

W. GAWROŃSKI (Gdańsk)

**MATRIX FUNCTIONS:
TAYLOR EXPANSION, SENSITIVITY AND ERROR ANALYSIS**

The paper presents the Taylor expansion of a matrix function, where its argument is a matrix itself. On this basis the sensitivity analysis of a matrix function is provided. For a matrix function of a matrix random argument, the error propagation law is derived. The derivation is provided on the basis of the presented Taylor expansion and a new definition of the correlation matrix. The method is applied to system theory and to the analysis of computation errors of systems.

1. Introduction. The paper deals with matrix-valued functions of a matrix argument. The definition of these functions is the following.

Let A be a $(p \times q)$ -matrix, and B an $(s \times t)$ -matrix. If the entries of A depend on B , i.e.,

$$a_{ij} = a_{ij}(B) \quad (i = 1, \dots, p; j = 1, \dots, q),$$

then the function ⁽¹⁾

$$A = [a_{ij}(B)] = A(B) \quad (i = 1, \dots, p; j = 1, \dots, q)$$

is called a *matrix function of a matrix argument*.

In many technical problems, the Taylor expansion is used, e.g., in sensitivity analysis [5], [7], [10], in optimum system theory [1], in stability [13], in identification of systems [8], as well as in the analysis of data errors in computation of large systems [4], [6], [7]. For some of these problems the notation of the Taylor expansion of a matrix function of a matrix argument is required. This paper gives the Taylor expansion of such a function, and some new matrix operations are presented along with it. The notation of Taylor series for such a function has not often appeared in the literature. Flaming [3] and Deutsch [2] have considered a vector function of a vector argument, and Vetter [17] has considered

⁽¹⁾ Descriptions of the symbols, definitions and basic relationships can be found in Appendix A.

a matrix function of a vector argument. These Taylor expansions can be used for a matrix function of a matrix argument if all objects of the function are represented in the vector form. In many cases it can be done by the column or row transformation (see [1], [17], and also Appendix A); these transformations are not convenient, however, because the two notations are used for one mathematical object (for matrix notation and vector notation — see [4]-[7]).

The notation of the Taylor expansion of a matrix function of a matrix argument proposed in this paper is compact and does not change the form of the objects. On this basis, the sensitivity analysis of matrix functions is presented.

In the statistical analysis of random vectors the correlation matrix is used as a measure of errors of the vectors. The correlation matrix of random matrices is not defined. In this paper two definitions of the correlation matrix of random matrices are given. On this basis, two methods of notation of the error propagation law are derived.

2. Matrix derivative, differential and integral. The derivative of a $(p \times q)$ -matrix function A with respect to the $(s \times t)$ -matrix argument B is defined by Vetter [17] by

$$(1) \quad \frac{\partial A(B)}{\partial B} = \left[\frac{\partial A}{\partial b_{kl}} \right] \quad (k = 1, \dots, s; l = 1, \dots, t),$$

where

$$(2) \quad \frac{\partial A}{\partial b_{kl}} = \left[\frac{\partial a_{ij}}{\partial b_{kl}} \right] \quad (i = 1, \dots, p; j = 1, \dots, q).$$

If the operator $\partial/\partial B$ is defined by

$$\frac{\partial}{\partial B} = \left[\frac{\partial}{\partial b_{kl}} \right] \quad (k = 1, \dots, s; l = 1, \dots, t),$$

the derivative of the matrix A with respect to the matrix B can be presented in the form

$$\frac{\partial A}{\partial B} = \frac{\partial}{\partial B} \otimes A.$$

This description of the matrix derivative leads to the following conclusions:

$$\frac{\partial^{\otimes k} A}{(\partial B)^{\otimes k}} = \frac{\partial}{\partial B} \otimes \frac{\partial}{\partial B} \otimes \dots \otimes \frac{\partial}{\partial B} \otimes A \quad (k \text{ factors}),$$

$$\frac{\partial^{\otimes 3} A}{\partial B \otimes \partial C \otimes \partial D} = \frac{\partial}{\partial B} \otimes \frac{\partial}{\partial C} \otimes \frac{\partial}{\partial D} \otimes A,$$

$$\left(\frac{\partial A}{\partial B} \right)^T = \left(\frac{\partial}{\partial B} \otimes A \right)^T = \frac{\partial}{\partial B^T} \otimes A^T = \frac{\partial A^T}{\partial B^T}.$$

Notice that a derivative of a matrix A with respect to a matrix B can be given as

$$\frac{\partial_R A}{\partial_R B} = \left[\frac{\partial a_{ij}}{\partial B} \right] \quad (i = 1, \dots, p; j = 1, \dots, q),$$

where

$$\frac{\partial a_{ij}}{\partial B} = \left[\frac{\partial a_{ij}}{\partial b_{kl}} \right] \quad (k = 1, \dots, s; l = 1, \dots, t).$$

This definition gives

$$\frac{\partial_R A}{\partial_R B} = A \otimes \frac{\partial}{\partial B}.$$

This matrix derivative is called the *right derivative* of A with respect to B . Later we use only the *left derivative* (1), since the relationships for the right derivative are similar (although they may be more convenient).

