Prepared for submission to JHEP

Notes of matrix derivatives

Yu Chen[a]

[a] *Tencent Company*

E-mail: alexychen@tencent.com

Abstract: This note gives some hints for canonical results of matrix derivatives.

Contents

1 Notations

Note In this note, I may use Einstein summation convention, where we will sum all repeated index.

2 Basic items

2.1 Definition

$$\mathbf{K}^{(m,n)} := \sum_i^m \sum_j^n \mathbf{E}_{ij}^{(m,n)} \otimes \mathbf{E}_{ji}^{(n,m)} = \mathbf{E}_{ij}^{(m,n)} \otimes \mathbf{E}_{ji}^{(n,m)} \tag{2.1}$$

$$\bar{\mathbf{K}}^{(m,n)} := \sum_i^m \sum_j^n \mathbf{E}_{ij}^{(m,n)} \otimes \mathbf{E}_{ji}^{(n,m)} = \mathbf{E}_{ij}^{(m,n)} \otimes \mathbf{E}_{ij}^{(m,nl)} \tag{2.2}$$

$$\mathrm{vec}(\mathbf{A}) = \begin{pmatrix} \mathbf{A}_{\cdot 1} \\ \mathbf{A}_{\cdot 2} \\ \vdots \end{pmatrix} \tag{2.3}$$

$$\frac{\partial \mathbf{A}}{\partial \mathbf{B}} := \sum_{r,s} \mathbf{E}_{rs}^{(k,l)} \otimes \frac{\partial \mathbf{A}}{b_{rs}} := E_{rs}^{(k,l)} \otimes \frac{\partial \mathbf{A}}{b_{rs}}, \forall \mathbf{A} \in \mathbb{F}^{m\times n}, B \in \mathbb{F}^{k\times l}. \tag{2.4}$$

2.2 Proofs

T1.5

$$\mathbf{K}^{(m,n)T} = \mathbf{K}^{(n,m)} \tag{2.5}$$

Proof

$$\mathbf{K}^{(m,n)T} = \mathbf{E}_{ji}^{(n,m)} \otimes \mathbf{E}_{ij}^{(m,n)} \tag{2.6}$$

$$= \mathbf{E}_{ij}^{(n,m)} \otimes \mathbf{E}_{ji}^{(m,n)} = \mathbf{K}^{(n,m)} \tag{2.7}$$

Table 1: Notation

Symbols:	
$\mathbf{A}$	$\mathbf{A}$ is a matrix
$\mathbf{A}_{ij}$	i, j element of matrix A
$[\mathbf{A}]_{ij}$	i, j element of matrix A
$:=$	definition
$\|x\rangle$	column vector
$\langle x\|$	(conjugate) transpose of column vector
$\mathbf{X}^{-}$	pseudo inverse of matrix $\mathbf{X}$
$\mathbf{X}^{\dagger}$	conjugate transpose of matrix $\mathbf{X}$
Operators:	
$\otimes$	Kronecker product
$\mathbf{A}^{T}$	transpose of matrix $\mathbf{A}$
$\dagger$	conjugate transpose of matrix $\mathbf{A}$
$\oplus$	direct sum
$\oplus_K$	Kronecker direct sum
$\mathrm{vec}(\mathbf{A})$	vectorized representation of matrix $\mathbf{A}$
det	determinant
Tr	trace of a matrix
Tr_X	partial trace in the corresponding Hilbert space
∂	an abstract for partial derivative
$\mathbf{K}^{(m,n)}$	commutation matrix
$\bar{\mathbf{K}}^{(n,m)}$	partial transpose of commutation matrix
$\mathbf{I}$	identity matrix
$\mathbf{e}_j^{(n)}$	n dimensional zero vector with one at index j
$\mathbf{E}_{i,j}^{(m,n)}$	(m, n) dimensional zero matrix with one at index (i, j)

