SOME REMARKS ABOUT RECENT NOTATIONS IN MATRIX ANALYSIS

By

F. Csáki

Department of Automation, Technical University, Budapest

(Received January 15, 1970)

The symbolism of vector analysis concepts is more than one hundred years old. Such notions as scalar product, vectorial product, gradient, divergence, curl are common not only for the physicist but also for engineers. Recently, with the extending of modern control theory based on state space techniques, especially with the application of the optimum control theory of Pontryagin and Bellman, as well as with the stability theorems of Lyapunov, n-dimensional vector spaces are often used. Instead of the classical vector notations the matrix notations are introduced and preferred.

The present paper has the aim to show how the introduction of the differential operator, similar to the classical Hamiltonian nabla operator, leads to a systematical treatment of some problems encountered in control theory. By the way it is also shown how the classical notations of vector analysis can be replaced by this modern symbolism.

Fundamentals of vector analysis

In the vector analysis of the Euclidean space, there are introduced, —in addition to the so-called scalar-scalar functions

$$g = g(x) \tag{1}$$

(that is, scalar functions with scalar argument), — also scalar-vector functions

$$g = g(\mathbf{x}) \tag{2}$$

(that is, scalar functions with vector argument), and vector-vector functions

$$\mathbf{g} = \mathbf{g}(\mathbf{x}) \tag{3}$$

(that is, vector functions with vector argument). All these represent special cases of the most general, → fairly rare, matrix-matrix or tensor-tensor functions

$$\mathbf{G} = \mathbf{G}(\mathbf{X}) \tag{4}$$

(that is, matrix functions with matrix argument, or tensor functions with tensor argument). Scalar-scalar, scalar-vector, or vector-vector functions are often multivariable, such as

$$f = f(x, u) \qquad f = f(\mathbf{x}, \mathbf{u}) \qquad \mathbf{f} = \mathbf{f}(\mathbf{x}, \mathbf{u}) \tag{5}$$

or may involve also the independent time-variable t:

$$f = f(x, u, t) \qquad f = (\mathbf{x}, \mathbf{u}, t) \qquad \mathbf{f} = \mathbf{f}(\mathbf{x}, \mathbf{u}, t) \tag{6}$$

It is essential to establish certain rules of differentiation. A derivative with respect to a scalar is quite simple, for example

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}(t) = \left[\frac{\mathrm{d}x_1}{\mathrm{d}t}, \quad \frac{\mathrm{d}x_2}{\mathrm{d}t}, \ldots, \frac{\mathrm{d}x_n}{\mathrm{d}t}\right]^{\mathrm{T}} = [\dot{x}_1, \dot{x}_2, \ldots, \dot{x}_n]^{\mathrm{T}} = \dot{\mathbf{x}}(t) \tag{7}$$

or

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{A}(t) \triangleq \begin{bmatrix} \dot{a}_{11}(t), & \dot{a}_{12}(t), \ldots, \dot{a}_{1n}(t) \\ \dot{a}_{21}(t), & \dot{a}_{22}(t), \ldots, \dot{a}_{2n}(t) \\ \cdots & \cdots \\ \dot{a}_{m1}(t), & \dot{a}_{m2}(t), \ldots, \dot{a}_{mn}(t) \end{bmatrix} = \dot{\mathbf{A}}(t) \tag{8}$$

In physics, however, the concepts of gradient, divergence and in three-dimensional space, of curl are widely used. The gradient of a scalar-vector function is the column vector

$$\mathbf{grad}\, g(\mathbf{x}) \triangleq \left[\frac{\partial g(\mathbf{x})}{\partial x_1}, \quad \frac{\partial g(\mathbf{x})}{\partial x_2}, \ldots, \frac{\partial g(\mathbf{x})}{\partial x_n}\right]^{\mathrm{T}} \tag{9}$$

whereas the divergence of a vector-vector function is the scalar

$$\mathrm{div}\, \mathbf{g}(\mathbf{x}) \triangleq \frac{\partial g_1(\mathbf{x})}{\partial x_1} + \frac{\partial g_2(\mathbf{x})}{\partial x_2} + \ldots + \frac{\partial g_n(\mathbf{x})}{\partial x_n} \tag{10}$$

