Differentiating a matrix expression with respect to a matrix.

To differentiate $\boldsymbol{Y = GAG^T}$ with respect to the matrix $\boldsymbol{G}$. Where $\boldsymbol{A}$ is a symmetric matrix and all matrices are m-square. Our matrices are all m-square in this paper which simplifies the notation.[1]

Let $\boldsymbol{E_{ij}}$ be an elementary matrix of order mxm defined as a matrix which has unity in the (i,j) element and is zero elsewhere.

Then the commutation matrix is defined as a square matrix[1-4]

$$\boldsymbol{T_{mm} = \sum_{i=1}^{m}\sum_{j=1}^{m} E_{ij} \otimes E_{ji}}$$

Where $\otimes$ denotes Kronecker product[5]

eg for m=2 the matrix is 4X4

$$\boldsymbol{T_{22}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The commutation matrix is a special form of permutation matrix. It has the property

$$\boldsymbol{Vec(G^{\mathrm{T}}) = T_{\mathrm{mm}}Vec(G)}$$

$$\boldsymbol{(I_m \otimes GA)T_{mm} = T_{mm}(GA \otimes I_m)}$$

and $\boldsymbol{T}_{\mathrm{mm}}^{-1} = \boldsymbol{T}_{\mathrm{mm}}^{\mathrm{T}} = \boldsymbol{T}_{\mathrm{mm}}$

Consider first a subset of this problem namely the matrix

$$\boldsymbol{X = I_m GA}$$

By definition of the Vec notation we have for matrices $\boldsymbol{A,B,C}$

$$Vec(ABC) = (C^T \otimes A)Vec(B)$$

We can write the using Vec notation

$$Vec(X) = (A^T \otimes I_m)Vec(G)$$

So that by differentiating a vector wrt another vector

$$dVec(X) / dVec(G) = (A^T \otimes I_m)$$

Therefore

$$\frac{dGA}{dG} = A^T \otimes I_m$$

It also follows that when **A=I** that

$$\frac{dG}{dG} = I_{m^2}$$

Using a similar method we have

$$Vec(I_m G^T I_m) = (I_m^T \otimes I_m)Vec(G^T)$$

$$= (I_m^T \otimes I_m)T_{mm}Vec(G)$$

Therefore

$$\frac{dG^T}{dG} = (I_m \otimes I_m)T_{mm} = T_{mm}$$

Now use the product rule

$$Df(x)g(x) = (g(x)^T \otimes I_m)f'(x) + (I_m \otimes f(x))g'(x)$$

Where D is the differential operator. And apply this to the original problem

$$\frac{dGAG^T}{dG} = (G \otimes I_m)\frac{dGA}{dG} + (I_m \otimes GA)\frac{dG^T}{dG}$$

and obtain

$$\frac{dGAG^T}{dG} = (G \otimes I_m)(A^T \otimes I_m) + (I_m \otimes GA)T_{mm}$$

But $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$

Therefore $(G \otimes I_m)(A^T \otimes I_m) = (GA^T \otimes I_m)$

$$\frac{dGAG^T}{dG} = (GA^T \otimes I_m) + (I_m \otimes GA)T_{mm}$$

It is well known that $(I_m \otimes GA)T_{mm} = T_{mm}(GA \otimes I_m)$ and since A is symmetric

we have the final result

$$\frac{dGAG^T}{dG} = (I_{m^2} + T_{mm})(GA \otimes I_m)$$

Another result which is required can be found using these same techniques.

$$Vec\left(AG^T\right) = \left(I_m \otimes A\right)Vec\left(G^T\right)$$
$$= \left(I \otimes A\right)T_{mm}Vec\left(G\right)$$

Therefore

$$\frac{dAG^T}{dG} = (I_m \otimes A)T_{mm}$$

Summarising the above four results, for m-square matrices G and A and where A is symmetric.

$$\frac{dGA}{dG} = A^T \otimes I_m$$

$$\frac{dG}{dG} = I_{m^2}$$

$$\frac{dG^T}{dG} = T_{mm}$$

$$\frac{dGAG^T}{dG} = (I_{m^2} + T_{mm})(GA \otimes I_m)$$

And a few more without proof[2]

$$\frac{dAG}{dG} = I_m \otimes G \qquad \frac{dAGB}{dG} = B^T \otimes A$$

References

[1] A. Graham, *Kronecker Products and Matrix Calculus with Applications*. UK: Ellis Horwood ltd, 1981.

[2] J. E. Gentle, *Matrix algebra. Theory,Computations and applications in Statistics*: Springer, 2007.

[3] J. W. Brewer, "Kronecker products and matrix calculus in system theory," *IEEE Trans. Circuits Systems,,* vol. cas-25, pp. 772-781, 1978.

[4] K. M. Abadir and J. R. Magnus, *Matrix Algebra*. Cambridge UK: Cambridge University Press, 2005.

[5] J. R. Magnus and H. Neudecker, *Matrix differential calculus with applications in statistics and economics*. New York: John Wiley and Sons 1999.