Matrix - Trace

Trace definition
Determinant is defined for square matrix A as sum of diagonal elements :

$$tr(A) = \sum_{n=1}^{N} a_{n,n}$$

Here are the properties.

$tr(A+B)$	=	$tr(A)+tr(B)$	(This is obvious, the proof is omitted.)
$tr(AB)$	=	$tr(BA)$	
$tr(cA)$	=	$tr(A)\times c$	(This is obvious, the proof is omitted.)
$tr(A^T)$	=	$tr(A)$	(This is obvious, the proof is omitted.)
$tr(A^{-1})$	=	no simplication	
$tr(A)$	=	$tr(QAQ^{-1})$	(Trace is invariant to similarity transformation.)
$tr(A)$	=	$\sum_{n=1}^{N} \Lambda_{n,n}$	(eigen decomposition)

Lets derive the trace of product, suppose A and B are N×M and M×N matrices, hence AB is N×N while B is M×M.

$$\begin{aligned} tr(AB) &= \sum_{n=1}^{N} \sum_{m=1}^{M} a_{n,m} b_{m,n} \\ &= \sum_{m=1}^{M} \sum_{n=1}^{N} b_{m,n} a_{n,m} \\ &= tr(BA) \end{aligned}$$

Hence trace of matrix product is invariant to cyclic permutation of matrix.

i.e.

$$\begin{aligned} tr(ABC) &= tr(A(BC)) = tr(BCA) \\ tr(ABC) &= tr((AB)C) = tr(CAB) \\ tr(ABC) &\neq tr(CBA) = tr(BAC) = tr(ACB) \end{aligned}$$

Lets derive the trace of similarity transformation, making use of the trace of product.

$$\begin{aligned} tr(QAQ^{-1}) &= tr(Q(AQ^{-1})) \\ &= tr(AQ^{-1}Q) \\ &= tr(A) \end{aligned}$$

$tr(A)$	=	$tr(Q\Lambda Q^{-1})$	(eigen decomposition)
	=	$tr(\Lambda)$	(trace is invariant to similarity transformation)
	=	$\sum_{n=1}^{N} \Lambda_{n,n}$	

Determinant vs Trace + Transpose vs Inverse

$(A^T)^T$	=	A	$(A^{-1})^{-1}$	=	A
$(A+B)^T$	=	$A^T + B^T$	$(A+B)^{-1}$	=	$A^{-1} - (I + A^{-1}B)^{-1} A^{-1} B A^{-1}$
$(AB)^T$	=	$B^T A^T$	$(AB)^{-1}$	=	$B^{-1}A^{-1}$
$(cA)^T$	=	$c(A^T)$	$(cA)^{-1}$	=	A^{-1}/c
$(A^T)^{-1}$	=	$(A^{-1})^T$			
			$tr(A+B)$	=	$tr(A)+tr(B)$
			$tr(AB)$	=	$tr(BA)$
			$tr(cA)$	=	$tr(A)\times c$
$\det(A^T)$	=	$\det(A)$	$tr(A^T)$	=	$tr(A)$
$\det(A^{-1})$	=	$1/\det(A)$	$tr(A^{-1})$	=	no simplication
$\det(A)$	=	$\det(QAQ^{-1})$	$tr(A)$	=	$tr(QAQ^{-1})$
$\det(A)$	=	$\prod_{n=1}^{N} \Lambda_{n,n}$	$tr(A)$	=	$\sum_{n=1}^{N} \Lambda_{n,n}$

As determinant denotes the product of eigenvalues (i.e. volume of hypercube enclosing all data), trace denotes the sum of eigenvalues (i.e.half of sum of edge length of hypercube enclosing all data).

Derivative of trace

Show that $\frac{\partial}{\partial A} tr(ABA^T) = 2AB$.

$$\left[\frac{\partial}{\partial A} tr(ABA^T)\right]_{y,x} = \frac{\partial}{\partial a_{y,x}} tr(ABA^T)$$

$$= \frac{\partial}{\partial a_{y,x}} \sum_{n=1}^N (ABA^T)_{n,n}$$

$$= \frac{\partial}{\partial a_{y,x}} \sum_{n=1}^N \sum_{m=1}^N a_{n,m} (BA^T)_{m,n}$$

$$= \frac{\partial}{\partial a_{y,x}} \sum_{n=1}^N \sum_{m=1}^N a_{n,m} \sum_{k=1}^N (b_{m,k} a_{n,k})$$

$$= \frac{\partial}{\partial a_{y,x}} \sum_{m=1}^N a_{y,m} \sum_{k=1}^N (b_{m,k} a_{y,k})$$ remove all terms with $n \neq y$

$$= \frac{\partial}{\partial a_{y,x}} \left[a_{y,x} \sum_{k=1}^N (b_{x,k} a_{y,k}) + \sum_{\substack{m=1 \\ m \neq x}}^N a_{y,m} \sum_{k=1}^N (b_{m,k} a_{y,k}) \right]$$ separate into 2 parts : m = x and $m \neq x$

$$= \frac{\partial}{\partial a_{y,x}} \left[a_{y,x} \sum_{\substack{k=1 \\ k \neq x}}^N (b_{x,k} a_{y,k}) + b_{x,x} a_{y,x}^2 + \sum_{\substack{m=1 \\ m \neq x}}^N a_{y,m} b_{m,x} a_{y,x} \right]$$ remove all terms with $m \neq x$ in 2nd term

$$= \sum_{\substack{k=1 \\ k \neq x}}^N (a_{y,k} b_{x,k}) + 2b_{x,x} a_{y,x} + \sum_{\substack{m=1 \\ m \neq x}}^N a_{y,m} b_{m,x}$$

$$= \sum_{k=1}^N (a_{y,k} b_{x,k}) + \sum_{m=1}^N a_{y,m} b_{m,x}$$

$$\frac{\partial}{\partial A} tr(ABA^T) = AB + AB^T$$

$$= 2AB$$ if B is symmetric

Frobenius norm

There are different definitions of matrix norm, among which, the Frobenius norm is related to trace. Suppose A is a N×M matrix, its Frobenius norm is defined as :

$$\|A\|_F = \sqrt{\sum_{n=1}^N \sum_{m=1}^M a_{n,m}^2}$$

$$= \sqrt{tr(A^T A)}$$

$$= \sqrt{\sum_{n=1}^{\min(N,M)} s_n^2}$$ where s_n are singular values of A

(proof) $$tr(A^T A) = \sum_{m=1}^M (A^T A)_{m,m}$$

$$= \sum_{m=1}^M (\sum_{n=1}^N a_{n,m} a_{n,m})$$

$$tr(A^T A) = tr((USV^T)^T (USV^T))$$ where $A = USV^T$

$$= tr(VSU^T USV^T)$$

$$= tr(VS^2 V^T)$$

$$= tr(S^2 V^T V)$$ since $tr(ABC) = tr(BCA) = tr(CAB)$

$$= tr(S^2)$$

$$= \sum_{n=1}^{\min(N,M)} s_n^2$$

Hence : $$\sqrt{tr(A^T A)} = \sqrt{\sum_{n=1}^N \sum_{m=1}^M a_{n,m}^2}$$

and $$\sqrt{tr(A^T A)} = \sqrt{\sum_{n=1}^{\min(N,M)} s_n^2}$$