Finally, the curl of a vector-vector function is the column vector

$$\mathbf{curl}\, \mathbf{g}(\mathbf{x}) \triangleq \left[\frac{\partial g_3}{\partial x_2} - \frac{\partial g_2}{\partial x_3}, \quad \frac{\partial g_1}{\partial x_3} - \frac{\partial g_3}{\partial x_1}, \quad \frac{\partial g_2}{\partial x_1} - \frac{\partial g_1}{\partial x_2}\right]^{\mathrm{T}} \tag{11}$$

The Jacobian derivative matrix

$$\mathbf{J} = \mathbf{J}(\mathbf{g}, \mathbf{x}) \triangleq \begin{bmatrix} \frac{\partial g_1}{\partial x_1}, & \frac{\partial g_1}{\partial x_2}, \ldots, \frac{\partial g_1}{\partial x_n} \\ \frac{\partial g_2}{\partial x_1}, & \frac{\partial g_2}{\partial x_2}, \ldots, \frac{\partial g_2}{\partial x_n} \\ \cdots & \cdots \\ \frac{\partial g_m}{\partial x_1}, & \frac{\partial g_m}{\partial x_2}, \ldots, \frac{\partial g_m}{\partial x_n} \end{bmatrix} \tag{12}$$

is often encountered; and so is the so-called *nabla operator* proposed by HAMILTON

$$\nabla \triangleq \left[\frac{\partial}{\partial x_1}, \quad \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n}\right]^{\mathrm{T}} \tag{13}$$

Making use of the nabla operator, the above definitions may be expressed as follows:

$$\mathbf{grad}\, g(\mathbf{x}) = \nabla g(\mathbf{x}) \tag{14}$$

$$\operatorname{div} \mathbf{g}(\mathbf{x}) = \nabla \cdot [\mathbf{g}(\mathbf{x})] = \nabla^{\mathrm{T}}[\mathbf{g}(\mathbf{x})] = [\mathbf{g}(\mathbf{x})]^{\mathrm{T}}\nabla \tag{15}$$

$$\mathbf{curl}\, \mathbf{g}(\mathbf{x}) = \nabla_3 \times [\mathbf{g}(\mathbf{x})] = -[\mathbf{g}(\mathbf{x})] \times \nabla_3 \tag{16}$$

$$\mathbf{J}(\mathbf{g}, \mathbf{x}) = [\mathbf{g}(\mathbf{x})]\nabla^{\mathrm{T}} \tag{17}$$

where the dot $\cdot$ denotes a scalar or inner product, and the mark $\times$ denotes a cross or outer product.

The vector product is defined exclusively in three-dimensional space and hence, the nabla operator expressing a vector product has three partial-derivative operators. Vectors encountered in control engineering do, however, usually belong to the n-dimensional Euclidean space E^n and, therefore, we shall desist from any further discussion of the curl operator.

Lately, since division by a vector is not defined and, hence, no misunderstanding is possible, the nabla operator is frequently replaced by the equivalent differential operator

$$\frac{\mathrm{d}}{\mathrm{d}\mathbf{x}} = \left[\frac{\partial}{\partial x_1}, \quad \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n}\right]^{\mathrm{T}}. \tag{18}$$