T1.6

$$\mathbf{K}^{(m,n)-1} = \mathbf{K}^{(n,m)} \tag{2.8}$$

Proof

$$\mathbf{K}^{(m,n)}\mathbf{K}^{(n,m)} = \mathbf{E}_{ij}^{(m,n)} \otimes \mathbf{E}_{ji}^{(n,m)}\mathbf{E}_{i'j'}^{(n,m)} \otimes \mathbf{E}_{j'i'}^{(m,n)} \tag{2.9}$$

$$= \delta_{ji'}\mathbf{E}_{ij'}^{(m,m)} \otimes \mathbf{E}_{j,i'}^{(n,n)}\delta_{ij'} \tag{2.10}$$

$$= \mathbf{E}_{ii}^{(m,m)} \otimes \mathbf{E}_{jj}^{(n,n)} = \mathbf{I}. \tag{2.11}$$

T2.13

$$\mathrm{vec}(\mathbf{ADB}) = \mathbf{B}^{T} \otimes \mathbf{A}\mathrm{vec}(\mathbf{D}), \mathbf{A} \in \mathbb{F}^{m\times n}, \mathbf{D} \in \mathbb{F}^{n\times p}, \mathbf{B} \in \mathbb{F}^{p\times q}. \tag{2.12}$$

Proof We show the equality point-wisely.

$$[\mathrm{vec}(\mathbf{ADB})]_{(j-1)m+i} = \mathbf{A}_{is}\mathbf{D}_{st}\mathbf{B}_{tj}. \tag{2.13}$$

$$[\mathbf{B}^T \otimes \mathbf{A}\mathrm{vec}(\mathbf{D})]_{(j-1)m+i} = [\mathbf{B}^T \otimes \mathbf{A}]_{(j-1)m+i,s}[\mathrm{vec}(\mathbf{D})]_s \tag{2.14}$$
$$= [\mathbf{B}^T]_{j,t}[\mathbf{A}]_{i,s}[\mathrm{vec}(\mathbf{D})]_{(s,t)} \tag{2.15}$$
$$= \mathbf{A}_{is}\mathbf{D}_{s,t}\mathbf{B}_{tj} \tag{2.16}$$
$$= [\mathrm{vec}(\mathbf{ADB})]_{(j-1)m+i} \tag{2.17}$$

T2.5

$$\mathbf{B} \otimes \mathbf{A} = \mathbf{K}^{(k,m)}\mathbf{A} \otimes \mathbf{B}\mathbf{K}^{(n,l)}, \forall \mathbf{A} \in \mathbb{F}^{m\times n}, \mathbf{B} \in \mathbb{F}^{k\times l} \tag{2.18}$$

Proof Firstly, we prove $\forall \mathbf{X} \in \mathbb{F}^{m\times n}$, we have

$$\mathbf{K}^{(m,n)}\mathrm{vec}(\mathbf{X}) = \mathrm{vec}(\mathbf{X}^{\mathbf{T}}) \tag{2.19}$$

As we known, X_{ij} appears at position $m(j-1)+i$ of $\mathrm{vec}(\mathbf{X})$ and at position $n(i-1)+j$, so we just need to prove that $\mathbf{K}^{(m,n)}$ permute element at $m(j-1)+i$ to $n(i-1)+j$ for all i, j. To this ends, we just need to show,

$$[\mathbf{K}^{(m,n)}]_{n(i-1)+j,\cdot} \cdot \mathrm{vec}(\mathbf{X}) = X_{ij}. \tag{2.20}$$

Imagine Kronecker product formula in your brain, we can find any contribution to $n(i-1)+j$ row of $\mathbf{E}^{(m,n)} \otimes \mathbb{E}^{(n,m)}$ comes from i row of $\mathbf{E}^{(m,n)}$ and j row of $\mathbf{E}^{(n,m)}$, which means only $\mathbf{E}_{ij}^{(m,n)}$ and $\mathbf{E}_{ji}^{(n,m)}$ should be kept. Obviously, $\mathbf{E}_{ij}^{(m,n)} \otimes \mathbf{E}_{ji}^{(n,m)}$ is non-zero only at index $((i-1)n+j, (j-1)m+i)$. Up to now, we can claim the correctness of Eq.(2.20) and Eq.(2.19).

