

PARTIAL DERIVATIVE OF MATRIX FUNCTIONS WITH RESPECT TO A VECTOR VARIABLE

Nguyen Van Khang

Hanoi University of Technology, Vietnam

Abstract. The partial derivatives of scalar functions and vector functions with respect to a vector variable are defined and used in dynamics of multibody systems. However the partial derivative of matrix functions with respect to a vector variable is also still limited. In this paper firstly the definitions of partial derivatives of scalar functions, vector functions and matrix functions with respect to a vector variable are represented systematically. After an overview of the matrix calculus related to Kronecker products is presented. Two theorems which specify the relationship between the time derivative of a matrix and its partial derivative with respect to a vector, and the partial derivative of product of two matrices with respect to a vector, are then proved.

1. INTRODUCTION

The partial derivatives with respect to a vector variable of scalar functions, vector functions and matrix functions have many practical applications in dynamics and control of mechanical systems [1-10]. The partial derivatives with respect to a vector variable of scalar functions and vector functions are defined and used in dynamics of multibody systems [1- 5] also in robot dynamics [6-10]. However, the investigation of the partial derivative of matrix functions with respect to a vector variable is still limited [6].

The purpose of this paper is to review the definitions of partial derivatives of scalar function and vector function with respect to a vector variable. Based on these definitions, the concept of partial derivative of matrix function with respect to a vector variable is defined in Sec. 2. The matrix product and the Kronecker product are reviewed in Sec. 3. The proofs of two theorems which specify the relationship between the time derivative of a matrix and its partial derivative with respect to a vector, and the partial derivative of product of two matrices with respect to a vector, are represented in Secs. 4 and 5.

2. PARTIAL DERIVATIVES OF A SCALAR, A VECTOR AND A MATRIX WITH RESPECT TO A VECTOR

2.1. Partial derivatives of scalar with respect to a vector

In the paper, vectors are represented in column forms. For example, a vector $\mathbf{x} \in \mathbb{R}^n$ is an n -dimensional column vector

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \mathbf{x} \in \mathbb{R}^n \quad (1)$$

and its row form is represented by \mathbf{x}^T , where $(*)^T$ indicates the transpose of vector or matrix $(*)$

$$\mathbf{x}^T = [x_1 \ x_2 \ \dots \ x_n]. \quad (2)$$

Let a scalar $\alpha(\mathbf{x})$ be a function of vector \mathbf{x} . Namely,

$$\alpha = \alpha(\mathbf{x}) = \alpha(x_1, x_2, \dots, x_n) \quad (3)$$

Definition 1. Partial derivative of a scalar $\alpha(\mathbf{x})$ with respect to vector \mathbf{x} is defined by [1-4]

$$\frac{\partial \alpha}{\partial \mathbf{x}} = \partial \alpha \frac{1}{\partial \mathbf{x}^T} = \left[\frac{\partial \alpha}{\partial x_1}, \frac{\partial \alpha}{\partial x_2}, \dots, \frac{\partial \alpha}{\partial x_n} \right]. \quad (4)$$

2.2. Partial derivatives of a vector with respect to a vector

Let vector $\mathbf{a} = \mathbf{a}(\mathbf{x}) \in \mathbb{R}^m$ be a function of vector $\mathbf{x} \in \mathbb{R}^n$. Namely,

$$\mathbf{a} = \begin{bmatrix} a_1(\mathbf{x}) \\ a_2(\mathbf{x}) \\ \vdots \\ a_m(\mathbf{x}) \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad (5)$$

Definition 2. Partial derivative of vector $\mathbf{a}(\mathbf{x})$ with respect to a vector \mathbf{x} is defined by [1-4]

$$\frac{\partial \mathbf{a}}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial a_1}{\partial \mathbf{x}} \\ \frac{\partial a_2}{\partial \mathbf{x}} \\ \vdots \\ \frac{\partial a_m}{\partial \mathbf{x}} \end{bmatrix} = \begin{bmatrix} \frac{\partial a_1}{\partial x_1} & \frac{\partial a_1}{\partial x_2} & \dots & \frac{\partial a_1}{\partial x_n} \\ \frac{\partial a_2}{\partial x_1} & \frac{\partial a_2}{\partial x_2} & \dots & \frac{\partial a_2}{\partial x_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial a_m}{\partial x_1} & \frac{\partial a_m}{\partial x_2} & \dots & \frac{\partial a_m}{\partial x_n} \end{bmatrix}. \quad (6)$$

Example 1. Calculate $\frac{\partial \mathbf{x}}{\partial \mathbf{x}}$ and $\frac{\partial \mathbf{x}^T}{\partial \mathbf{x}}$, where vector $\mathbf{x} \in \mathbb{R}^n$. Using the definition (6) we get

$$\frac{\partial \mathbf{x}}{\partial \mathbf{x}} = \mathbf{I}_n, \quad \frac{\partial \mathbf{x}^T}{\partial \mathbf{x}} = \left[\frac{\partial x_1}{\partial \mathbf{x}}, \frac{\partial x_2}{\partial \mathbf{x}}, \dots, \frac{\partial x_n}{\partial \mathbf{x}} \right] = [\mathbf{e}_1^T, \mathbf{e}_2^T, \dots, \mathbf{e}_n^T]$$

where \mathbf{I}_n is the $n \times n$ identity matrix, and

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \dots \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \dots \\ 0 \end{bmatrix}, \quad \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ \dots \\ 1 \end{bmatrix}.$$

2.3. Partial derivatives of a matrix with respect to a vector

Let matrix $\mathbf{A} = \mathbf{A}(\mathbf{x}) \in \mathbb{R}^{m \times p}$ be a function of vector $\mathbf{x} \in \mathbb{R}^n$, $a_{ij} = a_{ij}(\mathbf{x})$. Namely,