Denoting the transpose of the differential operator (its row matrix) in the form

$$\left[\frac{\partial}{\partial x_1}, \quad \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n}\right] = \left[\frac{\mathrm{d}}{\mathrm{d}\mathbf{x}}\right]^{\mathrm{T}} = \frac{\mathrm{d}}{\mathrm{d}\mathbf{x}^{\mathrm{T}}} \tag{19}$$

we may rewrite (14), (15) and (17) to read

$$\mathbf{grad}\, g(\mathbf{x}) = \frac{\mathrm{d}}{\mathrm{d}\mathbf{x}}\, g(\mathbf{x}) = \frac{\mathrm{d}g(\mathbf{x})}{\mathrm{d}\mathbf{x}} \tag{20}$$

$$\operatorname{div} \mathbf{g}(\mathbf{x}) = \frac{\mathrm{d}}{\mathrm{d}\mathbf{x}^{\mathrm{T}}}\, \mathbf{g}(\mathbf{x}) = \mathbf{g}^{\mathrm{T}}(\mathbf{x})\, \frac{\mathrm{d}}{\mathrm{d}\mathbf{x}} \tag{21}$$

$$\mathbf{J}(\mathbf{g}, \mathbf{x}) = \mathbf{g}(\mathbf{x})\, \frac{\mathrm{d}}{\mathrm{d}\mathbf{x}^{\mathrm{T}}} \tag{22}$$

In order to avoid the somewhat awkward notation, the JACOBIan matrix is symbolically expressed as

$$\mathbf{J}(\mathbf{g}, \mathbf{x}) \triangleq \frac{\mathrm{d}\mathbf{g}(\mathbf{x})}{\mathrm{d}\mathbf{x}^{\mathrm{T}}} \tag{23}$$

and its transpose is written as

$$\mathbf{J}^{\mathrm{T}}(\mathbf{g}, \mathbf{x}) = \frac{\mathrm{d}\mathbf{g}^{\mathrm{T}}(\mathbf{x})}{\mathrm{d}\mathbf{x}} \tag{24}$$

The transpose of the gradient vector is

$$[\mathbf{grad}\, g(\mathbf{x})]^{\mathrm{T}} = \left[\frac{\mathrm{d}g(\mathbf{x})}{\mathrm{d}\mathbf{x}}\right]^{\mathrm{T}} = \frac{\mathrm{d}g(\mathbf{x})}{\mathrm{d}\mathbf{x}^{\mathrm{T}}} = \mathbf{J}(g, \mathbf{x}) \tag{25}$$

whereas the divergence may be denoted by either (21) or by the trace of the JACOBIan matrix

$$\operatorname{div} \mathbf{g}(\mathbf{x}) = \operatorname{tr} \mathbf{J}(\mathbf{g}, \mathbf{x}) = \operatorname{tr} \frac{\mathrm{d}\mathbf{g}(\mathbf{x})}{\mathrm{d}\mathbf{x}^{\mathrm{T}}} \tag{26}$$

When applied to multivariable vector functions with vector arguments such as $\mathbf{f}(\mathbf{x}, \mathbf{u})$, the nabla operator is distinguished by a subscript, or partial differential operators are introduced:

$$\nabla_{\mathrm{x}} = \frac{\partial}{\partial \mathbf{x}} = \left[\frac{\partial}{\partial x_1}, \quad \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n}\right]^{\mathrm{T}} = [\nabla_{\mathrm{x}}^{\mathrm{T}}]^{\mathrm{T}} = \left[\frac{\partial}{\partial \mathbf{x}^{\mathrm{T}}}\right]^{\mathrm{T}} \tag{27}$$

$$\nabla_{\mathrm{u}} = \frac{\partial}{\partial \mathbf{u}} = \left[\frac{\partial}{\partial u_1}, \quad \frac{\partial}{\partial u_2}, \ldots, \frac{\partial}{\partial u_r}\right]^{\mathrm{T}} = [\nabla_{\mathrm{u}}^{\mathrm{T}}]^{\mathrm{T}} = \left[\frac{\partial}{\partial \mathbf{u}^{\mathrm{T}}}\right]^{\mathrm{T}} \tag{28}$$