With this lemma, we can prove Eq.(2.18) more easily. Choosing an arbitrary matrix $\mathbf{X} \in \mathbb{F}^{n\times l}$ $X \in \mathbb{F}^{n\times l}$, to prove Eq.(2.18),

$$\mathbf{K}^{(m,k)}\mathbf{B} \otimes \mathbf{A}\mathrm{vec}(\mathbf{X}) = \mathbf{A} \otimes \mathbf{B}\mathbf{K}^{(n,l)}\vec{X} \tag{2.21}$$
$$\Leftrightarrow \mathbf{K}^{(m,k)}\mathbf{B} \otimes \mathbf{A}\mathrm{vec}(\mathbf{X}) = \mathbf{A} \otimes \mathbf{B}\mathrm{vec}(\mathbf{X}^{\mathbf{T}}) \tag{2.22}$$
$$\Leftrightarrow \mathbf{K}^{(m,k)}\mathrm{vec}(\mathbf{AXB}^{\mathbf{T}}) = \mathrm{vec}(\mathbf{BX}^{\mathbf{T}}\mathbf{A}^{\mathbf{T}}) \tag{2.23}$$
$$\Leftrightarrow \mathrm{vec}((\mathbf{AXB}^{\mathbf{T}})^{\mathbf{T}}) = \mathrm{vec}(\mathbf{BX}^{\mathbf{T}}\mathbf{A}^{\mathbf{T}}) \tag{2.24}$$

Since X is arbitrary, we justify the Eq.(2.18).

Notes: Although proof of such equality by careful computation is strict, it can not bring us intuition. Hence, I use tensor graph language to prove these theorems, which is shown in App.(3)

T4.2

$$(\frac{\partial \mathbf{A}}{\partial \mathbf{B}})^T = \frac{\partial \mathbf{A}^T}{\partial \mathbf{B}^T} \tag{2.25}$$

Proof

$$\frac{\partial \mathbf{A}^T}{\partial \mathbf{B}^T} = \mathbf{E}_{rs} \otimes \frac{\partial \mathbf{A}^T}{\partial b_{sr}} \tag{2.26}$$
$$= \left(\mathbb{E}_{sr} \otimes \frac{\partial A}{\partial b_{sr}}\right)^T \tag{2.27}$$
$$= (\frac{\partial \mathbf{A}}{\partial \mathbf{B}})^T. \tag{2.28}$$

T4.3

$$\frac{\partial \mathbf{AC}}{\partial \mathbf{B}} = \frac{\partial \mathbf{A}}{\partial \mathbf{B}} I_l \otimes \mathbf{C} + I_k \otimes \mathbf{A}\frac{\partial \mathbf{C}}{\partial \mathbf{B}}, \forall \mathbf{C} \in \mathbb{F}^{n\times p}. \tag{2.29}$$

Proof

$$\begin{aligned}
\frac{\partial \mathbf{AC}}{\partial \mathbf{B}} &= \mathbf{E}_{rs}^{(k,l)} \otimes \frac{\mathbf{AC}}{b_{rs}} && (2.30)\\
&= \mathbf{E}_{rs}^{(k,l)} \otimes (\frac{\partial \mathbf{A}}{\partial b_{rs}}\mathbf{C} + \mathbf{A}\frac{\partial \mathbf{C}}{\partial b_{rs}}) && (2.31)\\
&= \mathbf{E}_{rs}^{(k,l)} I_l \otimes \frac{\partial \mathbf{A}}{\partial b_{rs}}\mathbf{C} + I_k \mathbf{E}_{rs}^{(k,l)} \otimes \mathbf{A}\frac{\partial \mathbf{C}}{\partial b_{rs}}) && (2.32)\\
&= \mathbf{E}_{rs}^{(k,l)} \otimes \frac{\partial \mathbf{A}}{\partial b_{rs}} \cdot I_l \otimes \mathbf{C} + I_k \otimes \mathbf{A} \cdot \mathbf{E}_{rs}^{(k,l)} \otimes \frac{\partial \mathbf{C}}{\partial b_{rs}} && (2.33)\\
&= \frac{\partial \mathbf{A}}{\partial \mathbf{B}} I_l \otimes \mathbf{C} + I_k \otimes \mathbf{A}\frac{\partial \mathbf{C}}{\partial \mathbf{B}} && (2.34)
\end{aligned}$$

