

2006

## The Kronecker Product

Bobbi Jo Broxson  
*University of North Florida*

Follow this and additional works at: <https://digitalcommons.unf.edu/etd>

 Part of the [Mathematics Commons](#)

---

### Suggested Citation

Broxson, Bobbi Jo, "The Kronecker Product" (2006). *UNF Graduate Theses and Dissertations*. 25.  
<https://digitalcommons.unf.edu/etd/25>

This Master's Thesis is brought to you for free and open access by the Student Scholarship at UNF Digital Commons. It has been accepted for inclusion in UNF Graduate Theses and Dissertations by an authorized administrator of UNF Digital Commons. For more information, please contact [Digital Projects](#).  
© 2006 All Rights Reserved

**THE KRONECKER PRODUCT**

by

**Bobbi Jo Broxson**

**A thesis submitted to the Department of Mathematics and Statistics  
in partial fulfillment of the requirements for the degree of**

**Master of Science in Mathematics**

**UNIVERSITY OF NORTH FLORIDA**

**COLLEGE OF ARTS AND SCIENCES**

**Spring 2006**

The thesis of Bobbi Jo Broxson is approved:

(Date)

**Signature Deleted**

7/19/06

**Signature Deleted**

7/19/06

**Signature Deleted**

7/20/06

\_\_\_\_\_  
Committee Chairperson

Accepted for the Department:

**Signature Deleted**

7/20/06

\_\_\_\_\_  
Chairperson

Accepted for the College:

**Signature Deleted**

7/21/06

\_\_\_\_\_  
Dean

Accepted for the University:

**Signature Deleted**

21 JUL 2006

\_\_\_\_\_  
Dean of Graduate Studies

## Table of Contents

Abstract .....	ii
Introduction .....	1
Section 1 – Definition and Properties .....	2
Section 2 – Matrix Equations .....	24
Section 3 – Permutation Matrices and Jordan - Canonical Form .....	41
Section 4 – Large Least Squares Problems .....	55
Conclusion .....	61
Bibliography .....	65

## Abstract

This paper presents a detailed discussion of the Kronecker product of matrices. It begins with the definition and some basic properties of the Kronecker product. Statements will be proven that reveal information concerning the eigenvalues, singular values, rank, trace, and determinant of the Kronecker product of two matrices. The Kronecker product will then be employed to solve linear matrix equations. An investigation of the commutativity of the Kronecker product will be carried out using permutation matrices. The Jordan - Canonical form of a Kronecker product will be examined. Variations such as the Kronecker sum and generalized Kronecker product will be introduced. The paper concludes with an application of the Kronecker product to large least squares approximations.

# The Kronecker Product

## Introduction

When most people multiply two matrices together, they generally use the conventional multiplication method. This type of matrix multiplication is taught in algebra courses and represents the composition of two linear transformations that are represented by the two matrices. There is a size restriction when performing this type of matrix multiplication. The number of columns in the first matrix must be the same as the number of rows in the second matrix. Also, this multiplication is not commutative. While this type of matrix multiplication is very common and useful, it is not the only one. The Hadamard product, represented by the symbol  $\circ$ , is another type of matrix multiplication. In this case, the two matrices are multiplied in the same way that they are added in conventional matrix addition. For this multiplication, the two matrices are required to be of the same size. The resulting matrix product is formed by multiplying corresponding entries of the two matrices together. One useful fact about this type of matrix multiplication is that it is commutative. This product is useful in several areas of study. Two such areas are the study of association schemes in combinatorial theory and the weak minimum principle in partial differential equations.

This paper is focused on yet another form of matrix multiplication – the Kronecker product, represented by the symbol  $\otimes$ . It is also known as the direct product or the tensor product. The Kronecker product has an interesting advantage over the previously discussed matrix products. The dimensions of the two matrices being multiplied together do not need to have any relation to each other. Many important

properties of this product will be discussed throughout this paper. Most of the results in Sections 1 – 3 came from statements and exercises in the two books by Horn and Johnson ([5],[6]).

The Kronecker product is useful in many areas of study. Specifically, Kronecker products are applied in the areas of signal processing, image processing, semidefinite programming, and quantum computing ([7]). The Kronecker product is also proving to be an effective way to look at fast linear transforms. This paper will look at applications of the Kronecker product in solving linear matrix equations and large least squares problems.

## **Section 1 – Definitions and Properties**

The Kronecker product has some of the same properties as conventional matrix multiplication. Both products follow the same properties for multiplication with a scalar. Also, both products are associative and they share the distributive property with conventional matrix addition. Furthermore, multiplying any matrix by the zero matrix yields the zero matrix. However, these two types of multiplication have many distinctions, such as results associated with taking transposes and inverses. Specifically, when taking the transpose or inverse of a conventional product of two matrices, the order of the matrices is reversed. In contrast, the transpose or inverse of a Kronecker product preserves the order of the two matrices. In order to begin the discussion, a definition of the Kronecker product of two matrices is needed. This is followed by a specific example

of this matrix multiplication. Shortly thereafter, a few of these characteristics of the Kronecker product are explained.

**Definition 1:** Let  $\mathbb{F}$  be a field. The *Kronecker product* of  $A = [a_{ij}] \in M_{m,n}(\mathbb{F})$  and

$B = [b_{ij}] \in M_{p,q}(\mathbb{F})$  is denoted by  $A \otimes B$  and is defined to be the block matrix

$$A \otimes B \equiv \begin{pmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{pmatrix} \in M_{mp,nq}(\mathbb{F}).$$

**Example 2:** Let  $A = \begin{pmatrix} 0 & -2 \\ 3 & -1 \end{pmatrix}$  and  $B = \begin{pmatrix} 2 & 1 & 5 & 0 \\ -4 & -2 & 6 & 3 \\ -3 & 2 & -1 & 4 \end{pmatrix}$ . Then,

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

In general,  $A \otimes B \neq B \otimes A$ , even though both products result in matrices of the same size. This is seen in the next example where  $B \otimes A$  is calculated using the same matrices  $A$  and  $B$  from Example 2.



**Example 3:**

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

It is interesting to note that for  $x, y \in \mathbb{F}^n$ ,  $xy^T = x \otimes y^T$  and  $xy^* = x \otimes y^*$ . That is, in special cases, the product that arises from the conventional matrix multiplication is the same as the resulting Kronecker product. Additionally, for two identity matrices  $I \in M_m$  and  $I \in M_n$ , the Kronecker product is also an identity matrix,  $I \in M_{mn}$ . The next definition provides information about exponents for the Kronecker product.

**Definition 4:** Let  $A \in M_{m,n}(\mathbb{F})$ . The  $k^{\text{th}}$  Kronecker power  $A^{\otimes k}$  is defined inductively for all positive integers  $k$  by

$$A^{\otimes 1} \equiv A \text{ and } A^{\otimes k} \equiv A \otimes A^{\otimes (k-1)} \text{ for } k = 2, 3, \dots.$$

This definition implies that for  $A \in M_{m,n}(\mathbb{F})$ , the matrix  $A^{\otimes k} \in M_{m^k, n^k}(\mathbb{F})$ .

The following theorem states some of the basic properties of the Kronecker product that were discussed earlier. They will turn out to be quite useful.

**Theorem 5:** Let  $A \in M_{m,n}(\mathbb{F})$ . Then,

- a)  $(\alpha A) \otimes B = \alpha(A \otimes B) = A \otimes (\alpha B)$  for all  $\alpha \in \mathbb{F}$  and  $B \in M_{p,q}(\mathbb{F})$
- b)  $(A \otimes B)^T = A^T \otimes B^T$  for  $B \in M_{p,q}(\mathbb{F})$
- c)  $(A \otimes B)^* = A^* \otimes B^*$  for  $B \in M_{p,q}(\mathbb{F})$
- d)  $(A \otimes B) \otimes C = A \otimes (B \otimes C)$  for  $B \in M_{p,q}(\mathbb{F})$  and  $C \in M_{r,s}(\mathbb{F})$

e)  $(A+B) \otimes C = A \otimes C + B \otimes C$  for  $B \in M_{m,n}(\mathbb{F})$  and  $C \in M_{r,s}(\mathbb{F})$

f)  $A \otimes (B+C) = A \otimes B + A \otimes C$  for  $B, C \in M_{p,q}(\mathbb{F})$

g)  $A \otimes 0 = 0 \otimes A = 0$

h)  $I_m \otimes I_n = I_{mn}$

Proof:

$$\begin{aligned} \text{a) } (\alpha A) \otimes B &= \left( \alpha \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \right) \otimes B = \begin{pmatrix} \alpha a_{11} & \cdots & \alpha a_{1n} \\ \vdots & \ddots & \vdots \\ \alpha a_{m1} & \cdots & \alpha a_{mn} \end{pmatrix} \otimes B \\ &= \begin{pmatrix} \alpha a_{11} B & \cdots & \alpha a_{1n} B \\ \vdots & \ddots & \vdots \\ \alpha a_{m1} B & \cdots & \alpha a_{mn} B \end{pmatrix} = \alpha \begin{pmatrix} a_{11} B & \cdots & a_{1n} B \\ \vdots & \ddots & \vdots \\ a_{m1} B & \cdots & a_{mn} B \end{pmatrix} = \alpha (A \otimes B) \\ &= \begin{pmatrix} \alpha a_{11} B & \cdots & \alpha a_{1n} B \\ \vdots & \ddots & \vdots \\ \alpha a_{m1} B & \cdots & \alpha a_{mn} B \end{pmatrix} = \begin{pmatrix} a_{11} \alpha B & \cdots & a_{1n} \alpha B \\ \vdots & \ddots & \vdots \\ a_{m1} \alpha B & \cdots & a_{mn} \alpha B \end{pmatrix} = A \otimes (\alpha B) \end{aligned}$$

$$\text{b) } (A \otimes B)^T = \begin{pmatrix} a_{11} B & \cdots & a_{1n} B \\ \vdots & \ddots & \vdots \\ a_{m1} B & \cdots & a_{mn} B \end{pmatrix}^T = \begin{pmatrix} a_{11} B^T & \cdots & a_{m1} B^T \\ \vdots & \ddots & \vdots \\ a_{1n} B^T & \cdots & a_{mn} B^T \end{pmatrix} = A^T \otimes B^T$$

$$\text{c) } (A \otimes B)^* = \begin{pmatrix} a_{11} B & \cdots & a_{1n} B \\ \vdots & \ddots & \vdots \\ a_{m1} B & \cdots & a_{mn} B \end{pmatrix}^* = \begin{pmatrix} \bar{a}_{11} B^* & \cdots & \bar{a}_{m1} B^* \\ \vdots & \ddots & \vdots \\ \bar{a}_{1n} B^* & \cdots & \bar{a}_{mn} B^* \end{pmatrix} = A^* \otimes B^*$$

$$\begin{aligned} \text{d) } (A \otimes B) \otimes C &= \begin{pmatrix} a_{11} B & \cdots & a_{1n} B \\ \vdots & \ddots & \vdots \\ a_{m1} B & \cdots & a_{mn} B \end{pmatrix} \otimes C = \begin{pmatrix} (a_{11} B) \otimes C & \cdots & (a_{1n} B) \otimes C \\ \vdots & \ddots & \vdots \\ (a_{m1} B) \otimes C & \cdots & (a_{mn} B) \otimes C \end{pmatrix} \\ &= \begin{pmatrix} a_{11} (B \otimes C) & \cdots & a_{1n} (B \otimes C) \\ \vdots & \ddots & \vdots \\ a_{m1} (B \otimes C) & \cdots & a_{mn} (B \otimes C) \end{pmatrix} = A \otimes (B \otimes C) \end{aligned}$$

$$\text{e) } (A+B) \otimes C = \left( \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \dots & b_{mn} \end{pmatrix} \right) \otimes C = \begin{pmatrix} a_{11}+b_{11} & \dots & a_{1n}+b_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1}+b_{m1} & \dots & a_{mn}+b_{mn} \end{pmatrix} \otimes C$$

$$= \begin{pmatrix} (a_{11}+b_{11})C & \dots & (a_{1n}+b_{1n})C \\ \vdots & \ddots & \vdots \\ (a_{m1}+b_{m1})C & \dots & (a_{mn}+b_{mn})C \end{pmatrix} = \begin{pmatrix} a_{11}C+b_{11}C & \dots & a_{1n}C+b_{1n}C \\ \vdots & \ddots & \vdots \\ a_{m1}C+b_{m1}C & \dots & a_{mn}C+b_{mn}C \end{pmatrix}$$

$$= \begin{pmatrix} a_{11}C & \dots & a_{1n}C \\ \vdots & \ddots & \vdots \\ a_{m1}C & \dots & a_{mn}C \end{pmatrix} + \begin{pmatrix} b_{11}C & \dots & b_{1n}C \\ \vdots & \ddots & \vdots \\ b_{m1}C & \dots & b_{mn}C \end{pmatrix} = (A \otimes C) + (B \otimes C)$$

$$\text{f) } A \otimes (B+C) = \begin{pmatrix} a_{11}(B+C) & \dots & a_{1n}(B+C) \\ \vdots & \ddots & \vdots \\ a_{m1}(B+C) & \dots & a_{mn}(B+C) \end{pmatrix} = \begin{pmatrix} a_{11}B+a_{11}C & \dots & a_{1n}B+a_{1n}C \\ \vdots & \ddots & \vdots \\ a_{m1}B+a_{m1}C & \dots & a_{mn}B+a_{mn}C \end{pmatrix}$$

$$= \begin{pmatrix} a_{11}B & \dots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \dots & a_{mn}B \end{pmatrix} + \begin{pmatrix} a_{11}C & \dots & a_{1n}C \\ \vdots & \ddots & \vdots \\ a_{m1}C & \dots & a_{mn}C \end{pmatrix} = (A \otimes B) + (A \otimes C)$$

$$\text{g) } A \otimes 0 = \begin{pmatrix} a_{11}0 & \dots & a_{1n}0 \\ \vdots & \ddots & \vdots \\ a_{m1}0 & \dots & a_{mn}0 \end{pmatrix} = \begin{pmatrix} 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{pmatrix} = 0 \text{ and}$$

$$0 \otimes A = \begin{pmatrix} 0A & \dots & 0A \\ \vdots & \ddots & \vdots \\ 0A & \dots & 0A \end{pmatrix} = \begin{pmatrix} 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{pmatrix} = 0$$

$$\text{h) } I_m \otimes I_n = \begin{pmatrix} 1I_n & \dots & 0I_n \\ \vdots & \ddots & \vdots \\ 0I_n & \dots & 1I_n \end{pmatrix} = \begin{pmatrix} I_n & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & I_n \end{pmatrix} = I_{mn}$$

□

The next result provides a startling contrast to conventional matrix multiplication. It is

easy to find a nonzero square matrix  $A$  such that  $A^2 = 0$ . Such a matrix is  $A = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}$ .

This cannot happen when using the Kronecker product.

**Corollary 6:** Let  $A \in M_{m,n}$  and  $B \in M_{p,q}$ . Then,  $A \otimes B = 0$  if and only if either  $A = 0$  or  $B = 0$ .

*Proof:* First, let either  $A = 0$  or  $B = 0$ . Then, by property (g) in Theorem 5,  $A \otimes B = 0$ .

Next, let  $A \otimes B = 0$ . This implies that 
$$\begin{pmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{pmatrix} = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}.$$
 In order for

this to be true, either  $B = 0$  or  $a_{ij} = 0$  for all  $i = 1, 2, \dots, m$  and  $j = 1, 2, \dots, n$ . Thus, either

$B = 0$  or  $A = 0$ . □

The next theorem, called the mixed product property, provides a very important and useful fact regarding the interchangeability of the conventional matrix product and the Kronecker product. This property will be used to prove:  $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$  when  $A$  and  $B$  are invertible matrices, the Kronecker product of two unitary matrices is a unitary matrix, and other results. In between these results and this theorem will be a corollary that generalizes the mixed product property.

**Theorem 7:** Let  $A \in M_{m,n}$ ,  $B \in M_{p,q}$ ,  $C \in M_{n,k}$ , and  $D \in M_{q,r}$ . Then,

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

*Proof:* Let  $A = [a_{ih}]$  and  $C = [c_{hj}]$ . By the definition of Kronecker product,

$A \otimes B = [a_{ih}B]$  and  $C \otimes D = [c_{hj}D]$ . The  $i,j^{th}$  block of  $(A \otimes B)(C \otimes D)$  is

$\sum_{h=1}^n (a_{ih}B)(c_{hj}D) = \sum_{h=1}^n (a_{ih}c_{hj})(BD) = \left[ \sum_{h=1}^n (a_{ih}c_{hj}) \right] BD$ . Also, the  $i,j^{th}$  entry of  $AC$  is

$\sum_{h=1}^n a_{ih}c_{hj}$  which implies that the  $i,j^{th}$  block of  $(AC) \otimes (BD)$  is  $\left[ \sum_{h=1}^n (a_{ih}c_{hj}) \right] (BD)$ .

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

**Corollary 8:** Theorem 7 can be generalized in the following two ways:

$$\text{a) } (A_1 \otimes A_2 \otimes \cdots \otimes A_k)(B_1 \otimes B_2 \otimes \cdots \otimes B_k) = A_1B_1 \otimes A_2B_2 \otimes \cdots \otimes A_kB_k$$

$$\text{b) } (A_1 \otimes B_1)(A_2 \otimes B_2) \cdots (A_k \otimes B_k) = (A_1A_2 \cdots A_k) \otimes (B_1B_2 \cdots B_k).$$

*Proof:* (by induction)

First, part a) is proved. By Theorem 7, the mixed product property,

$(A_1 \otimes A_2)(B_1 \otimes B_2) = A_1B_1 \otimes A_2B_2$ . Assume a) is true for  $k = n$ . Let  $k = n+1$ .

$$\begin{aligned} & (A_1 \otimes A_2 \otimes \cdots \otimes A_n \otimes A_{n+1})(B_1 \otimes B_2 \otimes \cdots \otimes B_n \otimes B_{n+1}) \\ &= [(A_1 \otimes A_2 \otimes \cdots \otimes A_n) \otimes A_{n+1}][[(B_1 \otimes B_2 \otimes \cdots \otimes B_n) \otimes B_{n+1}]] \\ &= [(A_1 \otimes A_2 \otimes \cdots \otimes A_n)(B_1 \otimes B_2 \otimes \cdots \otimes B_n)] \otimes [A_{n+1}B_{n+1}] \text{ (mixed product property)} \\ &= [A_1B_1 \otimes A_2B_2 \otimes \cdots \otimes A_nB_n] \otimes [A_{n+1}B_{n+1}] \text{ (induction hypothesis)} \\ &= A_1B_1 \otimes A_2B_2 \otimes \cdots \otimes A_nB_n \otimes A_{n+1}B_{n+1} \end{aligned}$$

Now part b) is proved. By Theorem 7, the mixed product property,

$(A_1 \otimes B_1)(A_2 \otimes B_2) = (A_1A_2) \otimes (B_1B_2)$ . Assume b) is true for  $k = n$ . Let  $k = n+1$ .

$$\begin{aligned} & (A_1 \otimes B_1)(A_2 \otimes B_2) \cdots (A_n \otimes B_n)(A_{n+1} \otimes B_{n+1}) \\ &= [(A_1 \otimes B_1)(A_2 \otimes B_2) \cdots (A_n \otimes B_n)](A_{n+1} \otimes B_{n+1}) \\ &= [(A_1A_2 \cdots A_n) \otimes (B_1B_2 \cdots B_n)](A_{n+1} \otimes B_{n+1}) \text{ (induction hypothesis)} \end{aligned}$$

$$\begin{aligned}
&= \left[ (A_1 A_2 \cdots A_n) A_{n+1} \right] \otimes \left[ (B_1 B_2 \cdots B_n) B_{n+1} \right] \text{ (mixed product property)} \\
&= (A_1 A_2 \cdots A_n A_{n+1}) \otimes (B_1 B_2 \cdots B_n B_{n+1}) \quad \square
\end{aligned}$$

**Corollary 9:** Let  $A \in M_m$  and  $B \in M_n$ . Then,  $(A \otimes I)^k = A^k \otimes I$  and  $(I \otimes B)^k = I \otimes B^k$

for  $k = 1, 2, \dots$ . Also, for any polynomial  $p(t)$ ,  $p(A \otimes I) = p(A) \otimes I$  and

$$p(I \otimes B) = I \otimes p(B) \text{ where } A^0 = B^0 = I.$$

*Proof:* By Corollary 8,

$$\begin{aligned}
(A \otimes I)^k &= (A \otimes I)(A \otimes I) \cdots (A \otimes I) = (AA \cdots A) \otimes (II \cdots I) = A^k \otimes I \text{ and} \\
(I \otimes B)^k &= (I \otimes B)(I \otimes B) \cdots (I \otimes B) = (II \cdots I) \otimes (BB \cdots B) = I \otimes B^k. \text{ Now, let} \\
p(t) &= \alpha_j t^j + \alpha_{j-1} t^{j-1} + \cdots + \alpha_1 t + \alpha_0 t^0. \text{ It then follows that } p(A \otimes I) \\
&= \alpha_j (A \otimes I)^j + \alpha_{j-1} (A \otimes I)^{j-1} + \cdots + \alpha_1 (A \otimes I) + \alpha_0 (A \otimes I)^0 \\
&= \alpha_j (A^j \otimes I) + \alpha_{j-1} (A^{j-1} \otimes I) + \cdots + \alpha_1 (A \otimes I) + \alpha_0 (A^0 \otimes I) \\
&= (\alpha_j A^j + \alpha_{j-1} A^{j-1} + \cdots + \alpha_1 A + \alpha_0 I) \otimes I = p(A) \otimes I. \text{ Similarly, } p(I \otimes B) \\
&= \alpha_j (I \otimes B)^j + \alpha_{j-1} (I \otimes B)^{j-1} + \cdots + \alpha_1 (I \otimes B) + \alpha_0 (I \otimes B)^0 \\
&= \alpha_j (I \otimes B^j) + \alpha_{j-1} (I \otimes B^{j-1}) + \cdots + \alpha_1 (I \otimes B) + \alpha_0 (I \otimes B^0) \\
&= I \otimes (\alpha_j B^j + \alpha_{j-1} B^{j-1} + \cdots + \alpha_1 B + \alpha_0 I) = I \otimes p(B). \quad \square
\end{aligned}$$

The following corollary presents the inverse of a Kronecker product with respect to conventional multiplication. All references to an inverse in this paper are with respect to the conventional matrix multiplication.