For example, the partial vector derivatives of the scalar function $f(\mathbf{x}, \mathbf{u}, t)$ or vector function $\mathbf{f}(\mathbf{x}, \mathbf{u}, t)$ are the column matrices

$$\frac{\partial f(\mathbf{x}, \mathbf{u}, t)}{\partial \mathbf{x}}, \quad \frac{\partial f(\mathbf{x}, \mathbf{u}, t)}{\partial \mathbf{u}} \text{ resp.}$$

or the JACOBIan matrices

$$\frac{\partial \mathbf{f}(\mathbf{x}, \mathbf{u}, t)}{\partial \mathbf{x}^{\mathrm{T}}}, \quad \frac{\partial \mathbf{f}(\mathbf{x}, \mathbf{u}, t)}{\partial \mathbf{u}^{\mathrm{T}}}$$

The total derivatives with respect to time

$$\frac{\mathrm{d}f(\mathbf{x}, \mathbf{u}, t)}{\mathrm{d}t} = \frac{\partial f}{\partial \mathbf{x}^{\mathrm{T}}} \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} + \frac{\partial f}{\partial \mathbf{u}^{\mathrm{T}}} \frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} + \frac{\partial f}{\partial t} \tag{29}$$

and

$$\frac{\mathrm{d}\mathbf{f}(\mathbf{x}, \mathbf{u}, t)}{\mathrm{d}t} = \frac{\partial \mathbf{f}}{\partial \mathbf{x}^{\mathrm{T}}} \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} + \frac{\partial \mathbf{f}}{\partial \mathbf{u}^{\mathrm{T}}} \frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} + \frac{\partial \mathbf{f}}{\partial t} \tag{30}$$

are formally similar (the first relationship can be derived directly from the second if the vector $\mathbf{f}$ is replaced by the scalar f). This is the great advantage of this notation.

Incidentally, for a function $g(\mathbf{x})$ not explicitly dependent on $\mathbf{u}$ and t, (29) and (30) give the often applied expressions

$$\frac{\mathrm{d}g(\mathbf{x})}{\mathrm{d}t} = \frac{\mathrm{d}g}{\mathrm{d}\mathbf{x}^{\mathrm{T}}} \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} \tag{31}$$

and

$$\frac{\mathrm{d}\mathbf{g}(\mathbf{x})}{\mathrm{d}t} = \frac{\mathrm{d}\mathbf{g}}{\mathrm{d}\mathbf{x}^{\mathrm{T}}} \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} \tag{32}$$

The transpose of the latter is

$$\frac{\mathrm{d}\mathbf{g}^{\mathrm{T}}(\mathbf{x})}{\mathrm{d}t} = \frac{\mathrm{d}t}{\mathrm{d}\mathbf{x}^{\mathrm{T}}} \frac{\mathrm{d}\mathbf{g}^{\mathrm{T}}}{\mathrm{d}\mathbf{x}} \tag{33}$$

Let us point out that the second partial derivatives may be written, for example, in the form

$$\frac{\partial^2 f(\mathbf{x}, \mathbf{u}, t)}{\partial \mathbf{x} \partial \mathbf{u}^{\mathrm{T}}}; \quad \frac{\partial^2 f(\mathbf{x}, \mathbf{u}, t)}{\partial \mathbf{u} \partial \mathbf{x}^{\mathrm{T}}}$$

The notation suggests that in the first case, for example, we have the partial derivative of the row vector $\partial f/\partial \mathbf{u}^{\mathrm{T}}$ with respect to the column vector $\mathbf{x}$, but the reverse sequence is also legitimate, that is, we may form the partial derivative of the column vector $\partial f/\partial \mathbf{x}$ with respect to the row vector $\mathbf{u}^{\mathrm{T}}$ as well. Thus the above notation does not record the sequence of partial derivations.

