BEYOND THIRD-ORDER EFFICIENCY: Supplement

For those who want to read Kano (1993) without pencil.

By Yutaka Kano
University of Osaka Prefecture

The present note gives elementary proofs of the propositions and some other results described in Kano (1993). We take over the definition and notations of his paper and do not duplicate here.

Proposition 2.1. Let A, B and $\mathbf{b}$ be of order $p_1 \times p_2$, $q_1 \times q_2$ and $q_1 \times 1$.

i) $K_{pq}^{-1} = K'_{pq} = K_{qp}$, $K_{p1} = K_{1p} = I_{p^2}$, and K_{pp} is symmetric;

ii) $K_{q_1 p_1}(A \otimes B)K_{p_2 q_2} = B \otimes A$, and $K_{q_1 p_1}(A \otimes \mathbf{b}) = \mathbf{b} \otimes A$;

iii) $K_{p,qr}(I_q \otimes K_{rp}) = K_{pq} \otimes I_r$, and $K_{pq,r}(K_{rp} \otimes I_q) = I_p \otimes K_{qr}$.

Magnus and Neudecker [(1988), page 47] and Magnus and Wansbeek [(1983), formula (1.8)] present concise proofs. So we omit most of them. We just note a short proof of iii) here. Let $\mathbf{p}$, $\mathbf{q}$ and $\mathbf{r}$ be p-, q- and r-vectors. We have

$$\begin{aligned} K_{p,qr}(I_q \otimes K_{rp})(\mathbf{q} \otimes \mathbf{p} \otimes \mathbf{r}) =& K_{p,qr}(\mathbf{q} \otimes \mathbf{r} \otimes \mathbf{p}) = K_{p,qr}\{(\mathbf{q} \otimes \mathbf{r}) \otimes \mathbf{p}\} \\ =& \mathbf{p} \otimes \mathbf{q} \otimes \mathbf{r} = (K_{pq} \otimes I_r)(\mathbf{q} \otimes \mathbf{p} \otimes \mathbf{r}), \end{aligned}$$

which means that $K_{p,qr}(I_q \otimes K_{rp}) = K_{pq} \otimes I_r$. The second formula can be shown in the same manner. *q.e.d.*

Proposition 2.2.

i) $N_{p^{<k>}}$ is symmetric, idempotent and $\mathrm{rank}(N_{p^{<k>}}) = \binom{p+k-1}{k}$;

ii) $N_{p^{<k>}}\mathbf{a}^{<k>} = \mathbf{a}^{<k>}$ for p-vector $\mathbf{a}$;

iii) $N_{p^{<k>}}A^{<k>} = A^{<k>}N_{q^{<k>}} = N_{p^{<k>}}A^{<k>}N_{q^{<k>}}$ for A of $p \times q$;

iv) $(I_{p^{k_1}} \otimes K_{p^{k_2}p^{k_3}} \otimes I_{p^{k_4}})N_{p^{<k>}} = N_{p^{<k>}}$ for any $k_j \geq 0$ with $\sum_j^4 k_j = k$;

$(I_{p^{k_1}} \otimes N_{p^{<k_2>}} \otimes I_{p^{k_3}})N_{p^{<k>}} = N_{p^{<k>}}$ for any $k_j \geq 0$ with $\sum_j^3 k_j = k$.

Let $\sigma = \binom{1 \cdots k}{\sigma(1) \cdots \sigma(k)} \in S_k$ be a permutation. The inverse of σ is denoted as $\sigma^{-1} = \binom{\sigma(1) \cdots \sigma(k)}{1 \cdots k} = \binom{1 \cdots k}{\sigma^{-1}(1) \cdots \sigma^{-1}(k)}$. We first note

$$\begin{aligned} N_{p^{<k>}}(\mathbf{a}_1 \otimes \cdots \otimes \mathbf{a}_p) &= \sum_{\sigma \in S_k} (\mathbf{a}_{\sigma(1)} \otimes \cdots \otimes \mathbf{a}_{\sigma(k)})/k! \\ &= \sum_{\sigma \in S_k} (\mathbf{a}_{\sigma^{-1}(1)} \otimes \cdots \otimes \mathbf{a}_{\sigma^{-1}(k)})/k!. \end{aligned}$$

