Kronecker product as well as the Hadamard product (also known as the Schur product or the entrywise product) while for the block structured matrices, the Khatri-Rao and Tracey-Singh products generalize these concepts [6]. When dealing with linear operations, the natural product is the conventional matrix multiplication while the tensor product (i.e. the Kronecker product for matrices) is the natural product for bilinear operations. To formally explore different kinds of matrix products, we first introduce a block matrix.

A block matrix is a matrix whose entries are matrices [5]. For example,

$$A = \begin{bmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \dots & A_{mn} \end{bmatrix}$$

with each matrix A_{ij} of order $p \times q$ where $i = 1, \dots, m$ and $j = 1, \dots, n$ [5]. Here A_{ij} is called a submatrix in the i-th row and j-th column of A. Therefore A is an $(mp) \times (nq)$ matrix [5]. Block matrices are perceived to have two structures, namely *balanced* and *unbalanced* partitions. In the case where all block submatrices have the same size, we say A is partitioned in a *balanced* way. An *unbalanced* partitioning has block submatrices A_{ij} of order $p_i \times q_j$ where $p_i \neq p_k$ and $q_j \neq q_l$ in general, for all $i \neq k$ and $j \neq l$ [5]. A block matrix can be visually represented with horizontal and vertical lines to distinguish the submatrices. Properties and standard operations of the conventional matrices also hold for block structured matrices.

We now introduce four matrix products collected from [6] starting with the Kronecker product which is the main product in this work, since the other matrix products considered here follow from it and it will be used throughout the dissertation.

Definition 1 *Let $A = (a_{ij})_{ij}$ be an $m \times n$ matrix and $B = (b_{ij})_{ij}$ be a $p \times q$ matrix over a field $\mathbb{F}$. The Kronecker product of A and B is defined as the $(mp) \times (nq)$ matrix [9]*

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \dots & a_{1n}B \\ a_{21}B & a_{22}B & \dots & a_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}B & a_{m2}B & \dots & a_{mn}B \end{bmatrix}.$$

Thus $A \otimes B$ is an $m \times n$ block matrix that contains $p \times q$ submatrices.

Tensor products obey the universal factorization property as follows [7]. Let V_1, V_2 and W be vector spaces. For every bilinear map $f : V_1 \times V_2 \to W$, there exists a unique linear map $\phi_f : V_1 \otimes V_2 \to W$ associated with the bilinear map $V_1 \times V_2 \to V_1 \otimes V_2$ where $\phi_f \circ \otimes = f$. Thus the following diagram commutes

$$\begin{tikzcd} V_1 \times V_2 \arrow[r, "\text{bilinear} f"] \arrow[d, "\otimes"'] & W \\ V_1 \otimes V_2 \arrow[ur, dotted, "!\phi_f"'] & \end{tikzcd}$$

For example, given a 2×3 matrix

$$A = \begin{pmatrix} 1 & 3 & 0 \\ 5 & 2 & 4 \end{pmatrix}$$

and a 3×2 matrix

$$B = \begin{pmatrix} 2 & 6 \\ 1 & 2 \\ 3 & 1 \end{pmatrix}.$$

Then

$$A \otimes B = \begin{bmatrix} 1\begin{pmatrix}2 & 6\\ 1 & 2\\ 3 & 1\end{pmatrix} & 3\begin{pmatrix}2 & 6\\ 1 & 2\\ 3 & 1\end{pmatrix} & 0\begin{pmatrix}2 & 6\\ 1 & 2\\ 3 & 1\end{pmatrix} \\ 5\begin{pmatrix}2 & 6\\ 1 & 2\\ 3 & 1\end{pmatrix} & 2\begin{pmatrix}2 & 6\\ 1 & 2\\ 3 & 1\end{pmatrix} & 4\begin{pmatrix}2 & 6\\ 1 & 2\\ 3 & 1\end{pmatrix} \end{bmatrix} = \begin{bmatrix} 2 & 6 & 6 & 18 & 0 & 0\\ 1 & 2 & 3 & 6 & 0 & 0\\ 3 & 1 & 9 & 3 & 0 & 0\\ 10 & 30 & 4 & 12 & 8 & 24\\ 5 & 10 & 2 & 4 & 4 & 8\\ 15 & 5 & 6 & 2 & 12 & 4 \end{bmatrix}$$

is a 6×6 matrix which can also be written as a 2×3 block matrix as follows

$$\left[\begin{array}{cc|cc|cc} 2 & 6 & 6 & 18 & 0 & 0\\ 1 & 2 & 3 & 6 & 0 & 0\\ 3 & 1 & 9 & 3 & 0 & 0\\ \hline 10 & 30 & 4 & 12 & 8 & 32\\ 5 & 10 & 2 & 4 & 4 & 8\\ 15 & 5 & 3 & 2 & 12 & 4 \end{array}\right].$$

Here each block or submatrix is a matrix of order 3×2. Unlike the conventional matrix multiplication, there is no restriction on the sizes of matrices when calculating the Kronecker product. In some cases, the Kronecker product is referred to as the tensor product or the direct product [9]. The Kronecker product is defined for standard matrices as shown above, we can also define it for block matrices. This introduces the next matrix product, the Tracey-Singh product.

The Tracey-Singh product generalizes the Kronecker product in the sense that [6]

$$[A \circ B]_{ij} = [[A_{ij} \otimes B_{kl}]_{kl}]_{ij}.$$

This is the Kronecker product of pairs of submatrices of each block matrix A and B that are partitioned (*balanced* or *unbalanced*) block structured matrices. For example, consider a 2×2 block matrix

$$A = \begin{bmatrix} A_{11} & A_{12}\\ A_{21} & A_{22} \end{bmatrix},$$

and a 1×2 block matrix

$$B = \begin{bmatrix} B_{11} & B_{12} \end{bmatrix},$$

then

$$A \circ B = \left[\begin{array}{cc|cc} A_{11} \otimes B_{11} & A_{11} \otimes B_{12} & A_{12} \otimes B_{11} & A_{12} \otimes B_{12} \\ \hline A_{21} \otimes B_{11} & A_{21} \otimes B_{12} & A_{22} \otimes B_{11} & A_{22} \otimes B_{12} \end{array}\right].$$

This matrix product is regarded as the block Kronecker product and is discussed fully in Chapter 3 (where it is called the balanced partitioned block Kronecker product).

The Hadamard product on the other hand takes two matrices of the same order and produces a third matrix of the same order and is denoted by $\odot$ [6]. Let A and B be $m \times n$ matrices, the Hadamard product of A and B is defined entrywise by [7]

$$(A \odot B)_{ij} = (a_{ij} b_{ij})_{ij}.$$

For example, if A and B both are 3×3 matrices, with

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix},$$

then

$$A \odot B = \begin{bmatrix} a_{11}b_{11} & a_{12}b_{12} & a_{13}b_{13} \\ a_{21}b_{21} & a_{22}b_{22} & a_{23}b_{23} \\ a_{31}b_{31} & a_{32}b_{32} & a_{33}b_{33} \end{bmatrix}.$$

Similar to the Kronecker product, we can generalize the Hadamard product to produce another matrix product defined for partitioned matrices [6]. This product is called the Khatri-Rao product (for *balanced* and *unbalanced* partitioning) and it generalizes the Hadamard product in the following way [6]

$$[A * B]_{ij} = (A_{ij} \otimes B_{ij})_{ij}.$$

For example, if A and B both are 2×2 partitioned block matrices,

$$A = \left[\begin{array}{cc|cc} 1 & 2 & 3 & 6 \\ 2 & 4 & 9 & 3 \\ \hline 7 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{array}\right] \quad \text{and} \quad B = \left[\begin{array}{cc|cc} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 3 \\ \hline 2 & 2 & 7 & 1 \\ 1 & 5 & 0 & 0 \end{array}\right],$$

then

$$A * B = \begin{bmatrix} A_{11} \otimes B_{11} & A_{12} \otimes B_{12} \\ A_{21} \otimes B_{21} & A_{22} \otimes B_{22} \end{bmatrix}$$

$$= \begin{bmatrix} \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \otimes \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & \begin{pmatrix} 3 & 6 \\ 9 & 3 \end{pmatrix} \otimes \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix} \\ \begin{pmatrix} 7 & 1 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 2 & 2 \\ 1 & 5 \end{pmatrix} & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 7 & 1 \\ 0 & 0 \end{pmatrix} \end{bmatrix}$$

$$= \left[\begin{array}{cccc|cccc} 0 & 1 & 0 & 2 & 3 & 3 & 6 & 6 \\ 1 & 0 & 2 & 0 & 3 & 9 & 6 & 18 \\ 0 & 2 & 0 & 4 & 9 & 9 & 3 & 3 \\ 2 & 0 & 4 & 0 & 9 & 27 & 3 & 9 \\ \hline 14 & 14 & 2 & 2 & 7 & 1 & 0 & 0 \\ 7 & 35 & 1 & 5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 2 & 0 & 0 & 7 & 1 \\ 0 & 0 & 1 & 5 & 0 & 0 & 0 & 0 \end{array}\right].$$

This matrix product produces a block structured matrix that consists of block matrices as its submatrices. It is important to note that the Kronecker product is not commutative while the Hadamard product is commutative. The Tracey-Singh product and the Khatri-Rao product do not commute also since they are defined in terms of the Kronecker product. These matrix products are connected since some properties can be derived from other products, for example the Hadamard product of two matrices can be written in terms of the Kronecker product as follows

$$A \odot B = M^T (A \otimes B) N$$

where M and N are selection matrices of order $m^2 \times m$ and $n^2 \times n$, respectively [6]. The matrix products are useful tools in areas of multilinear algebra, as well as in econometrics and statistics [6]. We now define the vec operator which is closely related to the Kronecker product.

Definition 2 *The vec operator is an operation that converts an $m \times n$ matrix into an $(mn) \times 1$ column vector by stacking each column of the matrix underneath the other [9].*

For example, given a 2×3 matrix

$$A = \begin{bmatrix} 1 & 3 & 0 \\ 0 & 4 & 9 \end{bmatrix}.$$

Then

$$\operatorname{vec}(A) = \begin{bmatrix} 1 \\ 0 \\ 3 \\ 4 \\ 0 \\ 9 \end{bmatrix}$$

is a 6×1 column vector. The vec operator also provides the relation between the usual matrix product and the Kronecker product, and we see this in detail in Chapter 2. The vec operator is a vector space homomorphism and is compatible with the Hadamard product in the sense that [6]

$$\operatorname{vec}(A \odot B) = \operatorname{vec}(A) \odot \operatorname{vec}(B).$$

A generalization of the vec operator is the block vec operator, the vectorization of a block matrix, denoted by vecb [5]. There exists two kinds of vecb operator, namely $\operatorname{vecb}_{\mathrm{c}}$ which stacks the vec of each submatrix of a block matrix, and $\operatorname{vecb}_{\mathrm{r}}$ which stacks the vec of each submatrix of the block transpose of the block matrix. For example, consider a 2×3 block matrix

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{bmatrix}.$$

Then we can vectorize A column-blockwise, i.e.

$$\operatorname{vecb}_{\mathrm{c}}(A) = \begin{bmatrix} \operatorname{vec}(A_{11}) \\ \operatorname{vec}(A_{21}) \\ \operatorname{vec}(A_{12}) \\ \operatorname{vec}(A_{22}) \\ \operatorname{vec}(A_{13}) \\ \operatorname{vec}(A_{23}) \end{bmatrix}. \tag{1.1}$$

We can also vectorize A row-blockwise denoted

$$\mathrm{vecb_r}(A) = \begin{bmatrix} \mathrm{vec}(A_{11}) \\ \mathrm{vec}(A_{12}) \\ \mathrm{vec}(A_{13}) \\ \mathrm{vec}(A_{21}) \\ \mathrm{vec}(A_{22}) \\ \mathrm{vec}(A_{23}) \end{bmatrix}. \tag{1.2}$$

Similar to the block vec operator, we have the block tensor unfolding. Here, a block tensor is defined in a similar way to the block matrix and is a tensor that consists of tensors as it entries [8]. We can transform a tensor into a matrix and this process is known as the tensor unfolding. A generalization of tensor unfolding is called the block tensor unfolding. The mathematical definition is provided in Chapter 4. The main focus of this dissertation will be on the block vec operator and the block tensor unfolding.

Before we begin with the next chapter, we first define the commutation matrix which will be useful in defining some of the properties throughout this dissertation.

Definition 3 *The commutation matrix K_{mn} is given by*

$$K_{mn} = \sum_{j=1}^{m} \sum_{k=1}^{n} E_{jk} \otimes E_{jk}^T,$$

where E_{jk} is an $m \times n$ matrix.

For example,

$$K_{23} = \left[\begin{array}{cc|cc|cc} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ \hline 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right].$$

The commutation matrix is sometimes referred to as the permutation matrix P (perfect shuffle and is defined in Van Loan [11], Schäcke [9] and Henderson and Searle's [3] vec permutation article) and is useful for transforming the vec operator to the vec of the transpose. In other words,

$$\mathrm{vec}(A^T) = K_{mn}\,\mathrm{vec}(A)$$

for every $m \times n$ matrix A [3]. Notice that $\mathrm{vec}(A)$ and $\mathrm{vec}(A^T)$ have the same number of entries but in a different order [3]. The other main function of the commutation matrix is to reverse the order of the Kronecker product. As stated above, the Kronecker product is not commutative, so we can use the commutation matrix to reverse the order. In particular, let A and B be matrices of order $m \times n$ and $p \times q$, respectively. Then [3, 9]

$$B \otimes A = K_{pm}(A \otimes B)K_{qn}^T = K_{mp}^T(A \otimes B)K_{nq}.$$

These two properties of both the Kronecker product and the vec operator appear again in Chapter 2 as part of the main discussion.

Chapter 2

The vec operator and Kronecker product

This chapter provides properties of the vec operator and the Kronecker product. The Kronecker product arises when applying the vec operator to matrix products and it has an important role in solving linear matrix equations [9]. We use both the Kronecker product and the vec operator to obtain equivalent equations which are provided in the example below, collected from [9].

Example 1 (Matrix equations) *The following are linear equations of unknown matrices X, Y and their equivalent systems of equations defined direclty from the mixed product property (Theorem 1, below).*

1. $AX = B \iff (I \otimes A)\operatorname{vec}(X) = \operatorname{vec}(B)$
2. $AX + XB = C \iff [(I \otimes A) + (B^T \otimes I)]\operatorname{vec}(X) = \operatorname{vec}(C)$
3. $AXB = C \iff (B^T \otimes A)\operatorname{vec}(X) = \operatorname{vec}(C)$
4. $AX + YB = C \iff (I \otimes A)\operatorname{vec}(X) + (B^T \otimes I)\operatorname{vec}(Y) = \operatorname{vec}(C)$

The basic properties of the vec operator collected from [3, 9] are stated below and we will refer to them later in the dissertation.

Properties of the vec **operator**

1. $\operatorname{vec}(ab^T) = b \otimes a$, where a and b are column vectors.
2. There exists a permutation matrix K_{mn} such that $\operatorname{vec}(A^T) = K_{mn}\operatorname{vec}(A)$ for all $m \times n$ matrices A.
3. $\operatorname{vec}(A + B) = \operatorname{vec}(A) + \operatorname{vec}(B)$.

4. For all $\alpha \in \mathbb{F}$, $\operatorname{vec}(\alpha A) = \alpha \operatorname{vec}(A)$.

5. $\operatorname{tr}(A^T B) = \operatorname{vec}(A)^T \operatorname{vec}(B)$.
   The following results will be used to prove many properties throughout the dissertation.

   Theorem 1 (The mixed product property) *Let A, B, C, D be matrices of order $m \times n$, $p \times q$, $n \times s$ and $q \times t$, respectively. Then*

$$(A \otimes B)(C \otimes D) = (AC) \otimes (BD).$$

   Proof *Let $A = (a_{ik})_{ik}$ and $C = (c_{kj})_{kj}$ then, by the properties of matrix multiplication (blockwise)*

$$\begin{aligned}
&(A \otimes B)(C \otimes D) \\
&\quad = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{bmatrix} \begin{bmatrix} c_{11}D & c_{12}D & \cdots & c_{1s}D \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1}D & c_{n2}D & \cdots & c_{ns}D \end{bmatrix} \\
&\quad = \begin{bmatrix} (a_{11}c_{11}+ & \cdots & +a_{1n}c_{n1})BD & \cdots & (a_{11}c_{1s}+ & \cdots & +a_{1n}c_{ns})BD \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ (a_{m1}c_{11}+ & \cdots & +a_{mn}c_{n1})BD & \cdots & (a_{m1}c_{1s}+ & \cdots & +a_{mn}c_{ns})BD \end{bmatrix} \\
&\quad = \left[\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1s} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{ns} \end{pmatrix} \right] \otimes (BD) \\
&\quad = (AC) \otimes (BD).
\end{aligned}$$

The next theorem provides an important application of the vec operator which enables us to express matrix multiplication as a linear transformation on (vectorized) matrices in terms of the Kronecker product.

Theorem 2 *If A, B, C are matrices of order $m \times n$, $n \times q$, $q \times t$, respectively. Then [3]*

$$\operatorname{vec}(ABC) = (C^T \otimes A) \operatorname{vec}(B).$$

Proof *Let $B = ab^T$ where a is an $n \times 1$ column vector and b^T a $1 \times q$ row* *vector. Then by property* 1 *and Theorem 1 above*

$$\begin{aligned}
\mathrm{vec}(ABC) &= \mathrm{vec}(Aab^T C) \\
&= \mathrm{vec}((Aa)(C^T b)^T) \\
&= C^T b \otimes Aa \\
&= (C^T \otimes A)(b \otimes a) \\
&= (C^T \otimes A)\,\mathrm{vec}(ab^T) \\
&= (C^T \otimes A)\,\mathrm{vec}(B).
\end{aligned}$$

The result follows by linearity of the vec *operator and writing B as the sum of rank 1 matrices.*

Corollary 1 *Let A be an $m \times n$ matrix and B be an $n \times q$, by Theorem 2*

$$\mathrm{vec}(AB) = \mathrm{vec}(ABI_q) = (I_q \otimes A)\,\mathrm{vec}(B)$$

and

$$\mathrm{vec}(AB) = \mathrm{vec}(I_m AB) = (B^T \otimes I_m)\,\mathrm{vec}(A).$$

Corollary 1 together with linearity of the vec yields the four matrix results in Example 1. Similarly, since $A = AI_nI_n$, then by Theorem 2, we obtain an explicit formula for the vec operator namely

$$\mathrm{vec}(A) = \sum_{i=1}^{n} e_i \otimes Ae_i.$$

This is the $(mn) \times 1$ column vector where $e_1, \ldots, e_n \in \mathbb{F}$ is the standard basis. A generalization of Theorem 2 will be provided for block matrices in the next chapter and we use the property to prove other common properties of the vecb operator.