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mp} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad (7)$$

Definition 3. Partial derivative of matrix $\mathbf{A}(\mathbf{x})$ with respect to a vector \mathbf{x} is defined by

$$\begin{aligned} \frac{\partial \mathbf{A}}{\partial \mathbf{x}} &= \begin{bmatrix} \frac{\partial \mathbf{a}_1}{\partial \mathbf{x}} & \frac{\partial \mathbf{a}_2}{\partial \mathbf{x}} & \dots & \frac{\partial \mathbf{a}_p}{\partial \mathbf{x}} \end{bmatrix} = \begin{bmatrix} \frac{\partial a_{11}}{\partial \mathbf{x}} & \frac{\partial a_{12}}{\partial \mathbf{x}} & \dots & \frac{\partial a_{1p}}{\partial \mathbf{x}} \\ \frac{\partial a_{21}}{\partial \mathbf{x}} & \frac{\partial a_{22}}{\partial \mathbf{x}} & \dots & \frac{\partial a_{2p}}{\partial \mathbf{x}} \\ \dots & \dots & \dots & \dots \\ \frac{\partial a_{m1}}{\partial \mathbf{x}} & \frac{\partial a_{m2}}{\partial \mathbf{x}} & \dots & \frac{\partial a_{mp}}{\partial \mathbf{x}} \end{bmatrix} \\ &= \begin{bmatrix} \frac{\partial a_{11}}{\partial x_1} \dots \frac{\partial a_{11}}{\partial x_n} & \frac{\partial a_{12}}{\partial x_1} \dots \frac{\partial a_{12}}{\partial x_n} & \dots & \frac{\partial a_{1p}}{\partial x_1} \dots \frac{\partial a_{1p}}{\partial x_n} \\ \frac{\partial a_{21}}{\partial x_1} \dots \frac{\partial a_{21}}{\partial x_n} & \frac{\partial a_{22}}{\partial x_1} \dots \frac{\partial a_{22}}{\partial x_n} & \dots & \frac{\partial a_{2p}}{\partial x_1} \dots \frac{\partial a_{2p}}{\partial x_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial a_{m1}}{\partial x_1} \dots \frac{\partial a_{m1}}{\partial x_n} & \frac{\partial a_{m2}}{\partial x_1} \dots \frac{\partial a_{m2}}{\partial x_n} & \dots & \frac{\partial a_{mp}}{\partial x_1} \dots \frac{\partial a_{mp}}{\partial x_n} \end{bmatrix}. \end{aligned} \quad (8)$$

In Eq. (8) matrix $\frac{\partial \mathbf{A}}{\partial \mathbf{x}}$ is a $m \times pn$ matrix.

Note that the partial derivative of scalar function $\alpha(\mathbf{x})$ with respect to a vector variable \mathbf{x} in Eq. (4) is a row matrix that is rarely used in the text. The common form used is column matrix. Note also that, according to Eq. (6) the partial derivative of vector function $\mathbf{a}(\mathbf{x})$ with respect to a vector \mathbf{x} , whose elements are scalar functions of the vector variable \mathbf{x} , is a matrix. Thus, the partial derivative of vector function $\mathbf{a}(\mathbf{x})$ with respect to vector \mathbf{x} is the column vector, whose elements are partial derivatives of scalar elements $a_i(\mathbf{x})$ of the vector $\mathbf{a}(\mathbf{x})$ with respect to vector \mathbf{x} . Based on the above definitions, the partial derivative of a matrix function $\mathbf{A}(\mathbf{x})$ with respect to a vector \mathbf{x} is defined by Eq. (8). This is a matrix, whose elements are partial derivatives of scalar elements $a_{ij}(\mathbf{x})$ of the matrix $\mathbf{A}(\mathbf{x})$ with respect to a vector \mathbf{x} . If $\mathbf{A}(\mathbf{x})$ is a $m \times p$ matrix and \mathbf{x} is a n -vector, then $\frac{\partial \mathbf{A}}{\partial \mathbf{x}}$ is a $m \times pn$ two-dimensional matrix.

However, according to the definition in Ref. [6] the partial derivative of matrix function $\mathbf{A}(\mathbf{x})$ with respect to a vector \mathbf{x} is a three-dimensional matrix. In viewpoint of the matrix calculus, the equation (2.15) in page 20 of this book is not correct.

3. AN OVERVIEW OF THE MATRIX PRODUCT AND THE KRONECKER PRODUCT OF TWO MATRICES

Let us review some basic concepts from matrix calculus that will enable us to better understand the proof of the following theorems which are useful to establish dynamical equations of mechanical systems.

3.1. Matrix product of two matrices

Let \mathbf{A} be a $m \times p$ matrix and let \mathbf{B} be a $p \times s$ matrix, which are written in the following form

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mp} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1s} \\ b_{21} & b_{22} & \dots & b_{2s} \\ \dots & \dots & \dots & \dots \\ b_{p1} & b_{p2} & \dots & b_{ps} \end{bmatrix}. \quad (9)$$

where elements of matrices \mathbf{A} and \mathbf{B} are functions of vector \mathbf{x} .

Definition 4. Matrix product of two matrices $\mathbf{A}(\mathbf{x}) \in \mathbb{R}^{m \times p}$ and $\mathbf{B}(\mathbf{x}) \in \mathbb{R}^{p \times s}$ is matrix $\mathbf{C}(\mathbf{x}) \in \mathbb{R}^{m \times s}$ defined by [11]

$$\mathbf{C} = \mathbf{A}\mathbf{B} = \begin{bmatrix} c_{11} & c_{12} & \dots & c_{1s} \\ c_{21} & c_{22} & \dots & c_{2s} \\ \dots & \dots & \dots & \dots \\ c_{m1} & c_{m2} & \dots & c_{ms} \end{bmatrix}, \quad c_{ij} = \sum_{k=1}^p a_{ik}b_{kj}. \quad (10)$$

Remark. By matrix multiplication, the number of columns in matrix \mathbf{A} must be equal to the number of rows in matrix \mathbf{B} . If \mathbf{A} is a $m \times p$ matrix and \mathbf{B} is a $p \times s$ matrix, then \mathbf{C} is a $m \times s$ matrix.