Let $I_p = [\mathbf{e}_1, \ldots, \mathbf{e}_p]$, where $\mathbf{e}_j$ is the j-th canonical base of $\mathbf{R}^p$, so that $I_p = \sum_{j=1}^p \mathbf{e}_j \mathbf{e}_j'$. We have

$$\begin{aligned} N_{p^{<k>}} &= N_{p^{<k>}}(\overbrace{I_p \otimes \cdots \otimes I_p}^{k\text{-fold}}) \\ &= \sum_{i_1=1}^{p} \cdots \sum_{i_k=1}^{p} N_{p^{<k>}}(\mathbf{e}_{i_1} \otimes \cdots \otimes \mathbf{e}_{i_k})(\mathbf{e}_{i_1} \otimes \cdots \otimes \mathbf{e}_{i_k})' \\ &= \sum_{i_1=1}^{p} \cdots \sum_{i_k=1}^{p} \sum_{\sigma \in S_k} (\mathbf{e}_{i_{\sigma(1)}} \otimes \cdots \otimes \mathbf{e}_{i_{\sigma(k)}})(\mathbf{e}_{i_1} \otimes \cdots \otimes \mathbf{e}_{i_k})'/k! \quad (1) \\ &= \sum_{i_1=1}^{p} \cdots \sum_{i_k=1}^{p} \sum_{\sigma \in S_k} (\mathbf{e}_{i_1} \otimes \cdots \otimes \mathbf{e}_{i_k})(\mathbf{e}_{i_{\sigma(1)}} \otimes \cdots \otimes \mathbf{e}_{i_{\sigma(k)}})'/k!, \quad (2) \end{aligned}$$

if $N_{p^{<k>}}$ is symmetric.

Proof of i). Symmetrization of the k-fold Kronecker product that has been symmetrized does not change anything, i.e., since for any permutation σ'

$$\begin{aligned} N_{p^{<k>}}(\mathbf{a}_{\sigma'(1)} \otimes \cdots \otimes \mathbf{a}_{\sigma'(k)}) &= \sum_{\sigma \in S_k} (\mathbf{a}_{\sigma\{\sigma'(1)\}} \otimes \cdots \otimes \mathbf{a}_{\sigma\{\sigma'(k)\}})/k! \\ &= \sum_{\sigma \in S_k} (\mathbf{a}_{\sigma(1)} \otimes \cdots \otimes \mathbf{a}_{\sigma(k)})/k!, \end{aligned}$$

we have

$$\begin{aligned} \left(N_{p^{<k>}}\right)^2 (\mathbf{a}_1 \otimes \cdots \otimes \mathbf{a}_k) &= N_{p^{<k>}} \left\{ \sum_{\sigma' \in S_k} (\mathbf{a}_{\sigma'(1)} \otimes \cdots \otimes \mathbf{a}_{\sigma'(k)})/k! \right\} \\ &= \sum_{\sigma \in S_k} (\mathbf{a}_{\sigma(1)} \otimes \cdots \otimes \mathbf{a}_{\sigma(k)})/k! = N_{p^{<k>}}(\mathbf{a}_1 \otimes \cdots \otimes \mathbf{a}_k) \end{aligned}$$

for any $\mathbf{a}_j'$s $\in \mathbf{R}^p$. Thus, $(N_{p^{<k>}})^2 = N_{p^{<k>}}$. *q.e.d.*

It is seen from (1) that the image of $N_{p^{<k>}}$ is spanned by

$$\left\{ \sum_{\sigma \in S_k} \mathbf{a}_{i_{\sigma(1)}} \otimes \cdots \otimes \mathbf{a}_{i_{\sigma(k)}} \,|\, i_1, \ldots i_k = 1, 2, \ldots, p \right\},$$

whose rank is equal to $\binom{p+k-1}{k}$, being the number of repeated combinations of p things, $\mathbf{e}_1, \ldots, \mathbf{e}_p$, taken k at a time. This proves that $\mathrm{rank}(N_{p^{<k>}}) = \binom{p+k-1}{k}$. *q.e.d.*