For the $(s \times t)$ -matrix B , the t -unit matrix function I_t and the s -unit matrix function I_s , from [17] we take useful relationships of the matrix derivative:

$$(3) \quad \frac{\partial(AC)}{\partial B} = \frac{\partial A}{\partial B} (I_t \otimes C) + (I_s \otimes A) \frac{\partial C}{\partial B},$$

$$(4) \quad \frac{\partial B}{\partial B} = E_{s \times t}^{s \times t},$$

$$(5)_i^j \quad \frac{\partial B^T}{\partial B} = E_{s \times t}^{t \times s}.$$

The *matrix differential* is defined by

$$(6) \quad dA = [da_{ij}] \quad (i = 1, \dots, p; j = 1, \dots, q).$$

The matrix row or column transformation for the $(p \times q)$ -matrix function A , the q -unit matrix function I_q and the p -unit matrix function I_p gives the differential of $A(B)$ in the form [17]

$$dA(B) = \frac{\partial A}{\partial_{rs} B} (d(rsB)^T \otimes I_q) = (d(csB)^T \otimes I_p) \frac{\partial A}{\partial_{cs} B}.$$

The matrix differential can be expressed avoiding the column or row transformation; in this case the differential is defined by (6), and

$$(7) \quad da_{ij} = \frac{\partial a_{ij}}{\partial B} \circ dB,$$

since

$$\frac{\partial a_{ij}}{\partial B} \circ dB = \text{tr} \left(\frac{\partial a_{ij}}{\partial B} dB^* \right) = \sum_{k=1}^s \sum_{l=1}^t \frac{\partial a_{ij}}{\partial b_{kl}} db_{kl} \stackrel{\text{df}}{=} da_{ij}(B).$$

For the matrix differential the following relationships are valid:

$$(8) \quad d(FG) = dFG + FdG,$$

$$(9) \quad d(F \otimes G) = dF \otimes G + F \otimes dG,$$

$$(10) \quad d(F \circ G) = dF \circ G + F \circ dG.$$

Equation (8) is well known (see, for example, [2]), equation (9) can be found in [17], and (10) is derived from the definition of the inner product and from (8):

$$d(F \circ G) = d(\text{tr}(FG^*)) = \text{tr}(d(FG^*)) = \text{tr}(dFG^* + FdG^*) = dF \circ G + F \circ dG.$$

A *matrix integral* is defined by

$$H = \int_{A_1}^{A_2} da = \left[\int_{a_{ij}^1}^{a_{ij}^2} da \right] = A_2 - A_1 \quad (i = 1, \dots, p; j = 1, \dots, q).$$

From the definition of the differentials (6) and (7) we obtain

$$h_{ij} = \int_{a_{ij}^1}^{a_{ij}^2} da = \int_{B_1}^{B_2} \frac{\partial a_{ij}}{\partial B} \circ dB.$$

The formula

$$H = \int_{B_1}^{B_2} F(B) dG(B) = \int_{B_1}^{B_2} F(B) \left[\frac{\partial g_{kl}(B)}{\partial B} \circ dB \right] \quad (k = 1, \dots, p; l = 1, \dots, q)$$

is a generalization of the latter integral formula.

The corollaries to formulas (8)-(10) are the following integration by part relationships:

$$(11) \quad \int_{B_1}^{B_2} F(B) dG(B) = F(B)G(B) \Big|_{B_1}^{B_2} - \int_{B_1}^{B_2} dF(B)G(B),$$

$$\int_{B_1}^{B_2} F(B) \otimes dG(B) = F(B) \otimes G(B) \Big|_{B_1}^{B_2} - \int_{B_1}^{B_2} dF(B) \otimes G(B),$$

$$(12) \quad \int_{B_1}^{B_2} F(B) \circ dG(B) = F(B) \circ G(B) \Big|_{B_1}^{B_2} - \int_{B_1}^{B_2} dF(B) \circ G(B).$$

3. Matrix Taylor expansion. Let $A = [a_{ij}]$ ($i = 1, \dots, p; j = 1, \dots, q$) be an $(N+1)$ -differentiable matrix-valued function of a matrix argument B . Then, for the $(p \times q)$ -matrix function A and the $(s \times t)$ -matrix argument B the Taylor expansion of $A(B)$ takes the form

$$(13) \quad a_{ij}(B) = a_{ij}(B_0) + \sum_{k=1}^N \frac{1}{k!} \frac{\partial^{\otimes k} a_{ij}}{(\partial B)^{\otimes k}} \Big|_{B=B_0} \circ (B - B_0)^{\otimes k} + r_{ij}^{(N+1)}$$

$$(i = 1, \dots, p; j = 1, \dots, q),$$

where

$$(14) \quad r_{ij}^{(N+1)} = \frac{-1}{(N+1)!} \int_{B_0}^B \frac{\partial^{\otimes(N+1)} a_{ij}}{(\partial Z)^{\otimes(N+1)}} \circ d(B - Z)^{\otimes(N+1)}$$

is the ij -th entry of the remainder matrix $R^{(N+1)}$,

$$R^{(N+1)} = [r_{ij}^{(N+1)}] \quad (i = 1, \dots, p; j = 1, \dots, q).$$

Proof. We prove this theorem by induction. At first we show that (13) is true for the sum consisting of one element ($N = 1$), and then, assuming that (13) is valid for the sum of $N - 1$ elements, we prove it for the sum of N elements.

In the integration formula (11) we put $F(B) = A(B)$, $G(B) = I$, $B_1 = B_0$, $B_2 = B$, and we get

$$a_{ij}(B) = a_{ij}(B_0) + \int_{B_0}^B da_{ij}(Z).$$

By (7) we have

$$da_{ij}(Z) = \frac{\partial a_{ij}}{\partial Z} \circ dZ = -\frac{\partial a_{ij}}{\partial Z} \circ d(B - Z),$$

therefore

$$a_{ij}(B) = a_{ij}(B_0) - \int_{B_0}^B \frac{\partial a_{ij}}{\partial Z} \circ d(B - Z).$$

Formula (12) gives

$$a_{ij}(B) = a_{ij}(B_0) - \frac{\partial a_{ij}}{\partial Z} \circ (B - Z) \Big|_{B_0}^B + \int_{B_0}^B d \left(\frac{\partial a_{ij}}{\partial Z} \right) \circ (B - Z)$$

or

$$a_{ij}(B) = a_{ij}(B_0) + \frac{\partial a_{ij}}{\partial B} \Big|_{B=B_0} \circ (B - B_0) + r_{ij}^{(2)},$$



where, from (B.9) in Appendix B, we find

$$r_{ij}^{(2)} = \frac{-1}{2!} \int_{B_0}^B \frac{\partial^{\otimes 2} a_{ij}}{(\partial Z)^{\otimes 2}} \circ d(B-Z)^{\otimes 2},$$

so that (13) is proved for $N = 2$.

