



ELSEVIER

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

SCIENCE @ DIRECT®

Linear Algebra and its Applications 388 (2004) 315–325

LINEAR ALGEBRA  
AND ITS  
APPLICATIONS

[www.elsevier.com/locate/laa](http://www.elsevier.com/locate/laa)

# Matrix integral

Margus Pihlak

*Institute of Mathematical Statistics, University of Tartu, Liivi-2, 50409 Tartu, Estonia*

Received 26 July 2002; accepted 17 February 2004

Submitted by G.P.H. Styan

---

## Abstract

In the paper matrix integral is introduced as the inverse operation of the matrix derivative. The definition of the integral is based on the star product of matrices introduced by MacRae [Ann. Stat. 7 (1974) 381]. Basic properties of the integral are presented and several examples given to demonstrate practical usage of the notion.

© 2004 Published by Elsevier Inc.

*AMS classification:* 15A69

*Keywords:* Matrix derivative; Kronecker product; Star product; Matrix integral

---

## 1. Introduction

Let  $X: p \times q$  and  $Y: r \times s$  be matrices. We denote the element in the  $i$ th row and  $j$ th column as  $(X)_{ij}$ .

Firstly we introduce notations from so called newer matrix algebra. The Kronecker product of matrices  $X: p \times q$  and  $Y: r \times s$  is defined as the  $pr \times qs$  partitioned matrix  $X \otimes Y$  consisting of  $r \times s$  blocks, where every block

$$(X)_{ij}(Y) = \begin{pmatrix} (X)_{ij}(Y)_{11} & \cdots & (X)_{ij}(Y)_{1s} \\ \vdots & \ddots & \vdots \\ (X)_{ij}(Y)_{r1} & \cdots & (X)_{ij}(Y)_{rs} \end{pmatrix},$$
$$i = 1, 2, \dots, p, j = 1, 2, \dots, q.$$

---

*E-mail address:* [margusi@ut.ee](mailto:margusi@ut.ee) (M. Pihlak).

0024-3795/\$ - see front matter © 2004 Published by Elsevier Inc.  
doi:10.1016/j.laa.2004.02.042

The  $r^k \times s^k$ -matrix  $Y^{\otimes k}$  is expressed as follows:

$$Y^{\otimes k} = \underbrace{Y \otimes Y \otimes \cdots \otimes Y}_{k \text{ times}}.$$

The vec-operator denoted by  $\text{vec}$  is defined as follows:

$$\text{vec}X = ((X)_{11}, \dots, (X)_{p1}, (X)_{12}, \dots, (X)_{p2}, \dots, (X)_{1q}, \dots, (X)_{pq})'.$$

Let us call the operator

$$\frac{d}{dX} = \begin{pmatrix} \frac{\partial}{\partial(X)_{11}} & \cdots & \frac{\partial}{\partial(X)_{1q}} \\ \vdots & \ddots & \vdots \\ \frac{\partial}{\partial(X)_{p1}} & \cdots & \frac{\partial}{\partial(X)_{pq}} \end{pmatrix}$$

the matrix differential operator and the operator

$$\frac{d}{s dX} = \sum_{i=1}^p \sum_{j=1}^q \frac{\partial}{\partial(X)_{ij}}$$

the scalar differential operator.

For notions and results from matrix algebra an interested reader is referred to Magnus and Neudecker [5], Schott [7] or Harville [1].

Matrix derivative has been a useful tool for statisticians in last 30 years. The notion “matrix derivative” has been used for different representations of the *Frechet*’ derivative. Neudecker [6] defined matrix derivative using  $\text{vec}$ -operator:

$$\frac{\partial Y}{\partial X} = \frac{d}{d\text{vec}'X} \otimes \text{vec}Y, \quad (1.1)$$

where

$$\frac{d}{d\text{vec}'X} = \left( \frac{\partial}{\partial(X)_{11}} \cdots, \frac{\partial}{\partial(X)_{p1}} \cdots, \frac{\partial}{\partial(X)_{1q}} \cdots, \frac{\partial}{\partial(X)_{pq}} \right).$$

Definition of MacRae in [4] was based on Kronecker product directly:

$$\frac{dY}{dX} = \frac{d}{dX} \otimes Y. \quad (1.2)$$

In the sense of metrics the definition of MacRae and Neudecker matrix derivative are equivalent. MacRae’s matrix derivative fits better presentation. In applications, however, the Neudecker’s matrix derivative is often used.