Finally, let us bear in mind that the second partial derivative of the vector function

$$\frac{\partial^2 \mathbf{f}(\mathbf{x}, \mathbf{u}, t)}{\partial \mathbf{x}\, \partial \mathbf{u}^{\mathrm{T}}}$$

etc. is not a rectangular matrix any more, but a parallel-epipedic one, just as the third partial derivative of a scalar function.

Some rules of differentiation

The derivative with respect to a scalar can be readily extended to products. It is important however to keep up the sequence of multiplications:

$$\frac{\mathrm{d}(\mathbf{u}^{\mathrm{T}}\mathbf{v})}{\mathrm{d}t} = \frac{\mathrm{d}\mathbf{u}^{\mathrm{T}}}{\mathrm{d}t}\mathbf{v} + \mathbf{u}^{\mathrm{T}}\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} \tag{34}$$

$$\frac{\mathrm{d}(\mathbf{A}\mathbf{z})}{\mathrm{d}t} = \frac{\mathrm{d}\mathbf{A}}{\mathrm{d}t}\mathbf{z} + \mathbf{A}\frac{\mathrm{d}\mathbf{z}}{\mathrm{d}t} \tag{35}$$

$$\frac{\mathrm{d}(\mathbf{A}\mathbf{B})}{\mathrm{d}t} = \frac{\mathrm{d}\mathbf{A}}{\mathrm{d}t}\mathbf{B} + \mathbf{A}\frac{\mathrm{d}\mathbf{B}}{\mathrm{d}t} \tag{36}$$

In order to define the rules of derivation with respect to a vector, let us first notice that

$$\frac{d\mathbf{x}}{d\mathbf{x}^T} = \mathbf{x}\,\frac{d}{d\mathbf{x}^T} = \mathbf{I} \tag{37}$$

On the other hand, the chain rule of the derivation of a compound function will be

$$\frac{d\mathbf{f}(\mathbf{g}(\mathbf{x}))}{d\mathbf{x}^T} = \frac{d\mathbf{f}}{d\mathbf{g}^T}\,\frac{d\mathbf{g}}{d\mathbf{x}^T} \tag{38}$$

since, similarly to (32),

$$\frac{\partial \mathbf{f}(\mathbf{g}(\mathbf{x}))}{\partial x_i} = \frac{d\mathbf{f}}{d\mathbf{g}^T}\,\frac{\partial \mathbf{g}}{dx_i}, \quad (i = 1, 2, \ldots, n)$$

yields the i-th column vector of the JACOBIan matrix. The result (38) can be arrived at also by forming, similarly to (31), the derivative

$$\frac{df_j}{d\mathbf{x}^T} = \frac{df_j}{d\mathbf{g}^T}\,\frac{d\mathbf{g}}{d\mathbf{x}^T}$$

which is the j-th row vector of the JACOBIan matrix $d\mathbf{f}/d\mathbf{x}^T$.

The chain rule applies, of course, also to the case where some of the vectors $\mathbf{f}$, $\mathbf{g}$, $\mathbf{x}$ degenerate to scalars. By the first rule (37), the derivative of the scalar product $\mathbf{x}^T\mathbf{c} = \mathbf{c}^T\mathbf{x}$ is

$$\frac{d\mathbf{c}^T\mathbf{x}}{d\mathbf{x}^T} = \mathbf{c}^T\,\frac{d\mathbf{x}}{d\mathbf{x}^T} = \mathbf{c}^T\mathbf{I} = \mathbf{c}^T \tag{39}$$

whereas that of the vector $\mathbf{A}\mathbf{x}$ is

$$\frac{d\mathbf{A}\mathbf{x}}{d\mathbf{x}^T} = \mathbf{A}\,\frac{d\mathbf{x}}{d\mathbf{x}^T} = \mathbf{A}\mathbf{I} = \mathbf{A} \tag{40}$$

In deriving both formulae we have taken into consideration that, on the one hand, the differential operator row matrix $d/d\mathbf{x}^T$ always multiplies from the right side and, on the other hand $\mathbf{c}^T$ and $\mathbf{A}$ do not depend on the vector $\mathbf{x}$.

