Journal of Theoretical and Applied Information Technology
10th February 2013. Vol. 48 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

DERIVATIVES OF KRONECKER PRODUCTS THEMSELVES BASED ON KRONECKER PRODUCT AND MATRIX CALCULUS

[1] XIAOFENG WANG, WENYAN YANG, BAOGUANG SUN
[1]College of Mathematical and Physical Sciences, Chongqing University of Science and Technology, Chongqing, China

ABSTRACT

In some fields, Kronecker product has been used extensively. In this paper, we first review properties and definitions of Kronecker. Then, we derive two properties of the derivatives of matrices with respect to matrices in terms of the proposed concept. Finally, this paper presents a novel method and new insight for proving derivatives of Kronecker product themselves by using the concepts and properties of Kronecker Products and Matrix Calculus.

Keywords: *Kronecker product; Matrix Calculus; Vec operator; Derivatives of Kronecker products*

1. INTRODUCTION

The interest in the Kronecker product has grown recently. Kronecker product which was named after German mathematician Leopold is a special operator for multiplication of two matrices. It is important for us that Kronecker product simplifies the notation of many algorithms.

Kronecker product arises in many different areas of science and engineering[4], which allows more elegant and compact derivations and has important applications in many fields, including computer vision[6,11], statistics[9], control and matrix equation[2,10]. Especially, the use of many information theories for Kronecker product is used widely. Several trends in the development of scientific computing suggest that Kronecker product operator will have a greater role to play in the future. But the rules and properties of Kronecker product are little discussed, even books on mathematical aspects little discussing the properties and applications of Kronecker product in very short without any explanations of its rules and properties. That is why we are discussing the rules and properties.

In this paper, we introduce some concepts of Kronecker product and its application in expressing, simplifying, and implementing Matrix Calculus. We discuss some results which will be found very useful for the development of the theory of both Kronecker product and matrix differentiation. Finally, we prove derivatives of Kronecker product themselves.

The paper is organized as follows: In Section 2, we review briefly some properties of Kronecker product and the vec operator, which together provide a compact notation. Section 3 then derive two properties of the derivatives of matrices with respect to matrices. In Section 4, we achieve derivatives of Kronecker product themselves. Section 5 concludes.

2. DEFINITIONS AND PROPERTIES OF KRONECKER PRODUCT

Let us review some basic concepts of Kronecker product and matrix calculus for understanding the proofs of the theorems presented in the following sections, which are useful to establish derivatives of Kronecker products themselves [12].

Definition 2.1. Let $A\in R^{m\times n}$, $B\in R^{p\times q}$, then Kronecker product of A and B is defined as the matrix[1,3], Kronecker product is also known as a direct product or a tensor product.

$$A\otimes B=\begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{bmatrix}\in R^{mp\times nq} \quad (1)$$

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Definition 2.2. Let $A \in R^{m\times n}$, then the operator $vec(A)$ is defined to create a vector by stringing together, one-by-one, the columns of matrix $A \in R^{m\times n}$ [5]

$$vec(A) = \begin{bmatrix} A(:,1) \\ \cdots \\ A(:,n) \end{bmatrix} \quad (2)$$

Theorem 2.1. For any three matrices $A \in R^{m\times n}$, $B \in R^{p\times q}$ and $X \in R^{n\times p}$

$$vec(AXB) = (B^T \otimes A)vec(X) \quad (3)$$

Theorem 2.2. Define the matrix $T_{m,n}$ as the matrix that transforms $vec(A)$ into $vec(A^T)$:

$$T_{m,n}vec(A) = vec(A^T) \quad (4)$$

3. DERIVATIVES OF KRONECKER PRO- DUCTS

Theorem 3.1. Determine the derivative of $dA^T A/dA$ where A is $m\times n$

$$\frac{dA^T A}{dA} = (I_n \otimes A^T) + (A^T \otimes I_n)T_{m,n} \quad (5)$$

Proof. Using the relationship between the vec and Kronecker product operators, this can be used to determine the derivative of $dA^T A/dA$, where A is $m\times n$.