3.2. Kronecker product of two matrices

Definition 5. Kronecker product of two matrices $\mathbf{A}(\mathbf{x}) \in \mathbb{R}^{m \times p}$ and $\mathbf{B}(\mathbf{x}) \in \mathbb{R}^{q \times s}$, denoted by $\mathbf{A} \otimes \mathbf{B}$, is a $mq \times ps$ matrix as [10, 12]

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \dots & a_{1p}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \dots & a_{2p}\mathbf{B} \\ \dots & \dots & \dots & \dots \\ a_{m1}\mathbf{B} & a_{m2}\mathbf{B} & \dots & a_{mp}\mathbf{B} \end{bmatrix}. \quad (11)$$

Example 2. we obtain the following expression for matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ and vector \mathbf{b}

$$\mathbf{A} \otimes \mathbf{b} = \begin{bmatrix} a_{11}\mathbf{b} & a_{12}\mathbf{b} & \dots & a_{1n}\mathbf{b} \\ a_{21}\mathbf{b} & a_{22}\mathbf{b} & \dots & a_{2n}\mathbf{b} \\ \dots & \dots & \dots & \dots \\ a_{n1}\mathbf{b} & a_{n2}\mathbf{b} & \dots & a_{nn}\mathbf{b} \end{bmatrix}.$$

It follows that

$$\mathbf{I}_n \otimes \mathbf{b} = \begin{bmatrix} \mathbf{b} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{b} & \dots & \mathbf{0} \\ \dots & \dots & \dots & \dots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{b} \end{bmatrix}.$$

If $\mathbf{b} \in \mathbb{R}^p$, then $\mathbf{I}_n \otimes \mathbf{b}$ is a $np \times n$ matrix. Using the definition in Eq. (11), it can be verified that

$$\mathbf{b} \otimes \mathbf{I}_n = \begin{bmatrix} b_1 \mathbf{I}_n \\ b_2 \mathbf{I}_n \\ \vdots \\ b_p \mathbf{I}_n \end{bmatrix}.$$

That is a $np \times n$ matrix. However $\mathbf{I}_n \otimes \mathbf{b} \neq \mathbf{b} \otimes \mathbf{I}_n$.

4. RELATIONSHIPS BETWEEN THE TIME DERIVATIVE OF MATRIX AND ITS PARTIAL DERIVATIVE WITH RESPECT TO A VECTOR

Theorem 1. When matrix $\mathbf{A}(\mathbf{x}) \in \mathbb{R}^{m \times p}$ is a function of vector $\mathbf{x} \in \mathbb{R}^n$, and \mathbf{x} is a function of the time, we have the following rule

$$\frac{d\mathbf{A}(\mathbf{x})}{dt} = \frac{\partial \mathbf{A}(\mathbf{x})}{\partial \mathbf{x}} (\mathbf{I}_n \otimes \dot{\mathbf{x}}) \tag{12}$$

Proof. The time derivative of a matrix \mathbf{A} is given by

$$\begin{aligned} \frac{d\mathbf{A}(\mathbf{x})}{dt} &= \begin{bmatrix} \dot{a}_{11}(\mathbf{x}) & \dots & \dot{a}_{1p}(\mathbf{x}) \\ \dots & \dots & \dots \\ \dot{a}_{m1}(\mathbf{x}) & \dots & \dot{a}_{mp}(\mathbf{x}) \end{bmatrix} \\ &= \begin{bmatrix} \frac{\partial a_{11}}{\partial x_1} \dot{x}_1 + \dots + \frac{\partial a_{11}}{\partial x_n} \dot{x}_n & \dots & \frac{\partial a_{1p}}{\partial x_1} \dot{x}_1 + \dots + \frac{\partial a_{1p}}{\partial x_n} \dot{x}_n \\ \dots & \dots & \dots \\ \frac{\partial a_{m1}}{\partial x_1} \dot{x}_1 + \dots + \frac{\partial a_{m1}}{\partial x_n} \dot{x}_n & \dots & \frac{\partial a_{mp}}{\partial x_1} \dot{x}_1 + \dots + \frac{\partial a_{mp}}{\partial x_n} \dot{x}_n \end{bmatrix}. \end{aligned} \tag{13}$$

The partial derivative of the matrix \mathbf{A} with respect to vector \mathbf{x} can be calculated as follows