Let (13) be valid for the sum of $N - 1$ elements, i.e., let

$$(15) \quad a_{ij}(B) = a_{ij}(B_0) + \sum_{k=1}^{N-1} \frac{1}{k!} \frac{\partial^{\otimes k} a_{ij}}{(\partial B)^{\otimes k}} \Big|_{B=B_0} \circ (B-B_0)^{\otimes k} + r_{ij}^{(N)},$$

where

$$(16) \quad r_{ij}^{(N)} = \frac{-1}{N!} \int_{B_0}^B \frac{\partial^{\otimes N} a_{ij}}{(\partial Z)^{\otimes N}} \circ d(B-Z)^{\otimes N}.$$

Applying formula (12) to (16) we obtain

$$r_{ij}^{(N)} = -\frac{1}{N!} \frac{\partial^{\otimes N} a_{ij}}{(\partial Z)^{\otimes N}} \circ (B-Z)^{\otimes N} + \frac{1}{N!} \int_{B_0}^B d \left(\frac{\partial^{\otimes N} a_{ij}}{(\partial Z)^{\otimes N}} \right) \circ (B-Z)^{\otimes N}$$

or

$$(17) \quad r_{ij}^{(N)} = \frac{1}{N!} \frac{\partial^{\otimes N} a_{ij}}{(\partial B)^{\otimes N}} \Big|_{B=B_0} \circ (B-B_0)^{\otimes N} + r_{ij}^{(N+1)},$$

where

$$(18) \quad r_{ij}^{(N+1)} = \frac{1}{N!} \int_{B_0}^B d \left(\frac{\partial^{\otimes N} a_{ij}}{(\partial Z)^{\otimes N}} \right) \circ (B-Z)^{\otimes N}.$$

By (B.9) we transform (18) to the form

$$(19) \quad r_{ij}^{(N+1)} = \frac{-1}{(N+1)!} \int_{B_0}^B \frac{\partial^{\otimes (N+1)} a_{ij}}{(\partial Z)^{\otimes (N+1)}} \circ d(B-Z)^{\otimes (N+1)}.$$

By (15), (17) and (19) we obtain finally

$$a_{ij}(B) = a_{ij}(B_0) + \sum_{k=1}^N \frac{1}{k!} \frac{\partial^{\otimes k} a_{ij}}{(\partial B)^{\otimes k}} \Big|_{B=B_0} \circ (B-B_0)^{\otimes k} + r_{ij}^{(N+1)},$$

where $r_{ij}^{(N+1)}$ is given by (19). Thus expansion (13) is valid.

4. Sensitivity analysis. The above-described Taylor expansion gives the possibility of calculating sensitivity functions of first and/or higher

orders for matrix relationships. Let us consider a first order sensitivity function, and let the matrix A depend on the matrix B , i.e.,

$$(20) \quad A = A(B).$$

The sensitivity function, which describes the relation between the variation ΔB of the matrix B with respect to a fixed value B_0 , and the variation of the matrix A with respect to A_0 ($A_0 = A(B_0)$), is given by

$$\Delta A = S(B_0, \Delta B).$$

We find this function expanding (20) in a Taylor series with respect to B_0 and neglecting the terms of second and higher orders. Thus we have

$$A(B) = A(B_0) + dA|_{B=B_0}.$$

Writing $\Delta A = A(B) - A(B_0)$, by (13) we obtain

$$(21) \quad \Delta A = S(B_0, \Delta B) = [S_{ij}^{aB} \circ \Delta B] \quad (i = 1, \dots, p; j = 1, \dots, q),$$

where

$$(22) \quad S_{ij}^{aB} = \left. \frac{\partial a_{ij}}{\partial B} \right|_{B=B_0}.$$

Example 1. Given the (3×2) -matrix function A of a (2×2) -matrix argument B represented by $A(B) = KBB^T$, where K is the constant matrix:

$$K = \begin{bmatrix} 1 & 0 \\ 2 & 3 \\ -1 & 2 \end{bmatrix}.$$

If the variation ΔB of B is

$$\Delta B = \begin{bmatrix} 0.03 & 0.10 \\ -0.20 & 0.05 \end{bmatrix}$$

with respect to $B = B_0$,

$$B_0 = \begin{bmatrix} 1 & -2 \\ 3 & -5 \end{bmatrix},$$

then the variation ΔA of A is found. The relation between ΔB and ΔA is given by (21) and (22). To use it, first we find the derivative $\partial A / \partial B$ from (3) for a 2-matrix I :

$$\frac{\partial A}{\partial B} = \frac{\partial(KBB^T)}{\partial B} = (I \otimes K) \frac{\partial(BB^T)}{\partial B}.$$

Moreover, for a 2-matrix I , equations (3)-(5) give

$$\begin{aligned} \frac{\partial(BB^T)}{\partial B} &= \frac{\partial B}{\partial B}(I \otimes B^T) + (I \otimes B) \frac{\partial B^T}{\partial B} = E_{2 \times 2}^2(I \otimes B^T) + (I \otimes B)E_{2 \times 2}^{2 \times 2} \\ &= \begin{bmatrix} 2b_{11} & b_{21} & 2b_{12} & b_{22} \\ b_{21} & 0 & b_{22} & 0 \\ 0 & b_{11} & 0 & b_{12} \\ b_{11} & 2b_{21} & b_{12} & 2b_{22} \end{bmatrix} \end{aligned}$$

or

$$\left. \frac{\partial A}{\partial B} \right|_{B=B_0} = (I \otimes K) \left. \frac{\partial(BB^T)}{\partial B} \right|_{B=B_0} = \begin{bmatrix} 2 & 3 & -4 & -5 \\ 13 & 6 & -23 & -10 \\ -4 & -3 & -2 & 5 \\ 0 & 1 & 0 & -2 \\ 3 & 20 & -6 & -34 \\ 2 & -11 & -4 & -18 \end{bmatrix}.$$