T4.4

$$\frac{\partial \mathbf{A}\otimes \mathbf{D}}{\partial \mathbf{B}} = \frac{\partial \mathbf{A}}{\partial \mathbf{B}} \otimes D + (I_k \otimes \mathbf{K}^{(m,p)})\frac{\partial \mathbf{D}}{\partial B} \otimes \mathbf{A}(I_l \otimes \mathbf{K}^{(q,n)}), \forall \mathbf{D} \in \mathbb{F}^{p\times q}. \tag{2.35}$$

$$\begin{aligned}
\frac{\partial \mathbf{A}\otimes \mathbf{D}}{\partial \mathbf{B}} &= \mathbf{E}_{rs} \otimes \frac{\partial \mathbf{A}\otimes \mathbf{D}}{\partial b_{rs}} && (2.36)\\
&= \mathbf{E}_{rs} \otimes \left(\frac{\partial \mathbf{A}}{\partial b_{rs}} \otimes \mathbf{D} + \mathbf{A} \otimes \frac{\partial \mathbf{D}}{\partial b_{rs}}\right) && (2.37)\\
&= \frac{\partial \mathbf{A}}{\partial \mathbf{B}} \otimes \mathbf{D} + \mathbf{E}_{rs} \otimes (A \otimes \frac{\partial \mathbf{D}}{\partial b_{rs}}) && (2.38)
\end{aligned}$$

By using Eq.(2.18), we can continue,

$$\begin{aligned}
\mathbf{E}_{rs} \otimes (A \otimes \frac{\partial \mathbf{D}}{\partial b_{rs}}) &= \mathbf{E}_{rs} \otimes \left(\mathbf{K}^{(m,p)}(\frac{\partial \mathbf{D}}{\partial b_{rs}} \otimes A)\mathbf{K}^{(q,n)}\right) && (2.39)\\
&= I_k \otimes \mathbf{K}^{(m,p)} \cdot \left(\mathbf{E}_{rs} \otimes \frac{\partial \mathbf{D}}{\partial b_{rs}} \otimes A \otimes \mathbf{K}^{(q,n)}\right) && (2.40)\\
&= I_k \otimes \mathbf{K}^{(m,p)} \cdot \left(\mathbf{E}_{rs} \otimes \frac{\partial \mathbf{D}}{\partial b_{rs}} \otimes \mathbf{A}\right) \cdot I_l \otimes K^{(q,n)} && (2.41)\\
&= I_k \otimes \mathbf{K}^{(m,p)} \cdot \left(\frac{\partial \mathbf{D}}{\partial \mathbf{B}} \otimes \mathbf{A}\right) \cdot I_l \otimes K^{(q,n)} && (2.42)
\end{aligned}$$

T4.6

$$\frac{\partial A(C(B))}{\partial B} = \left(I_k \otimes \frac{\mathbf{A}}{\partial \mathrm{vec}(\mathbf{C})}\right)\left(\frac{(\mathrm{vec}(\mathbf{C}^{\mathbf{T}}))^T}{\partial \mathbf{B}} \otimes I_n\right) = \left(\frac{\partial \mathbf{C}^T}{\partial \mathbf{B}} \otimes I_m\right)\left(I_l \otimes \frac{\partial \mathbf{A}}{\partial \mathbf{C}}\right). \tag{2.43}$$