The higher order derivatives have been defined recursively. In the case of definition (1.1) [2, p. 69]

$$\frac{\partial^k Y}{\partial X^k} = \frac{\partial}{\partial X} \left( \frac{\partial^{k-1}}{\partial X^{k-1}} \right) = \underbrace{\frac{d}{d\text{vec}'X} \otimes \frac{d}{d\text{vec}X} \otimes \cdots \otimes \frac{d}{d\text{vec}X}}_{k \text{ times}} \otimes \text{vec}Y \quad (1.3)$$

and

$$\frac{d^k Y}{dX^k} = \frac{d}{dX} = \left( \frac{d^{k-1} Y}{dX^{k-1}} \right), \quad k > 1, \quad k \in N$$

in the case of the derivative (1.2) [2, p. 70].

A stimulating result for our study has been a general relation between two density functions, derived in Kollo and von Rosen [3]. In their paper a general formal density expansion is presented where complicated density of interest is presented through the approximating density and cumulants of both distributions under consideration. The dimension of the approximating distribution can be higher than the dimension of the distribution of interest. This makes it possible to approximate the distribution of the sample correlation matrix through the Wishart distribution, for instance.

In applications approximation of the distribution function is at least as important as of the density function. In univariate case an expansion of the distribution function can be obtained from a density expansion by integration. In multivariate case the situation is more complicated. Kollo and von Rosen have approximated the density function  $f_Y(x)$  by the density function  $f_X(x)$  using matrix derivative (1.1). The notion “matrix integral” has been introduced in this paper to make it possible to integrate this relation between the two density function. In Section 2 we define indefinite and definite matrix integrals and give some simple examples. In Section 3 basic properties of this integral are examined. In Section 4 an application of the matrix integral is given and some comments for further development are made.

## 2. Definition and examples

Firstly we define the indefinite matrix integral. In definition we use the MacRae’s matrix derivative. The Neudecker’s matrix derivative we could be used in the same way.

**Definition 2.1.** The matrix  $Y: r \times s$  is called the matrix integral of  $Z = Z(X): pr \times sq$ , where  $X: p \times q$ , if

$$\frac{dY}{dX} = Z.$$

The fact, that matrix  $Y$  is the matrix integral of the matrix  $Z$  is denoted as

$$\int_{\mathfrak{R}^{pq}} Z \circ dX = Y.$$

If  $Y$  is the matrix integral of matrix  $Z$ , then also  $Y + C$  is a matrix integral of  $Z$ . Here  $C$  is a constant matrix of the same order as matrix  $Y$ . In Definition 2.1 we have introduced the undefinite matrix integral. It is used to define the definite matrix integral.

**Definition 2.2.** The difference  $Y(B) - Y(A)$  is called the definite matrix integral of matrix  $Z$  from  $A$  to  $B$ , where  $Y(A) = Y(X)|_{X=A}$  and  $Y(B) = Y(X)|_{X=B}$  are determined by Definition 2.1.

While a matrix derivative has larger dimensions than the differentiated matrix, then integrated matrix has smaller dimensions than the matrix to be integrated. Our further constructions are based on the star product of matrices introduced by MacRae [4]. This operation is denoted by  $*$ .

**Definition 2.3.** The star product of matrices  $A: p \times q$  and  $B: pr \times qs$  is the  $r \times s$ -matrix  $A * B$ , which is given by the equality

$$A * B = \sum_{i=1}^p \sum_{j=1}^q (A)_{ij} [B_{ij}],$$

where  $[B_{ij}]$  is  $r \times s$ -submatrix.

By means of the star product we shall now define the operator  $d = dX * \frac{d}{dX}$ . Let  $Z$  be  $pr \times qs$  and let us compose the matrix

$$D = dX * Z = \sum_{i=1}^p \sum_{j=1}^q (dX)_{ij} [Z_{ij}].$$

If every element of matrix  $D$  is the total derivative of some function depending on matrix  $X$ , then there exists matrix  $Y: r \times s$  so that  $dY = D$  and  $Z = \frac{dY}{dX}$ . Now we can define the matrix differential.