Therefore, we have

$$\begin{aligned} S_{11}^{aB} &= \left. \frac{\partial a_{11}}{\partial B} \right|_{B=B_0} = \begin{bmatrix} 2 & -4 \\ 0 & 0 \end{bmatrix}, & S_{12}^{aB} &= \left. \frac{\partial a_{12}}{\partial B} \right|_{B=B_0} = \begin{bmatrix} 3 & -5 \\ 1 & 2 \end{bmatrix}, \\ S_{21}^{aB} &= \left. \frac{\partial a_{21}}{\partial B} \right|_{B=B_0} = \begin{bmatrix} 13 & -23 \\ 3 & -6 \end{bmatrix}, & S_{22}^{aB} &= \left. \frac{\partial a_{22}}{\partial B} \right|_{B=B_0} = \begin{bmatrix} 6 & -10 \\ 20 & -34 \end{bmatrix}, \\ S_{31}^{aB} &= \left. \frac{\partial a_{31}}{\partial B} \right|_{B=B_0} = \begin{bmatrix} -4 & -2 \\ 2 & -4 \end{bmatrix}, & S_{32}^{aB} &= \left. \frac{\partial a_{32}}{\partial B} \right|_{B=B_0} = \begin{bmatrix} -3 & 5 \\ -11 & -18 \end{bmatrix}, \end{aligned}$$

and the variations of the entries of A are

$$\begin{aligned} \Delta a_{11} &= S_{11}^{aB} \circ \Delta B = -0.34, & \Delta a_{12} &= S_{12}^{aB} \circ \Delta B = -0.71, \\ \Delta a_{21} &= S_{21}^{aB} \circ \Delta B = -2.81, & \Delta a_{22} &= S_{22}^{aB} \circ \Delta B = -6.52, \\ \Delta a_{31} &= S_{31}^{aB} \circ \Delta B = -0.92, & \Delta a_{32} &= S_{32}^{aB} \circ \Delta B = 1.71, \end{aligned}$$

or

$$\Delta A = \begin{bmatrix} -0.34 & -0.71 \\ -2.81 & -6.52 \\ -0.92 & 1.71 \end{bmatrix}.$$

5. Error analysis. Error analysis for vector functions is provided on the basis of the Taylor expansion. Expanding the function $x = x(y)$ (x and y are vectors) in a Taylor series and neglecting the second and higher order terms, we obtain

$$x = x_0 + S(y - y_0),$$

where

$$S = \left. \frac{\partial x}{\partial y} \right|_{y=y_0},$$

and y is the random vector with the mean value y_0 and the correlation matrix R_y ,

$$(23) \quad R_y = E\{(y - y_0)(y - y_0)^*\}.$$

The mean value of the vector x is given by the relationship $x_0 = x(y_0)$, and the correlation matrix of x can be obtained from

$$(24) \quad R_x = E\{(x - x_0)(x - x_0)^*\} = SR_y S^*.$$

Definition (23) of the correlation matrix cannot be used for random matrices, but there are at least two methods of defining the correlation matrix for matrix random variables. The first method consists in the column or row transformation of the matrix. In this way, from a $(p \times q)$ -matrix A we obtain the vector $a = \text{cs}(A)$ of dimension pq . The correlation matrix of this vector is defined by

$$(25) \quad R_A^{(1)} = E\{(a - a_0)(a - a_0)^*\}, \quad \text{where } a_0 = E(a).$$

This matrix is called the *first form correlation matrix* of A . It is a square matrix of dimensions $pq \times pq$. The variance of a_{ij} is at the k -th place on the main diagonal of $R_A^{(1)}$, where

$$(26) \quad k = (j - 1)p + i.$$

The correlation coefficient of a_{ij} and a_{uv} is in the k -th row and the l -th column of $R_A^{(1)}$, where k is given by (26), and l by the equation $l = (v - 1)p + u$.

In the second method we may avoid the column transformation, i.e., renumbering the matrix entries, if we let the correlation matrix of the matrix A be in the form

$$(27) \quad R_A^{(2)} = E\{(A - A_0) \otimes (\bar{A} - \bar{A}_0)\}, \quad \text{where } A_0 = E(A).$$

This matrix is called the *second form correlation matrix* of A . It is a rectangular matrix of dimensions $p^2 \times q^2$. The difference between matrices $R_A^{(1)}$ and $R_A^{(2)}$ consists in the distribution of their entries. The correlation matrix $R_A^{(2)}$ contains pq blocks, each of which has dimensions $p \times q$. The variance of a_{ij} is in the ij -th block, at the ij -th position in it. In other words, variances of the entries of A are in the same places as non-zero elements in the permutation matrix $E_{p \times q}^{p \times q}$. The correlation coefficient of a_{ij} with a_{uv} is in the ij -th block at the uv -th position in it.

Note that, alternatively to (27), we can define the second form correlation matrix

$$(27a) \quad R_A^{(2)} = E\{(A - A_0) \otimes (A - A_0)^*\}.$$

The relationship between the first form correlation matrices is the same as between the correlation matrices of vectors (see (24)). This follows from definition (25), where the matrix A is represented as a vector a .