Let us show that $N_{p^{<k>}}$ is symmetric. The expression (1) is rewritten as

$$\begin{aligned}
N_{p^{<k>}} =& \frac{1}{k!}\sum_{\sigma\in S_k}\sum_{i_1=1}^{p}\cdots\sum_{i_k=1}^{p}(\mathbf{e}_{i_{\sigma(1)}}\otimes\cdots\otimes\mathbf{e}_{i_{\sigma(k)}})(\mathbf{e}_{i_1}\otimes\cdots\otimes\mathbf{e}_{i_k})' \\
=& \frac{1}{k!}\sum_{\sigma\in S_k}\sum_{i_{\sigma^{-1}(1)}=1}^{p}\cdots\sum_{i_{\sigma^{-1}(k)}=1}^{p}(\mathbf{e}_{i_1}\otimes\cdots\otimes\mathbf{e}_{i_k})(\mathbf{e}_{i_{\sigma^{-1}(1)}}\otimes\cdots\otimes\mathbf{e}_{i_{\sigma^{-1}(k)}})' \\
=& \frac{1}{k!}\sum_{\sigma\in S_k}\sum_{i_1=1}^{p}\cdots\sum_{i_k=1}^{p}(\mathbf{e}_{i_1}\otimes\cdots\otimes\mathbf{e}_{i_k})(\mathbf{e}_{i_{\sigma^{-1}(1)}}\otimes\cdots\otimes\mathbf{e}_{i_{\sigma^{-1}(k)}})' \\
=& \frac{1}{k!}\sum_{i_1=1}^{p}\cdots\sum_{i_k=1}^{p}\sum_{\sigma\in S_k}(\mathbf{e}_{i_1}\otimes\cdots\otimes\mathbf{e}_{i_k})(\mathbf{e}_{i_{\sigma^{-1}(1)}}\otimes\cdots\otimes\mathbf{e}_{i_{\sigma^{-1}(k)}})' \\
=& N'_{p^{<k>}}
\end{aligned}$$

in view of (2). *q.e.d.*

The result ii) is an immediate consequence of the definition of $N_{p^{<k>}}$. *q.e.d.*

Proof of iii). Let $\mathbf{b}_j$'s be q-vectors. Then

$$\begin{aligned}
N_{p^{<k>}}A^{<k>}(\mathbf{b}_1\otimes\cdots\otimes\mathbf{b}_k) =& N_{p^{<k>}}\left\{(A\mathbf{b}_1)\otimes\cdots\otimes(A\mathbf{b}_k)\right\} \\
=& \sum_{\sigma\in S_k}\left\{(A\mathbf{b}_{\sigma(1)})\otimes\cdots\otimes(A\mathbf{b}_{\sigma(k)})\right\}/k! \\
=& A^{<k>}\sum_{\sigma\in S_k}(\mathbf{b}_{\sigma(1)}\otimes\cdots\otimes\mathbf{b}_{\sigma(k)})/k! \\
=& A^{<k>}N_{q^{<k>}}(\mathbf{b}_1\otimes\cdots\otimes\mathbf{b}_k),
\end{aligned}$$

which proves that $N_{p^{<k>}}A^{<k>} = A^{<k>}N_{q^{<k>}}$. It follows

$$N_{p^{<k>}}A^{<k>}N_{q^{<k>}} = \left(N_{p^{<k>}}\right)^2 A^{<k>} = N_{p^{<k>}}A^{<k>}.$$

q.e.d.

Proof of iv). Since $N_{p^{<k>}}$ is symmetric, it suffices to show that $N_{p^{<k>}} = N_{p^{<k>}}(I_{p^{k_1}}\otimes K_{p^{k_3}p^{k_2}}\otimes I_{p^{k_4}})$ and $N_{p^{<k>}}(I_{p^{k_1}}\otimes N_{p^{<k_2>}}\otimes I_{p^{k_3}}) = N_{p^{<k>}}$. These formulas are obvious because symmetrization of $\mathbf{a}_1\otimes\cdots\otimes\mathbf{a}_k$ is identical with that of $\mathbf{a}_1\otimes\cdots\otimes\mathbf{a}_k$, some of which have been interchanged or symmetrized. *q.e.d.*

Taylor Expansion. Let $g(\mathbf{x})$ be a q-vector valued analytic function of p variates $\mathbf{x}$.

$$g(\mathbf{x}) = \sum_{k=0}^{\infty}\frac{G_k}{k!}(\mathbf{x}-\mathbf{x}_0)^{<k>} \quad \text{with} \quad G_k = \left(\frac{\mathrm{d}}{\mathrm{d}\mathbf{x}'}\right)^{<k>} g(\mathbf{x})\bigg|_{\mathbf{x}=\mathbf{x}_0}.$$