The following theorem provides basic properties of the Kronecker product collected from [9] and we will omit some of the proofs which are similar to the others.

Theorem 3 (Basic properties of the Kronecker product) *Let A, B and C be matrices of order $m \times n$, $p \times q$ and $k \times l$, respectively. Then [3]*

1. *$b^T \otimes a = a \otimes b^T = ab^T$ for column vectors a and b.*
2. *For all $\alpha \in \mathbb{F}$, $\alpha(A \otimes B) = (\alpha A) \otimes B = A \otimes (\alpha B)$.*

3. $(A \otimes B)^T = A^T \otimes B^T$.
4. $(A \otimes B)^* = A^* \otimes B^*$.
5. $A \otimes (B \otimes C) = (A \otimes B) \otimes C$.
6. $A \otimes (B + C) = (A \otimes B) + (A \otimes C)$ *when* $k = p$ *and* $l = q$.
7. $(A + B) \otimes C = (A \otimes C) + (B \otimes C)$ *when* $m = p$ *and* $n = q$.
8. $\operatorname{tr}(A \otimes B) = \operatorname{tr}(A)\operatorname{tr}(B) = \operatorname{tr}(B \otimes A)$ *where* $m = n$ *and* $p = q$.
9. $I_m \otimes I_n = I_{mn}$.
10. $B \otimes A = {K_{mp}}^T (A \otimes B) K_{nq}$.
11. $\det(A \otimes B) = \det(A)^p \det(B)^m$ *where* $m = n$ *and* $p = q$.

Proof

Proof of 1. Let a *and* b *be column vectors of order* $m \times 1$ *and* $n \times 1$, *respectively. Then*

$$\begin{aligned} b^T \otimes a &= \begin{pmatrix} b_1 & \cdots & b_n \end{pmatrix} \otimes \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} \\ &= \begin{pmatrix} b_1 a_1 & \cdots & b_n a_1 \\ \vdots & \ddots & \vdots \\ b_1 a_m & \cdots & b_n a_m \end{pmatrix} \\ &= \begin{pmatrix} a_1 b_1 & \cdots & a_1 b_n \\ \vdots & \ddots & \vdots \\ a_m b_1 & \cdots & a_m b_n \end{pmatrix} \\ &= \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} \begin{pmatrix} b_1 & \cdots & b_n \end{pmatrix} \\ &= a b^T. \end{aligned}$$

Proof of 2.

$$\begin{aligned}
\alpha(A \otimes B) &= \alpha \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{bmatrix} \\
&= \begin{bmatrix} \alpha a_{11}B & \alpha a_{12}B & \cdots & \alpha a_{1n}B \\ \alpha a_{21}B & \alpha a_{22}B & \cdots & \alpha a_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ \alpha a_{m1}B & \alpha a_{m2}B & \cdots & \alpha a_{mn}B \end{bmatrix} \\
&= \begin{bmatrix} a_{11}\alpha B & a_{12}\alpha B & \cdots & a_{1n}\alpha B \\ a_{21}\alpha B & a_{22}\alpha B & \cdots & a_{2n}\alpha B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}\alpha B & a_{m2}\alpha B & \cdots & a_{mn}\alpha B \end{bmatrix} \\
&= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \otimes (\alpha B) \\
&= A \otimes (\alpha B).
\end{aligned}$$

Proof of 3.

$$\begin{aligned}
(A \otimes B)^T &= \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}B & a_{m2}B & \dots & a_{mn}B \end{bmatrix}^T \\
&= \begin{bmatrix} (a_{11}B)^T & (a_{21}B)^T & \cdots & (a_{m1}B)^T \\ (a_{12}B)^T & (a_{22}B)^T & \cdots & (a_{m2}B)^T \\ \vdots & \vdots & \ddots & \vdots \\ (a_{1n}B)^T & (a_{2n}B)^T & \cdots & (a_{mn}B)^T \end{bmatrix} \\
&= \begin{bmatrix} a_{11}B^T & a_{21}B^T & \cdots & a_{m1}B^T \\ a_{12}B^T & a_{22}B^T & \cdots & a_{m2}B^T \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n}B^T & a_{2n}B^T & \cdots & a_{mn}B^T \end{bmatrix} \\
&= A^T \otimes B^T.
\end{aligned}$$

Proof of 4. The proof of 4 is similar to the proof of 3, by taking the complex conjugate throughout.

Proof of 5. Let $D = B \otimes C$, then

$$\begin{aligned}
A \otimes (B \otimes C) &= A \otimes D \\
&= \begin{bmatrix} a_{11}D & \cdots & a_{1n}D \\ \vdots & \ddots & \vdots \\ a_{m1}D & \cdots & a_{mn}D \end{bmatrix} \\
&= \begin{bmatrix} \begin{bmatrix} a_{11}b_{11}C & \cdots & a_{11}b_{1q}C \\ \vdots & \ddots & \vdots \\ a_{11}b_{p1}C & \cdots & a_{11}b_{pq}C \end{bmatrix} & \cdots & \begin{bmatrix} a_{1n}b_{11}C & \cdots & a_{1n}b_{1q}C \\ \vdots & \ddots & \vdots \\ a_{1n}b_{p1}C & \cdots & a_{1n}b_{pq}C \end{bmatrix} \\ \vdots & \ddots & \vdots \\ \begin{bmatrix} a_{m1}b_{11}C & \cdots & a_{m1}b_{1q}C \\ \vdots & \ddots & \vdots \\ a_{m1}b_{p1}C & \cdots & a_{m1}b_{pq}C \end{bmatrix} & \cdots & \begin{bmatrix} a_{mn}b_{11}C & \cdots & a_{mn}b_{1q}C \\ \vdots & \ddots & \vdots \\ a_{mn}b_{p1}C & \cdots & a_{mn}b_{pq}C \end{bmatrix} \end{bmatrix} \\
&= \begin{bmatrix} (a_{11}B)_{11}C & \cdots & (a_{1n}B)_{1q}C \\ \vdots & \ddots & \vdots \\ (a_{m1}B)_{p1}C & \cdots & (a_{mn}B)_{pq}C \end{bmatrix} \\
&= (A \otimes B) \otimes C.
\end{aligned}$$

Proof of 6.

$$\begin{aligned}
A \otimes (B + C) &= \begin{bmatrix} a_{11}(B+C) & a_{12}(B+C) & \cdots & a_{1n}(B+C) \\ a_{21}(B+C) & a_{22}(B+C) & \cdots & a_{2n}(B+C) \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}(B+C) & a_{m2}(B+C) & \cdots & a_{mn}(B+C) \end{bmatrix} \\
&= \begin{bmatrix} a_{11}B + a_{11}C & a_{12}B + a_{12}C & \cdots & a_{1n}B + a_{1n}C \\ a_{21}B + a_{21}C & a_{22}B + a_{22}C & \cdots & a_{2n}B + a_{2n}C \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}B + a_{m1}C & a_{m2}B + a_{m2}C & \cdots & a_{mn}B + a_{mn}C \end{bmatrix} \\
&= \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{bmatrix} + \begin{bmatrix} a_{11}C & a_{12}C & \cdots & a_{1n}C \\ a_{21}C & a_{22}C & \cdots & a_{2n}C \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}C & a_{m2}C & \cdots & a_{mn}C \end{bmatrix}
\end{aligned}$$

$$= (A \otimes B) + (A \otimes C).$$

Proof of 7. The proof of 7 is similar to the proof of 6, by taking the product of the sum of the two matrices and C.

Proof of 8. We have

$$\begin{aligned}
\operatorname{tr}(A \otimes B) &= \sum_{i=1}^{n} \operatorname{tr}(a_{ii} B) \\
&= \sum_{i=1}^{n} a_{ii} \operatorname{tr}(B) \\
&= \operatorname{tr}(A) \operatorname{tr}(B) \\
&= \operatorname{tr}(B) \operatorname{tr}(A) \\
&= \operatorname{tr}(B \otimes A).
\end{aligned}$$

Proof of 9.

$$\begin{aligned}
I_m \otimes I_n &= \sum_{i=1}^{m} E_{ii} \otimes \sum_{j=1}^{n} E_{jj} \\
&= \sum_{i=1}^{m} \sum_{j=1}^{n} (E_{ii} \otimes E_{jj}) \\
&= \sum_{i=1}^{m} \sum_{j=1}^{n} (e_i e_i^T) \otimes (e_j e_j^T) \\
&= \sum_{i=1}^{m} \sum_{j=1}^{n} (e_i \otimes e_j)(e_i \otimes e_j)^T \\
&= \sum_{k=1}^{mn} e_k e_k^T \\
&= I_{mn}.
\end{aligned}$$

Proof of 10. By Definition 3, we define the commutation matrices K_{mp}^T and K_{nq} as

$$K_{mp}^T = \sum_{i=1}^{m} \sum_{j=1}^{p} E_{ij}^T \otimes E_{ij} = \sum_{i=1}^{m} \sum_{j=1}^{p} e_j e_i^T \otimes e_i e_j^T$$

and

$$K_{nq} = \sum_{k=1}^{n} \sum_{l=1}^{q} E_{kl} \otimes E_{kl}^T = \sum_{k=1}^{n} \sum_{l=1}^{q} e_k e_l^T \otimes e_l e_k^T,$$

respectively. Using properties 1 and 2 of Theorem 3 as well as the mixed product property, we get

$$\begin{aligned}
K_{mp}^T(A\otimes B)K_{nq} &= \sum_{i=1}^{m}\sum_{j=1}^{p}\sum_{k=1}^{n}\sum_{l=1}^{q}(e_je_i^T\otimes e_ie_j^T)(A\otimes B)(e_ke_l^T\otimes e_le_k^T)\\
&= \sum_{i=1}^{m}\sum_{j=1}^{p}\sum_{k=1}^{n}\sum_{l=1}^{q} e_je_i^TAe_ke_l^T\otimes e_ie_j^TBe_le_k^T\\
&= \sum_{i=1}^{m}\sum_{j=1}^{p}\sum_{k=1}^{n}\sum_{l=1}^{q}(e_j\otimes e_ie_j^TBe_l)(e_i^TAe_ke_l^T\otimes e_k^T)\\
&= \sum_{i=1}^{m}\sum_{j=1}^{p}\sum_{k=1}^{n}\sum_{l=1}^{q}(e_j\otimes e_i\otimes e_j^TBe_l)(e_i^TAe_k\otimes e_l^T\otimes e_k^T)
\end{aligned}$$

and since $e_j^TBe_l$ *is a constant, by Property 1 of Theorem 3 we get*

$$\begin{aligned}
&\sum_{i=1}^{m}\sum_{j=1}^{p}\sum_{k=1}^{n}\sum_{l=1}^{q}(e_j\otimes e_j^TBe_l\otimes e_i)(e_l^T\otimes e_i^TAe_k\otimes e_k^T)\\
&= \sum_{i=1}^{m}\sum_{j=1}^{p}\sum_{k=1}^{n}\sum_{l=1}^{q}(e_je_j^TBe_l\otimes e_i)(e_l^T\otimes e_i^TAe_ke_k^T).
\end{aligned}$$

By Proof of Property 9 we get

$$\begin{aligned}
&\sum_{i=1}^{m}\sum_{l=1}^{q}\left(\sum_{j=1}^{p}E_{jj}Be_l\otimes e_i\right)\left(e_l^T\otimes e_i^TA\sum_{k=1}^{n}e_ke_k^T\right)\\
&= \sum_{i=1}^{m}\sum_{l=1}^{q}(Be_l\otimes e_i)(e_l^T\otimes e_i^TA)\\
&= \sum_{i=1}^{m}\sum_{l=1}^{q}(Be_le_l^T\otimes e_ie_i^TA)\\
&= B\otimes A.
\end{aligned}$$

Proof of 11. Since $A\otimes B=(A\otimes I_p)(I_m\otimes B)$ *it follows that*

$$\begin{aligned}
\det(A\otimes B) &= \det((A\otimes I_p)(I_m\otimes B))\\
&= \det(A\otimes I_p)\det(I_m\otimes B).
\end{aligned}$$

We use direct sums (Chapter 3) to calculate $\det(I_m\otimes B)$ *and* $\det(A\otimes I_p)$

follows in the same way. Consider the direct sum

$$I_m \oplus B = \begin{bmatrix} I_m & 0_{m\times p} \\ 0_{p\times m} & B \end{bmatrix}.$$

We know from the Leibniz formula that the determinant of a general $s \times s$ matrix is

$$\det(A) = \sum_{\sigma} sgn(\sigma) A_{1\sigma(1)} A_{2\sigma(2)} \cdots A_{s\sigma(s)}.$$

with the signature of the permutation σ given as

$$sgn(\sigma) = \begin{cases} 1 & \text{if } \sigma \text{is even} \\ -1 & \text{if } \sigma \text{is odd.} \end{cases}$$

As a result

$$\begin{aligned} \det(I_m \oplus B) &= \sum_{\substack{\sigma \\ \sigma(1)=1 \\ \vdots \\ \sigma(m)=m}} sgn(\sigma) 1 \cdot 1 \cdot \cdots \cdot 1 B_{1\sigma(m+1)} B_{2\sigma(m+2)} \cdots B_{p\sigma(m+p)} \\ &= \sum_{\sigma} sgn(\sigma) 1 \cdot 1 \cdot \cdots \cdot 1 B_{1\sigma(1)} B_{2\sigma(2)} \cdots B_{p\sigma(p)} \\ &= \det(B). \end{aligned}$$

Similarly we find

$$\det(B \oplus I_m) = \det(B).$$

We have that

$$\begin{aligned} \det(I_m \otimes B) &= \det(B \oplus (I_{m-1} \otimes B)) \\ &= \det(B \oplus I_{(m-1)p}) \det(I_p \oplus (I_{m-1} \otimes B)). \end{aligned}$$

Since

$$\begin{aligned} &\det(I_p \oplus (I_{m-1} \otimes B)) \\ &\quad = \det(I_{m-1} \otimes B) \\ &\quad = \det((B \oplus I_{(m-2)p})(I_p \oplus (I_{m-2} \otimes B))) \\ &\quad = \det(B \oplus I_{(m-2)p}) \det(I_{m-2} \otimes B) \\ &\quad = \det(B \oplus I_{(m-2)p}) \det(B \oplus I_{(m-3)p}) \cdots \det(B \oplus I_p) \det(I_p \oplus B) \end{aligned}$$

then

$$\begin{aligned}\det(I_m \otimes B) &= \det(B \oplus I_{(m-1)p}) \det(B \oplus I_{(m-2)p}) \cdots \det(I_p \oplus B) \\ &= \det(B) \det(B) \cdots \det(B) \\ &= \det(B)^m.\end{aligned}$$

In a similar way we can show that $\det(A \otimes I_p) = \det(A)^p$, *hence*

$$\begin{aligned}\det(A \otimes B) &= \det(A \otimes I_p) \det(I_m \otimes B) \\ &= \det(A)^p \det(B)^m.\end{aligned}$$

Theorem 4 *Given an* $n \times n$ *matrix* A *and a* $m \times m$ *matrix* B, *then* $(A \otimes B)^{-1}$ *exists if and only if* A^{-1} *and* B^{-1} *exist, and*

$$(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}.$$

Proof *From linear algebra we know that the inverse of matrices are unique (when they exist). Suppose that* A^{-1}, B^{-1} *exists. Then by Theorem 1,*

$$(A \otimes B)(A^{-1} \otimes B^{-1}) = (AA^{-1}) \otimes (BB^{-1})$$

and

$$\begin{aligned}(AA^{-1}) \otimes (BB^{-1}) &= (I_n \otimes I_m) \\ &= I_{nm}.\end{aligned}$$

It follows that $A \otimes B$ *is non-singular and we identify*

$$(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}.$$

Since inverses are unique, $(A \otimes B)^{-1}$ *exists if and only if* $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$ *so that* A^{-1} *and* B^{-1} *exist. Therefore both* A *and* B *are non-singular.*

Recall: We denote by $\sigma(A)$ the spectrum of a matrix which is the set of all eigenvalues. A square matrix A of order $m \times m$ has eigenvalue $\lambda \in \sigma(A)$ with corresponding eigenvector $x \neq 0$ if [1]

$$Ax = \lambda x.$$

Theorem 5 *Let A and B be square matrices of order $m \times m$ and $n \times n$, respectively. Then $\lambda\mu$ is an eigenvalue of $A \otimes B$ with the corresponding eigenvector $x \otimes y$ when $\lambda \in \sigma(A)$ and $\mu \in \sigma(B)$ and x and y are corresponding eigenvectors for λ and μ, respectively.*

Proof *From above,*

$$Ax = \lambda x, \qquad By = \mu y.$$

Then

$$Ax \otimes By = \lambda x \otimes \mu y.$$

By Theorem 1 and property 2 of Theorem 3, we get

$$(Ax \otimes By) = (A \otimes B)(x \otimes y) \qquad \text{and} \qquad (\lambda x) \otimes (\mu y) = \lambda(x \otimes \mu y).$$

Therefore,

$$\begin{aligned}(A \otimes B)(x \otimes y) &= \lambda(x \otimes \mu y)\\ &= \lambda\mu(x \otimes y).\end{aligned}$$

Consequently, $(A \otimes B)(x \otimes y) = \lambda\mu(x \otimes y)$, and $x \otimes y \neq 0$.

The Kronecker product also has many applications in areas of quantum computing, signal and image processing as well as semi-definite programming [9]. It is easy to see that the Kronecker product has properties which are similar to the conventional matrix multiplication. In particular, associativity, distributivity as well as the order independence of scalar multiplication holds for both [9]. These properties become useful throughout the dissertation as they will help us to prove properties of other operations. It is clear that the vec operator allows us to convert matrix equations into systems of equations using the Kronecker product that defines the coefficient matrix. The Kronecker product induces two additional operations, namely the Kronecker sum which arises naturally from the second matrix equation in Example 1 and is defined as

$$(A \oplus_K B^T)\,\mathrm{vec}(X) = \mathrm{vec}(C)$$

where the Kronecker sum

$$(A \oplus_K B) = (I_n \otimes A) + (B \otimes I_m)$$

is defined for square matrices A and B of order $m \times m$ and $n \times n$, respectively

[9]. In general, the distributive law does not hold for the Kronecker sum, i.e.

$$(A \oplus_K B) \otimes C \neq (A \otimes C) \oplus_K (B \otimes C)$$

and

$$A \otimes (B \oplus_K C) \neq (A \otimes B) \oplus_K (A \otimes C).$$

The second operation is called the symmetric product and is also known as the Jordan product where multiplication of matrices is the Kronecker product, i.e. $A \otimes B + B \otimes A$. The symmetric Kronecker product has properties that are similar to the Kronecker product and is discussed fully in [9]. The next chapter provides a detailed discussion on the generalized vec operator as well as relevant properties.