By the chain rule, the derivative of the quadratic form $Q = Q(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}$ can be determined. Let $\mathbf{y} = \mathbf{A}\mathbf{x}$ and $f(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T\mathbf{y} = \mathbf{y}^T\mathbf{x}$ whence, by the chain rule

$$\frac{dQ}{d\mathbf{x}^T} = \frac{\partial f}{\partial \mathbf{x}^T} + \frac{\partial f}{\partial \mathbf{y}^T}\,\frac{d\mathbf{y}}{d\mathbf{x}^T}$$

Similarly to (39):

$$\frac{\partial f}{\partial \mathbf{x}^{\mathrm{T}}} = \frac{\partial \mathbf{y}^{\mathrm{T}}\mathbf{x}}{\partial \mathbf{x}^{\mathrm{T}}} = \mathbf{y}^{\mathrm{T}}$$

$$\frac{\partial f}{\partial \mathbf{y}^{\mathrm{T}}} = \frac{\partial \mathbf{x}^{\mathrm{T}}\mathbf{y}}{\partial \mathbf{y}^{\mathrm{T}}} = \mathbf{x}^{\mathrm{T}}$$

and, by (40)

$$\frac{\mathrm{d}\mathbf{y}}{\mathrm{d}\mathbf{x}} = \frac{\mathrm{d}\mathbf{A}\mathbf{x}}{\mathrm{d}\mathbf{x}^{\mathrm{T}}} = \mathbf{A}$$

Hence

$$\frac{\mathrm{d}Q}{\mathrm{d}\mathbf{x}^{\mathrm{T}}} = \frac{\mathrm{d}\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}}{\mathrm{d}\mathbf{x}^{\mathrm{T}}} = \mathbf{y}^{\mathrm{T}} + \mathbf{x}^{\mathrm{T}}\mathbf{A} = \mathbf{x}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}} + \mathbf{x}^{\mathrm{T}}\mathbf{A} = \mathbf{x}^{\mathrm{T}}\,[\mathbf{A}^{\mathrm{T}} + \mathbf{A}] \tag{41}$$

If $\mathbf{A}$ is symmetric then

$$\frac{\mathrm{d}Q}{\mathrm{d}\mathbf{x}^{\mathrm{T}}} = 2\mathbf{x}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}} = 2\mathbf{x}^{\mathrm{T}}\mathbf{A} \tag{42}$$

Furthermore, if $\mathbf{A} = \mathbf{I}$, then

$$\frac{\mathrm{d}\mathbf{x}^{\mathrm{T}}\mathbf{x}}{\mathrm{d}\mathbf{x}^{\mathrm{T}}} = 2\mathbf{x}^{\mathrm{T}} \tag{43}$$

By the chain rule, or by (41), the derivative with respect to $\mathbf{x}^{\mathrm{T}}$ of the quadratic form $Q(\mathbf{u}) = \mathbf{u}^{\mathrm{T}}(\mathbf{x})\mathbf{R}\mathbf{u}(\mathbf{x})$ is, clearly,

$$\frac{\mathrm{d}\mathbf{u}^{\mathrm{T}}\mathbf{R}\mathbf{u}}{\mathrm{d}\mathbf{x}^{\mathrm{T}}} = \frac{\mathrm{d}\mathbf{u}^{\mathrm{T}}\mathbf{R}\mathbf{u}}{\mathrm{d}\mathbf{u}^{\mathrm{T}}}\,\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}\mathbf{x}^{\mathrm{T}}} = \mathbf{u}^{\mathrm{T}}\,[\mathbf{R}^{\mathrm{T}} + \mathbf{R}]\,\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}\mathbf{x}^{\mathrm{T}}} \tag{44}$$