**Corollary 10:** If  $A \in M_m$  and  $B \in M_n$  are nonsingular, then  $A \otimes B$  is also nonsingular

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

*Proof:* The following results from Theorem 7.

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

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

This implies that  $A^{-1} \otimes B^{-1}$  is the unique inverse of  $A \otimes B$  under conventional matrix multiplication. Therefore,  $A \otimes B$  is nonsingular.  $\square$

This result is completed in Corollary 18 where the converse is proved.

**Corollary 11:** If  $A \in M_n$  is similar to  $B \in M_n$  via  $S \in M_n$  and  $C \in M_m$  is similar to

$E \in M_m$  via  $T \in M_m$ , then  $A \otimes C$  is similar to  $B \otimes E$  via  $S \otimes T$ .

*Proof:* Assume that  $A = S^{-1}BS$  and  $C = T^{-1}ET$ . By the mixed product property,

$$(S \otimes T)^{-1}(B \otimes E)(S \otimes T) = (S^{-1} \otimes T^{-1})(BS \otimes ET) = (S^{-1}BS) \otimes (T^{-1}ET) = A \otimes C. \quad \square$$

**Corollary 12:** Let  $Q \in M_m$  and  $R \in M_n$  be orthogonal matrices. Then,  $Q \otimes R$  is an orthogonal matrix.

*Proof:*  $Q$  and  $R$  are orthogonal implies that  $QQ^T = I$  and  $RR^T = I$ . Using Theorem 7,

$$(Q \otimes R)(Q \otimes R)^T = (Q \otimes R)(Q^T \otimes R^T) = (QQ^T) \otimes (RR^T) = I \otimes I = I. \text{ Therefore, } Q \otimes R$$

is orthogonal.  $\square$

**Corollary 13:** Let  $U \in M_n$  and  $V \in M_m$  be unitary matrices. Then,  $U \otimes V$  is a unitary matrix.

*Proof:*  $U$  and  $V$  are unitary implies  $U^{-1} = U^*$  and  $V^{-1} = V^*$ . Using Corollary 8,

$$(U \otimes V)^{-1} = U^{-1} \otimes V^{-1} = U^* \otimes V^* = (U \otimes V)^*. \text{ Therefore, } U \otimes V \text{ is a unitary matrix. } \quad \square$$

The relationship between the eigenvalues of two matrices and the eigenvalues of the resulting Kronecker product is interesting and will be useful in proving the converse of Corollary 10 as well as determining the eigenvalues of some linear transformations. It will also be helpful in the discussion of the determinant and trace of a Kronecker product. But first, a lemma is needed to show that the Kronecker product of two upper triangular matrices is also upper triangular. This lemma can be duplicated to show the same property for lower triangular matrices.

**Lemma 14:** Let  $A \in M_n$  and  $B \in M_m$  be upper triangular. Then,  $A \otimes B$  is upper triangular.

**Proof:**  $A$  and  $B$  are upper triangular implies that  $A = [a_{i,j}]$  where  $a_{i,j} = 0$  for  $i > j$  and

$$B = [b_{p,q}] \text{ where } b_{p,q} = 0 \text{ for } p > q. \text{ By definition, } A \otimes B = \begin{pmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{n1}B & \cdots & a_{nn}B \end{pmatrix}$$

$$= \begin{pmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ 0 & \cdots & a_{nn}B \end{pmatrix}. \text{ All block matrices below the diagonal of block matrices will be}$$

zero since  $a_{i,j} = 0$  for  $i > j$ . Now examine the block matrices on the diagonal, or  $a_{ii}B$ .

These matrices are all upper triangular since  $B$  is upper triangular. Hence,  $A \otimes B$  is upper triangular. □

**Theorem 15:** Let  $A \in M_n$  and  $B \in M_m$ . If  $\lambda$  is an eigenvalue of  $A$  with corresponding eigenvector  $x \in \mathbb{F}^n$  and if  $\mu$  is an eigenvalue of  $B$  with corresponding eigenvector  $y \in \mathbb{F}^m$ , then  $\lambda\mu$  is an eigenvalue of  $A \otimes B$  with corresponding eigenvector  $x \otimes y \in \mathbb{F}^{mn}$ .

The set of eigenvalues of  $A \otimes B$  is  $\{\lambda_i \mu_j : i = 1, 2, \dots, n \text{ and } j = 1, 2, \dots, m\}$  where the set of



eigenvalues of  $A$  is  $\{\lambda_i : i = 1, 2, \dots, n\}$  and the set of eigenvalues of  $B$  is

$\{\mu_j : j = 1, 2, \dots, m\}$  (including algebraic multiplicities in all three cases). In particular,

the set of eigenvalues of  $A \otimes B$  is the same as the set of eigenvalues of  $B \otimes A$ .

*Proof:* Consider  $Ax = \lambda x$  and  $By = \mu y$  for  $x, y \neq \vec{0}$ . It follows from Theorem 7 that

$(A \otimes B)(x \otimes y) = (Ax) \otimes (By) = (\lambda x) \otimes (\mu y) = \lambda \mu (x \otimes y)$ . This shows that  $\lambda \mu$  is an

eigenvalue of  $A \otimes B$  with corresponding eigenvector  $x \otimes y$ . By the triangularization

theorem, there exist unitary matrices  $U \in M_n$  and  $V \in M_m$  such that  $U^*AU = \Delta_A$  and

$V^*BV = \Delta_B$  where  $\Delta_A$  and  $\Delta_B$  are upper triangular matrices. Consequently, by

Corollary 8,  $(U \otimes V)^*(A \otimes B)(U \otimes V) = (U^*AU) \otimes (V^*BV) = \Delta_A \otimes \Delta_B$ . From this, it

follows that  $\Delta_A \otimes \Delta_B$  is an upper triangular matrix (by Lemma 14) that is similar to

$A \otimes B$ . The eigenvalues of  $A$ ,  $B$ , and  $A \otimes B$  are the main diagonal entries of the upper

triangular matrices to which they are similar ( $\Delta_A$ ,  $\Delta_B$ , and  $\Delta_A \otimes \Delta_B$ , respectively). Since

$\Delta_A$  and  $\Delta_B$  are square matrices, it follows from the definition of the Kronecker product

that the entries of the main diagonal of  $\Delta_A \otimes \Delta_B$  are the  $nm$  pairwise products of the

entries on the main diagonals of  $\Delta_A$  and  $\Delta_B$ . Therefore, the eigenvalues of  $A \otimes B$  are

the  $nm$  pairwise products of the eigenvalues of  $A$  and  $B$ . Since the eigenvalues of

$B \otimes A$  are the  $mn$  pairwise products of the eigenvalues of  $B$  and  $A$ , they will be the same

as the eigenvalues of  $A \otimes B$ . □

**Corollary 16:** Let  $\hat{\lambda}_1, \hat{\lambda}_2, \dots, \hat{\lambda}_p$  be the distinct eigenvalues of  $A \in M_n$  with respective

algebraic multiplicities  $s_1, s_2, \dots, s_p$ , and let  $\hat{\mu}_1, \hat{\mu}_2, \dots, \hat{\mu}_q$  be the distinct eigenvalues of

$B \in M_m$  with respective algebraic multiplicities  $t_1, t_2, \dots, t_q$ . If  $\tau$  is an eigenvalue of

$A \otimes B$ , then the algebraic multiplicity of  $\tau$  is  $\sum_{\hat{\lambda}_i \hat{\mu}_j = \tau} s_i t_j$  where  $\hat{\lambda}_i \hat{\mu}_j = \tau$  includes all

possible ways that  $\tau$  arises as a product an eigenvalue of  $A$  and an eigenvalue of  $B$ .

**Proof**: Theorem 15 shows that the set of eigenvalues of  $A \otimes B$  is

$$\sigma(A \otimes B) = \{ \lambda_i \mu_j : i = 1, 2, \dots, n \text{ and } j = 1, 2, \dots, n \text{ (including multiplicities)} \}.$$

$\tau \in \sigma(A \otimes B)$ , then the algebraic multiplicity of  $\tau$  is the number of times a pairwise product  $\lambda_i \mu_j = \tau$  for  $i = 1, 2, \dots, n$  and  $j = 1, 2, \dots, n$ . If there exist integers  $k$  and  $l$  such that

$\hat{\lambda}_k \hat{\mu}_l = \tau$ , then this product will occur exactly  $s_k t_l$  times in  $\sigma(A \otimes B)$ . As a result, the

algebraic multiplicity of  $\tau$  is exactly  $\sum_{\hat{\lambda}_k \hat{\mu}_l = \tau} s_k t_l$ . □

**Corollary 17**: Let  $A \in M_n$  and  $B \in M_m$  be positive (semi) definite Hermitian matrices.

Then,  $A \otimes B$  is also positive (semi) definite Hermitian.

**Proof**: Since  $A$  and  $B$  are Hermitian matrices,  $A = A^*$  and  $B = B^*$ . Then,

$(A \otimes B)^* = A^* \otimes B^* = A \otimes B$ . Therefore,  $A \otimes B$  is Hermitian. Since  $A$  and  $B$  are positive (semi) definite, the eigenvalues of both  $A$  and  $B$  are all positive (non negative). The product of any eigenvalue from  $A$  with any eigenvalue from  $B$  will also be positive (non negative). Since the eigenvalues of  $A \otimes B$  are products of the eigenvalues of  $A$  and the eigenvalues of  $B$ , all of the eigenvalues of  $A \otimes B$  are positive (non negative). This makes  $A \otimes B$  positive (semi) definite. □

**Corollary 18**: Let  $A \in M_n$  and  $B \in M_m$ . Then,

$\det(A \otimes B) = (\det A)^m (\det B)^n = \det(B \otimes A)$ . Thus,  $A \otimes B$  (and consequently  $B \otimes A$ ) is nonsingular if and only if both  $A$  and  $B$  are nonsingular.

Proof: Since the determinant of a matrix is the product of the eigenvalues of the matrix,

$\det(A \otimes B) = \prod_{i=1}^{mn} \lambda_i$  where  $\lambda_i$  are the eigenvalues of  $A \otimes B$ . By Theorem 15, each

$\lambda_i = \alpha_j \beta_k$  where  $\alpha_j$  is an eigenvalue of  $A$  and  $\beta_k$  is an eigenvalue of  $B$ . As a result,

$$\det(A \otimes B) = \prod_{i=1}^{mn} \lambda_i = \prod_{j=1}^n \prod_{k=1}^m (\alpha_j \beta_k) = \left( \prod_{j=1}^n \alpha_j^m \right) \left( \prod_{k=1}^m \beta_k^n \right) = (\det A)^m (\det B)^n. \text{ Since the}$$

eigenvalues of  $A \otimes B$  and  $B \otimes A$  are the same,  $\det(A \otimes B) = \det(B \otimes A)$ . Now,  $A \otimes B$  is nonsingular if and only if  $\det(A \otimes B)$  is not zero. This happens if and only if

$(\det A)^m (\det B)^n$  is not equal to zero. By the zero product property, this is true if and only if  $\det A$  and  $\det B$  are both nonzero. Finally,  $\det A$  and  $\det B$  are both nonzero if and only if  $A$  and  $B$  are both nonsingular.  $\square$

**Corollary 19:** If  $A \in M_n$  and  $B \in M_m$  then,

$$\text{trace}(A \otimes B) = \text{trace}(A)\text{trace}(B) = \text{trace}(B \otimes A).$$

Proof: The trace of a matrix is equal to the sum of the eigenvalues of the matrix. If the eigenvalues of  $A$  are  $\alpha_i$  for  $i = 1, 2, \dots, n$  and the eigenvalues of  $B$  are  $\beta_j$  for

$$j = 1, 2, \dots, m, \text{ then } \text{trace}(A \otimes B) = \sum_{i=1}^n \sum_{j=1}^m (\alpha_i \beta_j) = \sum_{i=1}^n \alpha_i \sum_{j=1}^m \beta_j = \text{trace}(A)\text{trace}(B).$$

Consequently,  $\text{trace}(A \otimes B) = \text{trace}(A)\text{trace}(B) = \text{trace}(B)\text{trace}(A) = \text{trace}(B \otimes A)$ .  $\square$

Theorem 20 renders information regarding the singular values and the rank of the Kronecker product of two matrices. The relation between the singular values of two matrices and their Kronecker product is similar to the one for eigenvalues that was proven in Theorem 15. Theorem 20 will be useful when proving that  $\det(A \otimes B) = \det(B \otimes A)$  for certain non square matrices  $A$  and  $B$ . The rank of a

Kronecker product will be used to help develop properties about the Jordan - Canonical form of the Kronecker product of two matrices.

**Theorem 20:** Let  $A \in M_{m,n}$  and  $B \in M_{p,q}$  have singular value decompositions

$A = U_A \Sigma_A V_A^*$  and  $B = U_B \Sigma_B V_B^*$ . If  $\text{rank}(A) = r_1$  and  $\text{rank}(B) = r_2$  then, the nonzero singular values of  $A \otimes B$  are the  $r_1 r_2$  positive numbers

$\{\sigma_i(A) \sigma_j(B) : i = 1, 2, \dots, r_1 \text{ and } j = 1, 2, \dots, r_2\}$ , including multiplicities, where  $\sigma_i(A)$  is a nonzero singular value of  $A$  and  $\sigma_j(B)$  is a nonzero singular value of  $B$ . Zero is a singular value of  $A \otimes B$  with multiplicity  $\min\{mp, nq\} - r_1 r_2$ . In particular, the singular values of  $A \otimes B$  are the same as those of  $B \otimes A$  and  $\text{rank}(A \otimes B) = \text{rank}(B \otimes A) = r_1 r_2$ .

**Proof:** Since the singular value decompositions of  $A$  and  $B$  are  $A = U_A \Sigma_A V_A^*$  and

$B = U_B \Sigma_B V_B^*$  respectively, the matrices  $U_A \in M_m$ ,  $V_A \in M_n$ ,  $U_B \in M_p$ , and  $V_B \in M_q$  are unitary. Also,  $\Sigma_A \in M_{m,n}$  and  $\Sigma_B \in M_{p,q}$  are matrices where the elements

$\sigma_{ii}$  ( $1 \leq i \leq \text{rank}$ ) are the nonzero singular values of  $A$  and  $B$  respectively and all other elements of the matrix are zero. Using Corollary 8,

$A \otimes B = (U_A \Sigma_A V_A^*) \otimes (U_B \Sigma_B V_B^*) = (U_A \otimes U_B)(\Sigma_A \otimes \Sigma_B)(V_A \otimes V_B)^*$  where  $U_A \otimes U_B \in M_{mp}$

and  $V_A \otimes V_B \in M_{nq}$  are unitary by Corollary 13. The singular values of  $A \otimes B$  are the

square roots of the eigenvalues of  $(A \otimes B)^*(A \otimes B) = (A^* A) \otimes (B^* B)$ . By Theorem 15,

the eigenvalues of  $(A \otimes B)^*(A \otimes B)$  are the pairwise products of the eigenvalues of

$A^* A$  and  $B^* B$ . Since the singular values of  $A$  are the square roots of the eigenvalues of

$A^*A$  and the singular values of  $B$  are the square roots of the eigenvalues of  $B^*B$ , the singular values of  $A \otimes B$  are the pairwise products of the singular values of  $A$  and  $B$ . Also, since the rank of a matrix is equal to the number of nonzero singular values of the matrix, the number of nonzero singular values for  $A$  and  $B$  are  $r_1$  and  $r_2$  respectively. Therefore, the number of nonzero singular values of  $A \otimes B$  is  $r_1 r_2$  which means the rank of  $A \otimes B$  is  $r_1 r_2$ . Because the remaining singular values are zero, the number of zero singular values of  $A \otimes B$  is  $\min\{mp, nq\} - r_1 r_2$ . The singular values of  $B \otimes A$  are the pairwise products of the singular values of  $B$  and  $A$ . Therefore, the singular values of  $A \otimes B$  and  $B \otimes A$  are the same and  $\text{rank}(A \otimes B) = \text{rank}(B \otimes A) = r_1 r_2$ .  $\square$

**Corollary 21:** Let  $A \in M_{m,n}$  and  $B \in M_{p,q}$  be given with  $mp = nq$ . ( $A$  and  $B$  are not square matrices but, their resulting Kronecker product is.) Then,  
 $\det(A \otimes B) = \det(B \otimes A)$ .