Chapter 3

The vecb operator

In Chapter 2, we considered the Kronecker product and the vec operator both defined for conventional matrices of order $m \times n$, where $m, n \in \mathbb{N}$. In this chapter we discuss the generalization of these operations which are defined for block matrices, namely the block Kronecker product and the block vec operator. They follow the same operations and have similar properties to the Kronecker product and the vec operator defined for conventional matrices. Both the block Kronecker product and the block vec operator are presented in two parts with reference to the structure of our block matrix (the balanced and unbalanced partitioned block matrix). We denote by $A \boxtimes B$, the block Kronecker product defined for partitioned block matrices A and B (i.e. both A and B are partitioned in a balanced or unbalanced way). On the other hand $A \boxminus B$ denotes the block Kronecker product where only one block matrix is partitioned and the other is a general non-partitioned matrix. The algebra of the block Kronecker product $A \boxtimes B$ and the block vec operator for balanced partitioned matrices is already defined in [5]. We derive and introduce the algebra of the block Kronecker product $A \boxtimes B$ in the unbalanced case and one for $A \boxminus B$ defined in both cases following the work in [5]. We will also introduce the block vec operator for unbalanced partitioned matrices and discuss some of its properties. We denote the block vec operator by vecb which also arises in two parts, the $\mathrm{vecb_c}$ operator which induces the block Kronecker product $A \boxtimes B$ as well as the $\mathrm{vecb_r}$ operator which induces the block Kronecker product $A \boxminus B$.

Before we begin with the main discussion of the vecb operator and the block Kronecker product, we first define an unbalanced block matrix. Let A be an

$m \times n$ matrix where m and n are partitioned $m = m_1 + m_2 + \cdots + m_k$ and $n = n_1 + n_2 + \cdots + n_l$. Then we say A is an unbalanced partitioned block matrix where each submatrix A_{ij} may have a different size $m_i \times n_j$ [5], for example,

$$A = \left[\begin{array}{c|cc} 3 & 1 & 0 \\ \hline 4 & 1 & 5 \\ 0 & 2 & 6 \end{array}\right].$$

To introduce an algebraic expression for this matrix, we first introduce E-matrices which are used to identify the entry of the submatrix or the block in which each A_{ij} belongs. We denote such a matrix by E_{ij} and it is a matrix that consists of 1 in its i,j-position and 0 elsewhere. Let $\mathbf{m} = (m_1, \ldots, m_k)$ and $\mathbf{n} = (n_1, \ldots, n_l)$ be vectors of integers defined from the blocks of an $m \times n$ block matrix. We define the $k \times k$ block matrix $E_{e_p,\mathbf{m}}$ of order $m \times km_p$ that consists of the identity matrix I_{m_p} in its p-th row and column and a zero matrix $0_{m_q \times m_p}$ in its q-th row and r-th column. Here $e_p \in \mathbb{F}^k$ is the element of the standard basis in $\mathbb{F}^k$ with a 1 in entry p and 0 in all other entries. In particular,

$$E_{e_p,\mathbf{m}} = \left(\bigoplus_{i=1}^{p-1} 0_{m_i \times m_p} \right) \oplus I_{m_p} \oplus \left(\bigoplus_{i=p+1}^{k} 0_{m_i \times m_p} \right) \tag{3.1}$$

where the direct sum $A \oplus B$ of two matrices A and B of order $m \times n$ and $p \times q$, respectively is defined as

$$A \oplus B = \begin{bmatrix} A & 0_{m \times q} \\ 0_{p \times n} & B \end{bmatrix}.$$

Example 2 *Let* $\mathbf{m} = (1, 2)$. *We can see that* $k = 2$. *Choose* $p = 1$, *then*

$$\begin{aligned} E_{e_1,\mathbf{m}} &= I_{m_1} \oplus \left(\bigoplus_{i=2}^{2} 0_{m_i \times m_1} \right) = I_1 \oplus 0_{2\times 1} \\ &= \begin{bmatrix} I_1 & 0_{1\times 1} \\ 0_{2\times 1} & 0_{2\times 1} \end{bmatrix} \\ &= \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & 0 \\ 0 & 0 \end{array}\right]. \end{aligned}$$

Choose $p = 2$, *then*

$$\begin{aligned} E_{e_2,\mathbf{m}} &= \left(\bigoplus_{i=1}^{1} 0_{m_i \times m_2}\right) \oplus I_{m_2} = 0_{1\times 2} \oplus I_2 \\ &= \begin{bmatrix} 0_{1\times 2} & 0_{1\times 2} \\ 0_{2\times 2} & I_2 \end{bmatrix} \\ &= \left[\begin{array}{cc|cc} 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right]. \end{aligned}$$

The direct sum has some properties which are similar to the Kronecker product, e.g. the mixed product property seen above also arises in the case of direct sums for product-compatible matrices, that is

$$\begin{aligned} (A \oplus B)(C \oplus D) &= \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \begin{bmatrix} C & 0 \\ 0 & D \end{bmatrix} \\ &= \begin{bmatrix} AC & 0 \\ 0 & BD \end{bmatrix} \\ &= (AC) \oplus (BD). \end{aligned}$$

Further properties of the direct sum (in contrast with the properties of the Kronecker product on page 14) are

1. $(A \oplus B) + (C \oplus D) = (A + C) \oplus (B + D)$.
2. $(A \oplus B) \otimes C = (A \otimes C) \oplus (B \otimes C)$.
3. $(A \oplus B)^{-1} = A^{-1} \oplus B^{-1}$.
4. $\det(A \oplus B) = \det(A) \det(B)$.
5. $\operatorname{tr}(A \oplus B) = \operatorname{tr}(A) + \operatorname{tr}(B)$.

Remark: Property 2 follows from our definition of the Kronecker product, and $C \otimes (A \oplus B) \neq (C \otimes A) \oplus (C \otimes B)$ in general.

We will now introduce an algebraic formula of expressing a block matrix using the Kronecker product and discuss some of the basic properties.

Definition 4 *An unbalanced block matrix A is a block matrix of order $m \times n$*

$$\left[\begin{array}{c|c|c} A_{11} & \dots & A_{1q} \\ \hline \vdots & \ddots & \vdots \\ \hline A_{p1} & \dots & A_{pq} \end{array}\right]$$

where the submatrices A_{pq} need not have the same sizes and are of order $m_p \times n_q$ where $m = \sum_{p=1}^{k} m_p$ and $n = \sum_{q=1}^{l} n_q$. Using the matrices defined in (3.1) it follows that this definition is equivalent to expressing A as

$$A = \sum_{p=1}^{k} \sum_{q=1}^{l} E_{e_p,\mathbf{m}} (E_{pq} \otimes A_{pq}) E_{e_q,\mathbf{n}}^T.$$

For example, we can express the matrix

$$A = \left[\begin{array}{c|cc} 3 & 1 & 0 \\ \hline 4 & 1 & 5 \\ 0 & 2 & 6 \end{array}\right]$$

with $\mathbf{m} = \mathbf{n} = (1, 2)$ and $k = l = 2$ as

$$A = E_{e_1,\mathbf{m}} \left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 3 \end{pmatrix} \right) E_{e_1,\mathbf{n}}^T + E_{e_1,\mathbf{m}} \left(\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \end{pmatrix} \right) E_{e_2,\mathbf{n}}^T$$

$$+ E_{e_2,\mathbf{m}} \left(\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \otimes \begin{pmatrix} 4 \\ 0 \end{pmatrix} \right) E_{e_1,\mathbf{n}}^T + E_{e_2,\mathbf{m}} \left(\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 5 \\ 2 & 6 \end{pmatrix} \right) E_{e_2,\mathbf{n}}^T.$$

We say that A is an unbalanced block matrix of order 3×3. Here the E-matrices $E_{e_p,\mathbf{m}}$ and $E_{e_q,\mathbf{n}}$ reduce to one of

$$\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix},$$

respectively. Here I is an appropriate identity and the remaining entries are the appropriate zero matrices.

In Definition 4, we have introduced the matrices $E_{e_p,\mathbf{m}}$ (3.1) which, to some extent, allow us to regain the algebraic structure of balanced partitioned block matrices. This novel approach reflects the underlying algebraic structure via notation, as opposed to the informal approach presented in [5] and the indexed (modal) approach of [8].

We define a balanced block matrix from the unbalanced block matrix as follows,

$$A = \sum_{pq} (E_{pq} \otimes A_{pq}), \tag{3.2}$$

when $m_1 = m_2 = \cdots = m_k$ and $n_1 = n_2 = \cdots = n_l$. In this case, the matrices $E_{e_p,\mathbf{m}} = E_{pp} \otimes I_{m_1}$ and $E_{e_q,\mathbf{n}} = E_{qq} \otimes I_{n_1}$ so that

$$\begin{aligned} A &= \sum_{p=1}^{k} \sum_{q=1}^{l} E_{e_p,\mathbf{m}} (E_{pq} \otimes A_{pq}) E_{e_q,\mathbf{n}}^T \\ &= \sum_{p=1}^{k} \sum_{q=1}^{l} (E_{pp} \otimes I_{m_1})(E_{pq} \otimes A_{pq})(E_{qq} \otimes I_{n_1}) \\ &= \sum_{p=1}^{k} \sum_{q=1}^{l} E_{pp} E_{pq} E_{qq} \otimes I_{m_1} A_{pq} I_{n_1} \end{aligned}$$

which reduces to (3.2).

Definition 5 *A balanced partitioned matrix A is an $ms \times nt$ matrix with block enties A_{pq} each having the same order of $m \times n$ for all $p = 1, \ldots, s$ and $q = 1, \ldots, t$.*

For example, if

$$A = \left[\begin{array}{cc|cc} 1 & 2 & 3 & 6 \\ 2 & 4 & 9 & 3 \\ \hline 7 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{array}\right],$$

then A is said to be a 2×2 balanced block matrix of order 4×4 that consists of 2×2 blocks.

Matrix partitioning often arises naturally with large matrices since it allows for manipulation to be carried out on the submatrices or subblocks [4]. We can partition any matrix to make it a block matrix, for example, an $m \times n$ matrix can be partitioned into an $m \times n$ block matrix where each entry is a block of order 1×1, or as a 1×1 block matrix with an $m \times n$ block. For example, consider an $m \times n$ matrix A defined by

$$a_j = \sum_{i=1}^{m} \sum_{j=1}^{n} e_i e_i^T (A_{ij}) e_j \qquad \text{and} \qquad b_j = e_j,$$

where a_j and b_j are $m \times 1$ and $n \times 1$ column vectors, respectively. Using the fact that $e_j \in \mathbb{R}^n$, we can express A algebraically as follows

$$\begin{aligned}
A = \sum_j a_j b_j^T &= \sum_{ij} e_i \otimes e_i^T (A)_{ij} e_j e_j^T \\
&= \sum_{ij} e_i \otimes e_i^T (A)_{ij} e_j \otimes e_j^T \\
&= \sum_{ij} e_i \otimes e_j^T \otimes e_i^T (A)_{ij} e_j \\
&= \sum_{ij} E_{ij} \otimes (A)_{ij}.
\end{aligned}$$

Here, $(A)_{ij}$ is an $m \times n$ block matrix that consists of the entry a_{ij} in its i, j-position and 0 elsewhere. For example, if

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix},$$

then

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

Linear operations and matrix operations such as matrix multiplication, addition, transpose of a matrix, matrix conjugate are also applicable in both balanced and unbalanced block matrices.

3.1 Balanced partitioning

In this section we discuss the $\mathrm{vecb_c}$ operator and its properties which enable us to define the block Kronecker product $A \boxtimes B$. We consider balanced partitioned block matrices throughout this section. Properties of the vecb operator are also derived from the standard vec operator. We state and prove these properties below collected from [5]. This includes properties of the block Kronecker product which are a generalization of Theorem 3, defined for block matrices. We also consider the case where only one matrix is partitioned and describe the block Kronecker product in such a case.

Let A be a block matrix defined as in (3.2), we first show that block matrices obey the usual matrix product structure AB. Let

$$A = \sum_{pq}(E_{pq} \otimes A_{pq}) \qquad \text{and} \qquad B = \sum_{uv}(E_{uv} \otimes B_{uv})$$

be block matrices of order $ms \times nt$ and $nt \times rk$, respectively. Here A_{pq} has order $m \times n$ where $p = 1, \ldots, s, q = 1, \ldots, t$ and B_{uv} has order $n \times r$ with $u = 1, \ldots, t, v = 1, \ldots, k$. Then

$$\begin{aligned}
AB &= \left(\sum_{pq} E_{pq} \otimes A_{pq}\right)\left(\sum_{uv} E_{uv} \otimes B_{uv}\right) \\
&= \sum_{p=1}^{n}\sum_{q=1}^{t}\sum_{u=1}^{t}\sum_{v=1}^{k} E_{pq}E_{uv} \otimes A_{pq}B_{uv} \\
&= \sum_{p=1}^{n}\sum_{q=1}^{t}\sum_{v=1}^{k} E_{pv} \otimes A_{pq}B_{qv} \\
&= \sum_{p=1}^{n}\sum_{v=1}^{k} E_{pv} \otimes \left(\sum_{q=1}^{t} A_{pq}B_{qv}\right) \\
&= \left[\sum_{q=1}^{t} A_{pq}B_{qv}\right]_{pv}.
\end{aligned}$$

As mentioned in Chapter 2, the vec operator arises naturally when discussing the Kronecker product of matrices which is also the case for block structured matrices. We use (3.2) to define the concepts of the vecb operator.

Definition 6 *Let A be a balanced partitioned matrix defined as in* (3.2). *The column-blockwise vectorization* $\mathrm{vecb_c}(A)$ *is defined as follows,*

$$\mathrm{vecb_c}(A) = \sum_{pq}(\mathrm{vec}(E_{pq}) \otimes \mathrm{vec}(A_{pq})), \tag{3.3}$$

while the row-blockwise vectorization $\mathrm{vecb_r}(A)$ *is given by*

$$\mathrm{vecb_r}(A) = \sum_{pq}(\mathrm{vec}(E_{qp}) \otimes \mathrm{vec}(A_{pq})). \tag{3.4}$$

Equation (3.3) is the same as (1.1) and is connected to the block Kronecker product $A \boxtimes B$. Equation (1.2) is the same as (3.4) and is connected to the block Kronecker product $A \boxminus B$.

Similar to the vec operator, the vecb operator is also a linear operator, that is

$$\mathrm{vecb_c}(A+B) = \mathrm{vecb_c}(A) + \mathrm{vecb_c}(B),$$

$$\mathrm{vecb_c}(\alpha A) = \alpha\,\mathrm{vecb_c}(A).$$

The same results also hold for $\mathrm{vecb_r}$. Properties of the vecb operator enable us to define the above mentioned block Kronecker products. In a special case where A is not partitioned (partitioned into 1×1 submatrices), $\mathrm{vecb_c}(A) = \mathrm{vec}(A)$ and $\mathrm{vecb_r}(A) = \mathrm{vec}(A^T)$. The following lemma is given without proof in [5].

Lemma 1 *Let A be an $(ms) \times (nt)$ block matrix with $m \times n$ blocks and let K_{sn} be the commutation matrix. The permutation matrix $I_t \otimes K_{sn} \otimes I_m$ provides*

$$\mathrm{vecb_c}(A) = (I_t \otimes K_{sn} \otimes I_m)\,\mathrm{vec}(A).$$

Proof *By property 9 of Theorem 3 and Theorem 1, respectively*

$$\begin{aligned}
\mathrm{vecb_c}(A) &= \sum_{p=1}^{s}\sum_{l=1}^{n}(I_t \otimes e_p e_l^T \otimes e_l e_p^T \otimes I_m)\sum_{q=1}^{t}\sum_{k=1}^{n}(e_q \otimes e_k \otimes A(e_q \otimes e_k)) \\
&= \sum_{p=1}^{s}\sum_{l=1}^{n}\sum_{k=1}^{n}\sum_{q=1}^{t} e_q \otimes e_p e_l^T e_k \otimes (e_l e_p^T \otimes I_m)(A(e_q \otimes e_k)) \\
&= \sum_{p=1}^{s}\sum_{l=1}^{n}\sum_{q=1}^{t} e_q \otimes e_p \otimes (e_l e_p^T \otimes I_m)\left(\sum_{i=1}^{s}\sum_{j=1}^{t}(E_{ij} \otimes A_{ij})\right)(e_q \otimes e_l)) \\
&= \sum_{p=1}^{s}\sum_{l=1}^{n}\sum_{i=1}^{s}\sum_{j=1}^{t}\sum_{q=1}^{t} e_q \otimes e_p \otimes (e_l e_p^T e_i e_j^T e_q) \otimes (A_{ij} e_l) \\
&= \sum_{p=1}^{s}\sum_{l=1}^{n}\sum_{q=1}^{t} e_q \otimes e_p \otimes e_l \otimes A_{pq} e_l \\
&= \sum_{p=1}^{s}\sum_{q=1}^{t}(\mathrm{vec}(E_{pq}) \otimes \mathrm{vec}(A_{pq})).
\end{aligned}$$

We will use Lemma 1 to define the block Kronecker product by generalizing Theorem 2.

It is mentioned in Chapter 1 that we can use the commutation matrix K_{mn} to write $\mathrm{vec}(A)$ in terms of $\mathrm{vec}(A^T)$. Since we have two different block vectorizations, we introduce a way in which we can express $\mathrm{vecb_c}(A)$ in terms of

$\operatorname{vecb_r}(A)$ in the next result.