whereas the derivative of the more general bilinear form $\mathbf{u}^{\mathrm{T}}(\mathbf{x})\mathbf{R}\mathbf{v}(\mathbf{x})$ reads

$$\frac{\mathrm{d}\mathbf{u}^{\mathrm{T}}\mathbf{R}\mathbf{v}}{\mathrm{d}\mathbf{x}^{\mathrm{T}}} = \frac{\mathrm{d}\mathbf{u}^{\mathrm{T}}\mathbf{R}\mathbf{v}}{\mathrm{d}\mathbf{v}^{\mathrm{T}}}\,\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}\mathbf{x}^{\mathrm{T}}} + \frac{\mathrm{d}\mathbf{v}^{\mathrm{T}}\mathbf{R}^{\mathrm{T}}\mathbf{u}}{\mathrm{d}\mathbf{u}^{\mathrm{T}}}\,\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}\mathbf{x}^{\mathrm{T}}} = \mathbf{u}^{\mathrm{T}}\mathbf{R}\,\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}\mathbf{x}^{\mathrm{T}}} + \mathbf{v}^{\mathrm{T}}\mathbf{R}^{\mathrm{T}}\,\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}\mathbf{x}^{\mathrm{T}}} \tag{45}$$

On the other hand, putting $\mathbf{R} = \mathbf{I}$, we get

$$\frac{\mathrm{d}\mathbf{u}^{\mathrm{T}}\mathbf{v}}{\mathrm{d}\mathbf{x}^{\mathrm{T}}} + \mathbf{u}^{\mathrm{T}}\,\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}\mathbf{x}^{\mathrm{T}}} + \mathbf{v}^{\mathrm{T}}\,\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}\mathbf{x}^{\mathrm{T}}} \tag{46}$$

It is necessary to obtain derivatives also with respect to $\mathbf{x}$. In such cases, we multiply by the differential operator column matrix $\mathrm{d}/\mathrm{d}\mathbf{x}$ from the left side:

$$\frac{\mathrm{d}\mathbf{x}^{\mathrm{T}}}{\mathrm{d}\mathbf{x}} = \mathbf{I} \tag{47}$$

$$\frac{d\mathbf{f}^T(\mathbf{g}(\mathbf{x}))}{d\mathbf{x}} = \frac{d\mathbf{g}^T}{d\mathbf{x}} \frac{d\mathbf{f}^T}{d\mathbf{g}} \tag{48}$$

$$\frac{d\mathbf{x}^T\mathbf{c}}{d\mathbf{x}} = \mathbf{c} \tag{49}$$

$$\frac{d\mathbf{x}^T\mathbf{A}}{d\mathbf{x}} = \mathbf{A} \tag{50}$$

$$\frac{d\mathbf{x}^T\mathbf{A}\mathbf{x}}{d\mathbf{x}} = [\mathbf{A}^T + \mathbf{A}]\,\mathbf{x} \tag{51}$$

and for symmetric matrices **A**

$$\frac{d\mathbf{x}^T\mathbf{A}\mathbf{x}}{d\mathbf{x}} = 2\mathbf{A}\mathbf{x} = 2\mathbf{A}^T\mathbf{x} \tag{52}$$

Furthermore

$$\frac{d\mathbf{x}^T\mathbf{x}}{d\mathbf{x}} = 2\mathbf{x} \tag{53}$$

$$\frac{d\mathbf{u}^T\mathbf{R}\mathbf{u}}{d\mathbf{x}} = \frac{d\mathbf{u}^T}{d\mathbf{x}} [\mathbf{R}^T + \mathbf{R}]\,\mathbf{u} \tag{54}$$

$$\frac{d\mathbf{u}^T\mathbf{R}\mathbf{v}}{d\mathbf{x}} = \frac{d\mathbf{u}^T}{d\mathbf{x}} \mathbf{R}\mathbf{v} + \frac{d\mathbf{v}^T}{d\mathbf{x}} \mathbf{R}^T\mathbf{u} \tag{55}$$

$$\frac{d\mathbf{u}^T\mathbf{v}}{d\mathbf{x}} = \frac{d\mathbf{u}^T}{d\mathbf{x}} \mathbf{v} + \frac{d\mathbf{v}^T}{d\mathbf{x}} \mathbf{u} \tag{56}$$