Proof

$$\frac{\partial \mathbf{A}(\mathbf{C}(\mathbf{B}))}{\partial \mathbf{B}} = \mathbf{E}_{rs}^{(k,l)} \otimes \frac{\partial \mathbf{A}}{\partial b_{rs}} \tag{2.44}$$

$$= \mathbf{E}_{rs}^{(k,l)} \otimes \mathbf{E}_{r's'}^{(m,n)} \frac{\partial a_{r's'}}{\partial b_{rs}} \tag{2.45}$$

$$= \mathbf{E}_{rs}^{(k,l)} \otimes \mathbf{E}_{r's'}^{(m,n)} \frac{\partial a_{r's'}}{\partial c_{uv}} \frac{\partial c_{uv}}{\partial_{rs}} \tag{2.46}$$

$$= \frac{\partial c_{uv}}{\partial \mathbf{B}} \otimes \mathbf{E}_{r's'}^{(m,n)} \frac{\partial a_{r's'}}{\partial c_{uv}} \tag{2.47}$$

$$= \frac{\partial c_{uv}}{\partial \mathbf{B}} \otimes \frac{\partial \mathbf{A}}{\partial c_{uv}}. \tag{2.48}$$

T5.1, T5.2

$$\frac{\partial \mathbf{A}}{\partial \mathbf{A}} = \bar{\mathbf{K}}^{(m,n)}, \frac{\partial \mathbf{A}^T}{\partial \mathbf{A}} = \mathbf{K}^{(m,n)}. \tag{2.49}$$

Proof

$$\frac{\partial \mathbf{A}}{\partial \mathbf{A}} = \mathbf{E}_{rs}^{(m,n)} \otimes \mathbf{E}_{r's'}^{(m,n)} \frac{\partial a_{r's'}}{\partial a_{rs}} \tag{2.50}$$

$$= \mathbf{E}_{rs}^{(m,n)} \otimes \mathbf{E}_{r's'}^{(m,n)} \delta_{rr'}\delta_{ss'} = \bar{\mathbf{K}}^{(m,n)}. \tag{2.51}$$

T5.6, T5.7

$$\frac{\partial \mathbf{y}}{\partial \mathbf{y}} = \mathrm{vec}(\mathbf{I_n}) \forall y \in \mathbb{F}^{(n\times 1)}, \frac{\partial \mathbf{y}^T}{\partial \mathbf{y}} = I_n. \tag{2.52}$$

Proof, trivially.

T5.8,T5.9

$$\frac{\partial \mathbf{Ay}}{\partial \mathbf{y}} = \mathrm{vec}(\mathbf{A}), \frac{\partial \mathbf{Ay}}{\partial \mathbf{y}^T} = \mathbf{A}. \tag{2.53}$$

T5.10

$$\frac{\partial \mathbf{y} \otimes \mathbf{y}}{\partial \mathbf{y}^T} = I_n \otimes \mathbf{y} + \mathbf{y} \otimes I_n \tag{2.54}$$

Proof

$$\frac{\partial \mathbf{y} \otimes \mathbf{y}}{\partial \mathbf{y}^T} = \mathbf{e}_j^T \otimes \frac{\partial \mathbf{y} \otimes \mathbf{y}}{\partial y_j} \tag{2.55}$$

$$= \mathbf{e}_j^T \otimes \delta_{jj'} \mathbf{e}_{j'} \otimes \mathbf{y} + \mathbf{e}_j^T \otimes \mathbf{y} \delta_{jj'} \otimes \mathbf{e}_{j'} \tag{2.56}$$

$$= I_n \otimes \mathbf{y} + \mathbf{y} \otimes I_n. \tag{2.57}$$

T5.11

$$\frac{\partial \mathbf{y}^T \mathbf{Y} \mathbf{y}}{\partial \mathbf{y}} = (\mathbf{Y} + \mathbf{Y}^T)\mathbf{y}. \tag{2.58}$$

3 Proofs of equation in CookBook

Eq(41)

$$\partial \det \mathbf{X} = \mathrm{Tr}[\mathrm{adj}\mathbf{X})\partial \mathbf{X}] \tag{3.1}$$