*Proof:* Let  $A$  have singular value decomposition  $A = U_A \Sigma_A V_A^*$  and let  $B$  have singular value decomposition  $B = U_B \Sigma_B V_B^*$ . Thus,  $U_A \in M_m$ ,  $V_A \in M_n$ ,  $U_B \in M_p$ , and  $V_B \in M_q$  are unitary matrices. Also,  $\Sigma_A \in M_{m,n}$  and  $\Sigma_B \in M_{p,q}$  are diagonal matrices with the singular values of  $A$  and  $B$ , respectively, as the diagonal entries. As a result,  
 $A \otimes B = (U_A \Sigma_A V_A^*) \otimes (U_B \Sigma_B V_B^*) = (U_A \otimes U_B)(\Sigma_A \otimes \Sigma_B)(V_A \otimes V_B)^*$  and  
 $B \otimes A = (U_B \Sigma_B V_B^*) \otimes (U_A \Sigma_A V_A^*) = (U_B \otimes U_A)(\Sigma_B \otimes \Sigma_A)(V_B \otimes V_A)^*$  where all Kronecker product matrices, including  $\Sigma_A \otimes \Sigma_B, \Sigma_B \otimes \Sigma_A \in M_{mp, nq}$ , are square matrices of the same size, since  $mp = nq$ . Accordingly, all of their determinants are defined. Thus,

$$\det(A \otimes B) = \det(U_A \otimes U_B) \det(\Sigma_A \otimes \Sigma_B) \overline{\det(V_A \otimes V_B)} \text{ and}$$

$$\det(B \otimes A) = \det(U_B \otimes U_A) \det(\Sigma_B \otimes \Sigma_A) \overline{\det(V_B \otimes V_A)}. \text{ Since } U_A, U_B, V_A, \text{ and } V_B \text{ are}$$

square matrices, Corollary 18 implies that  $\det(U_A \otimes U_B) = \det(U_B \otimes U_A)$  and

$$\overline{\det(V_A \otimes V_B)} = \overline{\det(V_B \otimes V_A)}. \text{ Since } \Sigma_A \otimes \Sigma_B \text{ and } \Sigma_B \otimes \Sigma_A \text{ are diagonal matrices, their}$$

determinants are the products of their respective diagonal entries. By Theorem 20, their

diagonal entries, which are the singular values of  $A \otimes B$  and  $B \otimes A$ , are the same. Hence,

$$\det(\Sigma_A \otimes \Sigma_B) = \det(\Sigma_B \otimes \Sigma_A). \text{ Therefore, } \det(A \otimes B) = \det(B \otimes A). \quad \square$$

Corollary 21 showed that even if two matrices  $A$  and  $B$  are not square, the determinants of the Kronecker products  $A \otimes B$  and  $B \otimes A$  are still equal. However, this is not true for the trace of the Kronecker products. An example is shown below.

**Example 22:** Let  $A = \begin{bmatrix} 4 & -3 \end{bmatrix}$  and  $B = \begin{bmatrix} 2 & 6 \\ -7 & 0 \\ 5 & 9 \\ 1 & -1 \end{bmatrix}$ . As a result,

$$A \otimes B = \begin{bmatrix} 8 & 24 & -6 & -18 \\ -28 & 0 & 21 & 0 \\ 20 & 36 & -15 & -27 \\ 4 & -4 & -3 & 3 \end{bmatrix} \text{ and } B \otimes A = \begin{bmatrix} 8 & -6 & 24 & -18 \\ -28 & 21 & 0 & 0 \\ 20 & -15 & 36 & -27 \\ 4 & -3 & -4 & 3 \end{bmatrix}. \text{ Therefore,}$$

$\text{trace}(A \otimes B) = -4$  and  $\text{trace}(B \otimes A) = 68$ . Corollary 19 stated that if  $A$  and  $B$  are square matrices, the traces are equal. However, if they are not square, such as the  $A$  and  $B$  matrices in Example 22, the trace of  $A \otimes B$  is not necessarily equal to the trace of  $B \otimes A$ .

Corollary 13 stated that if  $U$  and  $V$  are unitary matrices, then  $U \otimes V$  is a unitary matrix. The following result explains when the converse is true as well.

**Theorem 23:** Let  $U = [u_{ij}] \in M_n$  and  $V = [v_{ij}] \in M_m$  be given matrices. Then,

$U \otimes V \in M_{mn}$  is unitary if and only if there is some  $r > 0$  such that  $rU$  and  $r^{-1}V$  are both unitary. Also,  $U \otimes V$  is unitary if and only if  $V \otimes U$  is unitary.

*Proof:* First, assume there exists an  $r > 0$  such that  $rU$  and  $r^{-1}V$  are both unitary matrices. By Corollary 13, this implies that  $(rU) \otimes (r^{-1}V) = (rr^{-1}U) \otimes V = U \otimes V$  is unitary. Now, assume that  $U \otimes V$  is given to be a unitary matrix. Hence,

$I_{mn} = (U \otimes V)^* (U \otimes V) = (U^*U) \otimes (V^*V)$ . Let  $W = U^*U = [w_{ij}]$ . Accordingly,

$$I_{mn} = W \otimes (V^*V) \text{ is equivalent to } \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix} = \begin{pmatrix} w_{11}V^*V & \cdots & w_{1n}V^*V \\ \vdots & \ddots & \vdots \\ w_{n1}V^*V & \cdots & w_{nn}V^*V \end{pmatrix} \text{ which implies}$$

that  $w_{ii}V^*V = I_m$  for all  $i$  and  $w_{ij}V^*V = 0$  for all  $i \neq j$ . This shows  $w_{ij} = 0$  for all  $i \neq j$

and  $w_{ii}$  is a nonzero constant for  $i = 1, 2, \dots, n$ . If  $w_{ii} = \alpha$ , then  $\alpha V^*V = I_m$  and

$U^*U = \alpha I_n$ . Also,  $U \otimes V$  being unitary gives  $I_{mn} = (U \otimes V)(U \otimes V)^* = (UU^*) \otimes (VV^*)$ .

Let  $B = UU^*$ . As before, this gives that  $b_{ii}VV^* = I_m$  for all  $i$  and  $b_{ij}VV^* = 0$  for all  $i \neq j$ .

Consequently,  $b_{ij} = 0$  for all  $i \neq j$  and  $b_{ii}$  is a nonzero constant for  $i = 1, 2, \dots, n$ . If

$b_{ii} = \beta$ , then  $\beta VV^* = I_m$  and  $UU^* = \beta I_n$ . Since  $U^*U = \alpha I_n$ , it follows that  $U$  is

invertible. If  $Q = \frac{1}{\alpha}U^*$ , then,  $QU = I_n$  implies that  $UQ = I_n$ . From this, it follows that

$U\left(\frac{1}{\alpha}U^*\right) = I_n$ . Hence,  $UU^* = \alpha I_n = \beta I_n$  and this implies that  $\alpha = \beta$ . Additionally,

$U^*U = \alpha I_n$  shows that  $\alpha = |u_{1k}|^2 + |u_{2k}|^2 + |u_{3k}|^2 + \cdots + |u_{nk}|^2$  for  $k = 1, 2, \dots, n$ . This implies

that  $\alpha \in \mathbb{R}$  and  $\alpha > 0$ . If  $r = \frac{1}{\sqrt{\alpha}}$  (which means that  $r \in \mathbb{R}$ ,  $r > 0$ , and  $\bar{r} = r$ ), then

$$(rU)^*(rU) = (\bar{r}U^*)(rU) = \frac{1}{\sqrt{\alpha}} \frac{1}{\sqrt{\alpha}} U^*U = \frac{1}{\alpha} U^*U = I_n \text{ and}$$

$$(rU)(rU)^* = (rU)(\bar{r}U^*) = \frac{1}{\sqrt{\alpha}} \frac{1}{\sqrt{\alpha}} UU^* = \frac{1}{\alpha} UU^* = \frac{1}{\beta} UU^* = I_n. \text{ The two preceding}$$

equations show that  $rU$  is unitary. Similarly,

$$(r^{-1}V)^*(r^{-1}V) = \left(\frac{1}{\bar{r}}V^*\right)\left(\frac{1}{r}V\right) = \sqrt{\alpha}\sqrt{\alpha}V^*V = \alpha V^*V = I_m \text{ and}$$

$$(r^{-1}V)(r^{-1}V)^* = \left(\frac{1}{r}V\right)\left(\frac{1}{\bar{r}}V^*\right) = \sqrt{\alpha}\sqrt{\alpha}VV^* = \alpha VV^* = \beta VV^* = I_m \text{ show that } r^{-1}V \text{ is}$$

unitary. Furthermore,  $V \otimes U$  is unitary because  $V \otimes U = (r^{-1}V) \otimes (rU)$  is unitary.  $\square$

As mentioned earlier, the Kronecker product is not commutative in general, even though the product creates a matrix that is the same size either way the two matrices are multiplied. The next theorem shows that requiring a matrix  $B$  to commute under the Kronecker product with a matrix  $A$  is quite limiting on  $B$ .

**Theorem 24:** Let  $A, B \in M_{m,n}$  be given. Then,  $A \otimes B = B \otimes A$  if and only if either  $A = cB$  or  $B = cA$  for some  $c \in \mathbb{F}$ .

**Proof:** First, assume that  $A = cB$  or  $B = cA$  for some  $c \in \mathbb{F}$ . If  $c$  is zero, then either  $A$  or  $B$  is the zero matrix which means that  $A \otimes B = 0 = B \otimes A$ . If  $c$  is nonzero, it suffices to consider just  $A = cB$  because  $B = c^{-1}A$ . If  $A = cB$ , then

$$A \otimes B = (cB) \otimes B = c(B \otimes B) = B \otimes (cB) = B \otimes A.$$

Now, assume that  $A \otimes B = B \otimes A$  with  $A, B \in M_{m,n}$ . By the definition of the Kronecker product, this implies that  $a_{ij}B = b_{ij}A$ . If  $a_{ij} = 0$  for all  $i = 1, 2, \dots, m$  and



$j = 1, 2, \dots, n$ , then  $A$  is the zero matrix and  $A = cB$  for  $c = 0$ . For each nonzero  $a_{ij}$ ,

$$B = \frac{b_{ij}}{a_{ij}} A \text{ or } B = cA \text{ where } c = \frac{b_{ij}}{a_{ij}}. \text{ Notice } \frac{b_{ij}}{a_{ij}} \text{ must be the same for all } i \text{ and } j \text{ whenever}$$

$$a_{ij} \neq 0.$$

□

Consideration needs to be given for the situation of  $A$  and  $B$  having different dimensions. For example, let  $A \in \mathbb{C}^2$  and  $B \in \mathbb{C}^3$ . If either of  $A$  or  $B$  are zero, then  $A \otimes B = 0 = B \otimes A$ . So, consider the case where both  $A$  and  $B$  are nonzero. Then,

$$A \otimes B = \begin{bmatrix} a_1 b_1 \\ a_1 b_2 \\ a_1 b_3 \\ a_2 b_1 \\ a_2 b_2 \\ a_2 b_3 \end{bmatrix} \text{ and } B \otimes A = \begin{bmatrix} b_1 a_1 \\ b_1 a_2 \\ b_2 a_1 \\ b_2 a_2 \\ b_3 a_1 \\ b_3 a_2 \end{bmatrix}. \text{ Setting these two equal produces the equations}$$

$a_1 b_2 = b_1 a_2$ ,  $a_1 b_3 = b_2 a_1$ ,  $a_2 b_1 = b_2 a_2$ , and  $a_2 b_2 = b_3 a_1$ . If  $a_1 \neq a_2$ , this would give the solution of  $B = 0$ . However,  $B$  is nonzero. So,  $a_1$  must be equal to  $a_2$ . Then, the set of four equations becomes  $b_2 = b_1$ ,  $b_3 = b_2$ ,  $b_1 = b_2$ , and  $b_2 = b_3$ . Hence, in order for these

two Kronecker products to be equal,  $A = \alpha \begin{bmatrix} 1 \\ 1 \end{bmatrix}$  and  $B = \beta \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ . Expanding this example

to  $A \in M_{m,n}$  and  $B \in M_{p,q}$  would show that  $A$  needs to be a multiple of

$$\begin{bmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{bmatrix} \in M_{m,n}, \text{ a matrix of all 1s. Similarly, } B \text{ would be equal to}$$

$$\beta \begin{bmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{bmatrix} \in M_{p,q}. \text{ Then,}$$

$$\begin{aligned} A \otimes B &= \left( \alpha \begin{bmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{bmatrix} \right) \otimes \left( \beta \begin{bmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{bmatrix} \right) = \alpha \beta \left( \begin{bmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{bmatrix} \otimes \begin{bmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{bmatrix} \right) \\ &= \alpha \beta \left( \begin{bmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{bmatrix} \right) \text{ where the resulting matrix is an } mp \times nq \text{ matrix of 1s. Writing out} \end{aligned}$$

$B \otimes A$  would produce the same result.

The next result concerns the normality of the Kronecker product of square matrices. Its proof is facilitated by the following lemma that establishes

$$\sigma(AB) = \sigma(BA).$$

**Lemma 25:** If  $A$  and  $B$  are  $n \times n$  matrices, then  $\sigma(AB) = \sigma(BA)$ .

$$\text{Proof: } \begin{bmatrix} AB & 0 \\ B & 0 \end{bmatrix} \begin{bmatrix} I & A \\ 0 & I \end{bmatrix} = \begin{bmatrix} AB & ABA \\ B & BA \end{bmatrix} \in M_{2n} \text{ and}$$

$$\begin{bmatrix} I & A \\ 0 & I \end{bmatrix} \begin{bmatrix} 0 & 0 \\ B & BA \end{bmatrix} = \begin{bmatrix} AB & ABA \\ B & BA \end{bmatrix} \in M_{2n}. \text{ Also } \begin{bmatrix} I & -A \\ 0 & I \end{bmatrix} \begin{bmatrix} I & A \\ 0 & I \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \in M_{2n} \text{ which}$$

$$\text{implies that } \begin{bmatrix} I & A \\ 0 & I \end{bmatrix}^{-1} = \begin{bmatrix} I & -A \\ 0 & I \end{bmatrix}. \text{ It now follows that } \begin{bmatrix} I & A \\ 0 & I \end{bmatrix}^{-1} \begin{bmatrix} AB & 0 \\ B & 0 \end{bmatrix} \begin{bmatrix} I & A \\ 0 & I \end{bmatrix}$$

$$= \begin{bmatrix} I & -A \\ 0 & I \end{bmatrix} \begin{bmatrix} AB & ABA \\ B & BA \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ B & BA \end{bmatrix}. \text{ That is, } C = \begin{bmatrix} AB & 0 \\ B & 0 \end{bmatrix} \text{ and } D = \begin{bmatrix} 0 & 0 \\ B & BA \end{bmatrix} \text{ are}$$

similar and  $\sigma(C) = \sigma(D)$ . Since the eigenvalues of  $C$  are the eigenvalues of  $AB$  together with  $n$  zeros and the eigenvalues of  $D$  are the eigenvalues of  $BA$  together with  $n$  zeros, the proof is complete.  $\square$

**Theorem 26:** Let  $A \in M_n$  and  $B \in M_m$  be nonzero matrices. Then,  $A \otimes B$  is normal if and only if both  $A$  and  $B$  are normal. Consequently,  $A \otimes B$  is normal if and only if  $B \otimes A$  is normal.

*Proof:* First, assume that both  $A \in M_n$  and  $B \in M_m$  are nonzero normal matrices. As a result,  $(A \otimes B)(A \otimes B)^* = (AA^*) \otimes (BB^*) = (A^*A) \otimes (B^*B) = (A \otimes B)^* (A \otimes B)$ .

Therefore,  $A \otimes B$  is normal. Now, assume that  $A$  and  $B$  are nonzero matrices and that  $A \otimes B$  is normal. From this it follows that  $(A \otimes B)(A \otimes B)^* = (A \otimes B)^* (A \otimes B)$ .

Consequently,  $(AA^*) \otimes (BB^*) = (A^*A) \otimes (B^*B)$ . Looking at the  $ii^{th}$  block gives

$$\left( \sum_{j=1}^n a_{ij} \bar{a}_{ij} \right) (BB^*) = \left( \sum_{j=1}^n \bar{a}_{ji} a_{ji} \right) (B^*B). \text{ This implies that}$$

$$\left( \sum_{j=1}^n |a_{ij}|^2 \right) (BB^*) = \left( \sum_{j=1}^n |a_{ji}|^2 \right) (B^*B). \text{ Since } A \text{ is nonzero, the matrix will have at least one}$$

element  $(a_{ki})$  that is nonzero. As a result, there exists positive real numbers  $c$  and  $d$  such

that  $cBB^* = dB^*B$ . The eigenvalues of  $BB^*$  and  $B^*B$  are nonnegative due to  $BB^*$  and  $B^*B$  being positive semidefinite. Also, since  $c > 0$  and  $d > 0$ , it follows that the largest eigenvalue of  $cBB^*$  is the product of  $c$  and the largest eigenvalue of  $BB^*$  and the largest eigenvalue of  $dB^*B$  is the product of  $d$  and the largest eigenvalue of  $B^*B$ .

Now, Lemma 25 tells us that  $\sigma(BB^*) = \sigma(B^*B)$ . Hence the largest eigenvalue of

$\sigma(BB^*)$  is the same as the largest eigenvalue of  $\sigma(B^*B)$ . Since  $cBB^* = dB^*B$ ,

$\sigma(cBB^*) = \sigma(dB^*B)$ . These last few observations imply that  $c = d$ . This gives that

$BB^* = B^*B$  and therefore  $B$  is normal. Now, let  $a_{ij}$  be the  $ij^{th}$  element of the matrix  $AA^*$

and  $\gamma_{ij}$  be the  $ij^{th}$  element of the matrix  $A^*A$ . Since it was shown

$$(AA^*) \otimes (BB^*) = (A^*A) \otimes (B^*B), \text{ this implies } \alpha_{ij}BB^* = \gamma_{ij}B^*B. \text{ Considering that } B \text{ is a}$$

normal matrix, this proves that  $\alpha_{ij} = \gamma_{ij}$  for all  $i = 1, 2, \dots, n$  and  $j = 1, 2, \dots, n$ . Thus,

$$AA^* = A^*A \text{ and } A \text{ is a normal matrix.} \quad \square$$

The next few results complete a collection of basic facts about the Kronecker product.

**Proposition 27:** Let  $A \in M_n$  and  $B \in M_m$ . Then,  $A \otimes B = I_{mn}$  if and only if  $A = \alpha I_n$  and

$$B = \frac{1}{\alpha} I_m.$$

Proof: First, assume that  $A = \alpha I_n$  and  $B = \frac{1}{\alpha} I_m$ . Consequently,

$$A \otimes B = (\alpha I_n) \otimes \left( \frac{1}{\alpha} I_m \right) = \left( \alpha \frac{1}{\alpha} \right) (I_n \otimes I_m) = I_{mn}. \text{ Now, assume that } A \otimes B = I_{mn}. \text{ Then,}$$

$$\begin{pmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{pmatrix} = \begin{pmatrix} a_{11}B & \dots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{n1}B & \dots & a_{nn}B \end{pmatrix}. \text{ This implies that } a_{ii}B = I_m \text{ for all } i = 1, 2, \dots, n \text{ and}$$

$a_{ij}B = 0$  for all  $i \neq j$ . This shows that  $a_{ij} = 0$  for all  $i \neq j$  and  $B = \frac{1}{a_{ii}} I$ . If  $\alpha = a_{ii}$ , then

$$A = \alpha I_n \text{ and } B = \frac{1}{\alpha} I_m. \quad \square$$

**Proposition 28:** Let  $A \in M_n$ ,  $B \in M_m$ , and  $C \in M_p$ . Then,

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

Proof:

$$(A \oplus B) \otimes C = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \otimes \begin{pmatrix} A \otimes C & 0 \otimes C \\ 0 \otimes C & B \otimes C \end{pmatrix} = \begin{pmatrix} A \otimes C & 0 \\ 0 & B \otimes C \end{pmatrix} = (A \otimes C) \oplus (B \otimes C) \quad \square$$

However, the distributive property in Proposition 28 does not hold true when the Kronecker product is to be distributed on the left side of the direct sum. In other words,  $A \otimes (B \oplus C) \neq (A \otimes B) \oplus (A \otimes C)$ . This is shown in Example 29.

