Matrix Calculus

Contents

1 Matrix transpose

- For matrix $\boldsymbol{A}$ and $\boldsymbol{B}$, and a scalar c,

$$(\boldsymbol{A}')' = \boldsymbol{A}$$
$$(\boldsymbol{A}+\boldsymbol{B})' = \boldsymbol{A}' + \boldsymbol{B}'$$
$$(c\boldsymbol{A})' = c\boldsymbol{A}'$$
$$(\boldsymbol{A}\boldsymbol{B})' = \boldsymbol{B}'\boldsymbol{A}'$$

- Given matrices $\boldsymbol{A}_1, \boldsymbol{A}_2, ..., \boldsymbol{A}_n$,

$$(\boldsymbol{A}_1\boldsymbol{A}_2\cdots\boldsymbol{A}_n)' = \boldsymbol{A}_n'\boldsymbol{A}_{n-1}'\cdots\boldsymbol{A}_1' \tag{1}$$

- $$det(\boldsymbol{A}') = det(\boldsymbol{A}) \tag{2}$$

- $$(\boldsymbol{A}')^{-1} = (\boldsymbol{A}^{-1})' \tag{3}$$

2 Derivatives of the trace of a matrix

2.1 Matrix and index notation

Given the product of matrix $\boldsymbol{A}$ and $\boldsymbol{B}$ is a $i \times k$ matrix, $[\boldsymbol{AB}]_{ik}$, we can write it in the index form as below,

$$[\boldsymbol{AB}]_{ik} = \sum_j A_{ij} B_{jk}. \tag{4}$$

And the matrix product $\boldsymbol{ABC'}$ can be written as

$$[\boldsymbol{ABC'}]_{il} = \sum_j A_{ij} [\boldsymbol{BC'}]_{jl} = \sum_j A_{ij} \sum_k B_{jk} C'_{kl} = \sum_j A_{ij} \sum_k B_{jk} C_{lk} = \sum_j \sum_k A_{ij} B_{jk} C_{lk} \tag{5}$$

2.2 First-order derivatives of the trace

The trace of a matrix, $tr(\cdot)$, is the summation of elements on the main diagonal of a square matrix, i.e., it must be a $n \times n$ matrix.

Consider this example:

$$f = tr[\boldsymbol{AXB}], \tag{6}$$

we can write this using index notation as below,

$$\begin{aligned} f &= \sum_i [\boldsymbol{AXB}]_{ii} \\ &= \sum_i \sum_j A_{ij} [\boldsymbol{XB}]_{ji} \\ &= \sum_i \sum_j A_{ij} \sum_k X_{jk} B_{ki} \\ &= \sum_i \sum_j \sum_k A_{ij} X_{jk} B_{ki}. \end{aligned}$$

Taking the derivative w.r.t. X_{jk}, we get

$$\frac{\partial tr[AXB]}{\partial X_{jk}} = \sum_i A_{ij} B_{ki} \tag{7}$$

$$= \sum_i B_{ki} A_{ij} \tag{8}$$

$$= [\boldsymbol{BA}]_{kj}, \text{ it is a } k \times j \text{ matrix} \tag{9}$$

The result has to be the same size as $\boldsymbol{X}$, i.e., a $j \times k$ matrix (We differentiate the trace w.r.t. X_{jk} in equation

(**??**)). Hence we have to transpose the result. Theresore,

$$\frac{\partial tr[\boldsymbol{A}\boldsymbol{X}\boldsymbol{B}]}{\partial \boldsymbol{X}} = \boldsymbol{A}'\boldsymbol{B}' \tag{10}$$