Let A be a general $ms \times nt$ block matrix,

$$A = \begin{bmatrix} A_{11} & \dots & A_{1q} \\ \vdots & \ddots & \vdots \\ A_{p1} & \dots & A_{pq} \end{bmatrix}$$

with each A_{pq} of order $m \times n$. Then

$$\operatorname{vecb_c}(A) = (\operatorname{vec}(E_{11}) \otimes \operatorname{vec}(A_{11})) + (\operatorname{vec}(E_{21}) \otimes \operatorname{vec}(A_{21})) + \cdots + (\operatorname{vec}(E_{pq}) \otimes \operatorname{vec}(A_{pq}))$$

and vectorizing row blockwise yields (3.4),

$$\begin{aligned}\operatorname{vecb_r}(A) &= (\operatorname{vec}(E_{11}) \otimes \operatorname{vec}(A_{11})) + (\operatorname{vec}(E_{21}) \otimes \operatorname{vec}(A_{12})) + \cdots + (\operatorname{vec}(E_{pq}) \otimes \operatorname{vec}(A_{pq})) \\ &= (\operatorname{vec}(E_{11}^T) \otimes \operatorname{vec}(A_{11})) + (\operatorname{vec}(E_{12}^T) \otimes \operatorname{vec}(A_{12})) + \cdots + (\operatorname{vec}(E_{pq}^T) \otimes \operatorname{vec}(A_{pq})).\end{aligned}$$

Using $\operatorname{vec}(A^T) = K_{st} \operatorname{vec}(A)$ where A is an $s \times t$ matrix gives us

$$\operatorname{vecb_r}(A) = (K_{st} \operatorname{vec}(E_{11}) \otimes \operatorname{vec}(A_{11})) + (K_{st} \operatorname{vec}(E_{12}) \otimes \operatorname{vec}(A_{12})) + \cdots .$$

We know that K_{st} is the $st \times st$ commutation matrix, $\operatorname{vec}(E_{pq})$ is an $st \times 1$ column vector and $\operatorname{vec}(A_{pq})$ is a $mn \times 1$ column vector. Therefore, by the mixed product property, we have

$$\begin{aligned}(K_{st} \operatorname{vec}(E_{pq})) \otimes \operatorname{vec}(A_{pq}) &= (K_{st} \operatorname{vec}(E_{pq})) \otimes (I_{mn} \operatorname{vec}(A_{pq})) \\ &= (K_{st} \otimes I_{mn})(\operatorname{vec}(E_{pq}) \otimes \operatorname{vec}(A_{pq})).\end{aligned}$$

Hence

$$\begin{aligned}\operatorname{vecb_r}(A) &= (K_{st} \otimes I_{mn})[(\operatorname{vec}(E_{11}) \otimes \operatorname{vec}(A_{11})) + (\operatorname{vec}(E_{12}) \otimes \operatorname{vec}(A_{12})) + \cdots] \\ &= (K_{st} \otimes I_{mn})[(\operatorname{vec}(E_{11}) \otimes \operatorname{vec}(A_{11})) + (\operatorname{vec}(E_{21}) \otimes \operatorname{vec}(A_{21})) + \cdots] \\ &= (K_{st} \otimes I_{mn}) \operatorname{vecb_c}(A).\end{aligned}$$

We have proved the following proposition.

Proposition 1 *Let A be an $ms \times nt$ block matrix with all the sub-matrices being of order $m \times n$ where $\operatorname{vecb_c}(A)$ is defined as in Definition 6, then*

$$\operatorname{vecb_r}(A) = (K_{st} \otimes I_{mn}) \operatorname{vecb_c}(A).$$

We now generalize Theorem 2 which will induce the definition of the block Kronecker product $A \boxtimes B$ using Lemma 1.

Theorem 6 *Let A be defined as in* (3.2)*. We define B and C in a similar way where both B and C are block matrices of order $(hj) \times (rk)$ and $(rk) \times (nt)$, respectively. Then there exists a unique bilinear matrix operation $A \boxtimes B$ such that*

$$\operatorname{vecb_c}(BCA^T) = (A \boxtimes B) \operatorname{vecb_c}(C),$$

for all C.

Proof *Using Lemma 1 and Theorem 2 respectively,*

$$\begin{aligned}
\operatorname{vecb_c}(BCA^T) &= (I_s \otimes K_{jm} \otimes I_h) \operatorname{vec}(BCA^T) \\
&= (I_s \otimes K_{jm} \otimes I_h)(A \otimes B) \operatorname{vec}(C) \\
&= (I_s \otimes K_{jm} \otimes I_h)(A \otimes B)(I_t \otimes K_{nk} \otimes I_r) \operatorname{vecb_c}(C) \\
&= (A \boxtimes B) \operatorname{vecb_c}(C)
\end{aligned}$$

where $K_{nx} = K_{xn}^{-1}$ and

$$(A \boxtimes B) = (I_s \otimes K_{jm} \otimes I_h)(A \otimes B)(I_t \otimes K_{nk} \otimes I_r). \tag{3.5}$$

Equation (3.5) provides a direct form for $A \boxtimes B$, which we give in the next corollary.

Corollary 2 *Consider the block Kronecker product $A \boxtimes B$ where A and B are block matrices of order $ms \times nt$ and $hj \times rk$, respectively. Then*

$$A \boxtimes B = \sum_{pquv} (E_{pq} \otimes E_{uv}) \otimes (A_{pq} \otimes B_{uv}).$$

We note that this is the Tracey-Singh product since it is given in the following way

$$A \boxtimes B = \left[\begin{array}{ccc|c|ccc} A_{11} \otimes B_{11} & \cdots & A_{11} \otimes B_{1v} & \cdots & A_{1q} \otimes B_{11} & \cdots & A_{1q} \otimes B_{1v} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ A_{11} \otimes B_{u1} & \cdots & A_{11} \otimes B_{uv} & \cdots & A_{1q} \otimes B_{u1} & \cdots & A_{1q} \otimes B_{uv} \\ \hline \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \hline A_{p1} \otimes B_{11} & \cdots & A_{p1} \otimes B_{1v} & \cdots & A_{pq} \otimes B_{11} & \cdots & A_{pq} \otimes B_{1v} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ A_{p1} \otimes B_{u1} & \cdots & A_{p1} \otimes B_{uv} & \cdots & A_{pq} \otimes B_{u1} & \cdots & A_{pq} \otimes B_{uv} \end{array}\right].$$

As seen in Chapter 2, the block Kronecker product also induces the block Kronecker sum, that is, the block Kronecker sum $\boxplus$ of block matrices A and B (denoted by $A \boxplus B$) is

$$A \boxplus B = I_n \boxtimes A + B \boxtimes I_m.$$

The following properties generalize the basic properties of the vec operator discussed in Chapter 1.

Properties of the $\mathrm{vecb_c}$ operator

1. Let $a = (a_p)$ and $b = (b_q)$ be block column vectors of order $m \times 1$ and $n \times 1$, respectively where $p = 1, \ldots, s$ and $q = 1, \ldots, t$. Then

$$\mathrm{vecb_c}(ab^T) = b \boxtimes a.$$

2. If A, B, D are block matrices of order $(ms)\times(ms), (nt)\times(nt), (ms)\times(nt)$, respectively and let $A \boxplus B$ be the block Kronecker sum of A and B, then

$$\mathrm{vecb_c}(AD + DB) = (A \boxplus B^T)\,\mathrm{vecb_c}(D).$$

3. Let A, B, C, D be partitioned block matrices of order $(ms) \times (nt), (nt) \times (rk), (rk) \times (wz), (ms) \times (wz)$, respectively. Then

$$\mathrm{tr}(ABCD^T) = (\mathrm{vecb_c}(A^T))^T (D \boxtimes B)\,\mathrm{vecb_c}(C).$$

4. Let $A \boxtimes B$ be defined as in Theorem 6. Then

$$\mathrm{vecb_c}(A \boxtimes B) = (I_t \otimes K_{ks} \otimes I_{jn} \otimes K_{rm} \otimes I_h)(\mathrm{vecb_c}(A) \boxtimes \mathrm{vecb_c}(B)).$$

5. Let A be an $m \times n$ matrix and B be a general partitioned matrix. Then considering $A \otimes B$ as a $m \times n$ balanced partitioned matrix,

$$\mathrm{vecb_c}(A \otimes B) = \mathrm{vec}(A) \otimes \mathrm{vec}(B).$$

Proof Some of the proofs are tedious (and proved by entrywise direct computation) and we have omitted them here.

Proof of 1.
Using Lemma 1 and Theorem 6,

$$\begin{aligned}\mathrm{vecb_c}(ab^T) &= (I_s \otimes K_{tm} \otimes I_n)\,\mathrm{vec}(ab^T)\\ &= (I_p \otimes K_{qm} \otimes I_n)(b \otimes a)\\ &= b \boxtimes a.\end{aligned}$$

Proof of 2.

$$\begin{aligned}\mathrm{vecb_c}(AD + DB) &= \mathrm{vecb_c}(AD) + \mathrm{vecb_c}(DB)\\ &= (I_t \otimes K_{sn} \otimes I_m)\,\mathrm{vec}(AD) + (I_t \otimes K_{sn} \otimes I_m)\,\mathrm{vec}(DB)\\ &= (I_t \otimes K_{sn} \otimes I_m)[\mathrm{vec}(AD) + \mathrm{vec}(DB)]\\ &= (I_t \otimes K_{sn} \otimes I_m)[(I \otimes A)\,\mathrm{vec}(D) + (B^T \otimes I)\,\mathrm{vec}(D)]\\ &= (I_t \otimes K_{sn} \otimes I_m)[(I \otimes A) + (B^T \otimes I)]\,\mathrm{vec}(D)\\ &= (I_t \otimes K_{sn} \otimes I_m)[(I \otimes A) + (B^T \otimes I)](I_t \otimes K_{sn} \otimes I_m)^T\,\mathrm{vecb_c}(D)\\ &= [I \boxtimes A + B^T \boxtimes I]\,\mathrm{vecb_c}(D)\\ &= (A \boxplus B^T)\,\mathrm{vecb_c}(D).\end{aligned}$$

Proof of 3. By property 4 of properties of the vec operator,

$$\begin{aligned}
\operatorname{tr}(ABCD^T) &= ((\operatorname{vec} A^T)^T)\operatorname{vec}(BCD^T) \\
&= ((I_s \otimes K_{tm}^T \otimes I_n)\operatorname{vecb_c}(A^T))^T \operatorname{vec}(BCD^T) \\
&= (\operatorname{vecb_c}(A^T))^T (I_s \otimes K_{tm} \otimes I_n)(D \otimes B)\operatorname{vec}(C) \\
&= (\operatorname{vecb_c}(A^T))^T (I_s \otimes K_{tm} \otimes I_n)(D \otimes B)(I_z \otimes K_{wk} \otimes I_r)\operatorname{vecb_c}(C) \\
&= ((\operatorname{vecb_c}(A))^T)^T (D \boxtimes B)\operatorname{vecb_c}(C).
\end{aligned}$$

The next theorem provides basic properties of the block Kronecker product collected from [4], [5] and [10]. These properties generalize Theorems 1 and 3 from Chapter 2 and we will use the results of Theorem 6 and Corollary 2 to prove them.

Theorem 7 (Basic properties of the block Kronecker product $A \boxtimes B$) *Let A, B, C, D be blockwise partitioned matrices of appropriate order. Then*

1. $(cA) \boxtimes B = A \boxtimes (cB) = c(A \boxtimes B)$
2. $(A \boxtimes B)^T = A^T \boxtimes B^T$
3. $(A \boxtimes B)(C \boxtimes D) = (AC \boxtimes BD)$
4. $A \boxtimes (B + C) = (A \boxtimes B) + (A \boxtimes C)$
5. $(A + B) \boxtimes C = (A \boxtimes C) + (B \boxtimes C)$
6. $\operatorname{tr}(A \boxtimes B) = \operatorname{tr}(A)\operatorname{tr}(B)$
7. $B \boxtimes A = K_1^T (A \boxtimes B) K_2$ *for some commutation K_1 and K_2 .*

Proof of 1. Let A and B be matrices of order $ms \times nt$ and $hj \times rk$, respectively. Using (3.5),

$$(cA) \boxtimes B = (I_s \otimes K_{jm} \otimes I_h)((cA) \otimes B)(I_t \otimes K_{nk} \otimes I_r),$$

then by property 2 of Theorem 3, we get

$$\begin{aligned}
&= (I_s \otimes K_{jm} \otimes I_h)(A \otimes (cB))(I_t \otimes K_{nk} \otimes I_r) \\
&= ((I_s \otimes K_{jm} \otimes I_h)(c(A \otimes B))(I_t \otimes K_{nk} \otimes I_r) \\
&= c((I_s \otimes K_{jm} \otimes I_h)(A \otimes B)(I_t \otimes K_{nk} \otimes I_r).
\end{aligned}$$

Proof of 2. Let A and B be matrices of order $ms \times nt$ and $hj \times rk$, respectively.

By (3.5) *and property 3 of Theorem 3,*

$$\begin{aligned}(A \boxtimes B)^T &= ((I_s \otimes K_{jm} \otimes I_h)(A \otimes B)(I_t \otimes K_{nk} \otimes I_r))^T \\ &= (I_t \otimes K_{nk} \otimes I_r)(A^T \otimes B^T)(I_s \otimes K_{mj} \otimes I_h) \\ &= A^T \boxtimes B^T.\end{aligned}$$

Proof of 3. Consider the balanced partitioned block matrices A, B, C, D of order $(ms) \times (nt), (hj) \times (rk), (nt) \times (wl), (rk) \times (yz)$, respectively. Let $A \boxtimes B$ be defined as in (3.5)*, and similarly $C \boxtimes D = (I_t \otimes K_{ln} \otimes I_r)(C \otimes D)(I_l \otimes K_{wz} \otimes I_y)$. Let $K_1 = (I_s \otimes K_{jm} \otimes I_h), K_2 = (I_t \otimes K_{nk} \otimes I_r)$ and $K_3 = (I_l \otimes K_{wz} \otimes I_y)$. By Theorem 1 and* (3.5)*, respectively.*

$$\begin{aligned}(A \boxtimes B)(C \boxtimes D) &= (K_1(A \otimes B)K_2)(K_2^T(C \otimes D)K_3) \\ &= K_1(A \otimes B)K_2K_2^T(C \otimes D)K_3 \\ &= K_1(A \otimes B)I(C \otimes D)K_3 \\ &= K_1(A \otimes B)(C \otimes D)K_3 \\ &= K_1(AC \otimes BD)K_3 \\ &= AC \boxtimes BD.\end{aligned}$$

Proof of 4. Let A be an $ms \times nt$ block matrix and B and C be matrices of order $hj \times rk$. Using Corollary 2 and property 6 of Theorem 3,

$$\begin{aligned}A \boxtimes (B + C) &= \sum_{pquv}(E_{pq} \otimes E_{uv}) \otimes (A_{pq} \otimes (B + C)_{uv}) \\ &= \sum_{pquv}(E_{pq} \otimes E_{uv}) \otimes (A_{pq} \otimes (B_{uv} + C_{uv})) \\ &= \sum_{pquv}(E_{pq} \otimes E_{uv}) \otimes ((A_{pq} \otimes B_{uv}) + (A_{pq} \otimes C_{uv})) \\ &= (A \boxtimes B) + (A \boxtimes C)\end{aligned}$$

by left distributivity of the Kronecker product.

Proof of 5. The proof of 5 is similar to the proof of 4, by right distributivity.

Proof of 6. Let A and B be balanced partitioned matrices of order $mp \times mp$ and $nq \times nq$, respectively. We define $A \boxtimes B$ using (3.5) *as follows*

$$(A \boxtimes B) = (I_p \otimes K_{qm} \otimes I_n)(A \otimes B)(I_p \otimes K_{mq} \otimes I_n),$$

then

$$\begin{aligned}
\operatorname{tr}(A \boxtimes B) &= \operatorname{tr}((I_p \otimes K_{qm} \otimes I_n)(A \otimes B)(I_p \otimes K_{mq} \otimes I_n)) \\
&= \operatorname{tr}((A \otimes B)(I_p \otimes K_{mq} \otimes I_n)(I_p \otimes K_{qm} \otimes I_n)) \\
&= \operatorname{tr}((A \otimes B)(I_p \otimes K_{mq} \otimes I_n)(I_p \otimes K_{mq}^T \otimes I_n)) \\
&= \operatorname{tr}(A \otimes B) \\
&= \operatorname{tr}(A)\operatorname{tr}(B)
\end{aligned}$$

using property 8 of Theorem 3.

Proof of 7. Let A and B be matrices of order $ms \times nt$ and $hj \times rk$, respectively. Let k_1 and k_2 be some appropriate commutation matrices satisfying

$$k_1^T(E_{pq} \otimes E_{uv})k_2 = E_{uv} \otimes E_{pq}.$$

By Corollary 2 and Theorem 1,

$$\begin{aligned}
B \boxtimes A &= \sum_{pquv}(E_{uv} \otimes E_{pq}) \otimes (B_{uv} \otimes A_{pq}) \\
&= \sum_{pquv} k_1^T(E_{pq} \otimes E_{uv})k_2 \otimes k_1^T(A_{pq} \otimes B_{uv})k_2 \\
&= \sum_{pquv}(k_1^T \otimes k_1^T)((E_{pq} \otimes E_{uv}) \otimes (A_{pq} \otimes B_{uv}))(k_2 \otimes k_2) \\
&= K_1^T(A \boxtimes B)K_2
\end{aligned}$$

for permutation matrices $K_1 = k_1 \otimes k_1$ and $K_2 = k_2 \otimes k_2$.

In the special case where only one matrix is partitioned, say B, a $hj \times rk$ matrix with the entry B_{uv} order $h \times r$ blocks and A is a general non-partitioned $m \times n$ matrix, we get

$$A \boxminus B = \sum_{uv} E_{uv} \otimes A \otimes B_{uv}. \tag{3.6}$$

We also write $A \boxminus B$ in matrix form as follows,

$$A \boxminus B = \begin{bmatrix} A \otimes B_{11} & \dots & A \otimes B_{1v} \\ \vdots & \ddots & \vdots \\ A \otimes B_{u1} & \dots & A \otimes B_{uv} \end{bmatrix}. \tag{3.7}$$

The same properties of the block Kronecker product stated above also hold for this product which arises naturally when we consider the operator $\mathrm{vecb_r}$ and is introduced in the next section.