The relationship between the correlation matrices $R_A^{(2)}$ and $R_B^{(2)}$, where $A = A(B)$, can be found as follows. From (21) we obtain

$$(28) \quad \Delta A = [S_{ij}^{aB} \circ \Delta B] = [\text{tr}(S_{ij}^{aB} \Delta B^*)] \quad (i = 1, \dots, p; j = 1, \dots, q),$$

where S_{ij}^{aB} is given by (22). Defining a matrix \mathcal{S}_{ijkl}^{aB} by

$$(29) \quad \mathcal{S}_{ijkl}^{aB} = S_{ij}^{aB} \otimes S_{kl}^{aB},$$

we get

$$(30) \quad R_A^{(2)} = [\mathcal{S}_{ijkl}^{aB} \circ R_B^{(2)}] \quad (i, k = 1, \dots, p; j, l = 1, \dots, q),$$

since the relationship (see [11] and [12])

$$\text{tr}(A \otimes B) = \text{tr}(A) \text{tr}(B)$$

implies

$$\begin{aligned} R_A^{(2)} &= E(\Delta A \otimes \Delta \bar{A}) = E[\text{tr}(S_{ij}^{aB} \Delta B^*) \text{tr}(\bar{S}_{kl}^{aB} \Delta \bar{B}^*)] \\ &= E[\text{tr}((S_{ij}^{aB} \Delta B^*) \otimes (\bar{S}_{kl}^{aB} \Delta \bar{B}^*))] = [\text{tr}(\mathcal{S}_{ijkl}^{aB} (R_B^{(2)})^*)] \\ &= [\mathcal{S}_{ijkl}^{aB} \circ R_B^{(2)}] \quad (i, k = 1, \dots, p; j, l = 1, \dots, q). \end{aligned}$$

Example 2. Let the argument B of the function $A(B)$ in Example 1 be a random matrix with the mean value

$$B_0 = \begin{bmatrix} 1 & -2 \\ 3 & -5 \end{bmatrix},$$

and with the second form correlation matrix

$$R_B^{(2)} = \begin{bmatrix} 0.4 & 0.0 & 0.0 & 0.1 \\ -0.02 & 0.1 & 0.02 & 0.0 \\ -0.02 & 0.02 & 0.1 & 0.0 \\ 0.2 & 0.0 & 0.0 & 0.5 \end{bmatrix}.$$

The mean value of A can be obtained from the relationship

$$A_0 = A(B_0) = K B_0 B_0^T = \begin{bmatrix} 5 & 13 \\ 49 & 128 \\ 19 & 55 \end{bmatrix},$$

and $R_A^{(2)}$ can be derived from (29) and (30). We compute two entries of $R_A^{(2)}$: $r_{1111} = \sigma_{a_{11}}^2$ (the variance of a_{11}) and r_{1232} (the correlation coefficient of a_{12} and a_{32}). From (29) we obtain

$$\mathcal{S}_{1111}^{aB} = S_{11}^{aB} \otimes \bar{S}_{11}^{aB} = \begin{bmatrix} 4 & -8 & -8 & 16 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\mathcal{S}_{1232}^{aB} = S_{12}^{aB} \otimes \bar{S}_{32}^{aB} = \begin{bmatrix} -9 & 15 & 15 & -25 \\ -33 & -54 & 55 & 90 \\ -3 & 5 & -6 & 10 \\ 11 & -18 & -22 & -36 \end{bmatrix},$$

and by (30) we have

$$\begin{aligned} r_{1111} &= \mathcal{S}_{1111}^{aB} \circ R_B^{(2)} = 1.76, \\ r_{1232} &= \mathcal{S}_{1232}^{aB} \circ R_B^{(2)} = -30.38. \end{aligned}$$

In the same way we find all entries of $R_A^{(2)}$, and we obtain

$$R_A^{(2)} = \begin{bmatrix} 1.76 & 4.68 & 4.68 & 9.18 \\ 18.04 & -0.40 & 21.30 & -30.32 \\ -3.44 & -6.68 & -9.68 & -30.38 \\ 18.04 & 21.30 & -0.40 & -30.32 \\ 120.38 & 105.04 & 105.04 & 628.80 \\ -5.48 & -3.60 & 79.76 & 255.92 \\ -3.44 & -9.68 & -6.68 & -30.38 \\ -5.48 & 79.76 & -3.60 & 255.92 \\ 18.96 & 43.66 & 43.66 & 200.38 \end{bmatrix}.$$

Example 3. Sensitivity and error of eigenvalues. Let us consider the eigenvalue problem

$$(31) \quad (A - \lambda_i I)x_i = 0,$$

where A is the symmetric $(n \times n)$ -matrix such that

1. A has the variation ΔA with respect to the fixed value A_0 . The variation $\Delta \lambda$ of the eigenvalue vector $\lambda = \text{col}(\lambda_i)$ ($i = 1, \dots, n$) with respect to $\lambda_0 = \lambda(A_0)$ should be found.

2. A is a random matrix with the mean value A_0 and with the second form correlation matrix $R_A^{(2)}$. The second form correlation matrix $R_\lambda^{(2)}$ of the eigenvalue vector should be found.

The problem of uncertainties of eigenvalues is often analyzed in the technical literature (see, for example, [4]-[6], [9], [14]-[16]). In [4]-[6] this problem is solved by the use of first form correlation matrices.