**Example 29:** Let  $A = \begin{pmatrix} -2 & 1 \\ 4 & 0 \end{pmatrix}$ ,  $B = \begin{pmatrix} 3 & 7 \\ -1 & -5 \end{pmatrix}$ , and  $C = \begin{pmatrix} 9 & -6 \\ 1 & 8 \end{pmatrix}$ . Then,

$$A \otimes (B \oplus C) = \begin{bmatrix} -2 & 1 \\ 4 & 0 \end{bmatrix} \otimes \begin{bmatrix} 3 & 7 & 0 & 0 \\ -1 & -5 & 0 & 0 \\ 0 & 0 & 9 & -6 \\ 0 & 0 & 1 & 8 \end{bmatrix} = \begin{bmatrix} -6 & -14 & 0 & 0 & 3 & 7 & 0 & 0 \\ 2 & 10 & 0 & 0 & -1 & -5 & 0 & 0 \\ 0 & 0 & -18 & 12 & 0 & 0 & 9 & -6 \\ 0 & 0 & -2 & -16 & 0 & 0 & 1 & 8 \\ 12 & 28 & 0 & 0 & 0 & 0 & 0 & 0 \\ -4 & -20 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 36 & -24 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & 32 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\text{but } (A \otimes B) \oplus (A \otimes C) = \begin{bmatrix} -6 & -14 & 3 & 7 \\ 2 & 10 & -1 & -5 \\ 12 & 28 & 0 & 0 \\ -4 & -20 & 0 & 0 \end{bmatrix} \oplus \begin{bmatrix} -18 & 12 & 9 & -6 \\ -2 & -16 & 1 & 8 \\ 36 & -24 & 0 & 0 \\ 4 & 32 & 0 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} -6 & -14 & 3 & 7 & 0 & 0 & 0 & 0 \\ 2 & 10 & -1 & -5 & 0 & 0 & 0 & 0 \\ 12 & 28 & 0 & 0 & 0 & 0 & 0 & 0 \\ -4 & -20 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -18 & 12 & 9 & -6 \\ 0 & 0 & 0 & 0 & -2 & -16 & 1 & 8 \\ 0 & 0 & 0 & 0 & 36 & -24 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4 & 32 & 0 & 0 \end{bmatrix}.$$

## Section 2 – Matrix Equations

One of the many uses of the Kronecker product is to solve or determine properties of solutions to linear matrix equations. These matrix equations can be converted into an equivalent system of equations where the coefficient matrix involves the Kronecker product. In this transition, matrices are also converted into vectors. The function that implements this conversion is the subject of the next definition.

**Definition 30:** With each matrix  $A = [a_{ij}] \in M_{m,n}(\mathbb{F})$ , we associate the vector

$\text{vec}(A) \in \mathbb{F}^{mn}$  defined by  $\text{vec}(A) \equiv [a_{11}, \dots, a_{m1}, a_{12}, \dots, a_{m2}, \dots, a_{1n}, \dots, a_{mn}]^T$  (a column vector of each sequential column of  $A$  stacked on top of one another).

**Proposition 31:** Let  $A \in M_{m,n}$ ,  $B \in M_{p,q}$ ,  $X \in M_{q,n}$ , and  $Y \in M_{p,m}$ . Then,

$$(\text{vec}(Y))^T (A \otimes B) \text{vec}(X) = \text{trace}(A^T Y^T B X).$$

*Proof:* First examine the left side of the equation.

$$(\text{vec}(Y))^T (A \otimes B) \text{vec}(X) = [y_{11}, \dots, y_{p1}, \dots, y_{1m}, \dots, y_{pm}] \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{bmatrix} \begin{bmatrix} x_{11} \\ \vdots \\ x_{q1} \\ \vdots \\ x_{1n} \\ \vdots \\ x_{qn} \end{bmatrix}$$

$$= \begin{bmatrix} y_{11}, \dots, y_{p1}, \dots, y_{1m}, \dots, y_{pm} \end{bmatrix} \begin{bmatrix} \sum_{l=1}^n \sum_{j=1}^q a_{1l} b_{1j} x_{jl} \\ \vdots \\ \sum_{l=1}^n \sum_{j=1}^q a_{1l} b_{pj} x_{jl} \\ \vdots \\ \sum_{l=1}^n \sum_{j=1}^q a_{ml} b_{1j} x_{jl} \\ \vdots \\ \sum_{l=1}^n \sum_{j=1}^q a_{ml} b_{pj} x_{jl} \end{bmatrix} = \sum_{i=1}^m \left[ \sum_{k=1}^p y_{ki} \left( \sum_{l=1}^n \sum_{j=1}^q a_{il} b_{kj} x_{jl} \right) \right]$$

$$= \sum_{i=1}^m \sum_{k=1}^p \sum_{l=1}^n \sum_{j=1}^q y_{ki} a_{il} b_{kj} x_{jl} . \text{ Now, examine the right side of the equation.}$$

$$\begin{aligned} A^T Y^T B X &= \begin{bmatrix} a_{11} & \dots & a_{m1} \\ \vdots & \ddots & \vdots \\ a_{1n} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} y_{11} & \dots & y_{p1} \\ \vdots & \ddots & \vdots \\ y_{1m} & \dots & y_{pm} \end{bmatrix} \begin{bmatrix} b_{11} & \dots & b_{1q} \\ \vdots & \ddots & \vdots \\ b_{p1} & \dots & b_{pq} \end{bmatrix} \begin{bmatrix} x_{11} & \dots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{q1} & \dots & x_{qn} \end{bmatrix} \\ &= \begin{bmatrix} \sum_{i=1}^m a_{i1} y_{1i} & \dots & \sum_{i=1}^m a_{i1} y_{pi} \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^m a_{in} y_{1i} & \dots & \sum_{i=1}^m a_{in} y_{pi} \end{bmatrix} \begin{bmatrix} \sum_{j=1}^q b_{1j} x_{j1} & \dots & \sum_{j=1}^q b_{1j} x_{jn} \\ \vdots & \ddots & \vdots \\ \sum_{j=1}^q b_{pj} x_{j1} & \dots & \sum_{j=1}^q b_{pj} x_{jn} \end{bmatrix} \\ &= \begin{bmatrix} \sum_{k=1}^p \left( \sum_{i=1}^m a_{i1} y_{ki} \right) \left( \sum_{j=1}^q b_{kj} x_{j1} \right) & \dots & \sum_{k=1}^p \left( \sum_{i=1}^m a_{i1} y_{ki} \right) \left( \sum_{j=1}^q b_{kj} x_{jn} \right) \\ \vdots & \ddots & \vdots \\ \sum_{k=1}^p \left( \sum_{i=1}^m a_{in} y_{ki} \right) \left( \sum_{j=1}^q b_{kj} x_{j1} \right) & \dots & \sum_{k=1}^p \left( \sum_{i=1}^m a_{in} y_{ki} \right) \left( \sum_{j=1}^q b_{kj} x_{jn} \right) \end{bmatrix} . \text{ Hence,} \end{aligned}$$

$$\text{trace}(A^T Y^T B X) = \sum_{l=1}^n \left( \sum_{k=1}^p \left( \sum_{i=1}^m a_{il} y_{ki} \right) \left( \sum_{j=1}^q b_{kj} x_{jl} \right) \right) = \sum_{i=1}^m \sum_{k=1}^p \sum_{l=1}^n \sum_{j=1}^q y_{ki} a_{il} b_{kj} x_{jl} \text{ and}$$

$$(\text{vec}(Y))^T (A \otimes B) \text{vec}(X) = \text{trace}(A^T Y^T B X) .$$

□

The mapping  $vec : M_{m,n} \rightarrow \mathbb{F}^{mn}$  is a linear transformation and an isomorphism.

Since any linear transformation between two vectors corresponds to a unique matrix, we have the following useful property.

**Theorem 32:** Let  $T : M_{m,n} \rightarrow M_{p,q}$  be a given linear transformation. Then, there exists a unique matrix  $K_T \in M_{pq,mn}$  such that  $vec[T(X)] = K_T vec(X)$  for all  $X \in M_{m,n}$ .

*Proof:* Since both  $T$  and  $vec$  are linear transformations with  $vec$  also being an isomorphism, it follows that  $vec(T(vec^{-1})) : \mathbb{F}^{mn} \rightarrow \mathbb{F}^{pq}$  is a linear transformation.

Consequently, there exists a unique  $K_T \in M_{pq,mn}$  such that  $vec(T(vec^{-1}(y))) = K_T y$  for all  $y \in \mathbb{F}^{mn}$ . Since  $vec : M_{m,n} \rightarrow \mathbb{F}^{mn}$  is an isomorphism, for each  $X \in M_{m,n}$ , there exists a unique  $y \in \mathbb{F}^{mn}$  such that  $vec(X) = y$ . Hence, there exists a unique  $K_T \in M_{pq,mn}$  such that  $vec[T(X)] = K_T vec(X)$  for all  $X \in M_{m,n}$ .  $\square$

**Proposition 33:** Let  $A \in M_{m,n}$  and  $B \in M_{n,p}$  be given. Then,  $vec(AB) = (I \otimes A)vec(B)$ .

*Proof:* Since  $A \in M_{m,n}$  and  $B \in M_{n,p}$ , it follows that  $AB \in M_{m,p}$ ,  $vec(AB) \in \mathbb{F}^{mp}$ ,

$I_p \otimes A \in M_{mp,np}$ ,  $vec(B) \in \mathbb{F}^{np}$ , and  $(I_k \otimes A)vec(B) \in \mathbb{F}^{mp}$ . Now, because

$$AB = \begin{pmatrix} \sum_{j=1}^n a_{1j} b_{j1} & \cdots & \sum_{j=1}^n a_{1j} b_{jp} \\ \vdots & \ddots & \vdots \\ \sum_{j=1}^n a_{mj} b_{j1} & \cdots & \sum_{j=1}^n a_{mj} b_{jp} \end{pmatrix}, \text{ the definition of } vec \text{ gives that}$$

$$vec(AB) = \left[ \sum_{j=1}^n a_{1j} b_{j1}, \cdots, \sum_{j=1}^n a_{mj} b_{j1}, \sum_{j=1}^n a_{1j} b_{j2}, \cdots, \sum_{j=1}^n a_{mj} b_{j2}, \cdots, \sum_{j=1}^n a_{1j} b_{jp}, \cdots, \sum_{j=1}^n a_{mj} b_{jp} \right]^T.$$



$$\text{Next, } I_p \otimes A = \begin{pmatrix} A & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & A \end{pmatrix} \text{ and } \text{vec}(B) = [b_{11}, \dots, b_{n1}, b_{12}, \dots, b_{n2}, \dots, b_{1p}, \dots, b_{np}]^T \text{ imply}$$

that  $(I_p \otimes A)\text{vec}(B)$

$$= \left[ \sum_{j=1}^n a_{1j} b_{j1}, \dots, \sum_{j=1}^n a_{mj} b_{j1}, \sum_{j=1}^n a_{1j} b_{j2}, \dots, \sum_{j=1}^n a_{mj} b_{j2}, \dots, \sum_{j=1}^n a_{1j} b_{jp}, \dots, \sum_{j=1}^n a_{mj} b_{jp} \right]^T. \text{ Therefore,}$$

$$\text{vec}(AB) = (I \otimes A)\text{vec}(B).$$

□

It will now be shown how the Kronecker product and  $\text{vec}$  can be used to rewrite linear matrix equations. The next theorem shows how these tools can be used to convert the equation  $AXB = C$  into a system of linear equations.

**Theorem 34:** Let  $A \in M_{m,n}$ ,  $B \in M_{p,q}$ , and  $C \in M_{m,q}$  be given and let  $X \in M_{n,p}$  be

unknown. Then,  $\text{vec}(AXB) = (B^T \otimes A)\text{vec}(X)$  and consequently, the matrix equation

$AXB = C$  is equivalent to the system of  $mq$  equations in  $np$  unknowns given by

$$(B^T \otimes A)\text{vec}(X) = \text{vec}(C).$$

Proof: For a given matrix  $Q$ , let  $Q_k$  denote the  $k^{\text{th}}$  column of  $Q$ . Then,

$(AXB)_k = A(XB)_k = AXB_k$ . This implies that

$$\begin{aligned} (AXB)_k &= A \begin{pmatrix} x_{11}b_{1k} + x_{12}b_{2k} + \cdots + x_{1p}b_{pk} \\ \vdots \\ x_{n1}b_{1k} + x_{n2}b_{2k} + \cdots + x_{np}b_{pk} \end{pmatrix} = A \left[ \begin{pmatrix} x_{11} \\ \vdots \\ x_{n1} \end{pmatrix} b_{1k} + \begin{pmatrix} x_{12} \\ \vdots \\ x_{n2} \end{pmatrix} b_{2k} + \cdots + \begin{pmatrix} x_{1p} \\ \vdots \\ x_{np} \end{pmatrix} b_{pk} \right] \\ &= b_{1k} A \begin{pmatrix} x_{11} \\ \vdots \\ x_{n1} \end{pmatrix} + b_{2k} A \begin{pmatrix} x_{12} \\ \vdots \\ x_{n2} \end{pmatrix} + \cdots + b_{pk} A \begin{pmatrix} x_{1p} \\ \vdots \\ x_{np} \end{pmatrix} = [b_{1k}A, b_{2k}A, \dots, b_{pk}A] \text{vec}(X) \end{aligned}$$

$= (B_k^T \otimes A) \text{vec}(X)$  for  $k = 1, 2, \dots, q$ . Therefore,

$$\text{vec}(C) = \text{vec}(AXB) = \begin{pmatrix} B_1^T \otimes A \\ \vdots \\ B_q^T \otimes A \end{pmatrix} \text{vec}(X) = (B^T \otimes A) \text{vec}(X).$$

□

Other linear matrix equations can be written in a similar manner using the Kronecker product and the  $\text{vec}$  notation. Some examples are:

$$AX = B \Leftrightarrow (I \otimes A) \text{vec}(X) = \text{vec}(B)$$

$$AX + XB = C \Leftrightarrow [(I \otimes A) + (B^T \otimes I)] \text{vec}(X) = \text{vec}(C)$$

$$A_1 X B_1 + \dots + A_k X B_k = C \Leftrightarrow [(B_1^T \otimes A_1) + \dots + (B_k^T \otimes A_k)] \text{vec}(X) = \text{vec}(C)$$

$$AX + YB = C \Leftrightarrow (I \otimes A) \text{vec}(X) + (B^T \otimes I) \text{vec}(Y) = \text{vec}(C)$$

In partitioned form, the previous equation can also be written as

$$\begin{bmatrix} I \otimes A & B^T \otimes I \end{bmatrix} \begin{bmatrix} \text{vec}(X) \\ \text{vec}(Y) \end{bmatrix} = \text{vec}(C).$$

Determining whether or not two matrices commute with each other using the conventional matrix multiplication can also be written

$$\text{in this way. That is, } AX = XA \Leftrightarrow AX - XA = 0 \Leftrightarrow [(I \otimes A) - (A^T \otimes I)] \text{vec}(X) = 0.$$

Theorem 34 also allows for an expansion of Proposition 33 and for an examination of a system of linear equations in parallel. This is the content of the next two corollaries.

**Corollary 35:** Let  $A \in M_{m,n}$  and  $B \in M_{n,p}$  be given. Then,

$$\text{vec}(AB) = (I_p \otimes A) \text{vec}(B) = (B^T \otimes A) \text{vec}(I_n) = (B^T \otimes I_m) \text{vec}(A).$$

Proof: The three forms of  $\text{vec}(AB)$  given in the corollary are developed by placing an identity matrix in different places in the conventional matrix multiplication. First,

$AB = ABI_p$ . Using Theorem 34, this is equivalent to  $\text{vec}(AB) = (I_p \otimes A)\text{vec}(B)$ . Next,

$AB = AI_n B$  is equivalent to  $\text{vec}(AB) = (B^T \otimes A)\text{vec}(I_n)$ . Finally,  $AB = I_m AB$  is

equivalent to  $\text{vec}(AB) = (B^T \otimes I_m)\text{vec}(A)$ .  $\square$

**Corollary 36:** Let  $A \in M_{m,p}$ ,  $B \in M_{n,q}$ ,  $x \in \mathbb{R}^{pq}$ , and  $t \in \mathbb{R}^{mn}$ .  $(A \otimes B)x = t$  if and only

if  $(B \otimes A)x_T = t_T$  where  $x = \text{vec}(X)$ ,  $t = \text{vec}(T)$ ,  $x_T = \text{vec}(X^T)$ , and  $t_T = \text{vec}(T^T)$ .

*Proof:* Using the definitions, it suffices to prove that  $(A \otimes B)\text{vec}(X) = \text{vec}(T)$  if and

only if  $(B \otimes A)\text{vec}(X^T) = \text{vec}(T^T)$ . Using Theorem 34,  $(A \otimes B)\text{vec}(X) = \text{vec}(T)$  if

and only if  $BXA^T = T$ . Taking the transpose of both sides gives  $AX^T B^T = T^T$ . The

proof is completed by applying Theorem 34 to this equation. Hence,  $AX^T B^T = T^T$  is true

if and only if  $(B \otimes A)x_T = t_T$ .  $\square$

As mentioned earlier, Kronecker products can help determine the solvability of a linear matrix equation. Kronecker products can also be used to discover properties of potential solutions. For example, are such solutions unique or nonsingular. The examination of solutions to linear matrix equations begins with  $AXB = C$ . The next theorem provides information regarding when this equation has a solution in general and when it is unique.

**Theorem 37:** Let  $A, B, C \in M_n$ . Then, the matrix equation  $AXB = C$  has a unique

solution  $X \in M_n$  for every given  $C$  if and only if both  $A$  and  $B$  are nonsingular. If either

$A$  or  $B$  is singular, then there exists a solution  $X \in M_n$  if and only if

$$\text{rank}(B^T \otimes A) = \text{rank}\left(\begin{bmatrix} B^T \otimes A & \text{vec}(C) \end{bmatrix}\right).$$

Proof: First, let  $A$  and  $B$  be nonsingular matrices. If the matrix equation  $AXB = C$  is multiplied on the left side by  $A^{-1}$  and on the right side by  $B^{-1}$ , the resulting equation is  $X = A^{-1}CB^{-1}$ . Thus, each solution  $X$  is unique for a given  $C$ . Next, assume that the matrix equation  $AXB = C$  has a unique solution  $X$  for every given  $C$ . Since the assumption is for all  $C \in M_n$ , let  $C$  be a nonsingular matrix. This implies that  $\det(A)\det(X)\det(B) = \det(C) \neq 0$ . By the zero product property, this gives that  $\det(A) \neq 0$  and  $\det(B) \neq 0$ . Hence, both  $A$  and  $B$  are nonsingular.

Now, Theorem 34 states that there exists a solution  $X$  to the equation  $AXB = C$  if and only if there exists a solution  $\text{vec}(X)$  to the linear system of equations

$(B^T \otimes A)\text{vec}(X) = \text{vec}(C)$ . Also, it is known that there exists a solution to the system of  $n^2$  linear equations  $(B^T \otimes A)\text{vec}(X) = \text{vec}(C)$  if and only if

$$\text{rank}(B^T \otimes A) = \text{rank}\left(\begin{bmatrix} B^T \otimes A & \text{vec}(C) \end{bmatrix}\right). \quad \square$$

By replacing the matrix  $B$  with an identity matrix, Theorem 37 can be used to determine conditions of solvability for the simpler matrix equation  $AX = C$ . This equation has a unique solution  $X$  if and only if  $A$  is nonsingular. If  $A$  is singular, then there exists a solution  $X$  if and only if  $\text{rank}(I_n \otimes A) = \text{rank}\left(\begin{bmatrix} I_n \otimes A & \text{vec}(C) \end{bmatrix}\right)$ . By Theorem 20,  $\text{rank}(I_n \otimes A) = n(\text{rank}(A))$ . Therefore, there exists a solution  $X$  if and only if  $\text{rank}\left(\begin{bmatrix} I_n \otimes A & \text{vec}(C) \end{bmatrix}\right) = n(\text{rank}(A))$ .

Theorem 34 will be used to obtain results concerning the eigenvalues of some interesting linear transformations in the following propositions. First, a lemma is needed

about the eigenvalues of a square matrix and its transpose that will be used later to prove Corollary 43.