$$\begin{aligned} \frac{\partial \mathbf{A}}{\partial \mathbf{x}} &= \begin{bmatrix} \frac{\partial a_{11}}{\partial \mathbf{x}} & \frac{\partial a_{12}}{\partial \mathbf{x}} & \dots & \frac{\partial a_{1p}}{\partial \mathbf{x}} \\ \frac{\partial a_{21}}{\partial \mathbf{x}} & \frac{\partial a_{22}}{\partial \mathbf{x}} & \dots & \frac{\partial a_{2p}}{\partial \mathbf{x}} \\ \dots & \dots & \dots & \dots \\ \frac{\partial a_{m1}}{\partial \mathbf{x}} & \frac{\partial a_{m2}}{\partial \mathbf{x}} & \dots & \frac{\partial a_{mp}}{\partial \mathbf{x}} \end{bmatrix} \\ &= \begin{bmatrix} \frac{\partial a_{11}}{\partial x_1} & \dots & \frac{\partial a_{11}}{\partial x_n} & \frac{\partial a_{12}}{\partial x_1} & \dots & \frac{\partial a_{12}}{\partial x_n} & \dots & \frac{\partial a_{1p}}{\partial x_1} & \dots & \frac{\partial a_{1p}}{\partial x_n} \\ \frac{\partial a_{21}}{\partial x_1} & \dots & \frac{\partial a_{21}}{\partial x_n} & \frac{\partial a_{22}}{\partial x_1} & \dots & \frac{\partial a_{22}}{\partial x_n} & \dots & \frac{\partial a_{2p}}{\partial x_1} & \dots & \frac{\partial a_{2p}}{\partial x_n} \\ \dots & \dots \\ \frac{\partial a_{m1}}{\partial x_1} & \dots & \frac{\partial a_{m1}}{\partial x_n} & \frac{\partial a_{m2}}{\partial x_1} & \dots & \frac{\partial a_{m2}}{\partial x_n} & \dots & \frac{\partial a_{mp}}{\partial x_1} & \dots & \frac{\partial a_{mp}}{\partial x_n} \end{bmatrix}. \end{aligned} \tag{14}$$

Using the definition of Kronecker product of two matrices, it can be verified that

$$\mathbf{I}_n \otimes \dot{\mathbf{x}} = \begin{bmatrix} \dot{\mathbf{x}} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \dot{\mathbf{x}} & \dots & \mathbf{0} \\ \dots & \dots & \dots & \dots \\ \mathbf{0} & \mathbf{0} & \dots & \dot{\mathbf{x}} \end{bmatrix}. \quad (15)$$

Using Eq. (14) and Eq. (15), one yields

$$\frac{\partial \mathbf{A}(\mathbf{x})}{\partial \mathbf{x}} (\mathbf{I}_n \otimes \dot{\mathbf{x}}) = \begin{bmatrix} \frac{\partial a_{11}}{\partial x_1} \dot{x}_1 + \dots + \frac{\partial a_{11}}{\partial x_n} \dot{x}_n, & \dots & \frac{\partial a_{1p}}{\partial x_1} \dot{x}_1 + \dots + \frac{\partial a_{1p}}{\partial x_n} \dot{x}_n \\ \dots & \dots & \dots \\ \frac{\partial a_{m1}}{\partial x_1} \dot{x}_1 + \dots + \frac{\partial a_{m1}}{\partial x_n} \dot{x}_n, & \dots & \frac{\partial a_{mp}}{\partial x_1} \dot{x}_1 + \dots + \frac{\partial a_{mp}}{\partial x_n} \dot{x}_n \end{bmatrix}. \quad (16)$$

Comparing Eq. (13) with Eq. (16) yields the rule expressed by Eq. (12).

5. PARTIAL DERIVATIVE OF THE PRODUCT OF TWO MATRICES WITH RESPECT TO A VECTOR

Let the matrices $\mathbf{A} \in \mathbb{R}^{m \times p}$ and $\mathbf{B} \in \mathbb{R}^{p \times s}$ be functions of a vector $\mathbf{x} \in \mathbb{R}^n$. Namely,

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mp} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1s} \\ b_{21} & b_{22} & \dots & b_{2s} \\ \dots & \dots & \dots & \dots \\ b_{p1} & b_{p2} & \dots & b_{ps} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}. \quad (17)$$

5.1. Theorem on partial derivative of product of two matrices with respect to a vector

a) *Lemma.* Partial derivative product of two scalar functions $\gamma(\mathbf{x}) = \alpha(\mathbf{x})\beta(\mathbf{x})$ with respect to vector \mathbf{x} of is defined by the following rule

$$\frac{\partial}{\partial \mathbf{x}} (\alpha(\mathbf{x})\beta(\mathbf{x})) = \frac{\partial \alpha(\mathbf{x})}{\partial \mathbf{x}} \beta(\mathbf{x}) + \alpha(\mathbf{x}) \frac{\partial \beta(\mathbf{x})}{\partial \mathbf{x}}. \quad (18)$$

Proof. By using the definition of the partial derivative of the scalar function with respect to a vector, we obtain

$$\begin{aligned} \frac{\partial \gamma(\mathbf{x})}{\partial \mathbf{x}} &= \left[\frac{\partial \gamma}{\partial x_1}, \frac{\partial \gamma}{\partial x_2}, \dots, \frac{\partial \gamma}{\partial x_n} \right] \\ &= \left[\frac{\partial \alpha}{\partial x_1} \beta + \alpha \frac{\partial \beta}{\partial x_1}, \frac{\partial \alpha}{\partial x_2} \beta + \alpha \frac{\partial \beta}{\partial x_2}, \dots, \frac{\partial \alpha}{\partial x_n} \beta + \alpha \frac{\partial \beta}{\partial x_n} \right] \\ &= \left[\frac{\partial \alpha}{\partial x_1} \beta, \frac{\partial \alpha}{\partial x_2} \beta, \dots, \frac{\partial \alpha}{\partial x_n} \beta \right] + \left[\alpha \frac{\partial \beta}{\partial x_1}, \alpha \frac{\partial \beta}{\partial x_2}, \dots, \alpha \frac{\partial \beta}{\partial x_n} \right] \\ &= \frac{\partial \alpha(\mathbf{x})}{\partial \mathbf{x}} \beta(\mathbf{x}) + \alpha(\mathbf{x}) \frac{\partial \beta(\mathbf{x})}{\partial \mathbf{x}}. \end{aligned}$$

b) *Theorem 2.* Partial derivative of the product of two matrices $\mathbf{A}(\mathbf{x})\mathbf{B}(\mathbf{x})$ with respect to a vector \mathbf{x} is defined by the following rule

$$\frac{\partial}{\partial \mathbf{x}}(\mathbf{A}(\mathbf{x})\mathbf{B}(\mathbf{x})) = \frac{\partial \mathbf{A}(\mathbf{x})}{\partial \mathbf{x}}(\mathbf{B}(\mathbf{x}) \otimes \mathbf{I}_n) + \mathbf{A} \frac{\partial \mathbf{B}}{\partial \mathbf{x}}, \quad (19)$$

where \mathbf{I}_n is the $n \times n$ identity matrix.