The eigenvalue problem can be presented as

$$(32) \quad \lambda_i = \frac{x_i^T A x_i}{x_i^T x_i}.$$

To solve both problems, stated above, we find the derivative $\partial \lambda_i / \partial A$. From (32) and (3) we obtain

$$\frac{\partial \lambda_i}{\partial A} = \frac{1}{x_i^T x_i} \frac{\partial (x_i^T A x_i)}{\partial A} + x_i^T A x_i \frac{\partial (x_i^T x_i)^{-1}}{\partial A},$$

but

$$\frac{\partial (x_i^T x_i)^{-1}}{\partial A} = - (x_i^T x_i)^{-2} \frac{\partial (x_i^T x_i)}{\partial A}.$$

Then

$$(33) \quad \frac{\partial \lambda_i}{\partial A} = (x_i^T x_i)^{-1} \left(\frac{\partial (x_i^T A x_i)}{\partial A} - \lambda_i \frac{\partial (x_i^T x_i)}{\partial A} \right).$$

We find the first derivative in (33) using formula (3):

$$\begin{aligned} \frac{\partial (x_i^T A x_i)}{\partial A} &= \frac{\partial x_i^T}{\partial A} (I \otimes A x_i) + (I \otimes x_i^T) \frac{\partial (A x_i)}{\partial A} \\ &= \frac{\partial x_i^T}{\partial A} (I \otimes A x_i) + (I \otimes x_i^T A) \frac{\partial x_i}{\partial A} + (I \otimes x_i^T) \frac{\partial A}{\partial A} (I \otimes x_i). \end{aligned}$$

From (4) we obtain

$$(I \otimes x_i^T) \frac{\partial A}{\partial A} (I \otimes x_i) = I \otimes x_i^T E_{n \times n}^{n \times n} (I \otimes x_i) = x_i x_i^T,$$

therefore

$$(34) \quad \frac{\partial (x_i^T A x_i)}{\partial A} = \frac{\partial x_i^T}{\partial A} (I \otimes A x_i) + (I \otimes x_i^T A) \frac{\partial x_i}{\partial A} + x_i x_i^T.$$

Similarly, we have

$$(35) \quad \frac{\partial (x_i^T x_i)}{\partial A} = \frac{\partial x_i^T}{\partial A} (I \otimes x_i) + (I \otimes x_i^T) \frac{\partial x_i}{\partial A}.$$

From (33), (34) and (35) we obtain

$$\frac{\partial \lambda_i}{\partial A} = (x_i^T x_i)^{-1} \left(\frac{\partial x_i^T}{\partial A} (I \otimes (A - \lambda_i I) x_i) + \left(I \otimes x_i^T (A - \lambda_i I) \frac{\partial x_i}{\partial A} + x_i x_i^T \right) \right).$$

By formula (31), the first two elements of the sum on the right-hand side of the above equation are equal to zero. Therefore

$$\frac{\partial \lambda_i}{\partial A} = \frac{x_i x_i^T}{x_i^T x_i}.$$

Putting

$$S_i^{\lambda A} = \frac{\partial \lambda_i}{\partial A} \Big|_{A=A_0} = \frac{x_{0i} x_{0i}^T}{g_{0i}},$$

where $g_{0i} = x_{0i}^T x_{0i}$ and $x_{0i} = x_i(A_0)$, we derive the variation of the eigenvector λ from formula (28):

$$\Delta \lambda = [S_i^{\lambda A} \circ \Delta A] = [g_{0i}^{-1} x_{0i} x_{0i}^T \circ \Delta A] \quad (i = 1, \dots, n).$$

From (29) we obtain the matrix $\mathcal{S}_{ij}^{\lambda A}$ in the form

$$\begin{aligned}\mathcal{S}_{ij}^{\lambda A} &= S_i^{\lambda A} \otimes S_j^{\lambda A} = (g_{0i}g_{0j})^{-1}(x_{0i}x_{0i}^T) \otimes (x_{0j}x_{0j}^T) \\ &= (g_{0i}g_{0j})^{-1}(x_{0i} \otimes x_{0j})(x_{0i}^T \otimes x_{0j}^T) = (g_{0i}g_{0j})^{-1}X_{0ij}X_{0ij}^T,\end{aligned}$$

where $X_{0ij} = x_{0i} \otimes x_{0j}$, and from (50) we derive the second form correlation matrix of the vector λ :

$$R_\lambda^{(2)} = [\mathcal{S}_{ij}^{\lambda A} \circ R_A^{(2)}] = [(g_{0i}g_{0j})^{-1}X_{0ij}X_{0ij}^T \circ R_A^{(2)}] \quad (i, j = 1, \dots, n).$$

6. Conclusions. The method of the Taylor expansion of a matrix function of a matrix argument, presented in this paper, has a compact notation. The application of this expansion to the analysis of the parametric sensitivity and the statistical analysis of matrix relationships is shown.

In the expansion the inner product of matrices is used. Notice that this product requires a small number of computations. If the components of the inner product are matrices of dimensions $p \times q$, then, to compute this product, pq multiplications and pq additions are required. The "regular" product requires p^2q multiplications and p^2q additions, and $(pq)^2$ multiplications and no additions are required for the direct product.

Appendix A. Some matrix notation.

$A = [a_{ij}]$ ($i = 1, \dots, p; j = 1, \dots, q$) — a rectangular ($p \times q$)-matrix;
 $\text{col}(x_i)$ ($i = 1, \dots, n$) — ($n \times 1$)-vector;
 \bar{A} — the complex conjugate matrix of A ;
 A^* — the complex conjugate transposition of A ;
 $\text{cs}(A)$ — the column transformation of a ($p \times q$)-matrix A ,

$$\text{cs}(A) = \text{col}(a_i) \quad (i = 1, \dots, q),$$

where a_i is the i -th column of A ;

$\text{rs}(A)$ — the row transformation of a ($p \times q$)-matrix A ,

$$\text{rs}(A) = [A_1, A_2, \dots, A_p],$$

where A_i is the i -th row of A ;

$\text{tr}(C) = \sum_{i=1}^p c_{ii}$ — trace of a square ($p \times p$)-matrix C ;

E^{kl} — the kl -th elementary ($p \times q$)-matrix, all zeros except 1 at the kl -th position;

$E_{p \times q}^{p \times q}$ — permutation matrix of dimensions $p^2 \times q^2$,

$$E_{p \times q}^{p \times q} = [E^{ij}] \quad (i = 1, \dots, p; j = 1, \dots, q);$$