Proof Define $\phi(\mathbf{X},\mathbf{Y}) = \frac{\mathrm{d}\det(\mathbf{X}+\alpha\mathbf{Y})}{\mathrm{d}\alpha}|_{\alpha=0}$, we can easily check that,

$$\phi(I,\mathbf{Y}) = \mathrm{Tr}(\mathbf{Y}). \tag{3.2}$$

Since $\det \mathbf{Y} = \det \mathbf{X} \det \mathbf{X}^{-1}\mathbf{Y}$, we obtain that,

$$\phi(\mathbf{Y},\mathbf{Z}) = \det \mathbf{X}\phi(\mathbf{X}^{-1}\mathbf{Y},\mathbf{X}^{-1}\mathbf{Z}), \tag{3.3}$$

Let $\mathbf{Y} = \mathbf{X}$, at once, $\phi(\mathbf{X},\mathbf{Z}) = \det \mathbf{X}\phi(I,\mathbf{X}^{-1}\mathbf{Z}) = \det \mathbf{X}\mathrm{Tr}[\mathbf{X}^{-1}\mathbf{Z}]$. Then if $\mathbf{Z} = \partial\mathbf{X}$,

$$\partial \det \mathbf{X} = \phi(\mathbf{X},\partial\mathbf{X}) = \det \mathbf{X}\mathrm{Tr}[\mathbf{X}^{-1}\partial\mathbf{X}] = \mathrm{Tr}[\mathrm{adj}X\partial\mathbf{X}] \tag{3.4}$$

Eq49

$$\frac{\det \mathbf{X}}{\partial \mathbf{X}} = \det \mathbf{X}(\mathbf{X}^{-1})^T. \tag{3.5}$$

Using Eq.(2.49), we can easily prove it.

$$\begin{aligned}
\frac{\partial \det \mathbf{X}}{\partial \mathbf{X}} &= \det \mathbf{X}\mathrm{Tr}_{\mathbf{X}}[I \otimes \mathbf{X}^{-1}\bar{\mathbf{K}}] && (3.6)\\
&= \det \mathbf{X}\mathrm{Tr}_X[I \otimes \mathbf{X}^{-1}\mathbf{E}_{ij} \otimes \mathbf{E}_{ij}] && (3.7)\\
&= \det \mathbf{X}\mathrm{Tr}_X[\mathbf{E}_{ij} \otimes \mathbf{X}^{-1}\mathbf{E}_{ij}] && (3.8)\\
&= \det \mathbf{X}\mathrm{Tr}_X[\mathbf{E}_{ij} \otimes \langle i'|\mathbf{X}^{-1}\mathbf{E}_{ij}|i'\rangle] && (3.9)\\
&= \det \mathbf{X}\mathbf{E}_{ij}\langle j|\mathbf{X}^{-1}|i\rangle && (3.10)\\
&= \det \mathbf{X}(\mathbf{X}^{-1})^T. && (3.11)
\end{aligned}$$

Eq51

$$\frac{\partial \det(\mathbf{AXB})}{\partial \mathbf{X}} = \det(\mathbf{AXB})(\mathbf{X}^{-1})^T \tag{3.12}$$

Proof:

$$\begin{aligned}
\frac{\partial \det(\mathbf{AXB})}{\partial \mathbf{X}} &= \det(\mathbf{AXB})\mathrm{Tr}_X[I \otimes (\mathbf{AXB})^{-1}\frac{\partial \mathbf{AXB}}{\partial \mathbf{X}}] && (3.13)\\
&= \det(\mathbf{AXB})\mathrm{Tr}_X[I \otimes (\mathbf{AXB})^{-1}(I \otimes A)(\mathbf{E}_{ij} \otimes \mathbf{E}_{ij})(I \otimes B)] && (3.14)\\
&= \det(\mathbf{AXB})\mathrm{Tr}_X[\mathbf{E}_{ij} \otimes \mathbf{B}^{-1}\mathbf{X}^{-1}\mathbf{A}^{-1}\mathbf{A}\mathbf{E}_{ij}\mathbf{B}] && (3.15)\\
&= \det(\mathbf{AXB})\mathrm{Tr}_X[\mathbf{E}_{ij} \otimes \mathbf{X}^{-1}\mathbf{E}_{ij}] && (3.16)\\
&= \det(\mathbf{AXB})(\mathbf{X}^{-1})^T && (3.17)
\end{aligned}$$

In the last second equality, we have used cyclic property of trace operator.