Proof. From the definition of the product of two matrices (10) we have

$$\mathbf{C}(\mathbf{x}) = \mathbf{A}(\mathbf{x})\mathbf{B}(\mathbf{x}) = \begin{bmatrix} c_{11} & c_{12} & \dots & c_{1s} \\ c_{21} & c_{22} & \dots & c_{2s} \\ \dots & \dots & \dots & \dots \\ c_{m1} & c_{m2} & \dots & c_{ms} \end{bmatrix}, \quad c_{ij} = \sum_{k=1}^p a_{ik}b_{kj}. \quad (20)$$

Using Eq. (8) and the lemma according to Eq.(18) we obtain

$$\begin{aligned} \frac{\partial(\mathbf{AB})}{\partial \mathbf{x}} &= \frac{\partial \mathbf{C}}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial c_{11}}{\partial \mathbf{x}} & \frac{\partial c_{12}}{\partial \mathbf{x}} & \dots & \frac{\partial c_{1s}}{\partial \mathbf{x}} \\ \dots & \dots & \dots & \dots \\ \frac{\partial c_{m1}}{\partial \mathbf{x}} & \frac{\partial c_{m2}}{\partial \mathbf{x}} & \dots & \frac{\partial c_{ms}}{\partial \mathbf{x}} \end{bmatrix} \\ \frac{\partial(\mathbf{AB})}{\partial \mathbf{x}} &= \begin{bmatrix} \sum_{k=1}^p \frac{\partial a_{1k}}{\partial \mathbf{x}} b_{k1}, & \dots, & \sum_{k=1}^p \frac{\partial a_{1k}}{\partial \mathbf{x}} b_{ks} \\ \dots & \dots & \dots \\ \sum_{k=1}^p \frac{\partial a_{mk}}{\partial \mathbf{x}} b_{k1}, & \dots, & \sum_{k=1}^p \frac{\partial a_{mk}}{\partial \mathbf{x}} b_{ks} \end{bmatrix} \\ &+ \begin{bmatrix} \sum_{k=1}^p a_{1k} \frac{\partial b_{k1}}{\partial \mathbf{x}}, & \dots, & \sum_{k=1}^p a_{1k} \frac{\partial b_{ks}}{\partial \mathbf{x}} \\ \dots & \dots & \dots \\ \sum_{k=1}^p a_{mk} \frac{\partial b_{k1}}{\partial \mathbf{x}}, & \dots, & \sum_{k=1}^p a_{mk} \frac{\partial b_{ks}}{\partial \mathbf{x}} \end{bmatrix} = \mathbf{K}_1 + \mathbf{K}_2 \end{aligned} \quad (21)$$

where symbols \mathbf{K}_1 and \mathbf{K}_2 denote the first term and the second term in the right-hand side of Eq. (21)

$$\mathbf{K}_1 = \begin{bmatrix} \frac{\partial a_{11}}{\partial x_1} b_{11} + \frac{\partial a_{12}}{\partial x_1} b_{21} + \dots + \frac{\partial a_{1p}}{\partial x_1} b_{p1} & \frac{\partial a_{11}}{\partial x_2} b_{11} + \frac{\partial a_{12}}{\partial x_2} b_{21} + \dots + \frac{\partial a_{1p}}{\partial x_2} b_{p1} \dots \\ \dots & \dots \\ \frac{\partial a_{m1}}{\partial x_1} b_{11} + \frac{\partial a_{m2}}{\partial x_1} b_{21} + \dots + \frac{\partial a_{mp}}{\partial x_1} b_{p1} & \frac{\partial a_{m1}}{\partial x_2} b_{11} + \frac{\partial a_{m2}}{\partial x_2} b_{21} + \dots + \frac{\partial a_{mp}}{\partial x_2} b_{p1} \dots \end{bmatrix}. \quad (22)$$

Note that

$$\mathbf{B} \otimes \mathbf{I}_n = \begin{bmatrix} b_{11}\mathbf{I}_n & b_{12}\mathbf{I}_n & \dots & b_{1s}\mathbf{I}_n \\ b_{21}\mathbf{I}_n & b_{22}\mathbf{I}_n & \dots & b_{2s}\mathbf{I}_n \\ \dots & \dots & \dots & \dots \\ b_{p1}\mathbf{I}_n & b_{p2}\mathbf{I}_n & \dots & b_{ps}\mathbf{I}_n \end{bmatrix}$$

$$\mathbf{B} \otimes \mathbf{I}_n = \begin{bmatrix} b_{11} & 0 & \dots & 0 & b_{12} & 0 & \dots & 0 & \dots & b_{1s} & 0 & \dots & 0 \\ 0 & b_{11} & \dots & 0 & 0 & b_{12} & \dots & 0 & \dots & 0 & b_{1s} & \dots & 0 \\ \dots & \dots \\ 0 & 0 & \dots & b_{11} & 0 & 0 & \dots & b_{12} & \dots & 0 & 0 & \dots & b_{1s} \\ \dots & \dots \\ b_{p1} & 0 & \dots & 0 & b_{p2} & 0 & \dots & 0 & \dots & b_{ps} & 0 & \dots & 0 \\ 0 & b_{p1} & \dots & 0 & 0 & b_{p2} & \dots & 0 & \dots & 0 & b_{ps} & \dots & 0 \\ \dots & \dots \\ 0 & 0 & \dots & b_{p1} & 0 & 0 & \dots & b_{p2} & \dots & 0 & 0 & \dots & b_{ps} \end{bmatrix} \quad (23)$$