3.2 Unbalanced partitioning

In this section we introduce the novel approach of the block Kronecker product $A \boxminus B$ which is compatible with $\mathrm{vecb_r}$. We have already defined an unbalanced block matrix which will be the centre of this subsection. Let A be defined as in Definition 4, and let B be the (unbalanced partitioned) matrix

$$B = \sum_{u=1}^{l} \sum_{v=1}^{r} E_{e_u,\mathbf{n}}(E_{uv} \otimes B_{uv}) E_{e_v,\mathbf{s}}^T.$$

The partitioned matrix product

$$AB = \left[\sum_{q=1}^{l} A_{pq} B_{qv} \right]_{pv}$$

for unbalanced partitioned block matrices obeys the usual matrix product formulae. This follows from the special case of the product

$$E_{e_p,\mathbf{m}}^T E_{e_q,\mathbf{m}} = \begin{cases} E_{pp} \otimes I_{m_p} & \text{if } p = q, \\ 0 & \text{if } p \neq q. \end{cases}$$

In particular,

$$\begin{aligned} AB &= \sum_{p=1}^{k} \sum_{q=1}^{l} \sum_{u=1}^{l} \sum_{v=1}^{r} E_{e_p,\mathbf{m}}(E_{pq} \otimes A_{pq}) E_{e_q,\mathbf{n}}^T E_{e_u,\mathbf{n}}(E_{uv} \otimes B_{uv}) E_{e_v,\mathbf{s}}^T \\ &= \sum_{p=1}^{k} \sum_{q=1}^{l} \sum_{v=1}^{r} E_{e_p,\mathbf{m}}(E_{pq} \otimes A_{pq})(E_{qq} \otimes I_{n_q})(E_{qv} \otimes B_{qv}) E_{e_v,\mathbf{s}}^T \\ &= \sum_{p=1}^{k} \sum_{q=1}^{l} \sum_{v=1}^{r} E_{e_p,\mathbf{m}}(E_{pq} E_{qq} E_{qv} \otimes A_{pq} I_{n_q} B_{qv}) E_{e_v,\mathbf{s}}^T \\ &= \sum_{p=1}^{k} \sum_{q=1}^{l} \sum_{v=1}^{r} E_{e_p,\mathbf{m}}(E_{pv} \otimes A_{pq} B_{qv}) E_{e_v,\mathbf{s}}^T \\ &= \sum_{p=1}^{k} \sum_{v=1}^{r} E_{e_p,\mathbf{m}} \left[\sum_{q=1}^{l} A_{pq} B_{qv} \right]_{pv} E_{e_v,\mathbf{s}}^T \end{aligned}$$

which completes the result.

Definition 7 *Let A be an $m \times n$ unbalanced block matrix where each A_{ij} block has order $m_i \times n_j$. We define the* vecb *operator as follows,*

$$\mathrm{vecb_c}(A) = \sum_{ij} E_{e_j \otimes e_i, \mathbf{n} \otimes \mathbf{m}} (\mathrm{vec}(E_{ij}) \otimes \mathrm{vec}(A_{ij})) \tag{3.8}$$

for column-blockwise vectorization and

$$\mathrm{vecb_r}(A) = \sum_{ij} E_{e_i \otimes e_j, \mathbf{m} \otimes \mathbf{n}} (\mathrm{vec}(E_{ji}) \otimes \mathrm{vec}(A_{ij})) \tag{3.9}$$

for row-blockwise vectorization.

Note that the equations (3.8) and (3.9) still yield the block vectorization shown in (1.1) and (1.2)

$$\mathrm{vecb_c}(A) = \begin{bmatrix} \mathrm{vec}(A_{11}) \\ \mathrm{vec}(A_{21}) \\ \vdots \\ \mathrm{vec}(A_{mn}) \end{bmatrix} \qquad \text{and} \qquad \mathrm{vecb_r}(A) = \begin{bmatrix} \mathrm{vec}(A_{11}) \\ \mathrm{vec}(A_{12}) \\ \vdots \\ \mathrm{vec}(A_{mn}) \end{bmatrix}$$

where the entries generally have different orders.

The above definition allows us to derive the results similar to Proposition 1 in the unbalanced case. We already know by definition, the matrix $E_{e_i,\mathbf{m}}$, where I_{m_i} is an unbalanced identity matrix. We define

$$K \otimes I_{m_i} = K$$

as a commutation matrix that contains the identity matrix I_{m_p} and satisfies the condition $K\,\mathrm{vec}(E_{ij}) = \mathrm{vec}(E_{ij}^T)$. By (3.8), we can vectorize an unbalanced block matrix as follows,

$$\begin{aligned} \mathrm{vecb_c}(A) &= \sum_{ij} E_{e_j \otimes e_i, \mathbf{n} \otimes \mathbf{m}}\,\mathrm{vec}(E_{ij}) \otimes \mathrm{vec}(A_{ij}) \\ &= E_{e_1 \otimes e_1, \mathbf{n} \otimes \mathbf{m}}\,\mathrm{vec}(E_{11}) \otimes \mathrm{vec}(A_{11}) + E_{e_1 \otimes e_2, \mathbf{n} \otimes \mathbf{m}}\,\mathrm{vec}(E_{21}) \otimes \mathrm{vec}(A_{21}) \\ &+ E_{e_1 \otimes e_3, \mathbf{n} \otimes \mathbf{m}}\,\mathrm{vec}(E_{31}) \otimes \mathrm{vec}(A_{31}) + \cdots \end{aligned}$$

By (3.9),

$$\begin{aligned}
\mathrm{vecb_r}(A) &= \sum_{ij} E_{e_i\otimes e_j,\mathbf{m}\otimes\mathbf{n}} \operatorname{vec}(E_{ji}) \otimes \operatorname{vec}(A_{ij}) \\
&= E_{e_1\otimes e_1,\mathbf{m}\otimes\mathbf{n}} \operatorname{vec}(E_{11}) \otimes \operatorname{vec}(A_{11}) + E_{e_1\otimes e_2,\mathbf{m}\otimes\mathbf{n}} \operatorname{vec}(E_{21}) \otimes \operatorname{vec}(A_{12}) \\
&+ E_{e_1\otimes e_3,\mathbf{m}\otimes\mathbf{n}} \operatorname{vec}(E_{31}) \otimes \operatorname{vec}(A_{13}) + \cdots \\
&= E_{e_1\otimes e_1,\mathbf{m}\otimes\mathbf{n}} \operatorname{vec}(E_{11}^T) \otimes \operatorname{vec}(A_{11}) + E_{e_1\otimes e_2,\mathbf{m}\otimes\mathbf{n}} \operatorname{vec}(E_{12}^T) \otimes \operatorname{vec}(A_{12}) \\
&+ E_{e_1\otimes e_3,\mathbf{m}\otimes\mathbf{n}} \operatorname{vec}(E_{13}^T) \otimes \operatorname{vec}(A_{13}) + \cdots
\end{aligned}$$

Then by definition of $\mathrm{vecb_r}$ and Proposition 1, we have

$$\begin{aligned}
\mathrm{vecb_r}(A) &= \sum_{ij} E_{e_i\otimes e_j,\mathbf{m}\otimes\mathbf{n}} \operatorname{vec}(E_{ij}^T) \otimes \operatorname{vec}(A_{ij}) \\
&= \sum_{ij} E_{e_i\otimes e_j,\mathbf{m}\otimes\mathbf{n}} K \operatorname{vec}(E_{ij}) \otimes \operatorname{vec}(A_{ij}) \\
&= \sum_{ij} E_{e_i\otimes e_j,\mathbf{m}\otimes\mathbf{n}} (K \otimes I_{m_i n_j}) \operatorname{vec}(E_{ij}) \otimes \operatorname{vec}(A_{ij}).
\end{aligned}$$

This result is reminiscent of the expression for $\mathrm{vecb_c}(A)$ in equation (3.8). The permutation matrix K factors out in the following way, and provides the unbalanced partitioning analogue of Proposition 1.

Proposition 2 *Let A be an $m \times n$ unbalanced block matrix of A_{ij} blocks.*

$$\mathrm{vecb_r}(A) = \tilde{K} \,\mathrm{vecb_c}(A)$$

where

$$\tilde{K} = \sum_{ij} E_{e_i\otimes e_j,\mathbf{m}\otimes\mathbf{n}} (K \otimes I_{m_i n_j}) E^T_{e_j\otimes e_i,\mathbf{n}\otimes\mathbf{m}}$$

is a permutation matrix.

Conversely,

$$\mathrm{vecb_c}(A) = \tilde{K}^T \,\mathrm{vecb_r}(A).$$

The proof is straightforward and follows from the fact that the commutation matrix K converts $\operatorname{vec}(E_{ij})$ into $\operatorname{vec}(E_{ji})$ which reflects the structure of vectorizing row-blockwise, thus we only show that $\tilde{K}\,\mathrm{vecb_c}(A) = \mathrm{vecb_r}(A)$. The converse follows since $\tilde{K}^{-1} = \tilde{K}^T$

Proof *If $E_{e_j,\boldsymbol{n}}$ and $E_{e_i,\boldsymbol{m}}$ are $n \times kn_j$ and $m \times km_i$ E matrices, respectively. Then the product $E^T_{e_j\otimes e_i,\mathbf{n}\otimes\mathbf{m}} E_{e_j\otimes e_i,\mathbf{n}\otimes\mathbf{m}}$ reduces to the Kronecker product*

$E_{(j-1)k+i,(j-1)k+i} \otimes I_{n_j m_i}$ *where*

$$\begin{aligned} E_{(j-1)k+i,(j-1)k+i} &= E_{jj} \otimes E_{ii} \\ &= e_j e_j^T \otimes e_i e_i^T \\ &= (e_j \otimes e_i)(e_j \otimes e_i)^T \\ &= \operatorname{vec}(E_{ij})(\operatorname{vec}(E_{ij}))^T. \end{aligned}$$

Therefore,

$$\begin{aligned} &(E_{(j-1)k+i,(j-1)k+i} \otimes I_{n_j m_i})(\operatorname{vec}(E_{ij}) \otimes \operatorname{vec}(A_{ij})) \\ &\quad = E_{(j-1)k+i,(j-1)k+i} \operatorname{vec}(E_{ij}) \otimes I_{n_j m_i} \operatorname{vec}(A_{ij}) \\ &\quad = \operatorname{vec}(E_{ij})(\operatorname{vec}(E_{ij}))^T \operatorname{vec}(E_{ij}) \otimes \operatorname{vec}(A_{ij}) \\ &\quad = \operatorname{vec}(E_{ij}) \otimes \operatorname{vec}(A_{ij}). \end{aligned}$$

As a result,

$$\begin{aligned} \tilde{K} \operatorname{vecb_c}(A) &= \sum_{ij} E_{e_i \otimes e_j, \mathbf{m} \otimes \mathbf{n}} K E^T_{e_j \otimes e_i, \mathbf{n} \otimes \mathbf{m}} E_{e_j \otimes e_i, \mathbf{n} \otimes \mathbf{m}} \operatorname{vec}(E_{ij}) \otimes \operatorname{vec}(A_{ij}) \\ &= \sum_{ij} E_{e_i \otimes e_j, \mathbf{m} \otimes \mathbf{n}} K \operatorname{vec}(E_{ij}) \otimes \operatorname{vec}(A_{ij}) \\ &= \sum_{ij} E_{e_i \otimes e_j, \mathbf{m} \otimes \mathbf{n}} \operatorname{vec}(E_{ji}) \otimes \operatorname{vec}(A_{ij}) \\ &= \operatorname{vecb_r}(A). \end{aligned}$$

We have introduced the block Kronecker product defined for balanced partitioned matrices A and B in Section 3.1 and we will now introduce one for unbalanced partitioned block matrices. Similar to Section 3.1, two block Kronecker products will be defined here, one where both matrices are partitioned and one where we partition only one block matrix.

Theorem 8 *Let A, B, C be unbalanced block matrices of order $m \times n, u \times v$ and $v \times n$, respectively where the submatrices A_{ij}, B_{kl} and C_{lj} are of order $m_i \times n_j, u_k \times v_l$ and $v_l \times n_j$, respectively. Then there exists a unique bilinear matrix operation $A \boxtimes B$ such that*

$$\operatorname{vecb_c}(BCA^T) = (A \boxtimes B) \operatorname{vecb_c}(C).$$

Proof *The matrix product BCA^T reduces to a block matrix of order $u \times m$*

where the submatrices are of order $u_k \times m_i$ as follows

$$\begin{aligned}BCA^T &= \sum_{ijkl} E_{e_k,\boldsymbol{u}}(E_{kl} \otimes B_{kl})E^T_{e_l,\boldsymbol{v}}E_{e_l,\boldsymbol{v}}(E_{lj} \otimes C_{lj})E^T_{e_j,\boldsymbol{n}}E_{e_j,\boldsymbol{n}}(E_{ji} \otimes A^T_{ij})E^T_{e_i,\boldsymbol{m}} \\ &= \sum_{ijkl} E_{e_k,\boldsymbol{u}}(E_{kl} \otimes B_{kl})(E_{ll} \otimes I_{v_l})(E_{lj} \otimes C_{lj})(E_{jj} \otimes I_{n_j})(E_{ji} \otimes A^T_{ij})E^T_{e_i,\boldsymbol{m}} \\ &= \sum_{ijkl} E_{e_k,\boldsymbol{u}}(E_{kl}E_{lj}E_{ji} \otimes B_{kl}C_{lj}A^T_{ij})E^T_{e_i,\boldsymbol{m}}.\end{aligned}$$

By (3.8),

$$\begin{aligned}&\operatorname{vecb_c}(BCA^T) \\ &\quad= \sum_{ik} E_{e_i\otimes e_k,\mathbf{m}\otimes\mathbf{u}} \operatorname{vec}(E_{kl}E_{lj}E_{ji}) \otimes \operatorname{vec}(B_{kl}C_{lj}A^T_{ij}) \\ &\quad= \sum_{ijkl} E_{e_i\otimes e_k,\mathbf{m}\otimes\mathbf{u}}(E_{ij} \otimes E_{kl}) \operatorname{vec}(E_{lj}) \otimes (A_{ij} \otimes B_{kl}) \operatorname{vec}(C_{lj}) \\ &\quad= \sum_{ijkl} E_{e_i\otimes e_k,\mathbf{m}\otimes\mathbf{u}}(E_{ij} \otimes E_{kl}) \otimes (A_{ij} \otimes B_{kl}) \operatorname{vec}(E_{lj}) \otimes \operatorname{vec}(C_{lj})\end{aligned}$$

and since $E^T_{e_j\otimes e_l,\mathbf{n}\otimes\mathbf{v}}E_{e_j\otimes e_l\mathbf{n}\otimes\mathbf{v}}$ is a projection onto the subspace spanned by $\{\operatorname{vec}(E_{lj}) \otimes x : x \in \mathbb{F}^{v_l n_j}\}$. *We have*

$$\begin{aligned}&\operatorname{vecb_c}(BCA^T) \\ &= \sum_{ijkl} E_{e_i\otimes e_k,\mathbf{m}\otimes\mathbf{u}}(E_{ij} \otimes E_{kl}) \otimes (A_{ij} \otimes B_{kl})E^T_{e_j\otimes e_l,\mathbf{n}\otimes\mathbf{v}}E_{e_j\otimes e_l,\mathbf{n}\otimes\mathbf{v}} \operatorname{vec}(E_{lj}) \otimes \operatorname{vec}(C_{lj}) \\ &= (A \boxtimes B) \operatorname{vecb_c}(C)\end{aligned}$$

where

$$A \boxtimes B := \sum_{ijkl} E_{e_i\otimes e_k,\mathbf{m}\otimes\mathbf{u}}(E_{ij} \otimes E_{kl}) \otimes (A_{ij} \otimes B_{kl})E^T_{e_j\otimes e_l,\mathbf{n}\otimes\mathbf{v}}.$$

Note that in the case of balanced partitioning we have $\mathbf{m}\otimes\mathbf{u} = (m_1u_1, \ldots, m_1u_1)$ and hence $E_{e_i\otimes e_k,\mathbf{m}\otimes\mathbf{u}} = E_{ii} \otimes E_{kk} \otimes I_{m_1u_1}$. By the same reasoning we obtain $E_{e_j\otimes e_l,\mathbf{n}\otimes\mathbf{v}} = E_{jj} \otimes E_{ll} \otimes I_{n_1v_1}$ and similar to the calculation on page 35 we obtain the product in Corollary 2 (in the balanced case). We extend the results of Theorem 8, using Proposition 2 to induce a different approach of vectorizing row-blockwise.

Corollary 3 *If the conditions of Theorem 8 hold, then there exists a unique* *product $A \boxminus B$ such that*

$$\mathrm{vecb_r}(BCA^T) = (A \boxminus B)\,\mathrm{vecb_r}(C).$$

Proof *By Proposition 2 and Theorem 8, respectively.*

$$\begin{aligned}\mathrm{vecb_r}(BCA^T) &= \tilde{K}\,\mathrm{vecb_c}(BCA^T)\\ &= \tilde{K}(A \boxtimes B)\,\mathrm{vecb_c}(C)\\ &= \tilde{K}(A \boxtimes B)\tilde{K}^T\,\mathrm{vecb_r}(C)\\ &= (A \boxminus B)\,\mathrm{vecb_r}(C)\end{aligned}$$

where

$$(A \boxminus B) := \tilde{K}(A \boxtimes B)\tilde{K}^T.$$

Remark: Since we know that the product $A \boxminus B$ is induced by $\mathrm{vecb_r}$ operator, Corollary 3 allows us to define the product in terms of $A \boxtimes B$ provided that both A and B are partitioned. We will elaborate on this in Definition 8.

To introduce the concept of the block Kronecker product $A \boxminus B$ in the case where we partition B only and A is a general non-partitioned matrix let $\mathbf{m}_B = (m_{B,1}, \ldots, m_{B,u})$ and $\mathbf{n}_B = (n_{B,1}, \ldots, n_{B,v})$. The matrix defined above in (3.7) also holds for the unbalanced case with the submatrices B_{uv} having different orders. We show that the equivalence in Definition 8 holds and the results are derived in the Appendix.

As a consequence of Corollary 3, we have the following definition.

Definition 8 *Let B be defined the same as in Definition 5, where $\mathbf{m}_B = (m_{B,1}, \ldots, m_{B,u})$ and $\mathbf{n}_B = (n_{B,1}, \ldots, n_{B,v})$. The Kronecker product $A \boxminus B$, where A is an $s \times t$ non-partitioned block matrix, is*

$$A \boxminus B = \sum_{uv} E_{e_u, s\mathbf{m}_B}(E_{uv} \otimes A \otimes B_{uv})E^T_{e_v, t\mathbf{n}_B}.$$

Remark: Since

$$\begin{aligned}A \boxminus B &= \sum_{pquv} E_{e_u \otimes e_p, \boldsymbol{m}_B \otimes \boldsymbol{m}_A}(E_{uv} \otimes (E_{pq} \otimes A_{pq}) \otimes B_{uv})E^T_{e_v \otimes e_q, \boldsymbol{n}_B \otimes \boldsymbol{n}_A}\\ &= \sum_{uv} E_{1 \otimes e_u, s \otimes \boldsymbol{m}_B}(1 \otimes E_{uv} \otimes A \otimes B_{uv})E^T_{1 \otimes e_v, t \otimes \boldsymbol{n}_B},\end{aligned}$$

it is clear that if we partition A here as a single block $A_{11} = A$, our product $A \boxminus B$ changes into $A \boxtimes B$. The same properties defined in the balanced case above are also applicable in the unbalanced case when partitioned in this way [5].