Some conclusions

In this paper it was shown how the proposed logical notation of the JACOBIan matrix and its inverse lead to expressions of the vector analysis, which are formally very similar to the expressions of the scalar analysis, the only difference being in the consistent application of the superscript **T**, which denotes the transpose of an $m \times n$ matrix. By this minor trick, the common rules of matrix multiplication can clearly be generalized also for the cases where the differential operators $d/d\mathbf{x}$ or $d/d\mathbf{x}^T$ are encountered, the former being a column matrix, whereas the latter a row matrix. The only fact to be reminded of is the rule that $d/d\mathbf{x}$ multiplies always from the left side and correspondingly, $d/d\mathbf{x}^T$ multiplies from the right side.

Some applications of the proposed method are also shown, for example the rule of the total derivatives, the chain rule, the differentiation of various quadratic forms, and so on.

Some elements of the proposed method can be found here and there in the technical literature dealing with control engineering problems, sometimes, however, the notation of the transpose is left out, giving rise to some misunderstanding. To the Author's best knowledge, this is the first time where this logical and systematic treatment of the problem is published in full detail.

Summary

In this paper a logical notation for the JACOBIan matrix, that is, $\mathrm{d}\mathbf{f}(\mathbf{x})/\mathrm{d}\mathbf{x}^T$ or $\partial\mathbf{f}(\mathbf{x}, \mathbf{u})/\partial\mathbf{x}^T$ as well as for its inverse $\mathrm{d}\mathbf{f}^T(\mathbf{x})/\mathrm{d}\mathbf{x}$ or $\partial\mathbf{f}^T(\mathbf{x}, \mathbf{u})/\partial\mathbf{x}$ etc. are proposed. It is shown how the expressions of the vector analysis are similar to the expressions of the common scalar analysis.

References

1. LEFSCHETZ, S.: Differential Equations. Geometric Theory. Interscience Publishers Inc. New York, 1957.
2. Гантмахер, Ф. Р.: Теория матриц. Гостехиздат, Москва, 1953.
3. GANTMACHER, F. R.: The Theory of Matrices. Chelsea Publishing Co. New York 1959.
4. BELLMAN, R.: Introduction to Matrix Analysis. McGraw Hill. 1960.
5. Гантмахер, Ф. Р.: Теория матриц. Наука, Москва, 1966.
6. DORF, R. C.: Time-domain analysis and design of control systems. Reading. Mass. Addison-Wesley (1965).
7. TIMOTHY, L. K.—BONA, B. E.: State Space Analysis: An Introduction. McGraw Hill Book Co. New York, 1968.
8. SCHULTZ, D. G.—MELSA, J. L.: State Functions and Linear Control Systems. McGraw Hill Book Co. New York 1967.
9. DERUSSO, P. M.—ROY, R. J.—CLOSE, CH. M.: State Variables for Engineers. John Wiley et Sons, Inc. New York, 1966.
10. SAATY, TH. L.—BRAM, J.: Nonlinear Mathematics. McGraw Hill Book Co. New York, 1964.
11. OGATA, K.: State Space Analysis of Control Systems. Englewood Cliffs, 1967.
12. GIBSON, J.E.: Nonlinear Automatic Control, McGraw Hill Book Co. New York, 1963.
13. Четаев, Н. Г.: Устойчивость движения. Наука, Москва, 1965.
14. Красовский, Н. Н.: Теория управления движением. Наука, Москва, 1968.

Prof. Dr. Frigyes CSÁKI, Budapest XI., Garami E. tér 3, Hungary