**Lemma 38:** Let  $A$  be a square matrix. Then, the eigenvalues of  $A$  are the same as the eigenvalues of  $A^T$ .

*Proof:* First, let  $A = U\Delta_A U^*$  where  $U$  is a unitary matrix and  $\Delta_A$  is an upper triangular matrix with the eigenvalues of  $A$  on the diagonal. Since  $U$  is a unitary matrix,

$$\det(UU^*) = \det(U)\det(U^*) = 1. \text{ Then, } \det(A) = \det(U\Delta_A U^*)$$

$$= \det(U)\det(\Delta_A)\det(U^*) = \det(U)\det(U^*)\det(\Delta_A) = \det(\Delta_A). \text{ Also,}$$

$$\det(A^T) = \det((U\Delta_A U^*)^T) = \det((U^*)^T \Delta_A^T U^T) = \det((U^*)^T)\det(\Delta_A^T)\det(U^T)$$

$$= \det((U^*)^T)\det(U^T)\det(\Delta_A^T) = \det(\Delta_A^T) \text{ where } \Delta_A^T \text{ is a lower triangular matrix with}$$

the eigenvalues of  $A$  on its diagonal. Since the determinant of an upper (or lower) triangular matrix is the product of its diagonal elements, it follows that

$$\det(\Delta_A^T) = \det(\Delta_A). \text{ This implies that } \det(A^T) = \det(A). \text{ Next, it is known that}$$

$$\det(A - \lambda I) = 0 \text{ if and only if } \lambda \in \sigma(A). \text{ Since } \det(A^T) = \det(A), \det(A - \lambda I) = 0 \text{ if}$$

$$\text{and only if } 0 = \det((A - \lambda I)^T) = \det(A^T - \lambda I). \text{ Therefore, } \lambda \in \sigma(A) \text{ if and only if}$$

$$\lambda \in \sigma(A^T). \text{ Thus, } \sigma(A^T) = \sigma(A). \quad \square$$

**Proposition 39:** Let  $S \in M_n$  be a given nonsingular matrix with eigenvalues

$\sigma(S) = \{\lambda_1, \dots, \lambda_n\}$ . Then, the similarity map  $T_s : M_n \rightarrow M_n$  given by  $T_s(X) \equiv S^{-1}XS$  is a

linear transformation with eigenvalues  $\sigma(T_s) = \left\{ \frac{\lambda_i}{\lambda_j} : \lambda_i, \lambda_j \in \sigma(S) \right\}$ .

Proof: Let  $\alpha \in \mathbb{F}$  and  $X, Y \in M_n$ . Then,  $T_s(\alpha X + Y) = S^{-1}(\alpha X + Y)S$

$= \alpha S^{-1}XS + S^{-1}YS = \alpha T_s(X) + T_s(Y)$ . Therefore,  $T_s$  is a linear transformation. If  $\beta$  is an eigenvalue of  $T_s$ , then  $T_s(X) = \beta X$  for some nonzero  $X$ . This gives that

$S^{-1}XS = \beta X$  and, consequently,  $(S^T \otimes S^{-1})\text{vec}(X) = \beta \text{vec}(X)$ . This means that  $\beta$  is also an eigenvalue of the matrix  $S^T \otimes S^{-1}$ . Since  $S$  is nonsingular, each of its eigenvalues

is nonzero. Also, since  $\sigma(S^T) = \sigma(S)$  and  $\sigma(S^{-1}) = \left\{ \frac{1}{\lambda_i} : \lambda_i \in \sigma(S) \right\}$ , it follows from

Theorem 15 that,  $\sigma(T_s) = \sigma(S^T \otimes S^{-1}) = \left\{ \frac{\lambda_i}{\lambda_j} : \lambda_i, \lambda_j \in \sigma(S) \right\}$ . □

The process of taking matrix equations and converting them to an equivalent Kronecker form shows that matrices of the form  $(I \otimes A) + (B \otimes I)$  arise naturally in the study of linear matrix equations. Properties of these matrices also turn out to be useful in determining solvability of matrix equations.

**Definition 40:** Let  $A \in M_n$  and  $B \in M_m$ . The square matrix  $(I_m \otimes A) + (B \otimes I_n) \in M_{mn}$  is called the *Kronecker sum* of  $A$  and  $B$ .

The next result shows that the summands in the Kronecker sum of  $A$  and  $B$  commute with respect to conventional matrix multiplication.

**Lemma 41:** Let  $A \in M_n$  and  $B \in M_m$  be given. Then,  $I_m \otimes A$  commutes with  $B \otimes I_n$  using conventional matrix multiplication.

Proof:  $(I_m \otimes A)(B \otimes I_n) = (I_m B) \otimes (A I_n) = B \otimes A = (B I_m) \otimes (I_n A)$   
 $= (B \otimes I_n)(I_m \otimes A)$ . □

Just as the eigenvalues of a Kronecker product have been very useful, so are the eigenvalues of a Kronecker sum. They will help determine if certain types of matrix equations have solutions.

**Theorem 42:** Let  $A \in M_n$  and  $B \in M_m$  be given matrices. If  $\lambda \in \sigma(A)$  and  $x \in \mathbb{C}^n$  is a corresponding eigenvector of  $A$ , and if  $\mu \in \sigma(B)$  and  $y \in \mathbb{C}^m$  is a corresponding eigenvector of  $B$ , then  $\lambda + \mu$  is an eigenvalue of the Kronecker sum  $(I_m \otimes A) + (B \otimes I_n)$  and  $y \otimes x \in \mathbb{C}^{mn}$  is a corresponding eigenvector of the Kronecker sum. Every eigenvalue of the Kronecker sum arises as such a sum of eigenvalues of  $A$  and  $B$ . If

$\sigma(A) = \{\lambda_1, \dots, \lambda_n\}$  and  $\sigma(B) = \{\mu_1, \dots, \mu_m\}$ , then

$\sigma((I_m \otimes A) + (B \otimes I_n)) = \{\lambda_i + \mu_j : i = 1, \dots, n; j = 1, \dots, m\}$ . In particular,

$$\sigma((I_m \otimes A) + (B \otimes I_n)) = \sigma((I_n \otimes B) + (A \otimes I_m)).$$

**Proof:** If  $\lambda \in \sigma(A)$  and  $x \in \mathbb{C}^n$  is a corresponding eigenvector of  $A$ , and if

$\mu \in \sigma(B)$  and  $y \in \mathbb{C}^m$  is a corresponding eigenvector of  $B$ , then  $Ax = \lambda x$  and  $By = \mu y$ .

$$\begin{aligned} \text{Then, } [(I_m \otimes A) + (B \otimes I_n)](y \otimes x) &= (I_m \otimes A)(y \otimes x) + (B \otimes I_n)(y \otimes x) \\ &= y \otimes (Ax) + (By) \otimes x = y \otimes (\lambda x) + (\mu y) \otimes x = \lambda(y \otimes x) + \mu(y \otimes x) = (\lambda + \mu)(y \otimes x). \end{aligned}$$

Therefore,  $\lambda + \mu$  is an eigenvalue of  $(I_m \otimes A) + (B \otimes I_n)$  with corresponding eigenvector  $y \otimes x$ .

To complete the proof, it will be shown that

$\sigma((I_m \otimes A) + (B \otimes I_n)) = \{\lambda_i + \mu_j : i = 1, \dots, n; j = 1, \dots, m\}$ . Let  $U \in M_n$  and  $V \in M_m$  be

unitary matrices such that  $U^*AU = \Delta_A$  and  $V^*BV = \Delta_B$  are upper triangular matrices

with the eigenvalues of  $A$  and  $B$  on the diagonals, respectively. Then,  $W = V \otimes U \in M_{mn}$

is a unitary matrix such that  $W^*(I_m \otimes A)W = (V^* \otimes U^*)(I_m \otimes A)(V \otimes U)$

$$= (V^* I_m V) \otimes (U^* A U) = I_m \otimes \Delta_A = \begin{pmatrix} \Delta_A & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \Delta_A \end{pmatrix} \text{ and}$$

$$W^*(B \otimes I_n)W = (V^* \otimes U^*)(B \otimes I_n)(V \otimes U) = (V^* B V) \otimes (U^* I_n U)$$

$$= \Delta_B \otimes I_n = \begin{pmatrix} \mu_1 I_n & \cdots & * \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \mu_m I_n \end{pmatrix} \text{ are upper triangular matrices. Hence,}$$

$W^*[(I_m \otimes A) + (B \otimes I_n)]W = (I_m \otimes \Delta_A) + (\Delta_B \otimes I_n)$  is upper triangular matrix with the eigenvalues of the Kronecker sum  $(I_m \otimes A) + (B \otimes I_n)$  on its diagonal. Inspection of this

diagonal shows that it consists of every possible sum of an eigenvalue of  $A$  with an

eigenvalue of  $B$ . Hence,  $\sigma((I_m \otimes A) + (B \otimes I_n)) = \{\lambda_i + \mu_j : i = 1, \dots, n; j = 1, \dots, m\}$ . This

also implies that  $\sigma((I_n \otimes B) + (A \otimes I_m)) = \{\mu_i + \lambda_j : i = 1, \dots, m; j = 1, \dots, n\}$

$$= \sigma((I_m \otimes A) + (B \otimes I_n)). \quad \square$$

**Corollary 43:** Let  $A \in M_n$  with eigenvalues  $\lambda_1, \dots, \lambda_n$ . Define the linear transformation

$T_A : M_n \rightarrow M_n$  to be  $T_A(X) \equiv AX - XA$ . Then,  $\sigma(T_A) = \{\lambda_i - \lambda_j : i, j = 1, \dots, n\}$ .

*Proof:* Let  $\mu$  be an eigenvalue of  $T_A$ . Then,  $T_A(X) = \mu X$  for some nonzero  $X \in M_n$ .

This implies that  $AX - XA = \mu X$ . Writing this as an equivalent system of equations

gives  $\mu$  is an eigenvalue of  $T$  if and only if  $[(I_n \otimes A) - (A^T \otimes I_n)]\text{vec}(X) = \mu \text{vec}(X)$ .

This shows that  $\mu$  is an eigenvalue of  $T$  if and only if  $\mu$  is an eigenvalue of



$(I_n \otimes A) - (A^T \otimes I_n)$ . That is,  $\sigma(T_A) = \sigma((I_n \otimes A) + (-A^T \otimes I_n))$ . Since

$\sigma(A) = \sigma(A^T) = \{\lambda_i : i = 1, 2, \dots, n\}$  by Lemma 38, it follows from Theorem 42 that

$$\sigma(T_A) = \{\lambda_i - \lambda_j : i, j = 1, 2, \dots, n\}.$$

□

Previously, the eigenvalues of some linear transformations were examined. Now that the Kronecker sum and some of its properties have been introduced, solutions of linear matrix equations will be investigated.

**Proposition 44:** Let  $A, B \in M_n$  be given matrices. Then, the equation  $AX - XB = 0$  has a nonsingular solution  $X \in M_n$  if and only if  $A$  and  $B$  are similar.

*Proof:* The matrices  $A$  and  $B$  are similar if and only if there exists a nonsingular matrix  $X \in M_n$  such that  $X^{-1}AX = B$ . This happens if and only if  $AX = XB$  or  $AX - XB = 0$  for some nonsingular  $X$ . □

Theorem 37 examined the matrix equation  $AXB = C$ . The following theorem looks at the equation  $AX + XB = C$  and explains when a solution  $X$  is unique. This is done using the eigenvalues of the Kronecker sum.

**Theorem 45:** Let  $A \in M_n$  and  $B \in M_m$ . The equation  $AX + XB = C$  has a unique solution  $X \in M_{n,m}$  for each  $C \in M_{n,m}$  if and only if  $\sigma(A) \cap \sigma(-B) = \emptyset$ .

*Proof:* By changing the equation  $AX + XB = C$  into its Kronecker form

$$[(I \otimes A) + (B^T \otimes I)] \text{vec}(X) = \text{vec}(C),$$

the matrix equation is replaced by a system of linear equations. This system of equations has a unique solution if and only if the coefficient matrix,  $[(I \otimes A) + (B^T \otimes I)]$ , does not have a zero eigenvalue. Since, by

Lemma 38, the eigenvalues of  $B^T$  are the same as the eigenvalues of  $B$ , it follows from

Theorem 42 that,  $\sigma\left((I \otimes A) + (B^T \otimes I)\right) = \{\lambda_i + \mu_j : i = 1, \dots, n; j = 1, \dots, m\}$  where  $\sigma(A) = \{\lambda_i : i = 1, \dots, n\}$  and  $\sigma(B) = \{\mu_j : j = 1, \dots, m\}$ . The matrix  $\left[(I \otimes A) + (B^T \otimes I)\right]$  will have zero as an eigenvalue if and only if there exists an  $i$  and  $j$  such that  $\lambda_i + \mu_j = 0$  or  $\lambda_i = -\mu_j$ . Therefore, all eigenvalues of  $\left[(I \otimes A) + (B^T \otimes I)\right]$  will be nonzero if and only if  $\sigma(A) \cap -\sigma(B) = \emptyset$  or  $\sigma(A) \cap \sigma(-B) = \emptyset$ .  $\square$

The generalized commutativity equation,  $AX - XA = C$ , is a special case of the equation in Theorem 45, where  $B = -A$ . It follows from Theorem 45 that  $\sigma(A) \cap \sigma(-(-A)) = \sigma(A) \cap \sigma(A) = \sigma(A) \neq \emptyset$ . Thus,  $AX - XA = C$  does not have a unique solution. This means that either the generalized commutativity equation has no solution  $X$  or there are infinitely many solutions  $X$ . For this equation, and other singular equations, it would be useful to determine the dimension of the nullspace of the linear transformation  $X \rightarrow AX + XB$ . That is, the dimension of the solution space of the homogeneous equation  $AX + XB = 0$ . The following lemma and theorem will provide this information.

The lemma introduces a new class of matrices. These are Toeplitz matrices. A Toeplitz matrix has the same values along each diagonal (from top-left to bottom-right), or  $[A]_{ij} = a_{i-j}$ . Numerical problems involving these matrices usually have faster solutions, since only  $2n - 1$  elements need to be found instead of  $n^2$  (for  $A \in M_n$ ). They are used in many areas of mathematics, such as finding numerical solutions to differential and integral equations, computation of splines, and signal and image processing.

Recall that the Jordan - Canonical form of a matrix  $A \in M_n$  is a block matrix with square matrices of various sizes along the diagonal and zeros elsewhere. These square matrices on the diagonal are called Jordan blocks. Each of these Jordan blocks,  $J_i(\lambda)$ , is associated with a  $\lambda$ , an eigenvalue of  $A$ , such that the values on the main diagonal of  $J_i(\lambda)$  are  $\lambda$ , the values on the diagonal directly above and to the left of the main diagonal are 1, and all other values in the matrix are 0.

**Lemma 46:** Let  $J_r(0) \in M_r$  and  $J_s(0) \in M_s$  be singular Jordan blocks. Then,  $X \in M_{r,s}$

is a solution to  $J_r(0)X - XJ_s(0) = 0$  if and only if  $X = \begin{bmatrix} 0 & Y \end{bmatrix}$ ,  $Y \in M_r$ ,  $0 \in M_{r,s-r}$  when

$r \leq s$  or  $X = \begin{bmatrix} Y \\ 0 \end{bmatrix}$ ,  $Y \in M_s$ ,  $0 \in M_{r-s,s}$  when  $r \geq s$  where, in either case,  $Y = [y_{ij}]$  is an

upper triangular Toeplitz matrix with  $y_{ij} = a_{i-j}$ . The dimension of the nullspace of the

linear transformation  $T(Z) = J_r(0)Z - ZJ_s(0)$  is  $\min\{r, s\}$ .

**Proof:** Let  $X = [x_{ij}]$ . This implies that  $[J_r(0)X]_{ij} = x_{i+1,j}$  and  $[XJ_s(0)]_{ij} = x_{i,j-1}$  for

$i = 1, 2, \dots, r$  and  $j = 1, 2, \dots, s$ , where it is defined that  $x_{r+1,j} = 0 = x_{i,0}$ . Then,

$[J_r(0)X - XJ_s(0)]_{ij} = 0$  for all  $i = 1, 2, \dots, r$  and  $j = 1, 2, \dots, s$  if and only if

$x_{i+1,j} - x_{i,j-1} = 0$  or  $x_{i+1,j+1} = x_{i,j}$  for  $i = 1, 2, \dots, r$  and  $j = 0, 1, 2, \dots, s-1$ . This implies that

a solution has the form  $X = \begin{bmatrix} 0 & Y \end{bmatrix}$  or  $X = \begin{bmatrix} Y \\ 0 \end{bmatrix}$  as described in the statement of the

theorem. The nullspace of the transformation  $T(Z) = J_r(0)Z - ZJ_s(0)$  consists of all

possible solutions  $X$ . The dimension of this nullspace would be the size of a basis for

these solutions  $X$ . Since  $Y$  is an upper triangular Toeplitz matrix and the remaining

portion of  $X$  is made up of zeros, the size of a basis for these solutions would be either  $r$  if

$r \leq s$  or  $s$  if  $r \geq s$ ; thus proving that the dimension of the nullspace is  $\min\{r, s\}$ .  $\square$

**Theorem 47:** Let  $A \in M_n$  and  $B \in M_m$ . Also, let  $SAS^{-1} = \begin{pmatrix} J_{n_1}(\lambda_1) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & J_{n_p}(\lambda_p) \end{pmatrix} = J_A$

with  $n_1 + n_2 + \cdots + n_p = n$  and  $RBR^{-1} = \begin{pmatrix} J_{m_1}(\mu_1) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & J_{m_q}(\mu_q) \end{pmatrix} = J_B$  with

$m_1 + m_2 + \cdots + m_q = m$  be the respective Jordan - Canonical forms of  $A$  and  $B$ . Then, the

dimension of the nullspace of the linear transformation  $L : M_{n,m} \rightarrow M_{n,m}$  given by

$L : X \rightarrow AX + XB$  is  $\sum_{i=1}^p \sum_{j=1}^q v_{ij}$  where  $v_{ij} = 0$  if  $\lambda_i \neq -\mu_j$  and  $v_{ij} = \min\{n_i, m_j\}$  if

$\lambda_i = -\mu_j$ .

Proof: Notice that the equation  $C = AX + XB$  is equivalent to  $SCR = SAXR + SXBR$

$= (SAS^{-1})(SXR) - (SXR)(R^{-1}(-B)R)$  for nonsingular  $S \in M_n$  and  $R \in M_m$ . Also, note

that  $\begin{pmatrix} J_{m_1}(-\mu_1) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & J_{m_q}(-\mu_q) \end{pmatrix} = J_{(-B)}$ . Thus, it suffices to determine the dimension of

the nullspace of the linear transformation  $T : X \rightarrow J_A X - X J_{(-B)}$ . Write the matrix  $X$  in

partitioned form as  $X = \begin{pmatrix} X_{11} & \cdots & X_{1q} \\ \vdots & \ddots & \vdots \\ X_{p1} & \cdots & X_{pq} \end{pmatrix}$  where  $X_{ij} \in M_{n_i, m_j}$  for  $i = 1, 2, \dots, p$  and

$j = 1, 2, \dots, q$ . The equation  $J_A X - X J_{(-B)} = 0$  is equivalent to the set of  $pq$  linear matrix

equations  $J_{n_i}(\lambda_i)X_{ij} - X_{ij}J_{m_j}(-\mu_j) = 0$  for  $i = 1, 2, \dots, p$  and  $j = 1, 2, \dots, q$ . By Theorem 45, each equation where  $\lambda_i + \mu_j \neq 0$  has only the trivial solution  $X_{ij} = 0$ . However, when  $\lambda_i + \mu_j = 0$ , use the identity  $J_k(\lambda) = \lambda I + J_k(0)$  to rewrite the equation as follows

$$0 = J_{n_i}(\lambda_i)X_{ij} - X_{ij}J_{m_j}(-\mu_j) = [\lambda_i I + J_{n_i}(0)]X_{ij} - X_{ij}[-\mu_j I + J_{m_j}(0)]$$

$$= \lambda_i X_{ij} + J_{n_i}(0)X_{ij} + \mu_j X_{ij} - X_{ij}J_{m_j}(0) = J_{n_i}(0)X_{ij} - X_{ij}J_{m_j}(0).$$

Using Lemma 46 with  $r = n_i$  and  $s = m_j$ , the dimension of the nullspaces of each of the  $pq$  matrix equations is exactly  $v_{ij}$  where  $v_{ij} = 0$  if  $\lambda_i \neq -\mu_j$  and  $v_{ij} = \min\{n_i, m_j\}$  if  $\lambda_i = -\mu_j$ . Therefore, the dimension of the nullspace of the linear transformation  $L: X \rightarrow AX + XB$  is  $\sum_{i=1}^p \sum_{j=1}^q v_{ij}$ .  $\square$

A very interesting result that follows from Theorem 47 concerns the dimension of the subspace of the matrices that commute with a given matrix.