Theorem 9 *Let A, B and D be block matrices and I be an appropriate identity matrix. Then*

$$\mathrm{vecb_r}(AD + BD) = (I \boxminus A + B^T \boxminus I)\,\mathrm{vecb_r}(D).$$

Proof *By Corollary 3,*

$$\begin{aligned}\mathrm{vecb_r}(AD + DB) &= \mathrm{vecb_r}(AD) + \mathrm{vecb_r}(DB)\\ &= \mathrm{vecb_r}(ADI^T) + \mathrm{vecb_r}(IDB)\\ &= (I \boxminus A)\,\mathrm{vecb_r}(D) + (B^T \boxminus I)\,\mathrm{vecb_r}(D)\\ &= (I \boxminus A + B^T \boxminus I)\,\mathrm{vecb_r}(D).\end{aligned}$$

We will now provide basic properties of the Kronecker product $A \boxminus B$ collected from [4, 5] which are proven in the appendix below.

Theorem 10 (Properties of the block Kronecker product $A \boxminus B$) *Let A, B, C and D be block matrices partitioned in an unbalanced way, then*

1. $A \boxminus B = {K_1}^T (B \otimes A) K_2$ *for some commutation matrices K_1 and K_2*
2. $A \boxminus (B + C) = (A \boxminus B) + (A \boxminus C)$
3. $(A + B) \boxminus C = (A \boxminus C) + (B \boxminus C)$
4. $(A \boxminus B)^T = A^T \boxminus B^T$
5. $(A \boxminus B)(C \boxminus D) = (AC \boxminus BD)$
6. $(I \boxminus A)(B \boxminus I) = (IB \boxminus AI) = (B \boxminus A)$
7. $(A \boxminus B)^{-1} = A^{-1} \boxminus B^{-1}$
8. $A \boxminus B = {K_1}^T (B \boxminus A) K_2$ *for some commutation matries K_1 and K_2*
9. $\mathrm{tr}(A \boxminus B) = \mathrm{tr}(A)\,\mathrm{tr}(B)$
10. $\det(A \boxminus B) = \det(A)^n \det(B)^m$
11. $\exp(A \boxminus I + I \boxminus B) = \exp A \boxminus \exp B$

Proof *See Appendix A1.*

We note that $A \boxminus B$ has the same properties as $A \otimes B$ and by uniqueness, there exists an isomorphism between these two tensor products which makes them equivalent. This is demonstrated the next commutative diagram.

$$\begin{tikzcd}
A \times B \arrow[r, "\boxminus", "\text{bilinear}"'] \arrow[d, "\otimes"', "\text{bilinear}"] & A \boxminus B \arrow[dl, dotted, shift left, "!\text{linear } \beta"] \\
A \otimes B \arrow[ur, dotted, shift left, "!\text{linear } \alpha"] \arrow[loop below, dotted, "!Id_{A\otimes B}"] &
\end{tikzcd}$$

As a result, uniqueness of the the isomorphism $A \otimes B \to A \boxminus B$ ensures that $\beta \circ \alpha = Id_{A\otimes B}$.

Theorem 11 *Let A and B be square (block) matrices where A has eigenvalue the α_i associated with eigenvector x_i and B has eigenvalue β_j associated with eigenvector y_j such that*

$$Ax_i = \alpha_i x_i, \qquad By_j = \beta_j y_j.$$

Then the Kronecker products $A \boxminus B$ and $B \otimes A$ have the same eigenvalue $\alpha_i\beta_j$ with corresponding eigenvector $(x_i \boxminus y_j)$ and $(y_j \otimes x_i)$, respectively.

Proof *By the Mixed product property,*

$$(A \boxminus B)(x \boxminus y) = Ax_i \boxminus By_j \qquad \textit{and} \qquad \alpha_i\beta_j(x_i \boxminus y_j) = \alpha_i x_i \boxminus \beta_j y_j,$$

$$(A \boxminus B)(x_i \boxminus y_j) = \alpha_i\beta_j(x_i \boxminus y_j).$$

Using property 1 of Theorem 10, we have

$$\begin{aligned}
K_1^T(B \otimes A)(y_j \otimes x_i)K_2 &= (By_j \otimes Ax_i) \\
&= (\beta_j y_j \otimes \alpha_i x_i) \\
&= \beta_j\alpha_i(y_j \otimes x_i).
\end{aligned}$$

Therefore,

$$(B \otimes A)(y_j \otimes x_i) = \beta_j\alpha_i(y_j \otimes x_i).$$

Hence the result holds.

Remark: The block Kronecker product $A \boxminus B$ defines a tensor product of matrix spaces and so by universality it is isomorphic to the algebra defined by $A \otimes B$. Since we know that $A \times B \to C$ is bilinear, we can use the universal factorization property to deduce existence of a unique linear ψ such that the following diagram commutes [7].

$$\begin{tikzcd}
A \times B \arrow[r, "\text{bilinear}"] \arrow[d, "\otimes"'] \arrow[dd, bend right] & C \\
A \otimes B \arrow[ur, dotted, "!\phi"] \arrow[d, shift left, "!"] & \\
A \boxminus B \arrow[u, shift left, "!"] \arrow[uur, dotted, bend right, "!\psi"'] &
\end{tikzcd}$$

It is clear that since we define the block Kronecker product in terms of the standard Kronecker product then operations that hold for the standard Kronecker product also hold for the block Kronecker product. As seen above, it is easier to show basic properties and operations for the block structures using the standard Kronecker product. An important observation is that in order to vectorize column-blockwise, all the matrices need to be partitioned. To vectorize row-blockwise all matrices do not need to be partitioned [5]. In other words,

$$A \boxminus B = \sum_{uv} E_{e_u, \boldsymbol{s}\boldsymbol{m}_B}(E_{uv} \otimes A \otimes B_{uv}) E^T_{e_v, \boldsymbol{t}\boldsymbol{n}_B}$$

is independent of any partitioning of A.

Chapter 4

Tensors and tensor unfoldings

4.1 Introduction to tensors

In this chapter, we introduce a novel approach of tensors and tensor unfoldings. A generalization of these concepts is discussed in the next section, namely block tensors and block tensor unfolding. We denote by $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ an N-th order tensor, which is an $I_1 \times I_2 \times \cdots \times I_N$ tensor and a block tensor will be denoted by $[\mathcal{A}]$. Tensors extend matrices to a higher order and we define a tensor as a higher order multidimensional array [2]. A multidimensional array is an array that consists of multiple arrays, for example a matrix is an array (a column) that contains arrays (rows) in each entry. We consider tensors whenever we cannot work with matrices since they yield specific mathematical framework for formulating and solving problems. Tensors are frequently used in various fields such as numerical multilinear algebra, physics, engineering and machine learning, signal processing, numerical analysis and chemometrics [11]. We can categorise tensors into different orders based on their respective sizes (the number of indices). For example, an $m \times 1$ column vector has one index and is denoted as $x = (x_i)$ for all $i \in \{1, \ldots, m\}$. This is classified as a first order tensor or a tensor of order one [2]. A second order tensor can be thought of as an $m \times n$ matrix $A = (a_{ij})$ where $i \in \{1, \ldots, m\}, j \in \{1, \ldots, n\}$ for all $m, n \in \mathbb{N}$ and is denoted

$$A = \sum_{ij} a_{ij} e_i e_j^T = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} e_i \otimes e_j^T \approx \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} e_i \otimes e_j. \tag{4.1}$$

An object with 3 indices is an $m \times n \times p$ (third order) tensor $\mathcal{A} = (a_{ijk})$ denoted by

$$\mathcal{A} = \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{k=1}^{p} a_{ijk} e_i \otimes e_j \otimes e_k. \tag{4.2}$$

Third order tensors and tensors of order greater than 3 are referred to as higher order tensors. A tensor of order N is an $I_1 \times I_2 \times \ldots \times I_N$ tensor with N indices and is denoted

$$\mathcal{A} = \sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \cdots \sum_{i_N=1}^{I_N} a_{i_1 i_2 \cdots i_N} e_{i_1} \otimes e_{i_2} \otimes \cdots \otimes e_{i_N}.$$

In this chapter we will focus on third order tensors to present most concepts since they are reasonable to visualize. In most cases, concepts generalize to higher order tensors in a straightforward way. Tensors are arrays that contain sub-arrays. These are referred to as fibres and slices. For a third order tensor, the sub-arrays are classified into three kinds based on which index is fixed. Slices are formed when we fix all but two indices of the given tensor indices. They can be thought of as a collection of matrices. For a third order tensor, slices are also divided into three modes, for example the frontal slices arise naturally from $\mathcal{A}_{::k}$. Here, " : " indicates that the corresponding index is not fixed. In the case where j is fixed, we get $\mathcal{A}_{:j:}$ which indicates a lateral slice. Horizontal slices are obtained by fixing i in $\mathcal{A}_{i::}$. For example, let a tensor of order three $\mathcal{A} \in \mathbb{R}^{3\times 4\times 2}$ be defined as

$$\mathcal{A} = \sum_{i=1}^{3} \sum_{j=1}^{4} \sum_{k=1}^{2} a_{ijk} e_i \otimes e_j \otimes e_k$$

where $a_{111} = 1$, $a_{121} = 2$, $a_{131} = 3$, $a_{141} = 4$, $a_{211} = 5$, $a_{221} = 6$, $a_{231} = 7$, $a_{241} = 8$, $a_{311} = 9$, $a_{321} = 10$, $a_{331} = 11$, $a_{341} = 12$, $a_{112} = 13$, $a_{122} = 14$, $a_{132} = 15$, $a_{142} = 16$, $a_{212} = 17$, $a_{222} = 18$, $a_{232} = 19$, $a_{242} = 20$, $a_{312} = 21$, $a_{322} = 22$, $a_{332} = 23$, and $a_{342} = 24$.

We write

$$A_{::1} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix} \quad \text{and} \quad A_{::2} = \begin{bmatrix} 13 & 14 & 15 & 16 \\ 17 & 18 & 19 & 20 \\ 21 & 22 & 23 & 24 \end{bmatrix},$$

then, $A_{::1}$ and $A_{::2}$ are called the frontal slices of the tensor. Lateral slices are defined when j is fixed and are expressed as

$$A_{:1:} = \begin{bmatrix} 1 & 13 \\ 5 & 17 \\ 9 & 21 \end{bmatrix}, \quad A_{:2:} = \begin{bmatrix} 2 & 14 \\ 6 & 18 \\ 10 & 22 \end{bmatrix}, \quad A_{:3:} = \begin{bmatrix} 3 & 15 \\ 7 & 19 \\ 11 & 23 \end{bmatrix} \quad \text{and} \quad A_{:4:} = \begin{bmatrix} 4 & 16 \\ 8 & 20 \\ 12 & 24 \end{bmatrix}.$$

The horizontal slices are described as

$$A_{1::} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 13 & 14 & 15 & 16 \end{bmatrix}, \quad A_{2::} = \begin{bmatrix} 5 & 6 & 7 & 8 \\ 17 & 18 & 19 & 20 \end{bmatrix}, \quad A_{3::} = \begin{bmatrix} 9 & 10 & 11 & 12 \\ 21 & 22 & 23 & 24 \end{bmatrix}$$

where i is fixed. We define fibres as an extension of matrix columns and rows defined by fixing all but one index. These are considered as a collection of vectors. From the example, the tensor $\mathcal{A} \in \mathbb{R}^{3\times4\times2}$ has the following mode-1 fibres,

$$A_{:11} = \begin{bmatrix} 1 \\ 5 \\ 9 \end{bmatrix}, \quad A_{:12} = \begin{bmatrix} 13 \\ 17 \\ 21 \end{bmatrix}, \quad A_{:21} = \begin{bmatrix} 2 \\ 6 \\ 10 \end{bmatrix}, \quad A_{:22} = \begin{bmatrix} 14 \\ 18 \\ 22 \end{bmatrix},$$

$$A_{:31} = \begin{bmatrix} 3 \\ 7 \\ 11 \end{bmatrix}, \quad A_{:32} = \begin{bmatrix} 15 \\ 19 \\ 23 \end{bmatrix}, \quad A_{:41} = \begin{bmatrix} 4 \\ 8 \\ 12 \end{bmatrix}, \quad A_{:42} = \begin{bmatrix} 16 \\ 20 \\ 24 \end{bmatrix}.$$

Each $A_{:jk}$ is regarded as a mode-1 fibre of the tensor where j, k are fixed. Mode-2 fibres are defined when i, k are fixed and expressed as

$$A_{1:1} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \quad A_{1:2} = \begin{bmatrix} 13 \\ 14 \\ 15 \\ 16 \end{bmatrix}, \quad A_{2:1} = \begin{bmatrix} 5 \\ 6 \\ 7 \\ 8 \end{bmatrix},$$

$$A_{2:2} = \begin{bmatrix} 17 \\ 18 \\ 19 \\ 20 \end{bmatrix}, \quad A_{3:1} = \begin{bmatrix} 9 \\ 10 \\ 11 \\ 12 \end{bmatrix}, \quad A_{3:2} = \begin{bmatrix} 21 \\ 22 \\ 23 \\ 24 \end{bmatrix}.$$

For mode-3 fibres, we get

$$A_{11:} = \begin{bmatrix} 1 \\ 13 \end{bmatrix}, \quad A_{12:} = \begin{bmatrix} 2 \\ 14 \end{bmatrix}, \quad A_{13:} = \begin{bmatrix} 3 \\ 15 \end{bmatrix}, \quad A_{14:} = \begin{bmatrix} 4 \\ 16 \end{bmatrix},$$

$$A_{21:} = \begin{bmatrix} 5 \\ 17 \end{bmatrix}, \quad A_{22:} = \begin{bmatrix} 6 \\ 18 \end{bmatrix}, \quad A_{23:} = \begin{bmatrix} 7 \\ 19 \end{bmatrix}, \quad A_{24:} = \begin{bmatrix} 8 \\ 20 \end{bmatrix},$$

$$A_{31:} = \begin{bmatrix} 9 \\ 21 \end{bmatrix}, \quad A_{32:} = \begin{bmatrix} 10 \\ 22 \end{bmatrix}, \quad A_{33:} = \begin{bmatrix} 11 \\ 23 \end{bmatrix}, \quad A_{34:} = \begin{bmatrix} 12 \\ 24 \end{bmatrix}$$

where i, j are fixed.

We can convert a tensor into a matrix by rearranging its entries into rows and columns. This is known as tensor unfolding or matricization [2]. We can transform a single tensor into different matrices depending on the order of the tensor by partitioning its index into different groups [2]. For example, a third order tensor $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ has three possible matrix outcomes (tensor unfoldings) as follows: mode-1 which is an $I_1 \times I_2 I_3$ matrix, mode-2 which is an $I_2 \times I_3 I_1$ matrix and mode-3 which is an $I_3 \times I_1 I_2$ matrix [2]. In particular, an N-th order tensor has mode-n unfolding denoted by $I_n \times (I_{n+1} I_{n+2} \cdots I_N I_1 I_2 \cdots I_{n-1})$. We note here that unfolding is defined differently in [2] and [8], the results which follow can easily be adapted to either definition. We will use the following definition [2].

Definition 9 (Tensor unfolding) *The mode-n tensor unfolding of an N-th order tensor $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ is the matrix $A_{(n)} \in \mathbb{R}^{I_n \times (I_{n+1} I_{n+2} \cdots I_N I_1 I_2 \cdots I_{n-1})}$ with the element $a_{i_1 i_2 \ldots i_N}$ at the position with row number i_n and column number*

$$\sum_{k=n+1}^{N} (i_k - 1) \prod_{s=k+1}^{N} I_s \prod_{j=1}^{n-1} I_j + \sum_{k=1}^{n-1} (i_k - 1) \prod_{j=2}^{n-1} I_j.$$

Tensor unfolding arises naturally from (4.1). In particular, the equation defines the mode-1 unfolding which is an $I_1 \times I_2$ matrix. The mode-2 unfolding is an $I_2 \times I_1$ matrix denoted

$$A_{(2)} = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} e_j \otimes e_i^T = A_{(1)}^T. \tag{4.3}$$

For a third order tensor, the mode-1 unfolding is defined as

$$\begin{aligned} A_{(1)} &= \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{k=1}^{p} a_{ijk} e_i \otimes e_j^T \otimes e_k^T \\ &= \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{k=1}^{p} a_{ijk} e_i \otimes (e_j \otimes e_k)^T. \end{aligned}$$

Mode-2 ($A_{(2)}$) and mode-3 ($A_{(3)}$) unfoldings are also defined similar to (4.3) by interchanging the indices with respect to their respective modes.

Example 3 *Let a tensor of order three $\mathcal{A} \in \mathbb{R}^{3\times4\times2}$ be given by*

$$A_{::1} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix} \quad \text{and} \quad A_{::2} = \begin{bmatrix} 13 & 14 & 15 & 16 \\ 17 & 18 & 19 & 20 \\ 21 & 22 & 23 & 24 \end{bmatrix},$$

where $A_{::1}$ and $A_{::2}$ are the frontal slices of $\mathcal{A}$. Then the mode-1 unfolding of $\mathcal{A}$ is the 3×8 matrix

$$A_{(1)} = \begin{bmatrix} 1 & 13 & 2 & 14 & 3 & 15 & 4 & 16 \\ 5 & 17 & 6 & 18 & 7 & 19 & 8 & 20 \\ 9 & 21 & 10 & 22 & 11 & 23 & 12 & 24 \end{bmatrix},$$

the mode-2 unfolding is the 4×6 matrix

$$A_{(2)} = \begin{bmatrix} 1 & 5 & 9 & 13 & 17 & 21 \\ 2 & 6 & 10 & 14 & 18 & 22 \\ 3 & 7 & 11 & 15 & 19 & 23 \\ 4 & 8 & 12 & 16 & 20 & 24 \end{bmatrix},$$

and the mode-3 unfolding is the 2×12 matrix

$$A_{(3)} = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 & 17 & 18 & 19 & 20 & 21 & 22 & 23 & 24 \end{bmatrix}.$$

For an N-th order tensor, the n-th unfolding is defined as

$$\begin{aligned} A_{(n)} &= \sum_{i_1 \ldots i_n}^{I_1 \ldots I_N} a_{i_1,i_2,\ldots,i_n} e_{i_n} \otimes e_{i_{n+1}}^T \otimes \cdots \otimes e_{i_N}^T \otimes e_{i_1}^T \otimes \cdots \otimes e_{i_{n-1}}^T \\ &= \sum_{i_1 \ldots i_n}^{I_1 \ldots I_N} a_{i_1,i_2,\ldots,i_n} e_{i_n} \otimes (e_{i_{n+1}} \otimes \cdots \otimes e_{i_N} \otimes e_{i_1} \otimes \cdots \otimes e_{i_{n-1}})^T \\ &= \sum_{i_1 \ldots i_n}^{I_1 \ldots I_N} a_{i_1,i_2,\ldots,i_N} e_{i_n} \otimes \left(\bigotimes_{k=n+1}^{N} e_{i_k} \right)^T \otimes \left(\bigotimes_{k=1}^{n-1} e_{i_k} \right)^T. \end{aligned}$$

As seen above with matrices and the vec operator, it is also possible to represent all the elements of an n-th order tensor as a vector. We vectorize tensors in a similar way seen above by vectorizing each array. In particular, if $\mathcal{A} \in \mathbb{R}^{I_1 \times \cdots \times I_n}$ then we define inductively the vec operator of the tensor denoted $\text{vec}(\mathcal{A}) \in \mathbb{R}^N$, where $N = I_1 \cdots I_n$, by

$$\text{vec}(\mathcal{A}) = \begin{bmatrix} \text{vec}(A_{::1}) \\ \text{vec}(A_{::2}) \\ \vdots \\ \text{vec}(A_{::I_n}) \end{bmatrix}$$

where $A_{::1}, A_{::2}, \ldots A_{::I_n}$ denote the mode-1 frontal slices. It is clear that the vec of a first order tensor (vector) is simply its column vector.