$$\frac{\partial \mathbf{A}}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial a_{11}}{\partial \mathbf{x}} & \frac{\partial a_{12}}{\partial \mathbf{x}} & \dots & \frac{\partial a_{1p}}{\partial \mathbf{x}} \\ \frac{\partial a_{21}}{\partial \mathbf{x}} & \frac{\partial a_{22}}{\partial \mathbf{x}} & \dots & \frac{\partial a_{2p}}{\partial \mathbf{x}} \\ \dots & \dots & \dots & \dots \\ \frac{\partial a_{m1}}{\partial \mathbf{x}} & \frac{\partial a_{m2}}{\partial \mathbf{x}} & \dots & \frac{\partial a_{mp}}{\partial \mathbf{x}} \end{bmatrix} \quad (24)$$

$$\mathbf{J} = \begin{bmatrix} \frac{\partial a_{11}}{\partial x_1} & \dots & \frac{\partial a_{11}}{\partial x_n} & \frac{\partial a_{12}}{\partial x_1} & \dots & \frac{\partial a_{12}}{\partial x_n} & \dots & \frac{\partial a_{1p}}{\partial x_1} & \dots & \frac{\partial a_{1p}}{\partial x_n} \\ \frac{\partial a_{21}}{\partial x_1} & \dots & \frac{\partial a_{21}}{\partial x_n} & \frac{\partial a_{22}}{\partial x_1} & \dots & \frac{\partial a_{22}}{\partial x_n} & \dots & \frac{\partial a_{2p}}{\partial x_1} & \dots & \frac{\partial a_{2p}}{\partial x_n} \\ \dots & \dots \\ \frac{\partial a_{m1}}{\partial x_1} & \dots & \frac{\partial a_{m1}}{\partial x_n} & \frac{\partial a_{m2}}{\partial x_1} & \dots & \frac{\partial a_{m2}}{\partial x_n} & \dots & \frac{\partial a_{mp}}{\partial x_1} & \dots & \frac{\partial a_{mp}}{\partial x_n} \end{bmatrix}.$$

Using Eqs. (23) and (24), one yields

$$\frac{\partial \mathbf{A}}{\partial \mathbf{x}} (\mathbf{B} \otimes \mathbf{I}_n) = \begin{bmatrix} \frac{\partial a_{11}}{\partial x_1} b_{11} + \frac{\partial a_{12}}{\partial x_1} b_{21} + \dots + \frac{\partial a_{1p}}{\partial x_1} b_{p1} & \dots & \frac{\partial a_{11}}{\partial x_n} b_{1s} + \frac{\partial a_{12}}{\partial x_n} b_{2s} + \dots + \frac{\partial a_{1p}}{\partial x_n} b_{ps} \\ \dots & \dots & \dots \\ \frac{\partial a_{m1}}{\partial x_1} b_{11} + \frac{\partial a_{m2}}{\partial x_1} b_{21} + \dots + \frac{\partial a_{mp}}{\partial x_1} b_{p1} & \dots & \frac{\partial a_{m1}}{\partial x_n} b_{1s} + \frac{\partial a_{m2}}{\partial x_n} b_{2s} + \dots + \frac{\partial a_{mp}}{\partial x_n} b_{ps} \end{bmatrix}. \quad (25)$$

Combining Eq. (22) with Eq. (25) we obtain

$$\mathbf{K}_1 = \frac{\partial \mathbf{A}}{\partial \mathbf{x}} (\mathbf{B} \otimes \mathbf{I}_n). \quad (26)$$

Similarly, we get the matrix \mathbf{K}_2

$$\mathbf{K}_2 = \begin{bmatrix} a_{11} \frac{\partial b_{11}}{\partial \mathbf{x}} + a_{12} \frac{\partial b_{21}}{\partial \mathbf{x}} + \dots + a_{1p} \frac{\partial b_{p1}}{\partial \mathbf{x}} & \dots & a_{11} \frac{\partial b_{1s}}{\partial \mathbf{x}} + a_{12} \frac{\partial b_{2s}}{\partial \mathbf{x}} + \dots + a_{1p} \frac{\partial b_{ps}}{\partial \mathbf{x}} \\ \dots & \dots & \dots \\ a_{m1} \frac{\partial b_{11}}{\partial \mathbf{x}} + a_{m2} \frac{\partial b_{21}}{\partial \mathbf{x}} + \dots + a_{mp} \frac{\partial b_{p1}}{\partial \mathbf{x}} & \dots & a_{m1} \frac{\partial b_{1s}}{\partial \mathbf{x}} + a_{m2} \frac{\partial b_{2s}}{\partial \mathbf{x}} + \dots + a_{mp} \frac{\partial b_{ps}}{\partial \mathbf{x}} \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{a}_{1r} \frac{\partial \mathbf{b}_1}{\partial \mathbf{x}} & \mathbf{a}_{1r} \frac{\partial \mathbf{b}_2}{\partial \mathbf{x}} & \cdots & \mathbf{a}_{1r} \frac{\partial \mathbf{b}_s}{\partial \mathbf{x}} \\ \cdots & \cdots & \cdots & \cdots \\ \mathbf{a}_{mr} \frac{\partial \mathbf{b}_1}{\partial \mathbf{x}} & \mathbf{a}_{mr} \frac{\partial \mathbf{b}_2}{\partial \mathbf{x}} & \cdots & \mathbf{a}_{mr} \frac{\partial \mathbf{b}_s}{\partial \mathbf{x}} \end{bmatrix} = \begin{bmatrix} \mathbf{a}_{1r} \\ \mathbf{a}_{2r} \\ \cdots \\ \mathbf{a}_{mr} \end{bmatrix} \begin{bmatrix} \frac{\partial \mathbf{b}_1}{\partial \mathbf{x}} & \frac{\partial \mathbf{b}_2}{\partial \mathbf{x}} & \cdots & \frac{\partial \mathbf{b}_s}{\partial \mathbf{x}} \end{bmatrix} \quad (27)$$

where

$$\mathbf{a}_{ir} = [a_{i1} \quad a_{i2} \quad \cdots \quad a_{ip}], \quad \mathbf{b}_j = \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{pj} \end{bmatrix}.$$

