-(9)

DOI: 10.32474/CTBB.2018.01.000110

Mini Review

On some Derivatives of Vector-Matrix Products Useful for Statistics

Michele Nichelatti*

ISSN: 2644-1381

Service of Biostatistics, Fondazione Malattie del Sangue, Milan, Italy

*Corresponding author: Michele Nichelatti, Service of Biostatistics, Fondazione Malattie del Sangue, Milan, Italy

Received:

December 17, 2018

Published:

December 20, 2018

Mini Review

In this brief description, we will use the numerator layout [1], and will tacitly assume that all products are conformable.

The derivative of the linear form $u^t v$ with respect to the vector v is given as

$$\begin{split} \frac{\partial (u^t v)}{\partial v} &= \left(\frac{\partial (\sum_{k=1}^n u_k v_k)}{\partial v_1} \cdots \frac{\partial (\sum_{k=1}^n u_k v_k)}{\partial v_n} \right) \\ &= \left(\frac{\partial (u_1 v_1 + \cdots + u_n v_n)}{\partial v_1} \cdots \frac{\partial (u_1 v_1 + \cdots + u_n v_n)}{\partial v_n} \right) \\ &= (u_1 \cdots u_n) \\ &= v^t \end{split}$$

and since u^tv is a scalar, we are facing a particular case of the derivative of a scalar λ with respect to a vector, e.g., $\partial_{\nu}\lambda=\left(\partial_{\nu_1}\lambda\cdots\partial_{\nu_n}\lambda\right)$ and it must also be $\partial_{\nu}(u^tv)=\partial_{\nu}(v^tu)$. Moreover, it is easy to demonstrate that using the denominator layout, the derivative would have been $\partial_{\nu}(u^tv)=u$.

If both u and v vectors are function of a third vector z, we get

$$\begin{split} &\frac{\partial (u^t v)}{\partial z} = \left(\frac{\partial (\sum_{k=1}^n u_k v_k)}{\partial z_1} \cdots \frac{\partial (\sum_{k=1}^n u_k v_k)}{\partial z_n}\right) \\ &= \left(\frac{\partial (u_1 v_1 + \dots + u_n v_n)}{\partial z_1} \cdots \frac{\partial (u_1 v_1 + \dots + u_n v_n)}{\partial z_n}\right) \\ &= \left(\sum_{k=1}^n \left(v_k \frac{\partial u_k}{\partial z_1} + u_k \frac{\partial v_k}{\partial z_1}\right) \cdots \sum_{k=1}^n \left(v_k \frac{\partial u_k}{\partial z_n} + u_k \frac{\partial v_k}{\partial z_n}\right)\right) \\ &= (v_1 \cdots v_n) \begin{pmatrix} \frac{\partial u_1}{\partial z_1} & \cdots & \frac{\partial u_1}{\partial z_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial u_n}{\partial z_1} & \cdots & \frac{\partial u_n}{\partial z_n} \end{pmatrix} + (u_1 \cdots u_n) \begin{pmatrix} \frac{\partial v_1}{\partial z_1} & \cdots & \frac{\partial v_1}{\partial z_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial v_n}{\partial z_1} & \cdots & \frac{\partial v_n}{\partial z_n} \end{pmatrix} \end{split}$$

$$=v^{t}\frac{\partial u}{\partial z}+u^{t}\frac{\partial v}{\partial z}$$

which, in the case u = v = w reduces to

$$\frac{\partial (w^t w)}{\partial z} = w^t \frac{\partial w}{\partial z} + w^t \frac{\partial w}{\partial z} = 2w^t \frac{\partial w}{\partial z}.$$