Eq54

$$\frac{\partial \det(\mathbf{X}^T\mathbf{AX})}{\partial \mathbf{X}} = \det(\mathbf{X}^T\mathbf{AX})\left(\mathbf{AX}(\mathbf{X}^T\mathbf{AX})^{-1} + \mathbf{A}^T\mathbf{X}(\mathbf{X}^T\mathbf{A}^T\mathbf{X})^{-1}\right) \tag{3.18}$$

$$\frac{\partial \det(\mathbf{X}^T\mathbf{A}\mathbf{X})}{\partial \mathbf{X}} = \det(\mathbf{X}^T\mathbf{A}\mathbf{X})\mathrm{Tr}_X\left[I\otimes(\mathbf{X}^T\mathbf{A}\mathbf{X})^{-1}(\frac{\partial \mathbf{X}^T}{\partial \mathbf{X}})I\otimes\mathbf{A}\mathbf{X}+I\otimes\mathbf{X}^T\frac{\partial \mathbf{A}\mathbf{X}}{\partial \mathbf{X}}\right] \tag{3.19}$$

$$= \det(\mathbf{X}^T\mathbf{A}\mathbf{X})\mathrm{Tr}_X\left[I\otimes(\mathbf{X}^T\mathbf{A}\mathbf{X})^{-1}(\mathbf{E}_{ij}\otimes\mathbf{E}_{ji}I\otimes\mathbf{A}\mathbf{X}+I\otimes\mathbf{X}^T\mathbf{A}\mathbf{E}_{ij}\otimes\mathbf{E}_{ij})\right] \tag{3.20}$$

$$= \det(\mathbf{X}^T\mathbf{A}\mathbf{X})\mathrm{Tr}_X\left[\mathbf{E}_{ij}\otimes(\mathbf{X}^T\mathbf{A}\mathbf{X})^{-1}\mathbf{E}_{ji}\mathbf{A}\mathbf{X}+\mathbf{E}_{ij}\otimes(\mathbf{X}^T\mathbf{A}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{A}\mathbf{E}_{ij}\right] \tag{3.21}$$

$$= \det(\mathbf{X}^T\mathbf{A}\mathbf{X})\left(\mathbf{A}\mathbf{X}(\mathbf{X}^T\mathbf{A}\mathbf{X})^{-1}+\mathbf{A}^T\mathbf{X}(\mathbf{X}^T\mathbf{A}^T\mathbf{X})^{-1}\right). \tag{3.22}$$

Eq55

$$\frac{\partial \ln\det(\mathbf{X}^T\mathbf{X})}{\partial \mathbf{X}} = 2(\mathbf{X}^-)^T \tag{3.23}$$

Proof By using Eq.(3.18)

$$\frac{\partial \ln\det(\mathbf{X}^T\mathbf{X})}{\partial \mathbf{X}} = (\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}+\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}) = 2(\mathbf{X}^-)^T. \tag{3.24}$$

Here, we should note that this equation assume row vectors of $\mathbf{X}$ span a full rank space, otherwise, we can not write $\mathbf{X}^-$ as this simple form.