**Definition 2.4.** The matrix  $dY: r \times s$  is called the matrix differential of matrix  $Y(X)$  if

$$dY = dX * \left( \frac{d}{dX} \otimes Y \right).$$

By means of operator  $d$  we can find the matrix differential  $dY$  as follows:

$$\begin{aligned}
 dY &= dX * \frac{dY}{dX} \\
 &= \begin{pmatrix} d(X)_{11} & \cdots & d(X)_{1q} \\ \vdots & \ddots & \vdots \\ d(X)_{p1} & \cdots & d(X)_{pq} \end{pmatrix} \\
 &\quad * \begin{pmatrix} \frac{\partial(Y)_{11}}{\partial(X)_{11}} & \cdots & \frac{\partial(Y)_{1s}}{\partial(X)_{11}} & \cdots & \frac{\partial(Y)_{11}}{\partial(X)_{1q}} & \cdots & \frac{\partial(Y)_{1s}}{\partial(X)_{1q}} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \frac{\partial(Y)_{r1}}{\partial(X)_{11}} & \cdots & \frac{\partial(Y)_{rs}}{\partial(X)_{11}} & \cdots & \frac{\partial(Y)_{r1}}{\partial(X)_{1q}} & \cdots & \frac{\partial(Y)_{rs}}{\partial(X)_{1q}} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \frac{\partial(Y)_{11}}{\partial(X)_{p1}} & \cdots & \frac{\partial(Y)_{1s}}{\partial(X)_{p1}} & \cdots & \frac{\partial(Y)_{11}}{\partial(X)_{pq}} & \cdots & \frac{\partial(Y)_{1s}}{\partial(X)_{pq}} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \frac{\partial(Y)_{r1}}{\partial(X)_{p1}} & \cdots & \frac{\partial(Y)_{rs}}{\partial(X)_{p1}} & \cdots & \frac{\partial(Y)_{r1}}{\partial(X)_{pq}} & \cdots & \frac{\partial(Y)_{rs}}{\partial(X)_{pq}} \end{pmatrix} \\
 &= \begin{pmatrix} \sum_{i,j=1}^{p,q} \frac{\partial(Y)_{11}}{\partial(X)_{ij}} d(X)_{i,j} & \cdots & \sum_{i,j=1}^{p,q} \frac{\partial(Y)_{1s}}{\partial(X)_{ij}} d(X)_{i,j} \\ \vdots & \ddots & \vdots \\ \sum_{i,j=1}^{p,q} \frac{\partial(Y)_{r1}}{\partial(X)_{ij}} d(X)_{i,j} & \cdots & \sum_{i,j=1}^{p,q} \frac{\partial(Y)_{rs}}{\partial(X)_{ij}} d(X)_{i,j} \end{pmatrix} \\
 &= \begin{pmatrix} d(Y)_{11} & \cdots & d(Y)_{1s} \\ \vdots & \ddots & \vdots \\ d(Y)_{r1} & \cdots & d(Y)_{rs} \end{pmatrix}.
 \end{aligned}$$

The matrix integral of  $Z$  can be presented of the form:

$$Y = \int_{\mathbb{R}^{pq}} dY = \int_{\mathbb{R}^{pq}} dX * Z,$$

where  $z = \frac{dY}{dX}$ .

**Example 2.1.** Let us demonstrate how to find an matrix integral. Let  $X = (x_1 \quad x_2)$  and let

$$Z = \begin{pmatrix} x_1 & x_2 \\ 3x_1 & x_2^2 \end{pmatrix}$$

be the matrix derivative of matrix  $Y$  by matrix  $X$ . We can find matrix  $Y$  as follows:

$$\begin{aligned} Y &= \int_{\mathbb{R}^2} dX * Z = \int_{\mathbb{R}^2} (dx_1 \quad dx_2) * \begin{pmatrix} x_1 & x_2 \\ 3x_1 & x_2^2 \end{pmatrix} \\ &= \int_{\mathbb{R}^2} \begin{pmatrix} x_1 dx_1 + x_2 dx_2 \\ 3x_1 dx_1 + x_2^2 dx_1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2}(x_1^2 + x_2^2) + c_1 \\ \frac{3}{2}x_1^2 + \frac{1}{3}x_2^3 + c_2 \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{2}(x_1^2 + x_2^2) \\ \frac{3}{2}x_1^2 + \frac{1}{3}x_2^3 \end{pmatrix} + C, \end{aligned}$$

where

$$C = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}.$$

Next an example of the definite matrix integral is given.