Eq. (27) can be rewritten in the following form

$$\mathbf{K}_2 = \mathbf{A} \frac{\partial \mathbf{B}}{\partial \mathbf{x}}. \quad (28)$$

Substituting Eqs. (26) and (28) into Eq. (21) we obtain the same rule as shown in Eq. (19)

$$\frac{\partial(\mathbf{A}\mathbf{B})}{\partial \mathbf{x}} = \frac{\partial \mathbf{A}}{\partial \mathbf{x}}(\mathbf{B} \otimes \mathbf{I}_n) + \mathbf{A} \frac{\partial \mathbf{B}}{\partial \mathbf{x}}. \quad (29)$$

5.2. Consequences

From the theorem on partial derivative of the product of two matrices $\mathbf{A}(\mathbf{x})\mathbf{B}(\mathbf{x})$ with to vector, Eq. (19), the following important consequences can be found:

a) *Consequence 1.* Let matrix $\mathbf{A} \in \mathbb{R}^{m \times p}$ and vector $\mathbf{b} \in \mathbb{R}^p$ be functions of the vector $\mathbf{x} \in \mathbb{R}^n$, we have

$$\frac{\partial}{\partial \mathbf{x}}(\mathbf{A}(\mathbf{x})\mathbf{b}(\mathbf{x})) = \frac{\partial \mathbf{A}(\mathbf{x})}{\partial \mathbf{x}}(\mathbf{b} \otimes \mathbf{I}_n) + \mathbf{A} \frac{\partial \mathbf{b}}{\partial \mathbf{x}}. \quad (30)$$

b) *Consequence 2.* Let vector $\mathbf{b} \in \mathbb{R}^p$ and scalar $\alpha(\mathbf{x})$ be functions of vector $\mathbf{x} \in \mathbb{R}^n$, it is follows from the rule (19)

$$\frac{\partial}{\partial \mathbf{x}}(\mathbf{b}(\mathbf{x})\alpha(\mathbf{x})) = \frac{\partial \mathbf{b}}{\partial \mathbf{x}}(\alpha \otimes \mathbf{I}_n) + \mathbf{b} \frac{\partial \alpha}{\partial \mathbf{x}}. \quad (31)$$

c) *Consequence 3.* Let scalar $\alpha(\mathbf{x})$ and scalar $\beta(\mathbf{x})$ be two functions of vector $\mathbf{x} \in \mathbb{R}^n$, we obtain

$$\frac{\partial}{\partial \mathbf{x}}(\alpha(\mathbf{x})\beta(\mathbf{x})) = \frac{\partial \alpha}{\partial \mathbf{x}}(\beta \otimes \mathbf{I}_n) + \alpha \frac{\partial \beta}{\partial \mathbf{x}}. \quad (32)$$

Remark. From the expression

$$\begin{aligned} \frac{\partial \alpha}{\partial \mathbf{x}}(\beta \otimes \mathbf{I}_n) &= \begin{bmatrix} \frac{\partial \alpha}{\partial x_1} & \frac{\partial \alpha}{\partial x_2} & \cdots & \frac{\partial \alpha}{\partial x_n} \end{bmatrix} \begin{bmatrix} \beta & 0 & \cdots & 0 \\ 0 & \beta & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \beta \end{bmatrix} \\ &= \begin{bmatrix} \beta \frac{\partial \alpha}{\partial x_1} & \beta \frac{\partial \alpha}{\partial x_2} & \cdots & \beta \frac{\partial \alpha}{\partial x_n} \end{bmatrix} = \beta \frac{\partial \alpha}{\partial \mathbf{x}}, \end{aligned}$$

it follows

$$\frac{\partial}{\partial \mathbf{x}}(\alpha(\mathbf{x})\beta(\mathbf{x})) = \alpha \frac{\partial \beta}{\partial \mathbf{x}} + \beta \frac{\partial \alpha}{\partial \mathbf{x}}. \quad (33)$$

d) *Consequence 4.* Let vector $\mathbf{a} \in \mathbb{R}^m$ and vector $\mathbf{b} \in \mathbb{R}^m$ be two vector functions of vector $\mathbf{x} \in \mathbb{R}^n$, we obtain

$$\frac{\partial(\mathbf{a}^T \mathbf{b})}{\partial \mathbf{x}} = \frac{\partial \mathbf{a}^T}{\partial \mathbf{x}} (\mathbf{b} \otimes \mathbf{I}_n) + \mathbf{a}^T \frac{\partial \mathbf{b}}{\partial \mathbf{x}}. \quad (34)$$

Note that the matrix $\frac{\partial(\mathbf{a}^T \mathbf{b})}{\partial \mathbf{x}}$ is a row matrix.