$$vec(A^T A) = (I_n \otimes A^T)vec(A) = (A^T \otimes I_n)vec(A^T) \quad (6)$$

That we define derivatives of matrices with respect to matrices is achieved by vectorizing the matrices, Therefore, derivatives of matrices $dA(X)/dX$ are the same as $dvec(A(X))/dvec(X)$, then we use the product rule and get

$$\frac{dA^T A}{dA} = \frac{dvec(A^T A)}{d(A)} = (I_n \otimes A^T) + (A^T \otimes I_n)T_{m,n} \quad (7)$$

This can be simplified

$$(A^T \otimes I_n)T_{m,n} = T_{n,n}(I_n \otimes A^T) \quad (8)$$

Thus

$$\frac{dA^T A}{dA} = (I_{n^2} + T_{m,n})(I_n \otimes A^T) \quad (9)$$

Theorem 3.2. Determine the derivative of $dA^{-1}A/dA$, where A is $m\times n$

$$\frac{dA^{-1}A}{dA} = (I_n \otimes A^{-1}) + (A^T \otimes I_n)\frac{dA^{-1}}{dA} \quad (10)$$

4. DERIVATIVES OF KRONECKER PRO− DUCTS THEMSELVES

In some fields, it is usually useful to establish derivatives of Kronecker products themselves [8]. Therefore, we achieve derivatives of Kronecker products themselves by using Kronecker Products and Matrix Calculus. In this section, we derive formulas about Kronecker products themselves as follow:

Theorem 4.1. For any two matrices $A \in R^{m\times n}$, $B \in R^{p\times q}$, the matrix product $\frac{dA\otimes B}{dA}$ is defined, as follow:

$$\frac{dA\otimes B}{dA} = \begin{bmatrix} \Psi_1 & 0 & \cdots & 0 \\ \Psi_2 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ \Psi_q & 0 & \cdots & 0 \\ 0 & \Psi_1 & \cdots & 0 \\ 0 & \Psi_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \Psi_q & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \Psi_1 \\ 0 & 0 & \cdots & \Psi_2 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \Psi_q \end{bmatrix} = I_n \otimes \begin{bmatrix} \Psi_1 \\ \Psi_2 \\ \cdots \\ \Psi_q \end{bmatrix}$$

where

Journal of Theoretical and Applied Information Technology
10th February 2013. Vol. 48 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

$$\Psi_i = \begin{bmatrix} B_{1i} & 0 & \cdots & 0 \\ B_{2i} & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ B_{pi} & 0 & \cdots & 0 \\ 0 & B_{1i} & \cdots & 0 \\ 0 & B_{2i} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & B_{pi} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & B_{1i} \\ 0 & 0 & \cdots & B_{2i} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & B_{pi} \end{bmatrix} = I_m \otimes B_{.i}$$

Proof. Let $A \in R^{m\times n}$, $B \in R^{p\times q}$, Then with the definition of Kronecker product, we vectorize the matrix. That is,

$$vec(A\otimes B) = [a_{11}B_{11} \quad \cdots \quad a_{11}B_{p1}, a_{21}B_{11} \quad \cdots \quad a_{21}B_{p1}, \cdots, a_{m1}B_{11} \quad \cdots \quad a_{m1}B_{p1}, a_{12}B_{11} \quad \cdots \quad a_{12}B_{p1}, \cdots, \quad a_{mn}B_{pq} \quad \cdots \quad a_{mn}B_{pq}]^T$$

and likewise

$$vec(A) = [a_{11} \quad a_{21} \quad \cdots \quad a_{m1}, a_{12} \quad a_{22} \quad \cdots \quad a_{m2}, \cdots, a_{1n} \quad a_{2n} \quad \cdots \quad a_{mn}]^T$$