Dealing with a linear transform u = Av, if A is $m \times m$ we have

$$\begin{split} \frac{\partial(Av)}{\partial v} &= \begin{pmatrix} \frac{\partial(a_{11}v_1 + \cdots a_{1n}v_n)}{\partial v_1} & \cdots & \frac{\partial(a_{11}v_1 + \cdots a_{1n}v_n)}{\partial v_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial(a_{m1}v_1 + \cdots a_{mn}v_n)}{\partial v_1} & \cdots & \frac{\partial(a_{m1}v_1 + \cdots a_{mn}v_n)}{\partial v_n} \end{pmatrix} \\ &= \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \\ &= A \end{split}$$

and if \boldsymbol{v} is a function of a vector \boldsymbol{w} , we get

$$\frac{\partial(Av)}{\partial w} = \begin{pmatrix} \frac{\partial(a_{11}v_1 + \cdots a_{1n}v_n)}{\partial w_1} & \cdots & \frac{\partial(a_{11}v_1 + \cdots a_{1n}v_n)}{\partial w_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial(a_{m1}v_1 + \cdots a_{mn}v_n)}{\partial w_1} & \cdots & \frac{\partial(a_{m1}v_1 + \cdots a_{mn}v_n)}{\partial w_n} \end{pmatrix}$$

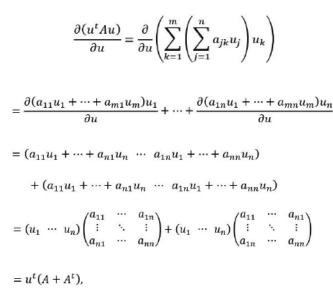
$$= \begin{pmatrix} a_{11} \frac{\partial v_1}{\partial w_1} + \dots + a_{1n} \frac{\partial v_n}{\partial w_1} & \dots & a_{11} \frac{\partial v_1}{\partial w_n} + \dots + a_{1n} \frac{\partial v_n}{\partial w_n} \\ \vdots & \ddots & \vdots \\ a_{m1} \frac{\partial v_1}{\partial w_1} + \dots + a_{mn} \frac{\partial v_n}{\partial w_1} & \dots & a_{m1} \frac{\partial v_1}{\partial w_n} + \dots + a_{mn} \frac{\partial v_n}{\partial w_n} \end{pmatrix}$$

$$= \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} \frac{\partial v_1}{\partial w_1} & \cdots & \frac{\partial v_1}{\partial w_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial v_n}{\partial w_1} & \cdots & \frac{\partial v_n}{\partial w_n} \end{pmatrix}$$
$$= A \frac{\partial v}{\partial w}.$$

From definition of bilinear form, we obtain, for $u^t A v$ the derivative

$$\begin{split} \frac{\partial (u^t A v)}{\partial v} &= \frac{\partial}{\partial v} \Biggl(\sum_{k=1}^m \Biggl(\sum_{j=1}^n a_{jk} u_j \Biggr) v_k \Biggr) \\ &= \frac{\partial (a_{11} u_1 + \dots + a_{m1} u_m) v_1}{\partial v} + \dots + \frac{\partial (a_{1n} u_1 + \dots + a_{mn} u_m) v_n}{\partial v} \\ &= (a_{11} u_1 + \dots + a_{m1} u_m \quad \dots \quad a_{1n} u_1 + \dots + a_{mn} u_m) \\ &= (u_1 \quad \dots \quad u_m) \left(\begin{matrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{matrix} \right) \\ &= u^t A_t \end{split}$$

while, for a quadratic form $u^t A u$ (where A is $n \times n$), we get

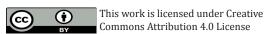


so that, if A is a symmetric matrix, say, for $A = X^t X$, then

$$\frac{\partial (u^t X^t X u)}{\partial u} = u^t (X^t X + X^t X) = 2u^t X^t X.$$

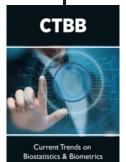
References

 Lütkepohl H (1997) Handbook of Matrices. John Wiley & Sons, New York, USA, pp. 320.



To Submit Your Article Click Here: Submit Article

DOI: 10.32474/CTBB.2018.01.000110



Current Trends on Biostatistics & Biometrics

Assets of Publishing with us

- Global archiving of articles
- Immediate, unrestricted online access
- Rigorous Peer Review Process
- Authors Retain Copyrights
- · Unique DOI for all articles