Example 3. Let us be a quadratic function of the component of vector $\mathbf{x} \in \mathbb{R}^n$ such that

$$Q(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x} = \mathbf{x}^T \mathbf{b}(\mathbf{x})$$

where $\mathbf{A} \in \mathbb{R}^{n \times n}$ is a symmetric matrix with constant elements. Show that

$$\frac{\partial Q(\mathbf{x})}{\partial \mathbf{x}} = \frac{\partial}{\partial \mathbf{x}} (\mathbf{x}^T \mathbf{A} \mathbf{x}) = 2\mathbf{x}^T \mathbf{A}.$$

From the rule (29) follows that

$$\begin{aligned} \frac{\partial}{\partial \mathbf{x}} (\mathbf{x}^T \mathbf{A} \mathbf{x}) &= \frac{\partial}{\partial \mathbf{x}} (\mathbf{x}^T \mathbf{b}) = \frac{\partial \mathbf{x}^T}{\partial \mathbf{x}} (\mathbf{b} \otimes \mathbf{I}_n) + \mathbf{x}^T \frac{\partial \mathbf{b}}{\partial \mathbf{x}} \\ &= [\mathbf{e}_1^T, \mathbf{e}_2^T, \dots, \mathbf{e}_n^T] \begin{bmatrix} b_1 \mathbf{I}_n \\ b_2 \mathbf{I}_n \\ \dots \\ b_n \mathbf{I}_n \end{bmatrix} + \mathbf{x}^T \frac{\partial}{\partial \mathbf{x}} (\mathbf{A} \mathbf{x}) \\ &= [b_1 \mathbf{e}_1^T \mathbf{I}_n + b_2 \mathbf{e}_2^T \mathbf{I}_n + \dots + b_n \mathbf{e}_n^T \mathbf{I}_n] + \mathbf{x}^T \left[\frac{\partial \mathbf{A}}{\partial \mathbf{x}} (\mathbf{x} \otimes \mathbf{I}_n) + \mathbf{A} \frac{\partial \mathbf{x}}{\partial \mathbf{x}} \right] \\ &= [b_1, b_2, \dots, b_n] + \mathbf{x}^T \mathbf{A} \mathbf{I}_n = \mathbf{b}^T + \mathbf{x}^T \mathbf{A} = (\mathbf{A} \mathbf{x})^T + \mathbf{x}^T \mathbf{A} = 2\mathbf{x}^T \mathbf{A}. \end{aligned}$$

6. CONCLUSIONS

The partial derivatives of scalar functions, vector functions and matrix functions with respect to a vector variable have an important role in computer aided kinematics and dynamics of mechanical systems. The partial derivatives of scalar functions and vector functions with respect to a vector variable are defined and used in dynamics of multibody systems. Eq. (8) is a generalization of the definitions of partial derivatives with respect to a vector variable of scalar function and vector function to matrix function. The theorems of the relationship between the time derivative of a matrix and its partial derivative with respect to a vector and the partial derivative of product of two matrices with respect to a vector according to Eqs. (12) and (19) are simple and very convenient to use.

The theoretical results investigated in this paper can also be applied to develop a new matrix form of Lagrangian equations and to derive the balancing conditions of spatial mechanisms that will be presented in other works.

ACKNOWLEDGEMENTS

This work is supported by the Vietnam Basic Research Program in Natural Science.

REFERENCES

- [1] W. Schiehlen, P. Eberhard, *Technische Dynamik (2. Auflage)*, B.G. Teubner, Stuttgart, 2004.
- [2] P. E. Nikravesh, *Computer Aided Analysis of Mechanical Systems*, Prentice-Hall, Englewood Cliffs, New Jersey, 1988.
- [3] E. J. Haug, *Computer Aided Kinematics and Dynamics of Mechanical Systems*, Vol.1: Basic Method, Allyn and Bacon, Boston, 1989.
- [4] A. A. Shabana, *Computational Dynamics (2.Edition)*, John Wiley & Sons, New York, 2001.
- [5] H. Josephs, R. L. Huston, *Dynamics of Mechanical Systems*, CRS Press, Boca Raton, 2002.
- [6] Y. Nakamura, *Advanced Robotics/ Redundancy and Optimization*, Addison-Wesley, Reading, 1991.
- [7] J. Angeles, *Fundamentals of Robotic Mechanical Systems (2. Edition)*, Springer-Verlag, New York, 2003.
- [8] M. W. Spong, M. Vidyasagar, *Robot Dynamics and Control*, John Wiley and Sons, New York, 1989.
- [9] R. M. Murray, Z. Li, S. S. Sastry, *A Mathematical Introduction to Robotic*, CRS Press, Boca Raton, 1994.
- [10] F. L. Lewis, D. M. Dawson, Ch.T. Abdallah, *Robot Manipulator Control, Theory and Practice (2.Edition)*, Marcel Dekker, New York, 2004.
- [11] F. Zhang, *Matrix Theory*, Springer, New York, 1999.
- [12] J. W. Brewer, Kronecker Products and Matrix Calculus in System Theory, *IEEE Transactions on Circuits and Systems* **CAS-25** (9) 772-781, September 1978.

Received June 17, 2009

ĐẠO HÀM RIÊNG THEO BIẾN VÉC TƠ CỦA HÀM MA TRẬN

Đạo hàm riêng theo biến véc tơ của hàm vô hướng và hàm véc tơ ĐÓ được định nghĩa và được sử dụng nhiều trong động lực học hệ nhiều vật. Tuy nhiên đạo hàm riêng theo biến véc tơ của hàm ma trận còn ít được nghiên cứu. Trong bài báo này, trước hết chúng tôi trình bày các định nghĩa đạo hàm riêng theo biến véc tơ của hàm vô hướng, hàm véc tơ và hàm ma trận một cách hệ thống. Sau đó giới thiệu tóm tắt phép nhân hai ma trận và phép nhân Kronecker hai ma trận. Cuối cùng chứng minh định lý về quan hệ giữa đạo hàm theo thời gian và đạo hàm riêng theo biến véc tơ của hàm ma trận và định lý về đạo hàm riêng theo biến véc tơ của tích hai ma trận.