Current Trends on Biostatistics & Biometrics

DOI: 10.32474/CTBB.2018.01.000110

ISSN: 2644-1381

Mini Review

On some Derivatives of Vector-Matrix Products Useful for Statistics

Michele Nichelatti*

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

$$\frac{\partial(u^t v)}{\partial v} = \left(\frac{\partial\left(\sum_{k=1}^{n} u_k v_k\right)}{\partial v_1} \cdots \frac{\partial\left(\sum_{k=1}^{n} u_k v_k\right)}{\partial v_n} \right)$$

$$= \left(\frac{\partial(u_1 v_1 + \cdots + u_n v_n)}{\partial v_1} \cdots \frac{\partial(u_1 v_1 + \cdots + u_n v_n)}{\partial v_n} \right)$$

$$= (u_1 \cdots u_n)$$

$$= u^t$$

and since $u^t v$ is a scalar, we are facing a particular case of the derivative of a scalar λ with respect to a vector, e.g., $\partial_v \lambda = (\partial_{v_1}\lambda \cdots \partial_{v_n}\lambda)$ and it must also be $\partial_v(u^t v) = \partial_v(v^t u)$. Moreover, it is easy to demonstrate that using the denominator layout, the derivative would have been $\partial_v(u^t v) = u$.

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

$$\frac{\partial(u^t v)}{\partial z} = \left(\frac{\partial\left(\sum_{k=1}^{n} u_k v_k\right)}{\partial z_1} \cdots \frac{\partial\left(\sum_{k=1}^{n} u_k v_k\right)}{\partial z_n} \right)$$

$$= \left(\frac{\partial(u_1 v_1 + \cdots + u_n v_n)}{\partial z_1} \cdots \frac{\partial(u_1 v_1 + \cdots + u_n v_n)}{\partial z_n} \right)$$

$$= \left(\sum_{k=1}^{n} \left(v_k \frac{\partial u_k}{\partial z_1} + u_k \frac{\partial v_k}{\partial z_1} \right) \cdots \sum_{k=1}^{n} \left(v_k \frac{\partial u_k}{\partial z_n} + u_k \frac{\partial v_k}{\partial z_n} \right) \right)$$

$$= (v_1 \cdots v_n) \begin{pmatrix} \frac{\partial u_1}{\partial z_1} & \cdots & \frac{\partial u_1}{\partial z_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial u_n}{\partial z_1} & \cdots & \frac{\partial u_n}{\partial z_n} \end{pmatrix} + (u_1 \cdots u_n) \begin{pmatrix} \frac{\partial v_1}{\partial z_1} & \cdots & \frac{\partial v_1}{\partial z_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial v_n}{\partial z_1} & \cdots & \frac{\partial v_n}{\partial z_n} \end{pmatrix}$$

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

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

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

$$= \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

$$= A$$

and if v is a function of a vector w, we get

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

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

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

$$= A\frac{\partial v}{\partial w}.$$

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

$$\frac{\partial (u^t A v)}{\partial v} = \frac{\partial}{\partial v}\left(\sum_{k=1}^{m}\left(\sum_{j=1}^{n} a_{jk}u_j\right)v_k\right)$$

$$= \frac{\partial (a_{11}u_1 + \cdots + a_{m1}u_m)v_1}{\partial v} + \cdots + \frac{\partial (a_{1n}u_1 + \cdots + a_{mn}u_m)v_n}{\partial v}$$

$$= (a_{11}u_1 + \cdots + a_{m1}u_m \quad \cdots \quad a_{1n}u_1 + \cdots + a_{mn}u_m)$$

$$= (u_1 \quad \cdots \quad u_m)\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

$$= u^t A,$$

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

$$\frac{\partial (u^t A u)}{\partial u} = \frac{\partial}{\partial u}\left(\sum_{k=1}^{m}\left(\sum_{j=1}^{n} a_{jk}u_j\right)u_k\right)$$

$$= \frac{\partial (a_{11}u_1 + \cdots + a_{m1}u_m)u_1}{\partial u} + \cdots + \frac{\partial (a_{1n}u_1 + \cdots + a_{mn}u_m)u_n}{\partial u}$$

$$= (a_{11}u_1 + \cdots + a_{n1}u_n \quad \cdots \quad a_{1n}u_1 + \cdots + a_{nn}u_n)$$

$$+ (a_{11}u_1 + \cdots + a_{n1}u_n \quad \cdots \quad a_{1n}u_1 + \cdots + a_{nn}u_n)$$

$$= (u_1 \quad \cdots \quad u_n)\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} + (u_1 \quad \cdots \quad u_n)\begin{pmatrix} a_{11} & \cdots & a_{n1} \\ \vdots & \ddots & \vdots \\ a_{1n} & \cdots & a_{nn} \end{pmatrix}$$

$$= u^t(A + A^t),$$

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

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

DOI: 10.32474/CTBB.2018.01.000110