**Example 2.2.** Let  $A = \begin{pmatrix} 1 & 1 \end{pmatrix}$  and  $B = \begin{pmatrix} 2 & 2 \end{pmatrix}$ . We shall find the definite matrix integral  $Y(B) - Y(A)$  of matrix  $Z$  given in Example 2.1. We obtain

$$Y(B) - Y(A) = \begin{pmatrix} 4 \\ 6 + \frac{8}{3} \end{pmatrix} - \begin{pmatrix} 1 \\ \frac{3}{2} + \frac{1}{3} \end{pmatrix} = \begin{pmatrix} 3 \\ \frac{41}{6} \end{pmatrix}.$$

### 3. Properties

Now we present some basic properties of the matrix integral. We assume without loss of generality that matrix  $C = 0$ . Let the dimensions of matrices  $X$ ,  $Y$  and  $Z$  be the same as in Definition 2.1. In these assumptions the following properties of the matrix integral are valid.

**Proposition 3.1.** Let  $X$  be a  $p \times q$ -matrix. Then

$$\int_{\mathbb{R}^{pq}} \text{vec} I_p \text{vec}' I_q \circ dX = X.$$

**Proof.** According to equality (1.2), matrix  $\text{vec} I_p \text{vec}' I_q$  is the matrix  $\frac{dX}{dX}$ . According to Definition 2.1 matrix  $X$  is the matrix integral of the matrix  $\text{vec} I_p \text{vec}' I_q$ .  $\square$

**Proposition 3.2.** Let  $Y = U + V$ , where  $U$  and  $V$  are  $r \times s$ -matrices. Then

$$\int_{\mathbb{R}^{pq}} Y \circ dX = \int_{\mathbb{R}^{pq}} U \circ dX + \int_{\mathbb{R}^{pq}} V \circ dX.$$

**Proof.** From distributivity of the Kronecker product:

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

we obtain

$$\frac{d}{dX} \otimes Y = \frac{d}{dX} \otimes (U + V) = \frac{d}{dX} \otimes U + \frac{d}{dX} \otimes V.$$

This proves the proposition.  $\square$

**Proposition 3.3.** Let  $A : pq \otimes pq$  be a constant diagonal matrix and let  $X : p \otimes q$ . Then

$$\int_{\mathfrak{R}_{pq}} A \circ dX = \begin{pmatrix} (A)_1(X)_{11} \\ (A)_2(X)_{21} \\ \vdots \\ (A)_{pq}(X)_{pq} \end{pmatrix},$$

where  $(A)_l = (A)_{ll}$ ,  $l = 1, \dots, pq$ .

**Proof.** The statement follows from the definition of the Kronecker product and from Definition 2.1.  $\square$

Let  $g(x)$  be a continuous function of  $p$ -vector  $x$  and let the derivative (1.3) of  $k$ th order of the function  $g(x)$  exist and be denoted by  $g^{(k)}(x)$ . Denote the inner product of  $p$ -vectors  $a$  ja  $b$  as  $(a, b) : (a, b) = a'b$ . Let us denote the  $i$ th element of vector  $x$  as  $(x)_i$ ,  $i = 1, 2, \dots, p$ . For notational conveniences we introduce the vector

$$e_p = \underbrace{(1, \dots, 1)'}_{p \text{ times}}$$

and the operator

$${}_s dx := \sum_{i=1}^p d(x)_i.$$

In the next properties we assume, that the matrix derivative is expressed as in Eq. (1.1).