Because $dA(X)/dX$ is the same thing as $dvec(A(X))/dvec(X)$. This is where the relationship between the vec operator and Kronecker products is useful [11]. That is

$$\frac{dA\otimes B}{dA} = \frac{d(vec(A\otimes B))}{d(vec(A))} \qquad (11)$$

With the properties of the Kronecker product, we derive a different expression. The derivative of $vec(A\otimes B)$ is

$$\frac{d(vec(A\otimes B))}{d(vec(A))} = \begin{bmatrix} \frac{\partial a_{11}B_{11}}{\partial a_{11}} & \frac{\partial a_{11}B_{11}}{\partial a_{21}} & \cdots & \frac{\partial a_{11}B_{11}}{\partial a_{mn}} \\ \frac{\partial a_{11}B_{21}}{\partial a_{11}} & \frac{\partial a_{11}B_{21}}{\partial a_{21}} & \cdots & \frac{\partial a_{11}B_{21}}{\partial a_{mn}} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial a_{11}B_{p1}}{\partial a_{11}} & \frac{\partial a_{11}B_{p1}}{\partial a_{21}} & \cdots & \frac{\partial a_{11}B_{p1}}{\partial a_{mn}} \\ \frac{\partial a_{21}B_{11}}{\partial a_{11}} & \frac{\partial a_{21}B_{11}}{\partial a_{21}} & \cdots & \frac{\partial a_{21}B_{11}}{\partial a_{mn}} \\ \frac{\partial a_{21}B_{21}}{\partial a_{11}} & \frac{\partial a_{21}B_{21}}{\partial a_{21}} & \cdots & \frac{\partial a_{21}B_{21}}{\partial a_{mn}} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial a_{21}B_{p1}}{\partial a_{11}} & \frac{\partial a_{21}B_{p1}}{\partial a_{21}} & \cdots & \frac{\partial a_{21}B_{p1}}{\partial a_{mn}} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial a_{m1}B_{11}}{\partial a_{11}} & \frac{\partial a_{m1}B_{11}}{\partial a_{21}} & \cdots & \frac{\partial a_{m1}B_{11}}{\partial a_{mn}} \\ \frac{\partial a_{m1}B_{21}}{\partial a_{11}} & \frac{\partial a_{m1}B_{21}}{\partial a_{21}} & \cdots & \frac{\partial a_{m1}B_{21}}{\partial a_{mn}} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial a_{m1}B_{p1}}{\partial a_{11}} & \frac{\partial a_{m1}B_{p1}}{\partial a_{21}} & \cdots & \frac{\partial a_{m1}B_{p1}}{\partial a_{mn}} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial a_{mn}B_{pq}}{\partial a_{11}} & \frac{\partial a_{mn}B_{pq}}{\partial a_{21}} & \cdots & \frac{\partial a_{mn}B_{pq}}{\partial a_{mn}} \end{bmatrix} \qquad (11)$$

Then, we have

$$\frac{dA\otimes B}{dA} = \frac{d(vec(A\otimes B))}{d(vec(A))} = \begin{bmatrix} B_{11} & 0 & \cdots \\ \cdots & \cdots & \cdots \\ B_{p1} & 0 & \cdots \\ 0 & B_{11} & \cdots \\ \cdots & \cdots & \cdots \\ 0 & B_{p1} & \cdots \\ \cdots & \cdots & \cdots \\ B_{11} & 0 & \cdots \\ \cdots & \cdots & \cdots \\ B_{p1} & 0 & \cdots \\ 0 & B_{11} & \cdots \\ \cdots & \cdots & \cdots \\ 0 & B_{p1} & \cdots \\ \cdots & \cdots & \cdots \end{bmatrix}$$