$E_{q \times p}^{p \times q}$ — permutation matrix of dimensions $pq \times pq$,

$$E_{q \times p}^{p \times q} = [E^{ji}] \quad (i = 1, \dots, p; j = 1, \dots, q);$$

$A \otimes B$ — the direct or Kronecker product (see [11] and [12]) of a $(p \times q)$ -matrix A and an $(s \times t)$ -matrix B ,

$$A \otimes B = [a_{ij}B] \quad (i = 1, \dots, p; j = 1, \dots, q);$$

$A^{\otimes k}$ — the k -th Kronecker power of A :

$$A^{\otimes k} = A \otimes A \otimes \dots \otimes A \quad (k \text{ factors}),$$

$$A^{\otimes 0} = 1, \quad A^{\otimes 1} = A;$$

$A \circ B$ — the inner product of $(p \times q)$ -matrices A and B ,

$$A \circ B = \text{tr}(AB^*);$$

$E(\cdot)$ — the mean value operator.

Appendix B. Some properties of the differential.

1. The differential of the Kronecker power has the property

$$(B.1) \quad d(Y^{\otimes(r+1)}) = \sum_{i=0}^r Y^{\otimes i} \otimes dY \otimes Y^{\otimes(r-i)},$$

since (9) implies the equality

$$\begin{aligned} d(Y^{\otimes(r+1)}) &= d(Y \otimes Y^{\otimes r}) = dY \otimes Y^{\otimes r} + Y \otimes d(Y^{\otimes r}) = dY \otimes Y^{\otimes r} + \\ &+ Y \otimes dY \otimes Y^{\otimes(r-1)} + Y^{\otimes 2} \otimes d(Y^{\otimes(r-1)}) = \dots \\ &= \sum_{i=0}^r Y^{\otimes i} \otimes dY \otimes Y^{\otimes(r-i)}. \end{aligned}$$

Putting $Y = B - Z$ we have

$$d(B - Z)^{\otimes(r+1)} = - \sum_{i=0}^r (B - Z)^{\otimes i} \otimes dZ \otimes (B - Z)^{\otimes(r-i)}.$$

2. Let Y and U be matrices of dimensions $p \times q$, and let X be of dimensions $s \times t$; then

$$(B.2) \quad \left(\frac{\partial^{\otimes 2} a}{\partial X \otimes \partial Y} \right) \circ (dX \otimes U) = \left(\frac{\partial^{\otimes 2} a}{\partial Y \otimes \partial X} \right) \circ (U \otimes dX).$$

This can be shown by noting that

$$\left(\frac{\partial^{\otimes 2} a}{\partial X \otimes \partial Y} \right) \circ (dX \otimes U) = \sum_{i,j} \sum_{k,l} \frac{\partial^2 a}{\partial x_{ij} \partial y_{kl}} dx_{ij} u_{kl},$$

$$\left(\frac{\partial^{\otimes 2} a}{\partial Y \otimes \partial X} \right) \circ (U \otimes dX) = \sum_{k,l} \sum_{i,j} \frac{\partial^2 a}{\partial y_{kl} \partial x_{ij}} u_{kl} dx_{ij}.$$

Since the derivative does not depend on the succession of derivation, and since the value of the product does not depend on the succession of its components, it follows from the above equation that (B.2) is valid.

As the corollary to (B.2) we obtain

$$(B.3) \quad \left(\frac{\partial^{\otimes 3} a}{\partial X \otimes \partial Y \otimes \partial Z} \right) \circ (dX \otimes U \otimes V) \\ = \left(\frac{\partial^{\otimes 3} a}{\partial Y \otimes \partial X \otimes \partial Z} \right) \circ (U \otimes dX \otimes V) = \dots \\ = \left(\frac{\partial^{\otimes 3} a}{\partial Z \otimes \partial Y \otimes \partial X} \right) \circ (V \otimes U \otimes dX).$$

3. Let matrices S and Y be of dimensions $s \times t$, let a matrix X be of dimensions $p \times q$, and let the derivative of S with respect to X exist. Then

$$(B.4) \quad dS \circ Y = \left(\frac{\partial S}{\partial X} \right) \circ (dX \otimes Y).$$

Let us consider both sides of (B.4) separately. For the left-hand side we have

$$(B.5) \quad L = dS \circ Y = \text{tr}(dS Y^*) = \sum_{i=1}^s \sum_{j=1}^t ds_{ij} \bar{y}_{ij} \\ = \sum_{i=1}^s \sum_{j=1}^t \left(\frac{\partial s_{ij}}{\partial X} \circ dX \right) \bar{y}_{ij} = \sum_{i=1}^s \sum_{j=1}^t \sum_{k=1}^p \sum_{l=1}^q \frac{\partial s_{ij}}{\partial x_{kl}} dx_{kl} \bar{y}_{ij}.$$

To consider the right-hand side of (B.4) we put

$$C = \frac{\partial S}{\partial X} \quad \text{and} \quad F = dX \circ Y.$$

Matrices C and F are of dimensions $ps \times qt$, and

$$(B.6) \quad c_{mn} = \frac{\partial s_{ij}}{\partial x_{kl}},$$

$$(B.7) \quad f_{mn} = y_{ij} dx_k,$$

where $m = (k-1) + i$, $n = (l-1) + j$, and $i = 1, \dots, s$; $j = 1, \dots, t$; $k = 1, \dots, p$; $l = 1, \dots, q$; $m = 1, \dots, ps$; $n = 1, \dots, qt$. Evidently, in this notation the right-hand side of (B.4) is presented as

$$R = C \circ F = \text{tr}(CF^*) = \sum_{m=1}^{ps} \sum_{n=1}^{qt} c_{mn} \bar{f}_{mn}.$$

Substituting c_{mn} and f_{mn} from (B.6) and (B.7) into the latter equation we obtain

$$R = \sum_{i=1}^s \sum_{j=1}^t \sum_{k=1}^p \sum_{l=1}^q \frac{\partial s_{ij}}{\partial x_{kl}} \bar{y}_{ij} dx_{kl}.$$

It follows from comparison of (B.5) and (B.6) that $R = L$, i.e., that (B.4) is valid.