**Proposition 3.4.** The next relation holds in the notations, given above

$$\int_{\mathfrak{R}_p} g^{(k)}(x) \circ dx = g^{(k-1)'}(x).$$

**Proof.** If  $k = 1$ , then the equality is obvious according to Definition 2.1. By means of Eq. (1.3) and Definition 2.1 the proposition is valid by every natural  $k$ .  $\square$

**Proposition 3.5.** Let  $A : n \times p^{k-1}$  be a constant matrix and  $A_l, l = 1, \dots, n$  the row vector of this matrix. Then

$$\int_{\mathfrak{M}_p} A_g^{(k)}(x) \circ dx' = \begin{pmatrix} (A_1, \text{vec}g^{(k-1)})(x) \\ \vdots \\ (A_n, \text{vec}g^{(k-1)})(x) \end{pmatrix}.$$

**Proof.** Apply the operator  $\frac{d}{dx'}$  to the right-hand side of the statement. It is easy to ensure that the matrices

$$\frac{d}{dx'} \otimes \begin{pmatrix} (A_1, \text{vec}g^{(k-1)})(x) \\ \vdots \\ (A_n, \text{vec}g^{(k-1)})(x) \end{pmatrix}$$

and  $A_g^{(k)}$  have the same dimensions. Now we must ensure that the corresponding elements are equal. We obtain that

$$\left( \frac{d}{dx'} \otimes \begin{pmatrix} (A_1, \text{vec}g^{(k-1)})(x) \\ \vdots \\ (A_n, \text{vec}g^{(k-1)})(x) \end{pmatrix} \right)_{lm} = \sum_{j=1}^{p^{k-1}} (A)_{lj} (g^{(k)})_{jm}(x) = (A_g^{(k)}(x))_{lm},$$

where  $l = 1, \dots, n$  and  $m = 1, \dots, p$ . Thus the corresponding elements of matrices are equal.  $\square$

**Proposition 3.6.** Let  $a$  be a constant not depending on vector  $x$ . Then

$$\int_{\mathfrak{M}_p} \left( \frac{d}{dx'}, e_p \right) ag(x) \circ_s dx = ag(x).$$

**Proof.** It is easy to check that the scalar differential operator can be presented as the inner product:

$$\frac{d}{s dx} = \left( \frac{d}{dx'}, e_p \right).$$

Thus

$$\begin{aligned} & \int_{\mathfrak{M}_p} \left( \frac{d}{dx'}, e_p \right) ag(x) *_s dx \\ &= ag(x) = \int_{\mathfrak{M}_p} \frac{d}{s dx} ag(x) *_s dx \\ &= \int_{\mathfrak{M}_p} dag(x) = ag(x). \quad \square \end{aligned}$$



**Proposition 3.7.** *Let functions  $g$  and  $G$  be such that*

$$g(x) = \frac{\partial G(x)}{\partial(x)_1 \dots \partial(x)_p}.$$

*Then the next relation is valid*

$$\underbrace{\int \dots \int}_p g(x) d(x)_1 \dots d(x)_p = \int_{\mathfrak{N}_p} \frac{d}{d_s} G(x) \circ_s dx.$$

**Proof.** By integrating the left-hand side of the statement we obtain

$$\underbrace{\int \dots \int}_p \frac{\partial G(x)}{\partial(x)_1 \dots \partial(x)_p} d(x)_1 \dots d(x)_p = G(x).$$

Applying the matrix integral to the right-hand side of the statement we obtain

$$\int_{\mathfrak{N}_p} \frac{d}{d_s} G(x) *_s dx = G(x).$$

Proposition 7 concludes the next property.  $\square$

**Proposition 3.8.** *Let  $a$  be a constant  $p$ -vector. Then*

$$\underbrace{\int \dots \int}_p \left( a, \frac{d}{dx}^{\otimes k} \right) g(x) d(x)_1 \dots d(x)_p = \left( a, \frac{d}{dx}^{\otimes k} \right) \int_{\mathfrak{N}_p} \frac{d}{d_s} G(x) \circ_s dx.$$

Now we shall use the considered properties in an example.