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Thus, it can be written compactly as

$$\frac{dA\otimes B}{dA}=\begin{bmatrix}\Psi_1 & 0 & \cdots & 0\\ \Psi_2 & 0 & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots\\ \Psi_q & 0 & \cdots & 0\\ 0 & \Psi_1 & \cdots & 0\\ 0 & \Psi_2 & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots\\ 0 & \Psi_q & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & \cdots & \Psi_1\\ 0 & 0 & \cdots & \Psi_2\\ \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & \cdots & \Psi_q\end{bmatrix}=I_n\otimes\begin{bmatrix}\Psi_1\\ \Psi_2\\ \cdots\\ \Psi_q\end{bmatrix}$$

where

$$\Psi_i=\begin{bmatrix}B_{1i} & 0 & \cdots & 0\\ B_{2i} & 0 & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots\\ B_{pi} & 0 & \cdots & 0\\ 0 & B_{1i} & \cdots & 0\\ 0 & B_{2i} & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots\\ 0 & B_{pi} & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & \cdots & B_{1i}\\ 0 & 0 & \cdots & B_{2i}\\ \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & \cdots & B_{pi}\end{bmatrix}=I_m\otimes B_{\cdot i}$$

5. CONCLUSION

The Kronecker producer, which is a very powerful matrix multiplication tool, has became more and more important pro-blems in various fields, such as 3D Computer Vision pro-blems, statistics, economics, control ,matrix equation, and so on. In this paper, we have presented a novel approach for acquiring derivatives of Kronecker products themselves.

We start by introducing the Kronecker product and describing the concepts and properties of the Kronecker product, and conclude by describing one application of the Kronecker product., then we derive two properties of the derivatives of matrices with respect to matrices and finally compute derivatives of Kronecker products themselves. This work provides new insights into acquiring derivatives of Kronecker products themselves.

ACKNOWLEDGEMENTS

This work is supported by the Science and Technology Foundation of the Education Department of Chongqing (KJ121404), Innovation team of Chongqing University of Science & Technology (1809013) and Research Foundation of Chongqing University of Science & Technology (CK2011Z15).

REFERENCES:

[1] H. C. Rews, J. Kane, “Kronecker matrices, computer implementation, and generalized spectra”, *J. Assoc. Comput. Mach.*, Vol.17, 1970 , pp. 260-268.

[2] Charles F.Van Loan. “Compulatational frame-works for the Fast Fourier Transform”. *SIAM*, 1992.

[3] Abadir, K.M , Magnus J.R., “Matrix Algebra”, *Cambridge University Press*, 2005.

[4] G.Strang, “Computational Science and Eng-ineering”, *Wellesley Cambridge Press*, 2007.

[5] P. Lancaster, M. Tismenetsky, “The Theory of Matrices”, second ed., *Academic Press*, New York, 1985.

[6] Andrea Fusiello, “A matter of notation: Several uses of the Kronecker product in 3D computer vision Molecular Evolution and Phylogenetics”, *Pattern Recognition Letters*, Vol.28, 2007, pp. 2127–2132.

[7] Fajun Yu, Li Li, “A new method to construct the integrable coupling system for discrete soliton equation with the Kronecker product”, *Physics Letters A*, Vol.372, 2008, pp. 3548–3554.

[8] Nguyen Van Khang, “Kronecker product and a new matrix form of Lagrangian equations with multipliers for constrained multibody systems”,

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Mechanics Research Communications, Vol.38, pp. 294–299, 2011.

[9] Amy N.Langville, William J.Stewart, "The Kronecker product and stochastic automata networks", *Journal of Computational and Applied Mathematics*, Vol.167, 2004, pp. 429–447.

[10] Donald W. Fausett, Charles T. Fulton, Hany Hashish, "Improved parallel QR method for large least squares problems involving Kronecker roducts", *Journal of Computational and Applied Mathematics*, Vol.78, 1997, pp. 63-78.

[11] Andrea Fusiello, "A matter of notation: Several uses of the Kronecker product in 3D computer vision", *Pattern Recognition Letters*, Vol.28, 2007, pp. 2127–2132.

[12] J. Nagy, M. Ng, L. Perrone, "Kronecker product approximations for image restoration with reflexive boundary conditions", *SIAM J. Matrix Anal. Appl.* Vol. 26, 2004, pp. 829–841.