A Curious Matrix Differential

Alexander Taskov

June 2021

Abstract

An interesting matrix differential, which behaves like a normal derivative for symmetric matrices, but with new behaviour for skew-symmetric matrices.

1 Consider

Consider the matrix

$$\partial_{ij} := (\partial_A)_{ij} := \frac{\partial}{\partial A_{ij}} \tag{1}$$

such that

$$(\partial_A F)_{ik} := \partial_{ij} F_{jk} \tag{2}$$

Obviously, the individual elements obey Leibnitz, and $\partial_{ij} A_{kl} = \delta_{ik}\delta_{jl}$. Then

$$\begin{aligned} \partial_{A^{(a)}} A^{(1)} \cdots A^{(a)} \cdots A^{(b)} &= \partial_{A^{(a)}_{i_0 i_1}} A^{(1)}_{i_1 i_2} \cdots A^{(a)}_{i_a i_{a+1}} \cdots A^{(b)}_{i_b i_{b+1}} \\ &= A^{(1)}_{i_1 i_2} \cdots A^{(a-1)}_{i_{a-1} i_0} A^{(a-1)}_{i_1 i_{a+1}} \cdots A^{(b)}_{i_b i_{b+1}} \\ &= (A^{(1)} \cdots A^{(a-1)})^T (A^{(a+1)} \cdots A^{(b)}) \end{aligned} \tag{3}$$

abusing indices. This is easier to see using link (Penrose) notation. This means it does not obey Leibnitz:

$$\partial_A (F_1 F_2) = \partial_A F_1 F_2 + F_1^T \partial_A F_2 \tag{4}$$

but it does mean

$$\partial_A \{A, B\} = B + B^T =: B_+ \tag{5}$$

$$\partial_A [A, B] = B - B^T =: B_- \tag{6}$$

but then

$$\partial_A A^n = A^{n-1} + A^T A^{n-2} + (A^T)^2 A^{n-3} + \cdots + (A^T)^{n-1} \tag{7}$$

which isn't trivial! If we constrain A such that $A = A^T$ then ∂_A does obeys Leibnitz, and $\partial_A A^n = nA^{n-1}$. In general, ∂_A acts as a normal derivative for symmetric A. But, if $A = -A^T$, then

$$\partial_A A^n = \frac{1}{2}(1 - (-1)^n) A^{n-1} \tag{8}$$

so this vanishes for n even and becomes A^{n-1} for n odd. If we let $F(A)$ map from $d \times d$ skew symmetric matrices to $d \times d$ general linear transformations, and be representable as

$$F(A) = \sum_{n=0} a_n A^n \tag{9}$$

then

$$A \partial_A F = \sum_{n \in odd} a_n A^n \tag{10}$$

$$\partial_A (AF) = \sum_{n \in even} a_n A^n \tag{11}$$

so for skew symmetric A, we get the operator expression

$$\{\partial_A, A\} = \mathbf{1} \tag{12}$$

where **1** is the identity functional on the set of functions like F. From its definition, we also have

$$\partial_{A^T} = (\partial_A)^T \tag{13}$$

and therefore

$$\partial_{A_+} = (\partial_A)_+ \tag{14}$$

$$\partial_{A_-} = (\partial_A)_- \tag{15}$$

which agrees if we use the chain rule treating (A, A^T) and (A_+, A_-) as two complete bases (NB: ∂_A does not enter into the change of variables, you need δ_A).