If in (B.4) we let $Y = (B-Z)^r$ and $X = Z$, then we get

$$dS \circ (B-Z)^{\otimes r} = \left(\frac{\partial S}{\partial Z} \right) \circ (dZ \otimes (B-Z)^{\otimes r}).$$

If, in addition, we put $S = \partial^{\otimes r} a / (\partial Z)^{\otimes r}$, then we obtain

$$d \left(\frac{\partial^{\otimes r} a}{(\partial Z)^{\otimes r}} \right) \circ (B-Z)^{\otimes r} = \frac{\partial^{\otimes (r+1)} a}{(\partial Z)^{\otimes (r+1)}} \circ (dZ \otimes (B-Z)^{\otimes r}).$$

From (B.3) it follows that

$$d \left(\frac{\partial^{\otimes r} a}{(\partial Z)^{\otimes r}} \right) \circ (B-Z)^{\otimes r} = \frac{\partial^{\otimes (r+1)} a}{(\partial Z)^{\otimes (r+1)}} \circ ((B-Z)^{\otimes i} \otimes dZ \otimes (B-Z)^{\otimes (r-i)})$$

$$(i = 0, 1, \dots, r),$$

whence

$$(B.8) \quad d \left(\frac{\partial^{\otimes r} a}{(\partial Z)^{\otimes r}} \right) \circ (B-Z)^{\otimes r}$$

$$= \frac{1}{r+1} \frac{\partial^{\otimes (r+1)} a}{(\partial Z)^{\otimes (r+1)}} \circ \left(\sum_{i=0}^r (B-Z)^{\otimes i} \otimes dZ \otimes (B-Z)^{\otimes (r-i)} \right).$$

Moreover, from (B.1) and (B.8) we obtain

$$(B.9) \quad d \left(\frac{\partial^{\otimes r} a}{(\partial Z)^{\otimes r}} \right) \circ (B-Z)^{\otimes r} = - \frac{1}{r+1} \frac{\partial^{\otimes (r+1)} a}{(\partial Z)^{\otimes (r+1)}} \circ d(B-Z)^{\otimes (r+1)}.$$

References

- [1] M. Athans, *The matrix minimum principle*, Information and Control 11 (1968), p. 592-606.
- [2] R. Deusch, *Estimation theory*, Prentice-Hall, Englewood Cliffs 1965.
- [3] W. H. Fleming, *Functions of several variables*, Addison-Wesley, Reading 1965.
- [4] W. Gawroński, *Błędy danych w obliczeniach układów mechanicznych*, Zeszyty Naukowe Politechniki Gdańskiej, Mechanika, 22 (1975), p. 1-102.

- [5] — *The effect of system parameter variations on natural frequencies*, Computers and Structures 5 (1975), p. 31-43.
- [6] — *Data errors in the computation of free motion*, ibidem 6 (1976), p. 17-27.
- [7] — *Two representations of sensitivity and error analysis of a dynamic system*, Archiwum Mechaniki Stosowanej 29 (1977), p. 53-67.
- [8] D. Graupe, *Identification of systems*, Van Nostrand, New York 1972.
- [9] G. C. Hart, *Eigenvalue uncertainty in stressed structures*, J. Eng. Mech. Div. EM3 (1973), p. 481-494.
- [10] J. C. Hsu and A. C. Meyer, *Modern control principles and applications*, McGraw-Hill, New York 1968.
- [11] P. Lancaster, *Theory of matrices*, Academic Press, New York 1969.
- [12] M. C. Pease III, *Methods of matrix algebra*, Academic Press, New York 1965.
- [13] W. R. Perkins and J. B. Cruz, Jr., *Engineering dynamics systems*, J. Wiley, New York 1969.
- [14] R. H. Plaut and K. Huseyin, *Derivatives of eigenvalues and eigenvectors in non-self-adjoint systems*, American Institute of Aeronautics and Astronautics Journal 11 (1973), p. 250-251.
- [15] A. R. Robinson and J. F. Harris, *Improving approximate eigenvalues and eigenvectors*, J. Eng. Mech. Div. EM2 (1971), p. 457-474.
- [16] A. Simpson, *Eigenvalue and vector sensitivities in Kron's method*, J. Sound and Vibration 31 (1973), p. 73-87.
- [17] W. J. Vetter, *Matrix calculus operations and Taylor expansions*, SIAM Review 15 (1973), p. 352-369.

INSTITUTE OF MECHANICS AND MACHINE DESIGN
TECHNICAL UNIVERSITY OF GDAŃSK
80-233 GDAŃSK

Received on 15. 1. 1976

W. GAWROŃSKI (Gdańsk)

FUNKCJE MACIERZOWE:

ROZWIĘCIE W SZEREG TAYLORA, ANALIZA WRAŻLIWOŚCI I BŁĘDÓW

STRESZCZENIE

W pracy przedstawiono rozwinięcie funkcji macierzowej argumentu macierzowego w szereg Taylora. Na tej podstawie przeprowadzono analizę wrażliwości funkcji macierzowych, a dla macierzowych funkcji losowego argumentu macierzowego wypro-
wadzono prawo propagacji błędów. Metoda ma zastosowanie w teorii systemów i w ana-
lizie błędów obliczeniowych w metodach komputerowych. Jako przykład podano ana-
lizę wrażliwości i błędów wartości własnych macierzy.