**Corollary 48:** Let  $A \in M_n$  be a given matrix. The set of matrices in  $M_n$  that commute with  $A$  is a subspace of  $M_n$  with dimension at least  $n$ . The dimension is equal to  $n$  if and only if each eigenvalue of  $A$  has geometric multiplicity one.

*Proof:* First, note that the set of matrices that commutes with  $A$  is a subspace of  $M_n$  since it is the nullspace of the linear transformation  $T_A(X) = AX - XA$ . A matrix that commutes with  $A$  will be a solution to the equation  $AX - XA = 0$ . If  $B = -A$  then, the dimension of the subspace will be equal to the dimension of the nullspace of the linear transformation  $L: X \rightarrow AX + XB$ . Theorem 47 proves that this dimension is  $\sum_{i=1}^p \sum_{j=1}^q v_{ij}$ .

Using the same notation as Theorem 47,  $B = -A$  gives  $p = q$ ,  $n_i = m_i$ , and  $\mu_i = -\lambda_i$  for

$i = 1, 2, \dots, p$ . As a result,  $\sum_{i=1}^p \sum_{j=1}^q v_{ij} = \sum_{i,j=1}^p v_{ij} \geq \sum_{i=1}^p v_{ii} = \sum_{i=1}^p n_{ii} = n$ . The inequality becomes

an equality if and only if  $\lambda_i \neq -\mu_j$  or  $\lambda_i \neq \lambda_j$  for all  $i \neq j$ . This happens if and only if each eigenvalue of  $A$  is associated with exactly one Jordan block which is true if and only if each eigenvalue of  $A$  has geometric multiplicity one.  $\square$

### Section 3 – Permutation Matrices and Jordan - Canonical Form

In this section, permutation matrices turn out to be key players in our results and their proofs. A permutation matrix  $P \in M_n$  is so named because the matrix is obtained by permuting the rows of an  $n \times n$  identity matrix according to some permutation of the numbers  $1, \dots, n$ . Therefore, each column and row contains exactly one 1 and zeros everywhere else. Using conventional matrix multiplication, the product of a square matrix  $A$  of the same dimensions as the permutation matrix,  $PA$ , creates a matrix  $B=PA$  such that  $B$  is a permutation of rows of  $A$  matching the same rearrangement as the permutation matrix. In Theorem 51, it is shown that permutation matrices are orthogonal and invertible. Two matrices  $X$  and  $Y$  are said to be permutation equivalent if  $X = YP$  for some permutation matrix  $P$ .

**Lemma 49:** Let  $A \in M_n$  be a matrix with exactly one 1 in each row and column and zeros elsewhere. Then,  $A$  is a permutation matrix.

**Proof:** Find the row whose first entry is 1. Switch this row with the first row. Next, find the row whose second entry is a 1. Switch this row with the second row. Continue this

process and it ends with the identity matrix  $I_n$ . Reversing this process turns  $I_n$  into  $A$ .

Hence,  $A$  is a permutation matrix.  $\square$

**Lemma 50:** Let  $A \in M_n$ . If  $A^T = A^{-1}$  and the entries of  $A$  are only ones and zeros, then,  $A$  is a permutation matrix.

*Proof (by contradiction):* Assume  $A$  is not a permutation matrix. Then,  $A$  has either a row (or column) with no entry equal to 1 or a row (or column) with more than one entry equal to 1. First, if the former situation is true, then  $A$  has a row (or column) whose entries are all zero and  $A$  would not be invertible, which is a contradiction. So, assume  $A$  has a row (or column),  $i$ , with more than one entry equal to 1. Then,  $A^T$  has a column (or row),  $i$ , with more than one entry equal to 1 in the same positions. Thus,  $AA^T$  (or  $A^T A$ ) will have an entry greater than 1 in the  $[AA^T]_{i,i}$  entry (or  $[A^T A]_{i,i}$  entry). The product of the two matrices is not the identity matrix, which is a contradiction to  $A^T = A^{-1}$ . Therefore,  $A$  is a permutation matrix.  $\square$

It has been shown that there are several instances in which some property of  $A \otimes B$  is the same as that of  $B \otimes A$ . This is primarily due to the fact, proven in Corollary 52, that  $A \otimes B$  is always permutation equivalent to  $B \otimes A$ .

**Theorem 51:** Let  $m, n$  be given positive integers. There exists a unique matrix

$P(m, n) \in M_{mn}$  such that  $\text{vec}(X^T) = P(m, n)\text{vec}(X)$  for all  $X \in M_{m,n}$ . (This is the  $K_T$

matrix from Theorem 32 for the linear transformation of transposing a matrix –

$T(X) = X^T$ .) The matrix  $P(m, n)$  depends only on the dimensions  $m$  and  $n$  and is given

by  $P(m, n) = \sum_{i=1}^m \sum_{j=1}^n E_{ij} \otimes E_{ij}^T$  where each  $E_{ij} \in M_{m,n}$  has a one in the  $i, j^{\text{th}}$  entry and every

other entry is zero. Moreover,  $P(m, n)$  is a permutation matrix and

$$P(m, n) = P(n, m)^T = P(n, m)^{-1}.$$

*Proof:* Let  $X = [x_{ij}] \in M_{m,n}$  and let  $E_{ij} \in M_{m,n}$  where the  $ij^{th}$  entry of  $E_{ij}$  is one and every other entry is zero. Accordingly,  $E_{ij}^T X E_{ij}^T \in M_{n,m}$  with the element  $x_{ij}$  in the  $j, i^{th}$  entry and zeros elsewhere for all  $i = 1, 2, \dots, m$  and  $j = 1, 2, \dots, n$ . Thus,

$$X^T = \sum_{i=1}^m \sum_{j=1}^n E_{ij}^T X E_{ij}^T. \text{ Hence, } \text{vec}(X^T) = \text{vec}\left(\sum_{i=1}^m \sum_{j=1}^n E_{ij}^T X E_{ij}^T\right) = \sum_{i=1}^m \sum_{j=1}^n \text{vec}(E_{ij}^T X E_{ij}^T)$$

$$= \sum_{i=1}^m \sum_{j=1}^n (E_{ij} \otimes E_{ij}^T) \text{vec}(X), \text{ by Theorem 34. Let } P(m, n) = \sum_{i=1}^m \sum_{j=1}^n E_{ij} \otimes E_{ij}^T. \text{ Then,}$$

$$\text{vec}(X^T) = P(m, n) \text{vec}(X). \text{ Now, if } Y = X^T, \text{ then } Y \in M_{n,m} \text{ and}$$

$$\text{vec}(Y^T) = P(n, m) \text{vec}(Y). \text{ This implies that } \text{vec}(X) = P(n, m) \text{vec}(X^T)$$

$$= P(n, m) P(m, n) \text{vec}(X) \text{ for all } X \in M_{m,n}. \text{ Similarly, } \text{vec}(Y) = P(m, n) \text{vec}(Y^T)$$

$$= P(m, n) P(n, m) \text{vec}(Y) \text{ for all } Y \in M_{n,m}. \text{ Consequently,}$$

$$P(m, n) P(n, m) = P(n, m) P(m, n) = I. \text{ That is, } P(m, n)^{-1} = P(n, m). \text{ If } F_{ji} = E_{ji}^T, \text{ then}$$

$$P(n, m) = \sum_{j=1}^n \sum_{i=1}^m E_{ji} \otimes E_{ji}^T = \sum_{j=1}^n \sum_{i=1}^m F_{ij}^T \otimes F_{ij} = \sum_{j=1}^n \sum_{i=1}^m F_{ji}^T \otimes F_{ji}$$

$$= \sum_{j=1}^n \sum_{i=1}^m (F_{ji} \otimes F_{ji}^T)^T = P(m, n)^T. \text{ Since } P(m, n) \text{ is a matrix consisting of ones and zeros}$$

for which  $P(m, n)^{-1} = P(m, n)^T$ , it is a permutation matrix by Lemma 50. Finally, assume

there exists  $A \in M_{mn}$  such that  $\text{vec}(X^T) = A \text{vec}(X)$  for all  $X \in M_{m,n}$ . Hence,

$A \text{vec}(X) = P(m, n) \text{vec}(X)$  implies that  $A \text{vec}(X) - P(m, n) \text{vec}(X) = 0$ . Consequently,



$(A - P(m, n))\text{vec}(X) = 0$  for all  $X \in M_{m,n}$ . Therefore,  $A - P(m, n) = 0$  and the matrix  $P(m, n)$  is unique.  $\square$

Note that when  $m = n$ ,  $P(m, n) = P(n, m)^T = P(n, m)^{-1}$  implies that  $P(n, n) = P(n, n)^T = P(n, n)^{-1}$ . Thus,  $P(n, n)$  is symmetric and its own inverse. Notice also that  $\text{vec}(v^T) = \text{vec}(v)$  for all  $v \in \mathbb{C}^n$  where  $\mathbb{C}^n = M_{n,1}$ . This shows that  $P(n, 1)$  is the identity matrix. Also,  $\text{vec}(v^T) = \text{vec}(v)$  for all  $v \in M_{1,n}$  which means that  $P(1, n)$  is the identity matrix as well.

**Corollary 52:** Let positive integers  $m, n, p$ , and  $q$  be given and let  $P(p, m) \in M_{pm}$  and  $P(n, q) \in M_{nq}$  be the permutation matrices defined in Theorem 51. Then,

$B \otimes A = P(m, p)^T (A \otimes B) P(n, q)$  for all  $A \in M_{m,n}$  and  $B \in M_{p,q}$ . That is,  $B \otimes A$  is always permutation equivalent to  $A \otimes B$ . When  $m = n$  and  $p = q$ ,

$B \otimes A = P(n, p)^{-1} (A \otimes B) P(n, p)$  for all  $A \in M_n$  and  $B \in M_p$ . That is, when  $A$  and  $B$  are square matrices,  $B \otimes A$  is always permutation similar to  $A \otimes B$ . More generally, let

$A_1, \dots, A_k \in M_{m,n}$  and  $B_1, \dots, B_k \in M_{p,q}$  be given. Then,

$$B_1 \otimes A_1 + \dots + B_k \otimes A_k = P(m, p)^T [A_1 \otimes B_1 + \dots + A_k \otimes B_k] P(n, q).$$

**Proof:** Let  $T : M_{n,q} \rightarrow M_{m,p}$  be defined by  $T(X) \equiv AXB^T = Y$ . Then,  $Y^T = BX^T A^T$ . By

Theorem 34, this gives  $\text{vec}(Y) = \text{vec}(AXB^T) = (B \otimes A)\text{vec}(X)$  and

$\text{vec}(Y^T) = \text{vec}(BX^T A^T) = (A \otimes B)\text{vec}(X^T)$  for all  $X \in M_{n,q}$ . Using Theorem 51, write

$\text{vec}(Y^T) = P(m, p)\text{vec}(Y)$  and  $\text{vec}(X^T) = P(n, q)\text{vec}(X)$  which gives that

$P(m, p)\text{vec}(Y) = (A \otimes B)P(n, q)\text{vec}(X)$ . As a result,

$\text{vec}(Y) = P(m, p)^T (A \otimes B) P(n, q) \text{vec}(X)$ . Since  $\text{vec}(Y) = \text{vec}(AXB^T) = (B \otimes A) \text{vec}(X)$ ,

this shows that  $(B \otimes A) \text{vec}(X) = P(m, p)^T (A \otimes B) P(n, q) \text{vec}(X)$  for all  $X \in M_{n, q}$ .

Hence,  $B \otimes A = P(m, p)^T (A \otimes B) P(n, q)$ . If  $A$  and  $B$  are square matrices, then  $m = n$  and

$p = q$ . Substitution shows  $B \otimes A = P(n, p)^T (A \otimes B) P(n, p) = P(n, p)^{-1} (A \otimes B) P(n, p)$ .

Finally, let  $A_1, \dots, A_k \in M_{m, n}$  and  $B_1, \dots, B_k \in M_{p, q}$ . Then,  $B_1 \otimes A_1 + \dots + B_k \otimes A_k =$

$$P(m, p)^T (A_1 \otimes B_1) P(n, q) + \dots + P(m, p)^T (A_k \otimes B_k) P(n, q)$$

$$= P(m, p)^T [A_1 \otimes B_1 + \dots + A_k \otimes B_k] P(n, q). \quad \square$$

Recall, Theorem 24 states that for  $A, B \in M_{m, n}$ ,  $A \otimes B = B \otimes A$  if and only if

either  $A = cB$  or  $B = cA$  for some  $c \in \mathbb{F}$ . In that theorem, the commuting matrices have

the restriction of being the same size. Since it was shown earlier that  $P(1, n)$  and

$P(n, 1)$  are identity matrices for any positive integer  $n$ , it then follows from Corollary 52

that for any given  $x \in \mathbb{C}^m$  and  $y \in \mathbb{C}^n$ ,  $x \otimes y^T = P(1, m)^T (y^T \otimes x) P(n, 1) = y^T \otimes x$ .

Thus, the Kronecker product also commutes for a column vector of any size with a row vector of any size.

**Corollary 53:** Let  $A \in M_{m, n}$ ,  $B \in M_{p, q}$ , and  $C \in M_{s, t}$ . Then,

$$C \otimes A \otimes B = P(mp, s)^T [A \otimes B \otimes C] P(nq, t) = P(p, ms)^T [B \otimes C \otimes A] P(q, nt).$$

Proof:  $A \otimes B \in M_{mp, nq}$  and  $C \otimes A \otimes B = C \otimes (A \otimes B)$ . By Corollary 52,

$$C \otimes (A \otimes B) = P(mp, s)^T [(A \otimes B) \otimes C] P(nq, t) = P(mp, s)^T [A \otimes B \otimes C] P(nq, t). \text{ Also,}$$

$C \otimes A \otimes B = (C \otimes A) \otimes B$  and  $C \otimes A \in M_{sm, tn}$ . Again by Corollary 52,

$$(C \otimes A) \otimes B = P(p, sm)^T [B \otimes (C \otimes A)] P(q, tn) = P(p, ms)^T [B \otimes C \otimes A] P(q, nt). \quad \square$$

**Proposition 54:** Let  $P(m, n) \in M_{mn}$  and  $P(k, l) \in M_{kl}$  be permutation matrices. Then,

$P(m, n) \otimes P(k, l) \in M_{mnkl}$  is a permutation matrix.

Proof: Let  $P(m, n) = [p_{i,j}]$  for  $i, j = 1, 2, \dots, mn$ . Since  $P(m, n)$  is a permutation matrix, for each  $i$  there exists a unique  $j$  such that  $p_{i,j} = 1$  and all other elements in the  $i^{th}$  row are zero. Also, for each  $j$  there exists a unique  $i$  such that  $p_{i,j} = 1$  and all other elements in the  $j^{th}$  column are zero. The same is true for  $P(k, l) = [q_{i,j}]$ .

$$P(m, n) \otimes P(k, l) = \begin{pmatrix} p_{1,1}P(k, l) & \cdots & p_{1,mn}P(k, l) \\ \vdots & \ddots & \vdots \\ p_{mn,1}P(k, l) & \cdots & p_{mn,mn}P(k, l) \end{pmatrix}. \text{ The } r^{th} \text{ column of}$$

$$P(m, n) \otimes P(k, l) \text{ is } \begin{pmatrix} p_{1,s}q_{1,t} \\ \vdots \\ p_{1,s}q_{kl,t} \\ \vdots \\ p_{mn,s}q_{1,t} \\ \vdots \\ p_{mn,s}q_{kl,t} \end{pmatrix}. \text{ Only one of } p_{1,s}, \dots, p_{mn,s} = 1 \text{ and all others are zero.}$$

Also, only one of  $q_{1,t}, \dots, q_{kl,t} = 1$  and all others are zero. Hence, the  $r^{th}$  column of  $P(m, n) \otimes P(k, l)$  has only one element equal to 1 and all others are zero. Since the  $r^{th}$  column was arbitrarily chosen, this holds true for all columns of  $P(m, n) \otimes P(k, l)$ .

Similarly, the  $r^{th}$  row of  $P(m, n) \otimes P(k, l)$  is

$(p_{s,1}q_{1,t}, \dots, p_{s,1}q_{t,kl}, \dots, p_{s,mn}q_{1,t}, \dots, p_{s,mn}q_{t,kl})$ . Only one of  $p_{s,1}, \dots, p_{s,mn} = 1$  and all others are zero. Also, only one of  $q_{1,t}, \dots, q_{t,kl} = 1$  and all others are zero. Hence, the  $r^{th}$  row of  $P(m, n) \otimes P(k, l)$  has only one element equal to 1 and all others are zero. Again, since

the  $r^{th}$  row was arbitrarily chosen, this holds true for all rows of  $P(m, n) \otimes P(k, l)$ .

Therefore,  $P(m, n) \otimes P(k, l)$  is a permutation matrix by Lemma 49 because it has exactly one element in each column is equal to 1 and exactly one element in each row is equal to 1.  $\square$

The next result uses permutation matrices to rewrite matrix equations.

**Corollary 55:** Let  $A \in M_{m, n}$ ,  $B, X \in M_{n, m}$ , and  $C \in M_m$ . Then, the matrix equation

$AX + X^T B = C$  can be rewritten as  $\left[ (I_m \otimes A) + (B^T \otimes I_m) P(n, m) \right] \text{vec}(X) = \text{vec}(C)$  or

$$\left[ (I_m \otimes A) + P(m, m) (B^T \otimes I_m) \right] \text{vec}(X) = \text{vec}(C).$$

**Proof:** Applying Corollary 35 and the fact that  $\text{vec}(\cdot)$  is a linear transformation, the

matrix equation  $AX + X^T B = C$  is equivalent to

$$(I_m \otimes A) \text{vec}(X) + (B^T \otimes I_m) \text{vec}(X^T) = \text{vec}(C). \text{ By Theorem 51, this is equivalent to}$$

$$(I_m \otimes A) \text{vec}(X) + (B^T \otimes I_m) P(n, m) \text{vec}(X) = \text{vec}(C). \text{ Factoring this equation gives}$$

$$\left[ (I_m \otimes A) + (B^T \otimes I_m) P(n, m) \right] \text{vec}(X) = \text{vec}(C). \text{ By Corollary 52, this is equivalent to}$$

$$\left[ (I_m \otimes A) + P(m, m)^T (I_m \otimes B^T) P(m, n) P(n, m) \right] \text{vec}(X) = \text{vec}(C) \text{ which, when}$$

$$\text{simplified, is the same as } \left[ (I_m \otimes A) + P(m, m) (B^T \otimes I_m) \right] \text{vec}(X) = \text{vec}(C). \quad \square$$

When  $A$  and  $B$  are both square matrices, the fact that  $A \otimes B$  is similar to  $B \otimes A$  means that  $A \otimes B$  and  $B \otimes A$  have the same Jordan - Canonical forms. Also, if

$$A = SJS^{-1} \text{ and } B = TKT^{-1}, \text{ where } J \text{ and } K \text{ are the Jordan - Canonical forms of } A \text{ and } B$$

respectively, then,  $A \otimes B = (SJS^{-1}) \otimes (TKT^{-1}) = (S \otimes T)(J \otimes K)(S \otimes T)^{-1}$ . This shows

that  $A \otimes B$  and  $J \otimes K$  are similar. Hence, the Jordan - Canonical form of  $A \otimes B$  is the

same as the form of  $J \otimes K$ . If the Jordan blocks in the Jordan - Canonical form of  $A \in M_n$  are  $J_1, J_2, \dots, J_p$  and those of  $B \in M_m$  are  $K_1, K_2, \dots, K_q$ , then the Jordan - Canonical form of  $A \otimes B$  is  $(J_1 \oplus \dots \oplus J_p) \otimes (K_1 \oplus \dots \oplus K_q)$ . Proposition 28 showed that a Kronecker product can be distributed through a direct sum when the sum is on the left side. Example 29 showed that this is not always true if the direct sum is on the right side. However, using permutation matrices and Corollary 52, the Kronecker product of two direct sums can be rewritten so that  $A \otimes B = (J_1 \oplus \dots \oplus J_p) \otimes (K_1 \oplus \dots \oplus K_q)$  is permutation similar to  $J_1 \otimes K_1 \oplus \dots \oplus J_1 \otimes K_q \oplus \dots \oplus J_p \otimes K_1 \oplus \dots \oplus J_p \otimes K_q$ . Since they are similar, their Jordan - Canonical forms are the same. By its very nature, the Jordan - Canonical form of  $J_1 \otimes K_1 \oplus \dots \oplus J_1 \otimes K_q \oplus \dots \oplus J_p \otimes K_1 \oplus \dots \oplus J_p \otimes K_q$  is itself. Hence, the Jordan - Canonical form of  $A \otimes B$  is the direct sum of the Jordan blocks in the Jordan - Canonical forms of  $J_i \otimes K_j$  for  $i = 1, \dots, p$  and  $j = 1, \dots, q$ . Each pair of Jordan blocks, one associated with an eigenvalue of  $A$  and one associated with an eigenvalue of  $B$ , contributes independently to the Jordan structure of  $A \otimes B$ .