Example 4 *Consider the tensor $\mathcal{A} \in \mathbb{R}^{3\times4\times2}$ defined above, the vectorization of $\mathcal{A}$ is given as follows*

$$\text{vec}(\mathcal{A}) = \begin{bmatrix} 1 \\ 5 \\ 9 \\ 13 \\ 17 \\ \vdots \\ 8 \\ 12 \\ 16 \\ 20 \\ 24 \end{bmatrix} \in \mathbb{R}.$$

In other cases, the vec of an n-th order tensor can be defined by taking the vec of the mode-1 unfolding of the same tensor, depending on the ordering. For example, the third order tensor defined above has the mode-1 unfolding

$$A_{(1)} = \sum_{i=1}^{3}\sum_{j=1}^{4}\sum_{k=1}^{2} a_{ijk} e_i \otimes (e_j \otimes e_k)^T.$$

We can derive the algebra of the unfolding vec of $A_{(1)}$ using the definition as follows,

$$\begin{aligned}
\mathrm{vec}(A_{(1)}) &= \mathrm{vec}\left(\sum_{i=1}^{3}\sum_{j=1}^{4}\sum_{k=1}^{2} a_{ijk} e_i \otimes (e_j \otimes e_k)^T\right) \\
&= \sum_{i=1}^{3}\sum_{j=1}^{4}\sum_{k=1}^{2} \mathrm{vec}(a_{ijk} e_i \otimes (e_j \otimes e_k)^T) \\
&= \sum_{i=1}^{3}\sum_{j=1}^{4}\sum_{k=1}^{2} (e_j \otimes e_k) \otimes a_{ijk} e_i,
\end{aligned}$$

which is a column vector in $\mathbb{R}^{24}$. The general equation of this vec is given by

$$\mathrm{vec}(A_{(n)}) = \sum_{i_1 \ldots i_n}^{I_1 \ldots I_N} \left(\bigotimes_{k=n+1}^{N} e_{i_k}\right) \otimes \left(\bigotimes_{k=1}^{n-1} e_{i_k}\right) \otimes a_{i_1, i_2, \ldots, i_N} e_{i_n}.$$

Matrix operations such as matrix products and matrix addition can be extended to tensors. We can add any two tensors entrywise with restriction that they have the same order and size, for example $\mathcal{C}_{ijk} = \mathcal{A}_{ijk} + \mathcal{B}_{ijk}$.

4.2 Tensor multiplications

We have already defined matrix products in Chapter 1, we will now describe the different tensor products which involves multiplying a tensor by a vector, matrix and by another tensor. We first demonstrate how the algebra of these multiplications by describing them using a first order tensor and a second order tensor, called the matrix-vector product.

Matrix-vector multiplication

We use matrix-vector multiplication as an inspiration for tensor-vector multiplication. Matrix-vector multiplication is modal, i.e. we distinguish between multiplying with row and column vectors.

Let A be an $m \times n$ matrix and v, w be vectors of order $n \times 1$ and $m \times 1$, respectively. Then we have

$$(A \times_2 v)_i = \sum_{J=1}^{n} A_{ij} v_j = (Av)_i,$$

$$(A \times_1 w)_j = \sum_{i=1}^{m} A_{ij}^T w_i = (w^T A)_j^T.$$

Tensor-vector multiplication

Let $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_n \times \cdots \times I_N}$ be a tensor of order N and let $v \in \mathbb{R}^{I_n}$. We define the tensor vector multiplication by

$$(\mathcal{A} \times_k v)_{i_1 i_2 \cdots i_{n-1} i_{n+1} \cdots i_N} = \sum_{i_n=1}^{N} a_{i_1 i_2 \cdots i_n \cdots i_N} v_{i_n}.$$

This product is of size $I_1 \times I_2 \times \cdots \times I_{n-1} \times I_{n+1} \times \cdots \times I_N$. If the matrix $A_{(n)}$ is the n-th unfolding of $\mathcal{A}$ then

$$\mathcal{A} \times_k v = \text{unvec}(v^T A_{(n)})$$

where unvec is the appropriate inverse of the vec operation. We find the following property which holds for all the tensor vector multiplications. Consider the tensor $\mathcal{A}$ above and let v and w be vectors of order $I_n \times 1$ and $I_m \times 1$, respectively. Then

$$(\mathcal{A} \times_n v) \times_m w = \begin{cases} \mathcal{A} \times_m w \times_n v, & n > m \\ \mathcal{A} \times_{m+1} w \times_n v, & n \leq m. \end{cases}$$

Tensor-matrix product

Consider an N-th order tensor $\mathcal{A} \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ and a matrix $B \in \mathbb{R}^{I_n \times K}$. We define the tensor matrix multiplication, denoted $\times_n$, as

$$(\mathcal{A} \times_n B)_{i_1 i_2 \cdots i_{n-1} k i_{n+1} \cdots i_N} = \sum_{i_n=1}^{I_N} a_{i_1 \cdots i_n \cdots i_N} B_{i_n k}.$$

Here we have a contraction over the index i_n, thus replaced by k. Equivalently, we can unfold a tensor $\mathcal{A}$ into a matrix $A_{(n)}$ and multiply with our matrix B to get a new matrix $A_{(n)}$ which can be folded back into the desired tensor matrix product. In other words, let $A_{(n)}$ be the mode-n unfolding of the tensor $\mathcal{A} \in \mathbb{R}^{I_1 \times \cdots \times I_n}$, then

$$(\mathcal{A} \times_n B) = \text{fold}_n(B^T A_{(n)}),$$

where fold_n denoted the folding of a matrix back to a tensor. The tensor matrix multiplication has the following properties (collected from [2])

1. Let B and C be matrices of order $I_n \times I_m$ and $I_m \times J_k$, respectively where $m \neq n$. Then

$$\mathcal{A} \times_n B \times_m C = \mathcal{A} \times_m C \times_n B.$$

2. Let B and C be matrices of order $I_n \times I_m$ and $I_m \times I_k$. Then

$$\begin{aligned}(\mathcal{A} \times_n B) \times_n C &= \mathcal{A} \times_n (BC) \\ &= \text{fold}_n((BC)^T A_{(n)}).\end{aligned}$$

Tensor-tensor product

This product involves a tensor multiplied by a tensor. Let $\mathcal{A} \in \mathbb{R}^{I_1 \times \cdots \times I_n}$ and $\mathcal{B} \in \mathbb{R}^{J_1 \times \cdots \times J_m}$ be a tensor of order n and m, respectively. Then

$$\mathcal{A} \circ \mathcal{B} = \mathcal{C}$$

where $\mathcal{C} \in \mathbb{R}^{I_1 \times \cdots \times I_n \times J_1 \times \cdots \times J_m}$ is called the outer product.. The order of the outer product of two tensors, is the sum of the order of the those tensors, for example $\mathcal{A} \circ B$ is of order $n + m$.

4.3 Block tensors

4.3.1 Balanced and unbalanced block tensors

We have seen block structured matrices in Chapter 3 called block matrices. We follow the same pattern in this section by introducing block tensors. Similar to block matrices, a block tensor is a tensor that consists of tensors as its entries [11]. Let $\mathbf{M} = \left(\mathbf{m}^{(1)}, \ldots, \mathbf{m}^{(N)}\right)$ where $\mathbf{m}^{(k)} = [m_1^{(k)}, \ldots, m_{b_k}^{(k)}]$ is a vector of positive integers, then we say $\mathbf{M}$ is a blocking of an N-th order tensor $\mathcal{A} \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ [8]. This is usually indicated by the dimension of the block tensor. We call $\mathcal{A}$ a $b_1 \times b_2 \times \cdots \times b_N$ block tensor with entries $\mathcal{A}_{i_1 \cdots i_N}$. If each $\mathcal{A}_{i_1 \cdots i_N}$ has the same size, then we say that the block tensor is balanced [8].

Example 5 *Let* $\mathbf{M} = \left(\mathbf{m}^{(1)}, \mathbf{m}^{(2)}, \mathbf{m}^{\ (3)}\right)$, *where* $\mathbf{m}^{(1)} = [m_1^{(1)}, m_2^{(1)}]$, $\mathbf{m}^{(2)} = [m_1^{(2)}, m_2^{(2)}, m_3^{(2)}]$ *and* $\mathbf{m}^{(3)} = [m_1^{(3)}, m_2^{(3)}]$. *Then we regard our tensor as a third order* $2 \times 3 \times 2$ *block tensor.*

A general balanced third order block tensor can be written in the form

$$[\mathcal{A}] = \sum_{ijk} e_i \otimes e_j \otimes e_k \otimes \mathcal{A}_{ijk},$$

where each $\mathcal{A}_{ijk}$ is a third order tensor (all of the same size). In general, an N-th order block tensor may be written

$$[\mathcal{A}] = \sum_{i_1 \cdots i_N} e_{i_1} \otimes e_{i_2} \otimes \cdots \otimes e_{i_N} \otimes \mathcal{A}_{i_1 \cdots i_N},$$

where each $\mathcal{A}_{i_1 \cdots i_N}$ is of the same size.

Definition 10 *A block tensor (unbalanced partitioning) may be written in the form*

$$[\mathcal{A}] = \sum_{i_1} \sum_{i_2} \cdots \sum_{i_N} (e_{i_1} \otimes e_{i_2} \otimes \cdots \otimes e_{i_N} \otimes \mathcal{A}_{i_1 \cdots i_N}) \times_1 E_{e_{i_1}, \boldsymbol{m}_1} \times_2 E_{e_{i_2}, \boldsymbol{m}_2} \times_3 \cdots .$$

Consider a third order tensor $\mathcal{A} \in \mathbb{R}^{9 \times 6 \times 8}$. It is inconvenient to write this tensor in a similar way to the one in Example 3, since it is large. So we use a number to express the column, row or fibre in which we partition. For example, a 2×3 matrix can be represented with rows as follows

$$1:2 = [1 \quad 2]$$

where the numbers $1, 2$, represent the first and second rows, and columns

$$1:3 = [1 \quad 2 \quad 3]$$

which represents the first, second and third column. We can partition the columns, rows and fibres as follow.

Example 6 *A* $9 \times 6 \times 8$ *tensor may be indexed as follows*

$$\begin{aligned}
[1:9] &= \left[\begin{array}{ccc|ccc|ccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \end{array}\right] = \left[\begin{array}{c|c|c} 1:3 & 4:6 & 7:9 \end{array}\right], \\
[1:6] &= \left[\begin{array}{cc|cccc} 1 & 2 & 3 & 4 & 5 & 6 \end{array}\right] = \left[\begin{array}{c|c} 1:2 & 3:6 \end{array}\right], \\
[1:8] &= \left[\begin{array}{cc|cc|cc|cc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{array}\right] = \left[\begin{array}{c|c|c|c} 1:2 & 3:4 & 5:6 & 7:8 \end{array}\right].
\end{aligned}$$

Here, the colons represent the blocks, as a result this reduces to a $3 \times 2 \times 4$

block tensor.

Example 7 *In the example above, the* $(1,2,4)$ *block, denoted as the tensor* $\mathcal{A}_{124}$, *is the subtensor* $\mathcal{A}(1:3, 3:6, 7:8)$ *with entries* a_{137}, a_{138}, a_{147},a_{148}, a_{157}, a_{158}, a_{167}, a_{168}, a_{237}, a_{238}, a_{247}, a_{248}, a_{257}, a_{258}, a_{267}, a_{268}, a_{337}, a_{338}, a_{347}, a_{348}, a_{357}, a_{358}, a_{367} *and* a_{368}. *This is a third order tensor that acts an an entry of the block tensor.*

4.3.2 Unfolding

Block tensor unfolding arises naturally from the definition of a balanced block matrix defined in Chapter 3, in particular, for a second order block tensor

$$\begin{aligned} A_{(1)} &= \sum_{ij} E_{ij} \otimes (A_{ij})_{(1)} \\ &= \sum_{ij} e_i e_j^T \otimes (A_{ij})_{(1)} \\ &= \sum_{ij} e_i \otimes e_j^T \otimes (A_{ij})_{(1)} \end{aligned}$$

which is called the mode-1 unfolding of a second order block tensor. Mode-2 unfolding becomes

$$A_{(2)} = \sum_{ij} e_j \otimes e_i^T \otimes (A_{ij})_{(2)}.$$

This is the case when our block tensor above is partitioned in a balanced way. For a third order tensor, the unfolding is defined in a similar way to Definition 10 and unfolding in the first mode yields

$$A_{(1)} = \sum_{ijk} e_i \otimes e_j^T \otimes e_k^T \otimes (A_{ijk})_{(1)}$$

For an N-th order block tensor, we define its n-th block unfolding as follows,

$$\begin{aligned} A_{(n)} &= \sum_{i_1 \cdots i_N} e_{i_n} \otimes e_{i_{n+1}}^T \otimes \cdots e_{i_N}^T \otimes e_{i_1}^T \otimes \cdots \otimes e_{i_{n-1}}^T \otimes (A_{i_1 \cdots i_N})_{(n)} \\ &= \sum_{i_1 \cdots i_N}^{I_1 \cdots I_N} e_{i_n} \otimes \left(\bigotimes_{k=n+1}^{N} e_{i_k} \right)^T \otimes \left(\bigotimes_{k=1}^{n-1} e_{i_k} \right)^T \otimes (A_{i_1 \cdots i_N})_{(n)}. \end{aligned}$$

Consider a case where our block tensor is partitioned in an unbalanced way. To introduce the algebra of an unbalanced block tensor and its unfolding, we use mode-n multiplication.

Since the unbalanced block tensor

$$[\mathcal{A}] = \sum_{i_1}\sum_{i_2}\cdots\sum_{i_N}(e_{i_1}\otimes e_{i_2}\otimes\cdots\otimes e_{i_N}\otimes\mathcal{A}_{i_1\cdots i_N})\times_1 E_{e_{i_1},\mathbf{m}_1}\times_2 E_{e_{i_2},\mathbf{m}_2}\times_3\cdots,$$

its unfolding retains the structure of the balanced unfoldings with the necessary modal tensor multiplications which express the unbalanced structure. In particular, an unfolded unbalanced block tensor of order 2 is defined also from the equation of the unbalanced block matrix as follows:

$$\begin{aligned}A_{(1)} &= \sum_{ij} E_{e_i,\mathbf{m}}(E_{ij}\otimes(A_{ij})_{(1)})E_{e_j,\mathbf{n}}^T\\ &= \sum_{ij} E_{e_i,\mathbf{m}}(e_i\otimes e_j^T\otimes(A_{ij})_{(1)})E_{e_j,\mathbf{n}}^T.\end{aligned}$$

This is a mode-1 unfolding. The mode-2 unfolding of $\mathcal{A}$ is given by

$$A_{(2)} = \sum_{ij} E_{e_j,\mathbf{n}}(E_{ji}\otimes(A_{ij})_{(2)})E_{e_i,\mathbf{m}}^T.$$

Definition 11 *Let*

$$\mathcal{A}\times_{23} E_{e_j\otimes e_k,\mathbf{n}\otimes\mathbf{p}} =: \mathcal{A}\times_2 E_{e_j,\mathbf{n}}\times_3 E_{e_k,\mathbf{p}},$$

whenever the products on the right hand side are defined.

Proposition 3 *A block tensor,*

$$\begin{aligned}[\mathcal{A}] &= \sum_{ijk}(e_i\otimes e_j\otimes e_k\otimes A_{ijk})\times_1 E_{e_i,\mathbf{m}}\times_2 E_{e_j,\mathbf{n}}\times_3 E_{e_k,\mathbf{p}}\\ &= \sum_{ijk}(e_i\otimes e_j\otimes e_k\otimes A_{ijk})\times_1 E_{e_i,\mathbf{m}}\times_{23} E_{e_j\otimes e_k,\mathbf{n}\otimes\mathbf{p}}\end{aligned}$$

has the mode-1 block unfolding

$$A_{(1)} = \sum_{ijk} E_{e_i,\boldsymbol{m}}(e_i\otimes e_j^T\otimes e_k^T\otimes(A_{ijk})_{(1)})E_{e_j\otimes e_k,\mathbf{n}\otimes\mathbf{p}}^T. \tag{4.4}$$

The operation $\times_{23}$ is well defined in the sense that $E_{e_j\otimes e_k,\mathbf{n}\otimes\mathbf{p}} = E_{e_u\otimes e_v,\mathbf{n}'\otimes\mathbf{p}'}$ if and only if $j=u$ and $k=v$, also $\mathbf{n}=\mathbf{n}'$ and $\mathbf{p}=\mathbf{p}'$.