Eq58

$$\frac{\partial \det \mathbf{X}^k}{\partial \mathbf{X}} = k\det\mathbf{X}^k(X^{-1})^T \tag{3.25}$$

Proof

$$\frac{\partial \det \mathbf{X}^k}{\partial \mathbf{X}} = \det\mathbf{X}\frac{\partial \det \mathbf{X}^{k-1}}{\partial \mathbf{X}} + \det\mathbf{X}^k(X^{-1})^T. \tag{3.26}$$

Eq61

$$\frac{\mathbf{a}^T\mathbf{X}^{-1}\mathbf{b}}{\partial \mathbf{X}} = -\mathbf{X}^{-T}\mathbf{a}\mathbf{b}^T\mathbf{X}^{-T} \tag{3.27}$$

Proof

$$\frac{\mathbf{a}^T\mathbf{X}^{-1}\mathbf{b}}{\partial \mathbf{X}} = I\otimes\mathbf{a}^T\frac{\partial \mathbf{X}^{-1}}{\partial X}I\otimes\mathbf{b} \tag{3.28}$$

$$= -(\mathbf{X}^{-1})_{ki}(\mathbf{X}^{-1})_{jl}\mathbf{E}_{ij}\otimes\mathbf{a}^T\mathbf{E}_{kl}\mathbf{b} \tag{3.29}$$

$$= -(\mathbf{X}^{-T})_{ik}\mathbf{a}_k\mathbf{b}_l(\mathbf{X}^{-T})_{lj}\mathbf{E}_{ij} \tag{3.30}$$

$$= -\mathbf{X}^{-T}\mathbf{a}\mathbf{b}^T\mathbf{X}^{-T} \tag{3.31}$$

Eq63

$$\frac{\mathrm{Tr}[\mathbf{A}\mathbf{X}^{-1}\mathbf{B}]}{\partial \mathbf{X}} = -(\mathbf{X}^{-1}\mathbf{B}\mathbf{A}X^{-1})^T \tag{3.32}$$

Proof

$$\frac{\mathrm{Tr}[\mathbf{A}\mathbf{X}^{-1}\mathbf{B}]}{\partial \mathbf{X}} = \frac{\mathrm{Tr}[\mathbf{X}^{-1}\mathbf{B}\mathbf{A}]}{\partial \mathbf{X}} \tag{3.33}$$

$$= -(\mathbf{X}^{-1})_{ki}(\mathbf{X}^{-1})_{jl}\mathrm{Tr}[\mathbf{E}_{ij}\otimes\mathbf{E}_{kl}\mathbf{B}\mathbf{A}] \tag{3.34}$$

$$= (\mathbf{X}^{-1})_{ki}(\mathbf{X}^{-1})_{jl}(\mathbf{B}\mathbf{A})_{lk}\mathbf{E}_{ij} \tag{3.35}$$

$$= -(\mathbf{X}^{-1}\mathbf{B}\mathbf{A}X^{-1})^T \tag{3.36}$$

$$\frac{\partial \mathbf{J}(\mathbf{X})}{\partial \mathbf{X}^{-1}} = -\mathbf{X}^T \frac{\partial \mathbf{J}(\mathbf{X})}{\partial \mathbf{X}^{-1}} \mathbf{X}^T. \tag{3.37}$$

By using Eq.(2.48), we can directly obtain this result.

Eq67,68 $\rightarrow$ trivially.

The whole equations in section 2.4 are trivial, by using the above equation.

For section 2.5, we just need to prove the following equation,

$$\frac{\partial \mathrm{Tr}[F(\mathbf{X})]}{\partial \mathbf{X}} = f(\mathbf{X})^T \tag{3.38}$$

Proof

$$\frac{\partial \mathrm{Tr}[F(\mathbf{X})]}{\partial \mathbf{X}} = \mathrm{Tr}[\frac{\partial F(\mathbf{X})}{\partial \mathbf{X}}] \tag{3.39}$$

$$= \mathrm{Tr}[I \otimes f(\mathbf{X})\mathbf{E}_{ij} \otimes \mathbf{E}_{ji}] \tag{3.40}$$

$$= \mathbf{E}_{ij} f(\mathbf{X})_{ij} = f(\mathbf{X})^T. \tag{3.41}$$

A Tensor Network

References

[1] Author, *Title*, *J. Abbrev.* **vol** (year) pg.

[2] Author, *Title*, arxiv:1234.5678.

[3] Author, *Title*, Publisher (year).