Our next result describes the Jordan structure of  $J_i \otimes K_j$ .

**Theorem 56:** Suppose that in the Jordan - Canonical form of  $A \in M_n$  there exists a Jordan block of size  $p$  associated with an eigenvalue  $\lambda \in \sigma(A)$ . Also suppose that in the Jordan - Canonical form of  $B \in M_m$  there exists a Jordan block of size  $q$  associated with an eigenvalue  $\mu \in \sigma(B)$ . Then, independent of any other eigenvalues or Jordan structure of  $A$  and  $B$ , the contribution of this pair of blocks to the Jordan - Canonical form of  $A \otimes B$  is as follows:

- a) If  $\lambda \neq 0$  and  $\mu \neq 0$ , then associated with the eigenvalue  $\lambda\mu$  there is one Jordan block of size  $p + q - (2k - 1)$  for each  $k = 1, 2, \dots, \min\{p, q\}$
- b) If  $\lambda \neq 0$  and  $\mu = 0$ , then associated with each eigenvalue  $\lambda\mu = 0$  there are  $p$  Jordan blocks of size  $q$ .
- c) If  $\lambda = 0$  and  $\mu \neq 0$ , then associated with each eigenvalue  $\lambda\mu = 0$  there are  $q$  Jordan blocks of size  $p$ .
- d) If  $\lambda = 0$  and  $\mu = 0$ , then associated with each eigenvalue  $\lambda\mu = 0$  there are two Jordan blocks of each size  $k = 1, 2, \dots, \min\{p, q\} - 1$  (which are absent if  $\min\{p, q\} = 1$ ), and  $|p - q| + 1$  blocks of size  $\min\{p, q\}$ .

Proof: Let  $J$  be the Jordan - Canonical form of  $A$  and let  $K$  be the Jordan - Canonical form of  $B$ . Then, the Jordan - Canonical form of  $A \otimes B$  is the same as the form for  $J \otimes K$ . Let  $J_\lambda$  be the Jordan block of  $A$  of size  $p$  associated with the eigenvalue  $\lambda$  and let  $K_\mu$  be the Jordan block of  $B$  of size  $q$  associated with the eigenvalue  $\mu$ . As it has already been established, examination of  $J_\lambda \otimes K_\mu$  will determine the contribution of this pair of blocks.

The first part a) requires a lengthy proof that will not be included here. The interested reader can find a proof of part a) in the article ‘Combinatorial Verification of the Elementary Divisors of Tensor Products’ by R. Brualdi ([1]).

For part b), where  $\lambda \neq 0$ , let  $J_\lambda \in M_p$  be the associated Jordan block of size  $p$  and for  $\mu = 0$ , let  $K_\mu \in M_q$  be the associated Jordan block of size  $q$ . It follows from the structure of a Jordan block that  $J_\lambda$  is of rank  $p$  and  $K_\mu$  is of rank  $q - 1$ . Also, since

$\mu = 0$ ,  $(K_\mu)^q$  has rank zero. From Theorem 20,  $\text{rank}\left((J_\lambda \otimes K_\mu)^q\right)$   
 $= \left(\text{rank}(J_\lambda^q)\right)\left(\text{rank}(K_\mu^q)\right) = (p)(0) = 0$ . Due to the fact that  $\text{rank}\left((J_\lambda \otimes K_\mu)^q\right) = 0$   
and  $\text{rank}\left((J_\lambda \otimes K_\mu)^{q-1}\right) \neq 0$ , the maximum size of the Jordan blocks of  $J_\lambda \otimes K_\mu$ ,  
associated with the eigenvalue  $\lambda\mu = 0$ , is  $q$ ; and there exists at least one Jordan block of  
size  $q$ . Now examine the rank of  $J_\lambda \otimes K_\mu$  raised to the  $q-1$  power. Accordingly,  
 $\text{rank}\left((J_\lambda \otimes K_\mu)^{q-1}\right) = \left(\text{rank}(J_\lambda^{q-1})\right)\left(\text{rank}(K_\mu^{q-1})\right) = (p)(1) = p$ . When going from  
raising  $J_\lambda \otimes K_\mu$  to the  $q-1$  power to raising it to the  $q$  power, the rank went from  $p$  to  
zero. This implies that  $(J_\lambda \otimes K_\mu)^{q-1}$  has  $p$  Jordan blocks of rank one, or, equivalently,  $p$   
Jordan blocks of size two. Hence,  $(J_\lambda \otimes K_\mu)^{q-2}$  has  $p$  Jordan blocks of size three,  
 $(J_\lambda \otimes K_\mu)^{q-3}$  has  $p$  Jordan blocks of size four, ...,  $(J_\lambda \otimes K_\mu)^2$  has  $p$  Jordan blocks of  
size  $q-1$ , and  $J_\lambda \otimes K_\mu$  has  $p$  Jordan blocks of size  $q$ . Since  $J_\lambda \otimes K_\mu \in M_{pq}$ , there can  
be no other blocks. Therefore,  $J_\lambda \otimes K_\mu$  has exactly  $p$  Jordan blocks of size  $q$   
corresponding to  $\lambda\mu = 0$ .

The proof of part c) is obtained by switching  $\lambda$  and  $\mu$ ,  $p$  and  $q$ , and  $J$  and  $K$  in the proof  
for part b).

To prove part d), assume  $\lambda = 0 = \mu$ , let  $J_\lambda \in M_p$  be the associated Jordan block of size  $p$   
and let  $K_\mu \in M_q$  be the associated Jordan block of size  $q$ . It follows from the structure of  
a Jordan block that  $J_\lambda$  is of rank  $p-1$  and  $K_\mu$  is of rank  $q-1$ . In this case,

$\text{rank}\left(J_\lambda \otimes K_\mu\right)^{\min\{p,q\}} = 0$  but  $\text{rank}\left(J_\lambda \otimes K_\mu\right)^{\min\{p,q\}-1} \neq 0$ . Without loss of generality, let  $\min\{p,q\} = p$ . It follows that  $\text{rank}\left(\left(J_\lambda \otimes K_\mu\right)^p\right) = 0(q-p) = 0$ . Hence, the maximum size of the Jordan blocks of  $J_\lambda \otimes K_\mu$ , associated with the eigenvalue  $\lambda\mu = 0$ , is  $p$  and there exists at least one Jordan block of size  $p = \min\{p,q\}$ . Now examine the rank of  $J_\lambda \otimes K_\mu$  raised to the  $p-1$  power. Accordingly,

$\text{rank}\left(\left(J_\lambda \otimes K_\mu\right)^{p-1}\right) = 1(q-p+1) = |p-q|+1$ . When going from raising  $J_\lambda \otimes K_\mu$  to the  $p-1$  power to raising it to the  $p$  power, the rank went from  $|p-q|+1$  to zero. This implies that  $\left(J_\lambda \otimes K_\mu\right)^{p-1}$  has  $|p-q|+1$  Jordan blocks of rank one, or, equivalently,  $|p-q|+1$  Jordan blocks of size two. Hence,  $\left(J_\lambda \otimes K_\mu\right)^{p-2}$  has  $|p-q|+1$  Jordan blocks of size three,  $\left(J_\lambda \otimes K_\mu\right)^{p-3}$  has  $|p-q|+1$  Jordan blocks of size four, ...,  $\left(J_\lambda \otimes K_\mu\right)^2$  has  $|p-q|+1$  Jordan blocks of size  $p-1$ , and  $J_\lambda \otimes K_\mu$  has  $|p-q|+1$  Jordan blocks of size  $p$ . Since  $J_\lambda \otimes K_\mu \in M_{pq}$ , if  $p=1$ , then there can be no other blocks. However, if  $p>1$ , there are other Jordan blocks whose sizes need to be determined. This is done by examining the rank of the remaining powers of  $J_\lambda \otimes K_\mu$ . Next,  $\text{rank}\left(\left(J_\lambda \otimes K_\mu\right)^{p-2}\right) = 2(q-p+2) = 2|p-q|+4$ . These previous  $|p-q|+1$  Jordan blocks now contribute  $2|p-q|+2$  to this rank. Since  $2|p-q|+4 - (2|p-q|+2) = 2$ , this implies that  $\left(J_\lambda \otimes K_\mu\right)^{p-2}$  has two Jordan blocks of rank one (size two). For  $J_\lambda \otimes K_\mu$ , this will result in two blocks of size  $p-1 = \min\{p,q\}-1$ . Now,



$\text{rank}\left(\left(J_\lambda \otimes K_\mu\right)^{p-3}\right) = 3(q-p+3) = 3|p-q|+9$ . The previously determined blocks will

contribute  $3|p-q|+3+(2)(2)$ . Since  $3|p-q|+9-(3|p-q|+3+(2)(2)) = 2$ , this

implies that  $\left(J_\lambda \otimes K_\mu\right)^{p-3}$  has two Jordan blocks of rank one (size two). For  $J_\lambda \otimes K_\mu$ ,

this will result in 2 blocks of size  $p-2 = \min\{p, q\} - 2$ . Now, assume this is true for

$\left(J_\lambda \otimes K_\mu\right)^{r+1}$  and show this is true for  $\left(J_\lambda \otimes K_\mu\right)^r$ . Examining the rank yields

$$\text{rank}\left(\left(J_\lambda \otimes K_\mu\right)^r\right) = (p-r)(q-r) = (p-r)(q-p+p-r) = (p-r)|p-q| + (p-r)^2.$$

The previously determined blocks contribute

$$\begin{aligned} & (p-r)|p-q| + (p-r) + 2(p-r-1 + p-r-2 + \cdots + 2) \\ &= (p-r)|p-q| + (p-r) + 2\left(\frac{(p-r-1)(p-r)}{2} - 1\right) \\ &= (p-r)|p-q| + (p-r) + (p-r-1)(p-r) - 2 = (p-r)|p-q| + (p-r)^2 - 2 \text{ to the rank.} \end{aligned}$$

This again leaves 2 in the rank. This implies that  $\left(J_\lambda \otimes K_\mu\right)^r$  has 2 Jordan blocks of rank

one, or size two. For  $J_\lambda \otimes K_\mu$ , this will result in 2 blocks of size  $r+1$ . Finally,

$$\text{rank}\left(J_\lambda \otimes K_\mu\right) = (p-1)(q-1) = (p-1)|p-q| + (p-1)^2. \text{ The previously determined}$$

blocks contribute  $(p-1)|p-q| + (p-1) + 2(p-2 + p-3 + \cdots + 2)$

$$= (p-1)|p-q| + (p-1)^2 - 2 \text{ to the rank, leaving 2. This implies that } J_\lambda \otimes K_\mu \text{ has 2}$$

Jordan blocks of rank one, or size two. Due to the size of  $J_\lambda \otimes K_\mu$ , there remain only

two Jordan blocks of size one. □

**Theorem 57:** Let  $A \in M_n$  and  $B \in M_m$  be given nonzero matrices. Then,  $A \otimes B$  is diagonalizable if and only if both  $A$  and  $B$  are diagonalizable. Consequently,  $A \otimes B$  is diagonalizable if and only if  $B \otimes A$  is diagonalizable.

*Proof:* First, let both  $A$  and  $B$  be diagonalizable matrices with eigenvalues

$\lambda_1, \dots, \lambda_n$  and  $\mu_1, \dots, \mu_m$  respectively. Then, the Jordan - Canonical forms of both  $A$  and  $B$

have blocks that are all of size one for each eigenvalue, meaning both  $p = 1$  and  $q = 1$ .

Using Theorem 56, if  $\lambda_i$  and  $\mu_j$  are both nonzero, then by part a)

$\min\{p, q\} = \min\{1, 1\} = 1$ . Thus, there will be one associated Jordan block of size

$1+1-(2-1)=1$  in  $A \otimes B$ . If  $\lambda_i$  and  $\mu_j$  are both zero, then by part d) of Theorem 56, there

will be  $|1-1|+1=1$  Jordan blocks of size  $\min\{p, q\} = \min\{1, 1\} = 1$  in  $A \otimes B$ . If one of

$\lambda_i$  and  $\mu_j$  is zero and the other is nonzero, then by parts b) and c) of Theorem 56, there

will be one Jordan block of size one in  $A \otimes B$ .

Since, all Jordan blocks of  $A \otimes B$  are of size one,  $A \otimes B$  is diagonalizable.

Now, the converse of the theorem will be proven. Let  $A \otimes B$  be a diagonalizable matrix with eigenvalues  $\alpha_1, \dots, \alpha_{mn}$ . It follows from Theorem 15 that each  $\alpha_r$  is a product of an

eigenvalue from  $A$  with an eigenvalue from  $B$ . Then, the Jordan - Canonical form of

$A \otimes B$  contains blocks that are all of size one. If  $\alpha_r$  is nonzero for some  $r$ , then there

exists a nonzero eigenvalue of  $A$ ,  $\lambda_i$ , and a nonzero eigenvalue of  $B$ ,  $\mu_j$ , such that

$\lambda_i \mu_j = \alpha_r$ . Hence, part a) of Theorem 56 applies. That is, for  $k = 1, 2, \dots, \min\{p, q\}$ ,

there is one Jordan block of size  $p + q - (2k - 1)$ . For  $k = 1$ , the size of the associated

Jordan block is one because  $\alpha_r$  is an eigenvalue. Consequently,  $1 = p + q - (2(1) - 1)$ .

This implies that  $p + q = 2$ . Since both  $p$  and  $q$  are integers greater than or equal to one,  $p$  and  $q$  have to both be equal to one in order to sum to two. Therefore, the size of the Jordan block of  $A$  that is associated with  $\lambda_i$  is one and the size of the Jordan block of  $B$  that is associated with  $\mu_j$  is one. Since  $k$  reaches a maximum of  $\min\{p, q\} = \min\{1, 1\} = 1$ , there are no other blocks to consider. If  $\alpha_r$  is zero, there are three possibilities for products of eigenvalues from  $A$  and  $B$  that will contribute to the Jordan - Canonical form of  $A \otimes B$ .

- a)  $0 = \alpha_r = \lambda_i \mu_j$  where  $\lambda_i$  is a nonzero eigenvalue of  $A$  and  $\mu_j$  is a zero eigenvalue of  $B$ . In this case, part b) of Theorem 56 applies. This means that there will be  $p$  Jordan blocks of size  $q$  associated with  $\alpha_r$ . Since  $A \otimes B$  is diagonalizable, there is one Jordan block of size one associated with  $\alpha_r$ , or  $p = 1$  and  $q = 1$ .
- b) The proof of part b) is similar to the proof of part a). Switching  $\mu$  and  $\lambda$ , and  $q$  and  $p$  will produce the desired result using part c) of Theorem 56.
- c)  $0 = \alpha_r = \lambda_i \mu_j$  where  $\lambda_i$  is a zero eigenvalue of  $A$  and  $\mu_j$  is a zero eigenvalue of  $B$ .

In this final case, part d) of Theorem 56 applies. Since  $A \otimes B$  is diagonalizable, there is one Jordan block of size one associated with  $\alpha_r$ . Hence, the first statement in part d) of Theorem 56 does not apply. According to the second statement, there are  $|p - q| + 1$  blocks of size  $\min\{p, q\} = 1$ . Since there is only one block of size one associated with  $\alpha_r$ , this means that  $|p - q| + 1 = 1$ . Hence,  $p = q$ . Also,  $1 = \min\{p, q\} = \min\{p, p\} = p$ . This shows that  $1 = p = q$ .

Therefore, both  $p$  and  $q$  are equal to one.

Since all blocks in the Jordan - Canonical form of both  $A$  and  $B$  are of size one, both  $A$  and  $B$  are diagonalizable. □

#### Section 4 – Large Least Squares Problems

In Section 2, the Kronecker product was used to solve linear matrix equations. In this section, it is used to find the least squares solution of  $(A \otimes B)x = t$ . These problems are found in the areas of signal and image processing, fast transform algorithms, and multidimensional approximation. For example, when data is collected by satellites, the matrices  $A$  and  $B$  can be very large, resulting in large systems of linear equations of this form.

The most common approach to solving equations of this type is to form the normal equation  $(A \otimes B)^T (A \otimes B)x = (A \otimes B)^T t$ . The main drawback of this method is the instability associated with the formation of  $A^T A$  and  $B^T B$ , where instability means that small changes in  $A$  or  $B$  would result in large changes in  $A^T A$  and  $B^T B$ , respectively. The method outlined in this section not only avoids the formation of  $A^T A$  and  $B^T B$ , but even the formation of  $(A \otimes B)$ . The majority of information in this section comes from the two articles by Fausett and Fulton ([3], [4]). First, information about permuted  $QR$  factorization is needed.

**Definition 58:** The  $QR$  factorization of a matrix  $A \in M_{m,p}$  decomposes the matrix into the product of an orthogonal matrix ( $Q$ ) and an upper triangular matrix ( $R$ ).

**Definition 59:** The permuted  $QR$  factorization of  $A$  is obtained by multiplying the matrix  $A$  by a permutation matrix on the right so that the diagonal elements of  $R$  are as far away from zero as possible.

Here, getting the diagonal elements of  $R$  as far away from zero is accomplished by performing a modified Householder reduction method on  $A$ . The  $QR$  factorization of  $A$  can be found using elementary reflector matrices of the form  $H = I - 2ww^T$  where  $w$  is a unit vector. First, the lengths of all columns of  $A$  are determined. The column with the largest length is switched with the first column (if it is not the first column) by multiplying a permutation matrix,  $P_1$ , on the right of  $A$ . Then, the matrix is multiplied by an elementary reflector matrix on the left,  $H_1 = I - 2w_1w_1^T$ , such that the first column of

$$H_1(AP_1) \text{ is } \begin{bmatrix} k_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \text{ where } k_1 \text{ is the length of the first column of } AP_1. \text{ Now, ignoring the}$$

first element of each column in the matrix  $H_1(AP_1)$ , the lengths of the columns are determined. Again, the one with the largest length is switched with the second column by multiplying the permutation matrix,  $P_2$ , on the right. Then, the matrix  $(H_1(AP_1))P_2$  is multiplied on the left by  $H_2 = I - 2w_2w_2^T$ , such that the first column remains the same

$$\text{and the second column becomes } \begin{bmatrix} c_{1,2} \\ k_2 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \text{ where } k_2 \text{ is the length of the second column of}$$

$(H_1(AP_1))P_2$  without the first element. This process is continued until an upper

triangular matrix,  $R$ , is produced such that  $R = H_p H_{p-1} \cdots H_2 H_1 A P_1 P_2 \cdots P_{p-1} P_p$   
 $= H_p H_{p-1} \cdots H_2 H_1 A P$  where the diagonal elements of  $R$  are as far away from zero as possible. Since  $H_i$  is a reflector matrix, this equation can be written to

$(H_p H_{p-1} \cdots H_2 H_1)^T R = A P$ . Letting  $Q = (H_p H_{p-1} \cdots H_2 H_1)^T$  gives the permuted  $QR$  factorization of  $A$ , or  $AP = QR$  ([2]).