Proof First, we note that

$$\begin{aligned} e_j \otimes e_k = e_u \otimes e_v &\iff \operatorname{vec}(E_{kj}) = \operatorname{vec}(E_{vu}) \\ &\iff E_{kj} = E_{vu} \\ &\iff j = u, k = v. \end{aligned}$$

In addition, suppose that $\sum_j n_j = n = \sum_j n_j^{'}$ and $\sum_k p_k = p = \sum_k p_k^{'}$ for all j, k. Then

$$\begin{aligned} \mathbf{n} \otimes \mathbf{p} = \mathbf{n}' \otimes \mathbf{p}' &\iff n_j p_k = n_j^{'} p_k^{'} \\ &\iff \sum_k n_j p_k = \sum_k n_j^{'} p_k^{'} \\ &\iff n_j p = n_j^{'} p \\ &\iff n_j = n_j^{'}. \end{aligned}$$

Consequently,

$$\begin{aligned} \sum_j n_j p_k = \sum_j n_j^{'} p_k^{'} & \\ &\iff n p_k = n p_k^{'} \\ &\iff p_k = p_k^{'} \end{aligned}$$

which implies $\mathbf{n} = \mathbf{n}'$ and $\mathbf{p} = \mathbf{p}'$. Equation (4.4) can be generalized for an N-th order tensor as the n-th block unfolding defined as

$$\begin{aligned} A_{(n)} = \sum_{i_1 \cdots i_N} E_{e_{i_n}, \mathbf{m}_n} (e_{i_n} \otimes e_{i_{n+1}}^T \otimes \cdots \otimes e_{i_N}^T \otimes e_{i_1}^T \otimes \cdots \otimes e_{i_{n-1}}^T \\ \otimes (A_w)_{(n)}) E^T_{e_{i_{n+1}} \otimes \cdots \otimes e_{i_N} \otimes \cdots \otimes e_{i_{n-1}}, \mathbf{m}_{n+1} \otimes \cdots \otimes \mathbf{m}_N \otimes \cdots \otimes \mathbf{m}_{n-1}}. \end{aligned}$$

where $((A_w)_{(n)})$ is the mode-n tensor unfolding for $w = i_1 i_2 \cdots i_N$. The same idea of tensor unfolding seen above also holds in this case, that is, the block tensor unfolding will be divided into different modes depending on the dimension as seen above in Example 3.

As seen with tensors above, we can also vectorize block tensors to result in block vectors. This is similar to the concept of block vectorization of matrices, i.e. the operator vecb_c, and we denote it by $\operatorname{vec}_{\mathbf{M}}(\mathcal{A})$.

Definition 12 *Let* $\mathbf{M} = \left(\mathbf{m}^{(1)}, \ldots, \mathbf{m}^{(N)}\right)$ *be a blocking of an* N*-th order ten-*

sor $\mathcal{A} \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ *denoted by* $\mathcal{A}_i$. *We define* $\text{vec}_{\mathbf{M}}(\mathcal{A})$ *by*

$$\text{vec}_{\mathbf{M}}(\mathcal{A}) = \sum_{i_1=1}^{I_1} \cdots \sum_{i_N=1}^{I_N} \left(\bigotimes_{j=1}^{n} e_{i_{n-j+1}} \right) \otimes \text{vec}(\mathcal{A}_i),$$

where $\boldsymbol{i} = (i_1 \cdots i_N)$.

Example 8 *Let* $\mathbf{M} = \left(\mathbf{m}^{(1)}, \mathbf{m}^{(2)}, \mathbf{m}^{(3)}\right)$, *where* $\mathbf{m}^{(1)} = [m_1^{(1)}, m_2^{(1)}]$, $\mathbf{m}^{(2)} = [m_1^{(2)}, m_2^{(2)}, m_3^{(2)}]$ *be the blocking of a tensor* $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2}$, *then our block tensor* $[\mathcal{A}]$ *is defined as a* 2×3 *block matrix that contains block tensors as its entries denoted as*

$$[\mathcal{A}] = \begin{bmatrix} \mathcal{A}_{11} & \mathcal{A}_{12} & \mathcal{A}_{13} \\ \mathcal{A}_{21} & \mathcal{A}_{22} & \mathcal{A}_{23} \end{bmatrix}.$$

Then

$$\text{vec}_{\mathbf{M}}(\mathcal{A}) = \begin{bmatrix} \text{vec}\left(A_{11}\right) \\ \text{vec}\left(A_{21}\right) \\ \text{vec}\left(A_{12}\right) \\ \text{vec}\left(A_{22}\right) \\ \text{vec}\left(A_{13}\right) \\ \text{vec}\left(A_{23}\right) \end{bmatrix}.$$

This allows us to extend the results as follows,

$$\text{vec}_{\mathbf{M}}(\mathcal{A}) = \sum_{i=1}^{2} \sum_{j=1}^{3} \text{vec}(E_{ij}) \otimes \text{vec}(\mathcal{A}_{ij}) = \text{vecb}_{\text{c}}(\mathcal{A}_{ij})$$

For a third order tensor, $\text{vec}_{\mathbf{M}}(\mathcal{A})$ acts as our block vec and can be algebraically defined as

$$\text{vec}_{\mathbf{M}}(\mathcal{A}) = \sum_{ijk} E_{e_k \otimes e_j \otimes e_i, \mathbf{m}^{(3)} \otimes \mathbf{m}^{(2)} \otimes \mathbf{m}^{(1)}} e_k \otimes e_j \otimes e_i \otimes \text{vec}(\mathcal{A}_{ijk}).$$

Notice that the $\text{vec}_{\mathbf{M}}$ is defined in the unbalanced case. We consider the unbalanced case since the balanced case can be derived directly from it, in particular the general equation for both balanced and unbalanced is given in the balanced case defined as follows:

$$\text{vec}_{\mathbf{M}}(\mathcal{A}) = \sum_{i_1 \cdots i_n} E_{e_{i_n} \otimes \cdots \otimes e_{i_1}, \mathbf{m}^{(n)} \otimes \cdots \otimes \mathbf{m}^{(1)}} e_{i_n} \otimes \cdots \otimes e_{i_1} \otimes \text{vec}(\mathcal{A}_{i_1 \cdots i_n})$$

for any given block tensor $\mathcal{A} \in \mathbb{R}^{I_1 \times \cdots \times I_n}$.

We see that it is almost impossible to present a tensor array in text whenever the size is too large. The algebra enables us to define such cases and allows for calculations as well. A number of properties in Chapter 3 can be generalized to tensors and tensor unfoldings. We have seen in Chapter 2 that we can transform matrices into vectors (i.e. the vec operator) and in the tensor case the common transformation is the tensor unfolding which takes us from a tensor to matrices as seen above. References [2] and [8] provide a broad discussion on tensor computations including its applications.

Chapter 5

Conclusion

In this dissertation we considered the work of [5] and [11] which describes some algebraic properties and relations of block tensors and block matrices. The Kronecker product realises the tensor product for matrix algebra and naturally provides a notion of balanced partitioned block matrices. For unbalanced partitioned block matrices and block tensors, we introduced a novel approach to consider the algebraic properties via the matrix $E_{e_i,\mathbf{m}}$ (Equation (3.1)) which reflects the algebraic properties of balanced block matrices and block tensors in the unbalanced case. It is clear that the balanced partitioning arises naturally from the unbalanced partitioning but it is usually treated separately because of the underlying algebraic structure. The results of both matrices and tensors show that both concepts can be unified. Some of the properties of block tensors mirror properties of block matrices. The matrix $E_{e_i,\mathbf{m}}$ also enabled us to define the block vec operator for unbalanced partitioned block matrices using the novel approach by reflecting on the block vec operator for balanced partitioned block matrices described in [5]. This further extends to block tensors and block tensor unfolding which reflects the connection of both the vec operator and the Kronecker product described in Chapter 2. A similar pattern on partitioning of block matrices is also formed in block tensors to showcase the balanced and unbalanced partitioning of block tensors. The concept of unbalanced block tensors is introduced via special matrices defined in collaboration with block matrices. The results extend further to the unfolding of block tensors which follows directly from the introduced algebraic equation of unbalanced block matrices.

Appendix A

Additional proofs

A1. The following results are proofs of properties of the block Kronecker product of the unbalanced partition matrices stated in Theorem 10.

Proof of 1. The Kronecker product

$$
\begin{aligned}
A \boxminus B &= \begin{bmatrix} A \otimes B_{11} & \dots & A \otimes B_{1t} \\ \vdots & \ddots & \vdots \\ A \otimes B_{s1} & \dots & A \otimes B_{st} \end{bmatrix} \\
&= \begin{bmatrix} K_{1,L}^T (B_{11} \otimes A) K_{1,R} & \dots & K_{n,L}^T (B_{1n} \otimes A) K_{n,R} \\ \vdots & \ddots & \vdots \\ K_{m,L}^T (B_{m1} \otimes A) K_{m,R} & \dots & K_{m,L}^T (B_{mn} \otimes A) K_{m,R} \end{bmatrix} \\
&= \begin{bmatrix} K_{1,L} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & K_{M,L} \end{bmatrix}^T \begin{bmatrix} (B_{11} \otimes A) & \dots & (B_{1n} \otimes A) \\ \vdots & \ddots & \vdots \\ (B_{m1} \otimes A) & \dots & (B_{mn} \otimes A) \end{bmatrix} \begin{bmatrix} K_{1,R} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & K_{n,R} \end{bmatrix}
\end{aligned}
$$

which follows directly from matrix multiplication. Hence

$$A \boxminus B = K_1^T (B \otimes A) K_2.$$

Proof of 2. By Property 1 and 6 of Theorems 10 and 3, respectively.

$$
\begin{aligned}
A \boxminus (B + C) &= K_1^T (A \otimes (B + C)) K_2 \\
&= K_1^T (A \otimes (B + C)) K_2 \\
&= K_1^T ((A \otimes B) + (A \otimes C)) K_2 \\
&= A \boxminus B + A \boxminus C.
\end{aligned}
$$

Proof of 3. Similar to the proof of 2.

Proof of 4.

$$\begin{aligned}(A \boxminus B)^T &= (K_1^T(B \otimes A)K_2)^T \\ &= K_1(B \otimes A)^T K_2^T \\ &= K_1(B^T \otimes A^T)K_2^T \\ &= K_2^T(B^T \otimes A^T)K_1 \\ &= A^T \boxminus B^T.\end{aligned}$$

Proof of 5. Let A, C and B, D be block product-compatible matrices. There exists permutation matrices K_1, K_2, K_3 such that $K_1(B \otimes A)K_2^T = (A \boxminus B)$ and $K_2(B \otimes A)K_3^T = (C \boxminus D)$. Then,

$$\begin{aligned}(A \boxminus B)(C \boxminus D) &= K_1(B \otimes A)K_2^T K_2(D \otimes C)K_3 \\ &= K_1(B \otimes A)I(D \otimes C)K_2^T \\ &= K_1(B \otimes A)(D \otimes C)K_2^T \\ &= K_1(BD \otimes AC)K_1^T \\ &= AC \boxminus BD.\end{aligned}$$

Proof of 6. Follows directly from Property 5 and multiplication with the identity matrix.

Proof of 7. Follows from Theorem 4.

Proof of 8. Using property 1 of Theorem 10,

$$\begin{aligned}(A \boxminus B) &= K_1^T(B \otimes A)K_2 \\ &= K_1^T(K_1^T(A \otimes B)K_2)K_2 \\ &= K_1^T(B \boxminus A)K_2.\end{aligned}$$

Proof of 9. *By property 1 above*

$$\begin{aligned}\operatorname{tr}(A \boxminus B) &= \operatorname{tr}(K^T(B \otimes A)K) \\ &= \operatorname{tr}(A \otimes B)KK^T \\ &= \operatorname{tr}(A \otimes B) \\ &= \operatorname{tr}(A)\operatorname{tr}(B).\end{aligned}$$

Proof of 10. *Using property 1,*

$$\begin{aligned}\det(A \boxminus B) &= \det(K^T(B \otimes A)K) \\ &= \det(A \otimes B) \\ &= \det(A)^n \det(B)^m.\end{aligned}$$

A2. *The following construction shows that the equivalence in Definition 9 holds.*

Let $\mathbf{m} = (m_1, \ldots, m_k)$ *for* $m \in \mathbb{N}$*, such that*

$$\sum_{j=1}^{k} m_j = m.$$

We define $|\mathrm{m}| := k$ *, where* $|s\mathbf{m}| = |\mathbf{m}|, |m_u\mathbf{s}| = |\mathbf{s}|$ *and* $|\mathbf{m} \otimes \mathbf{n}| = |\mathbf{m}||\mathbf{n}|$*. Let* A *and* B *be block matrices of order* $s \times t$ *and* $m \times n$ *blocks, respectively. Then the Kronecker product*

$$\begin{aligned}A \boxminus B &= \sum_{uv} E_{e_u,s\mathbf{m}}(E_{uv} \otimes A \otimes B_{uv})E^T_{e_v,t\mathbf{n}} \\ &= \sum_{jkuv} E_{e_u,s\mathbf{m}}(E_{uv} \otimes (E_{e_j,\mathbf{s}}(E_{jk} \otimes A_{jk})E^T_{e_u,\mathbf{t}} \otimes B_{uv})E^T_{e_v,t\mathbf{n}} \\ &= \sum_{jkuv} E_{e_u,s\mathbf{m}}(I_{|m|} \otimes E_{e_j,\mathbf{s}} \otimes I_{m_u})E_{uv} \otimes (E_{jk} \otimes A_{jk}) \\ &\qquad \otimes B_{uv}(I_{|n|} \otimes E^T_{e_k,\mathbf{t}} \otimes I_{n_v})E^T_{e_v,t\mathbf{n}}.\end{aligned}$$

Here

$$E_{e_u,s\mathbf{m}} = \left(\bigoplus_{j=1}^{u-1} 0_{sm_j \times sm_u}\right) \oplus I_{sm_u} \oplus \left(\bigoplus_{j=u+1}^{|m|} 0_{sm_j \times sm_u}\right),$$

and

$$I_{|m|} \otimes E_{e_j,\mathbf{s}} \otimes I_{m_u} = I_{|m|} \otimes E_{e_j,m_u\mathbf{s}} = \bigoplus_{i=1}^{|m|} E_{e_j,m_u\mathbf{s}}$$

where $E_{e_j,\mathbf{s}}$ *is an* $s \times |s|s_j$ E *matrix. The product*

$$\begin{aligned}
&E_{e_u,s\mathbf{m}} \left(\bigoplus_{k=1}^{|m|} E_{e_j,m_u\mathbf{s}} \right) \\
&\quad = \left(\left(\bigoplus_{j=1}^{u-1} 0_{sm_j \times sm_u} \right) \oplus I_{sm_u} \oplus \left(\bigoplus_{j=u+1}^{|m|} 0_{sm_j \times sm_u} \right) \right) \left(\bigoplus_{k=1}^{|m|} E_{e_j,m_u\mathbf{s}} \right) \\
&\quad = \left(\bigoplus_{j=1}^{u-1} 0_{sm_j \times s_j|s|m_u} \right) \oplus E_{e_j,m_u\mathbf{s}} \oplus \left(\bigoplus_{j=u+1}^{|m|} 0_{sm_j \times s_j|s|m_u} \right).
\end{aligned}$$

Note that

$$E_{e_j,m_u\mathbf{s}} = \left(\bigoplus_{k=1}^{j-1} 0_{m_u s_k \times m_u s_j} \right) \oplus I_{m_u s_j} \oplus \left(\bigoplus_{k=j+1}^{|s|} 0_{m_u s_k \times m_u s_j} \right).$$

Our product becomes

$$\begin{aligned}
E_{e_u,s\mathbf{m}} \left(\bigoplus_{k=1}^{|m|} E_{e_j,m_u\mathbf{s}} \right) &= \left(\bigoplus_{j=1}^{u-1} 0_{sm_j \times s_j|s|m_u} \right) \\
&\oplus \left(\left(\bigoplus_{k=1}^{j-1} 0_{m_u s_k \times m_u s_j} \right) \oplus I_{m_u s_j} \oplus \left(\bigoplus_{k=j+1}^{|s|} 0_{m_u s_k \times m_u s_j} \right) \right) \\
&\oplus \left(\bigoplus_{j=u+1}^{|m|} 0_{sm_j \times s_j|s|m_u} \right).
\end{aligned}$$

We have

$$0_{sm_j \times s_j|s|m_u} = \bigoplus_{l=1}^{|s|} 0_{m_j s_l \times m_u s_j}$$

Substituting back into the equation above yields

$$\underbrace{\left(\bigoplus_{j=1}^{u-1}\bigoplus_{l=1}^{|s|} 0_{m_j s_l \times m_u s_j}\right)}_{(u-1)|s|} \oplus \left(\underbrace{\left(\bigoplus_{l=1}^{j-1} 0_{m_u s_l \times m_u s_j}\right)}_{j-1} \oplus \underbrace{I_{m_u s_j}}_{(u-1)|s|+j} \oplus \left(\bigoplus_{l=j+1}^{|s|} 0_{m_u s_l \times m_u s_j}\right)\right)$$

$$\oplus \left(\bigoplus_{j=u+1}^{|m|}\bigoplus_{l=1}^{|s|} 0_{m_j s_l \times m_u s_j}\right)$$

which reduces to

$$E_{e_{(u-1)|s|+j},|\mathbf{m}||\mathbf{s}|} = E_{e_u \otimes e_j, \mathbf{m}\otimes\mathbf{s}}$$

by the definition of our E-matrix in Chapter 3.

Bibliography

[1] H. Anton and C. Rorres. *Elementary linear algebra: applications version.* John Wiley & Sons, 2013.

[2] L. De Lathauwer, B. De Moor, and J. Vandewalle. A multilinear singular value decomposition. *SIAM Journal on Matrix Analysis and Applications*, 21(4):1253–1278, 2000.

[3] H. V. Henderson and S. R. Searle. The vec-permutation matrix, the vec operator and Kronecker products: A review. *Linear and Multilinear Algebra*, 9(4):271–288, 1981.

[4] D. C. Hyland and E. G. Collins, Jr. Block Kronecker products and block norm matrices in large-scale systems analysis. *SIAM Journal on Matrix Analysis and Applications*, 10(1):18–29, 1989.

[5] R. H. Koning, H. Neudecker, and T. Wansbeek. Block Kronecker products and the vecb operator. *Linear Algebra and its Applications*, 149:165–184, 1991.

[6] S. Liu. Matrix results on the Khatri-Rao and Tracy-Singh products. *Linear Algebra and its Applications*, 289(1-3):267–277, 1999.

[7] R. Merris. *Multilinear algebra.* CRC Press, 1997.

[8] S. Ragnarsson and C. F. Van Loan. Block tensor unfoldings. *SIAM Journal on Matrix Analysis and Applications*, 33(1):149–169, 2012.

[9] K. Schäcke. On the Kronecker product. *Master's thesis, University of Waterloo*, 2004.

[10] D. S. Tracy and R. P. Singh. A new matrix product and its applications in partitioned matrix differentiation. *Statistica Neerlandica*, 26(4):143–157, 1972.

[11] C. F. Van Loan. Structured matrix problems from tensors. In *Exploiting Hidden Structure in Matrix Computations: Algorithms and Applications*, pages 1–63. Springer, 2016.