**Example 3.1.** Let  $x = ((x)_1 \quad (x)_2)'$ ,  $g(x) = (x)_1^2 + (x)_2^2$  and  $A = (1 \quad 2 \quad 3)'$ . Using the star product we obtain

$$\int_{\mathfrak{N}_p} g^{(2)}(x) \circ dx = \int_{\mathfrak{N}_p} dx * g^{(2)}(x) = \begin{pmatrix} 2(x)_1 \\ 2(x)_2 \end{pmatrix} = (g^{(1)})'.$$

So we have demonstrated the usage of Proposition 3.4.

**Example 3.2.** Let us apply the Proposition 3.5 to the matrices from the previous example. We obtain

$$\int_{\mathfrak{N}_p} Ag^{(1)}(x) \circ dx' = \int_{\mathfrak{N}_p} dx' * Ag^{(1)}(x) = \begin{pmatrix} (x)_1^2 + (x)_2^2 \\ 2(x)_1^2 + 2(x)_2^2 \\ 3(x)_1^2 + 3(x)_2^2 \end{pmatrix}.$$

#### 4. Application

We shall show how one can use matrix integral when integrating the relation between two density functions. Let  $X$  and  $Y$  be random  $p$ -vectors. The formula of approximation of the density function of  $f_Y(x)$  by  $f_X(x)$  in Kollo and von Rosen [3] has the form:

$$f_Y(x) \approx f_X(x) + A' \text{vec} f_X^{(1)}(x) + \text{vec}' B \text{vec} f_X^{(2)}(x) + \text{vec}' C \text{vec} f_X^{(3)}(x), \quad (4.1)$$

where  $A$  is a  $p$ -vector,  $B$  is a  $p \times p$ -matrix and  $C$  is a  $p^2 \times p$ -matrix. Now we apply matrix integral to the relation 4.1. Let  $F_Y(x)$  and  $F_X(x)$  be distribution functions corresponding to the density functions  $f_Y(x)$  and  $f_X(x)$  respectively. By definition

$$F_X(x_0) = \int_{-\infty}^{(x_0)_1} \int_{-\infty}^{(x_0)_2} \cdots \int_{-\infty}^{(x_0)_p} f_X(x) d(x)_1 d(x)_2 \cdots d(x)_p.$$

After using Propositions 7 and 8 we obtain the next relation between distribution functions  $F_Y(x_0)$  and  $F_X(x_0)$ :

$$\begin{aligned} F_Y(x_0) \approx F_X(x_0) + A' e_p \otimes f(x)(x_0) + \sum_{l=1}^p B_l' e_p \otimes f_{Xl}^1(x_0) \\ + \sum_{l=1}^p \sum_{j=1}^p C'_{(l-1)p+j} e_p \otimes f_{Xlj}^2(x_0), \end{aligned}$$

where  $B_l$  denotes the  $l$ th column vector of matrix  $B$ ,  $C_{(l-1)p+j}$  denotes the  $((l-1)p+j)$ th column vector of matrix  $C$  and  $f_{Xl}^1$  is the derivative of function  $f_X(x)$  by the  $l$ th component.

So we have demonstrated an application of the matrix integral for integrating a relation between two multivariate density functions.

#### Acknowledgements

The author is thankful to two referees whose comments and suggestions were helpful in preparation of the final version of the paper. The author is grateful to the Estonian Research Foundation for support through grant 4366 and an object-oriented grant 1776.

#### References

- [1] A. Harville, Matrix Algebra from a Statistician's Perspective, Springer, New York, 1997.
- [2] T. Kollo, Matrix Derivative in Multivariate Statistics, Tartu University Press, Tartu, 1991 (in Russian).

- [3] T. Kollo, D. von Rosen, A unified approach to the approximation of multivariate densities, *Scand. J. Statist.* 25 (1998) 93–109.
- [4] E.C. MacRae, Matrix derivatives with an applications to an adaptive linear decision problem, *Ann. Stat.* 7 (1974) 381–394.
- [5] J.R. Magnus, H. Neudecker, *Matrix Differential Calculus with Applications in Statistics and Econometrics*, Revised Edition, Wiley, New York, 1999.
- [6] H. Neudecker, Some theorems on matrix differentiations with special reference to Kronecker matrix products, *J. Amer. Stat. Assoc.* 64 (1969) 953–963.
- [7] R. Schott, *Matrix Analysis for Statistics*, Wiley, New York, 1997.