**Theorem 60:** Let  $A \in M_{m,p}$  with  $m > p$  have permuted  $QR$  factorization

$$AP_A = Q_A R_A = Q_A \begin{pmatrix} \tilde{R}_A \\ 0_A \end{pmatrix} \text{ where } P_A \in M_p \text{ is a permutation matrix, } Q_A \in M_m \text{ is orthogonal,}$$

$R_A \in M_{m,p}$  is upper triangular,  $\tilde{R}_A \in M_p$  is square upper triangular, and  $0_A \in M_{m-p,p}$  is a zero matrix. Also, let  $B \in M_{n,q}$  with  $n > q$  have permuted  $QR$  factorization

$$BP_B = Q_B R_B = Q_B \begin{pmatrix} \tilde{R}_B \\ 0_B \end{pmatrix} \text{ where } P_B \in M_q \text{ is a permutation matrix, } Q_B \in M_n \text{ is orthogonal,}$$

$R_B \in M_{n,q}$  is upper triangular,  $\tilde{R}_B \in M_q$  is square upper triangular, and  $0_B \in M_{n-q,q}$  is a zero matrix. Also, let  $\text{rank}(A) = p$  and  $\text{rank}(B) = q$ . Then,  $A \otimes B$  has permuted  $QR$

factorization  $(AP_A) \otimes (BP_B) = [(Q_A \otimes Q_B) P^T] [P(R_A \otimes R_B)]$  where  $P \in M_{mn}$  is a

permutation matrix defined by the requirement that  $P(R_A \otimes R_B) = \begin{pmatrix} \tilde{R}_A \otimes \tilde{R}_B \\ 0 \end{pmatrix}$  where

$$0 \in M_{mn-pq, pq}.$$

**Proof:**  $(AP_A) \otimes (BP_B) = (Q_A R_A) \otimes (Q_B R_B) = (Q_A \otimes Q_B)(R_A \otimes R_B)$  where  $Q_A \otimes Q_B$  is

orthogonal by Corollary 12. If  $\tilde{R}_A = [r_{i,j}^A]$  and  $\tilde{R}_B = [r_{i,j}^B]$ , then

$$(R_A \otimes R_B) = \begin{pmatrix} \tilde{R}_A \\ 0_A \end{pmatrix} \otimes R_B = \begin{pmatrix} r_{1,1}^A R_B & r_{1,2}^A R_B & \cdots & r_{1,p}^A R_B \\ 0 & r_{2,2}^A R_B & \cdots & r_{2,p}^A R_B \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & r_{p,p}^A R_B \\ 0 & \cdots & \cdots & 0 \\ \vdots & \cdots & \cdots & \vdots \\ 0 & \cdots & \cdots & 0 \end{pmatrix}. \text{ In each of the first } p \text{ block}$$

rows of the Kronecker product, the bottom  $n - q$  rows are zero. This is due to the fact that the bottom  $n - q$  rows of  $R_B$  are zero. Therefore,  $R_A \otimes R_B$  does not have the required upper triangular form for a  $QR$  factorization. This can be changed using a permutation matrix, specifically one that performs the necessary row changes to move the discussed zero rows below all nonzero rows. This permutation matrix,  $P$ , is the one

described in the statement of the theorem,  $P(R_A \otimes R_B) = \begin{pmatrix} \tilde{R}_A \otimes \tilde{R}_B \\ 0 \end{pmatrix}$ . Since both  $\tilde{R}_A$  and

$\tilde{R}_B$  are both square upper triangular, their Kronecker product is also square upper

triangular. As stated in Theorem 51,  $P^{-1} = P^T$ , which means  $I = P^T P$  and  $P$  is

orthogonal. Inserting  $I$  into the factorization yields  $(AP_A) \otimes (BP_B)$

$$= (Q_A \otimes Q_B)(R_A \otimes R_B) = (Q_A \otimes Q_B)I(R_A \otimes R_B) = (Q_A \otimes Q_B)P^T P(R_A \otimes R_B)$$

$$= [(Q_A \otimes Q_B)P^T][P(R_A \otimes R_B)] \text{ where } (Q_A \otimes Q_B)P^T \text{ is orthogonal and } P(R_A \otimes R_B) \text{ is}$$

upper triangular. □

**Corollary 61:** Let  $R = \tilde{R}_A \otimes \tilde{R}_B \in M_{pq}$ , where  $\tilde{R}_A$  and  $\tilde{R}_B$  are the matrices described

above in Theorem 60, be the  $pq \times pq$  upper triangular part of  $P(R_A \otimes R_B)$ . Then,

$$R^T R = [(AP_A) \otimes (BP_B)]^T [(AP_A) \otimes (BP_B)].$$

$$\begin{aligned}
\text{Proof: } & \left[ (AP_A) \otimes (BP_B) \right]^T \left[ (AP_A) \otimes (BP_B) \right] = \left[ (Q_A R_A) \otimes (Q_B R_B) \right]^T \left[ (Q_A R_A) \otimes (Q_B R_B) \right] \\
& = \left[ (Q_A \otimes Q_B) (R_A \otimes R_B) \right]^T \left[ (Q_A \otimes Q_B) (R_A \otimes R_B) \right] \\
& = (R_A \otimes R_B)^T (Q_A \otimes Q_B)^T (Q_A \otimes Q_B) (R_A \otimes R_B) = (R_A \otimes R_B)^T I (R_A \otimes R_B) \\
& = (R_A \otimes R_B)^T P^T P (R_A \otimes R_B) = \left[ P (R_A \otimes R_B) \right]^T \left[ P (R_A \otimes R_B) \right] = \begin{pmatrix} \tilde{R}_A \otimes \tilde{R}_B \\ 0 \end{pmatrix}^T \begin{pmatrix} \tilde{R}_A \otimes \tilde{R}_B \\ 0 \end{pmatrix} \\
& = \begin{pmatrix} (\tilde{R}_A \otimes \tilde{R}_B)^T & 0^T \end{pmatrix} \begin{pmatrix} \tilde{R}_A \otimes \tilde{R}_B \\ 0 \end{pmatrix} = (\tilde{R}_A \otimes \tilde{R}_B)^T (\tilde{R}_A \otimes \tilde{R}_B) = R^T R. \quad \square
\end{aligned}$$

Using the same matrices as in the previous two results, the rectangular least squares problem with the overdetermined system  $(A \otimes B)x = t$ , where  $A$  and  $B$  are real matrices with  $A \in M_{m,p}$  and  $m > p$  and  $B \in M_{n,q}$  and  $n > q$  and  $\text{rank}(A) = p$  and  $\text{rank}(B) = q$ , can be readily solved. Note,  $x \in \mathbb{R}^{pq}$  and  $t \in \mathbb{R}^{mn}$ . The least squares solution of  $(A \otimes B)x = t$  is accomplished by transforming this system into a square block upper triangular system of dimension  $pq \times pq$ , and then finding the exact solution to the new system.

Inserting  $I = (P_A \otimes P_B)(P_A \otimes P_B)^T$  into the system of equations produces

$$(A \otimes B)(P_A \otimes P_B)(P_A \otimes P_B)^T x = (AP_A \otimes BP_B)(P_A^T \otimes P_B^T)x = t. \text{ From Corollary 61,}$$

$$R^T R = \left[ (AP_A) \otimes (BP_B) \right]^T \left[ (AP_A) \otimes (BP_B) \right]. \text{ Hence, multiplying}$$

$$(AP_A \otimes BP_B)(P_A^T \otimes P_B^T)x = t \text{ on the left by } \left[ (AP_A) \otimes (BP_B) \right]^T \text{ generates}$$

$$R^T R (P_A^T \otimes P_B^T)x = \left[ (AP_A) \otimes (BP_B) \right]^T t. \text{ Letting } y = (P_A^T \otimes P_B^T)x \in \mathbb{R}^{pq} \text{ be a}$$

$$\text{permutation of the vector } x \text{ and using Theorem 60, } R^T R y = \left[ (AP_A) \otimes (BP_B) \right]^T t$$



$$= \left( \left[ (Q_A \otimes Q_B) P^T \right] \left[ P(R_A \otimes R_B) \right] \right)^T t = \left[ P(R_A \otimes R_B) \right]^T \left[ (Q_A \otimes Q_B) P^T \right]^T t. \text{ In Corollary}$$

61, it was stated that  $R$  is the square  $pq \times pq$  upper triangular part of  $P(R_A \otimes R_B)$ .

$$\text{Hence, it follows that } R^T R y = \begin{bmatrix} R \\ 0 \end{bmatrix}^T \left[ P(Q_A^T \otimes Q_B^T) \right] t = \begin{bmatrix} R^T & 0 \end{bmatrix} \left[ P(Q_A^T \otimes Q_B^T) \right] t$$

where  $R \in M_{pq}$  and  $0 \in M_{pq, mn-pq}$ . Letting  $h = \left[ P(Q_A^T \otimes Q_B^T) \right] t \in \mathbb{R}^{mn}$  gives

$$R^T R y = \begin{bmatrix} R^T & 0 \end{bmatrix} h. \text{ Since the last } mn - pq \text{ components of } h \text{ do not contribute to the}$$

final vector on the right hand side of the equation, define  $[h]_{pq} \in \mathbb{R}^{pq}$  to be the vector

containing the first  $pq$  components of  $h = \left[ P(Q_A^T \otimes Q_B^T) \right] t$ . Hence,  $R^T R y = R^T [h]_{pq}$ .

Since  $A$  and  $B$  were assumed full rank and each matrix has more rows than columns, it follows from Theorem 60 that all diagonal elements of  $\tilde{R}_A$  and  $\tilde{R}_B$  are nonzero. Hence,  $R = \tilde{R}_A \otimes \tilde{R}_B$  is nonsingular and, in turn,  $R^T$  is nonsingular.

Multiplying both sides of the equation  $R^T R y = R^T [h]_{pq}$  by  $(R^T)^{-1}$  yields

$$R y = [h]_{pq} = (\tilde{R}_A \otimes \tilde{R}_B) y. \text{ Recall that } \tilde{R}_A = [r_{i,j}^A]. \text{ This implies that}$$

$$\begin{pmatrix} r_{1,1}^A R^B & r_{1,2}^A R^B & \dots & r_{1,p}^A R^B \\ 0 & r_{2,2}^A R^B & \dots & r_{2,p}^A R^B \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & r_{p,p}^A R^B \end{pmatrix} \begin{pmatrix} y^{(1)} \\ y^{(2)} \\ \vdots \\ y^{(p)} \end{pmatrix} = \begin{pmatrix} h^{(1)} \\ h^{(2)} \\ \vdots \\ h^{(p)} \end{pmatrix} \text{ where } y^{(i)} = \begin{pmatrix} y_1^{(i)} \\ y_2^{(i)} \\ \vdots \\ y_q^{(i)} \end{pmatrix} \text{ for } i = 1, \dots, p \text{ and}$$

$$h^{(i)} = \begin{pmatrix} h_1^{(i)} \\ h_2^{(i)} \\ \vdots \\ h_q^{(i)} \end{pmatrix} \text{ for } i = 1, \dots, p. \text{ This equation can be solved by block back substitution to}$$

obtain  $y$ . The original least squares solution,  $x$ , is obtained by left multiplication by

$$P_A \otimes P_B. \text{ That is } (P_A \otimes P_B)y = (P_A \otimes P_B)(P_A^T \otimes P_B^T)x = [(P_A P_A^T) \otimes (P_B P_B^T)]x \\ = (I \otimes I)x = x.$$

Unfortunately, the vector  $h = P(Q_A^T \otimes Q_B^T)t$  is not easily determined because it is hard to calculate  $P$ . However,  $P$  can be avoided by noting that  $(Q_A^T \otimes Q_B^T)t$

$$= [(Q_A^T I_m) \otimes (I_n Q_B^T)]t = (Q_A^T \otimes I_n)(I_m \otimes Q_B^T)t = (Q_A^T \otimes I_n)(I_m \otimes Q_B^T)t \\ = (Q_A^T \otimes I_n) \begin{pmatrix} Q_B^T t^{(1)} \\ Q_B^T t^{(2)} \\ \vdots \\ Q_B^T t^{(m)} \end{pmatrix} \text{ where } t^{(i)} = \begin{pmatrix} t_1^{(i)} \\ t_2^{(i)} \\ \vdots \\ t_n^{(i)} \end{pmatrix} \text{ for } i = 1, \dots, m. \text{ Since } P \text{ is the permutation}$$

matrix such that  $P(R_A \otimes R_B)$  is upper triangular (Theorem 60), it follows that in the vector  $h = P(Q_A^T \otimes Q_B^T)t$ ,  $P$  performs the same interchanges on  $(Q_A^T \otimes Q_B^T)t$ . Hence, the elements in  $[h]_{pq}$  are those of  $(Q_A^T \otimes Q_B^T)t$  which are located in the same rows of  $R_A \otimes R_B$  that have nonzero elements. Now, once  $A$  and  $B$  are factored into their respective  $QR$  factorizations, the locations of the rows of  $R_A \otimes R_B$  with nonzero elements is known. Hence, the matrix  $P$  does not need to be formed.

The rows of  $R_A \otimes R_B$  that are zero (and omitted from  $\tilde{R}_A \otimes \tilde{R}_B$ ) along with the corresponding components of  $(Q_A^T \otimes Q_B^T)t$  represent  $p(n-q) + n(m-p)$  equations in the original linear system  $(A \otimes B)x = t$ . These equations usually will not be satisfied by the solution  $x$ . Hence,  $x$  is a least squares solution instead of an exact solution. Due to

the great computational effort needed to find  $(R^T)^{-1}$  in most cases, improvements have been made on this algorithm so that this inverse no longer needs to be calculated ( [4] ).

## Conclusion

The Kronecker product is a useful way to multiply matrices. It has several advantages over conventional matrix multiplication. For example, there are no dimension restrictions on the factors, it is relatively easy to find the eigenvalues of a product, and it is easy to know when a matrix product will be the zero matrix. The Kronecker product also, has properties in common with conventional matrix multiplication. Two such properties are it is not commutative and the distributive property holds for it. Also, qualities of its factors are carried over into their Kronecker product. For example, the Kronecker product of two upper triangular matrices is also upper triangular, the Kronecker product of two positive (semi) definite matrices is positive (semi) definite, and the Kronecker product of two unitary matrices is unitary.

Several important and beneficial statements were proven about the Kronecker product. Unlike conventional matrix multiplication, a Kronecker product of two matrices is equal to zero if and only if one of the factors is zero. The mixed product property,  $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$ , revealed how conventional multiplication and the Kronecker multiplication can be exchanged. The inverse, transpose, and conjugate transpose of a Kronecker product is the Kronecker product of the respective operation on each factor. For example,  $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$ . The set of eigenvalues of the Kronecker product of  $A$  and  $B$  are the set of all possible products of an eigenvalue of  $A$

and an eigenvalue of  $B$ , which implies that the eigenvalues of  $A \otimes B$  are the same as the eigenvalues of  $B \otimes A$ . This helped prove many interesting facts about determinants, traces, and singular values of Kronecker products. It also was used to explain properties and solutions of certain linear transformations. The Kronecker product is highly non-commutative. That is, severe restrictions need to be placed on  $A$  and  $B$  to get

$A \otimes B = B \otimes A$ . Specifically, when  $A$  and  $B$  have the same dimensions,  $A \otimes B = B \otimes A$  if and only if  $A$  is a multiple of  $B$  or  $B$  is a multiple of  $A$ , and when  $A$  and  $B$  have different dimensions,  $A \otimes B = B \otimes A$  if and only if all elements of  $A$  are equal to each other and all elements of  $B$  are equal to each other.

One of the many uses of the Kronecker product is to solve or determine properties of solutions to linear matrix equations and linear transformations. Matrix equations can be converted into an equivalent system of equations in which the coefficient matrix involves the Kronecker product. This transformation is accomplished using the natural  $\text{vec}(\cdot)$  function. Kronecker products can also be used to discover properties of potential solutions. For example, the Kronecker product was used to show that the linear matrix equation  $AXB = C$  has a unique solution  $X$  for every given  $C$  if and only if both  $A$  and  $B$  are nonsingular, and that the generalized commutativity equation,  $AX - XA = C$ , does not have a unique solution.

Although it is generally not the case that  $A \otimes B = B \otimes A$ , it is true that for any matrices  $A$  and  $B$ ,  $A \otimes B$  is permutation equivalent to  $B \otimes A$ . More is true when  $A$  and  $B$  are square matrices. In this case,  $A \otimes B$  is permutation similar to  $B \otimes A$ . This helps explain why many properties of  $A \otimes B$  are the same as those of  $B \otimes A$ . Permutation matrices were also used to rewrite matrix equations into a system of linear equations

where the unknown is transposed in the matrix equation. For example, the matrix equation  $AX + X^T B = C$  can be converted to a system of linear equations as

$$\left[ (I_m \otimes A) + (B^T \otimes I_m) P(n, m) \right] \text{vec}(X) = \text{vec}(C) \text{ for}$$

$$A \in M_{m,n}, B, X \in M_{n,m}, \text{ and } C \in M_m.$$

The Jordan - Canonical form of a Kronecker product was also examined. When  $A$  and  $B$  are both square matrices, the fact that  $A \otimes B$  is similar to  $B \otimes A$  means that  $A \otimes B$  and  $B \otimes A$  have the same Jordan - Canonical forms. Also, if the Jordan - Canonical form of  $A$  is  $J$  and the Jordan - Canonical form of  $B$  is  $K$ , then  $A \otimes B$  is similar to  $J \otimes K$ . This fact lead to the description of the Jordan - Canonical form of  $A \otimes B$  in terms of the Jordan blocks of  $J$  and  $K$ . Once the Jordan - Canonical form of a Kronecker product was described, it was shown that  $A \otimes B$  is diagonalizable if and only if both  $A$  and  $B$  are diagonalizable.

The Kronecker product is also used to find the least squares solution of a system of equations. These least squares problems are found in the areas of signal and image processing, fast transform algorithms, and multidimensional approximation. The overdetermined system examined in this thesis was  $(A \otimes B)x = t$ . This problem was transformed into a square block upper triangular system whose exact solution is the least squares solution of  $(A \otimes B)x = t$ .

## Bibliography

- 1) R. Brualdi, Combinatorial Verification of the Elementary Divisors of Tensor Products, *Linear Algebra Applications*. (1985).
- 2) C. Cullen, *An Introduction to Numerical Linear Algebra* (PWS Publishing Company, Boston, MA, 1994).
- 3) D.W. Fausett and C.T. Fulton, Large least squares problems involving Kronecker products, *SIAM J. Matrix Anal. Appl.* 15 (1994) 219-227.
- 4) D.W. Fausett, C.T. Fulton, and H. Hashish, Improved parallel QR method for large least squares problems involving Kronecker products, *J. Comput. Appl. Math.* 78 (1997) 63-78.
- 5) R.A. Horn and C.R. Johnson, *Matrix Analysis* (Cambridge University Press, Cambridge, 1985).
- 6) R.A. Horn and C.R. Johnson, *Topics in Matrix Analysis* (Cambridge University Press, Cambridge, 1991).
- 7) C.F. Van Loan, The ubiquitous Kronecker product, *J. Comput. Appl. Math.* 123 (2000) 85-100.