We take $q = 1$ without loss of generality, i.e., $g(\mathbf{x})$ is scalar-valued. The Taylor series of $g(\mathbf{x})$ is formally expressible as:

$$g(\mathbf{x}) = \sum_{k=0}^{\infty} \frac{1}{k!} \left(\sum_{j=1}^{p} (x_j - x_{j0}) \frac{\partial}{\partial x_j} \right)^k g(\mathbf{x}_0),$$

and hence it suffices to show

$$\left[\left(\frac{\mathrm{d}}{\mathrm{d}\mathbf{x}'} \right)^{<k>} g(\mathbf{x}) \right] (\mathbf{z} - \mathbf{x}_0)^{<k>} = \left(\sum_{j=1}^{p} (z_j - x_{j0}) \frac{\partial}{\partial x_j} \right)^k g(\mathbf{x})$$

for all $\mathbf{z} = [z_1, \ldots, z_p]'$ and $\mathbf{x} = [x_1, \ldots, x_p]'$. We prove it by induction:

$$\begin{aligned}
&\left[\left(\frac{\mathrm{d}}{\mathrm{d}\mathbf{x}'} \right)^{<k+1>} g(\mathbf{x}) \right] (\mathbf{z} - \mathbf{x}_0)^{<k+1>} \\
&\quad = \left[\frac{\partial}{\partial x_1} \left(\frac{\mathrm{d}}{\mathrm{d}\mathbf{x}'} \right)^{<k>} g(\mathbf{x}), \ldots, \frac{\partial}{\partial x_p} \left(\frac{\mathrm{d}}{\mathrm{d}\mathbf{x}'} \right)^{<k>} g(\mathbf{x}) \right] \begin{bmatrix} (z_1 - x_{10})(\mathbf{z} - \mathbf{x}_0)^{<k>} \\ \vdots \\ (z_p - x_{p0})(\mathbf{z} - \mathbf{x}_0)^{<k>} \end{bmatrix} \\
&\quad = \sum_{i=1}^{p} (z_i - x_{i0}) \frac{\partial}{\partial x_i} \left[\left(\frac{\mathrm{d}}{\mathrm{d}\mathbf{x}'} \right)^{<k>} g(\mathbf{x}) \right] (\mathbf{z} - \mathbf{x}_0)^{<k>} \\
&\quad = \sum_{i=1}^{p} (z_i - x_{i0}) \frac{\partial}{\partial x_i} \left(\sum_{j=1}^{p} (z_j - x_{j0}) \frac{\partial}{\partial x_j} \right)^k g(\mathbf{x}) \\
&\qquad\qquad\qquad\qquad\qquad\qquad \text{(by the assumption due to induction)} \\
&\quad = \left(\sum_{j=1}^{p} (x_j - x_{j0}) \frac{\partial}{\partial x_j} \right)^{k+1} g(\mathbf{x}).
\end{aligned}$$

q.e.d.

Next, let us show (2.1) in Kano (1993) and related consequences. Before this, we state a result on a linear combination of differential operators. Define $\mathbf{b}' \frac{\mathrm{d}}{\mathrm{d}\mathbf{x}} = \sum_{j=1}^{p} b_j \frac{\partial}{\partial x_j}$. Then it follows that

$$(3) \qquad (\mathbf{b}_k \otimes \cdots \otimes \mathbf{b}_1 \otimes \mathbf{a})' \left(\frac{\mathrm{d}}{\mathrm{d}\mathbf{x}} \right)^{<k>} A(\mathbf{x}) = \left(\mathbf{b}_k' \frac{\mathrm{d}}{\mathrm{d}\mathbf{x}} \right) \cdots \left(\mathbf{b}_1' \frac{\mathrm{d}}{\mathrm{d}\mathbf{x}} \right) [\mathbf{a}' A(\mathbf{x})],$$

where $A(\mathbf{x})$ is an $a_1 \times a_2$ matrix-valued function, $\mathbf{b}_j$'s are p-vectors, and $\mathbf{a}$ an a_1-vector.

This can be shown again by induction as follows:

$$
\begin{aligned}
&(\mathbf{b}_{k+1} \otimes \cdots \otimes \mathbf{b}_1 \otimes \mathbf{a})' \left(\frac{\mathrm{d}}{\mathrm{d}\mathbf{x}}\right)^{<k+1>} A(\mathbf{x}) \\
&\quad = \left[b_{k+1}^{(1)}(\mathbf{b}_k \otimes \cdots \otimes \mathbf{b}_1 \otimes \mathbf{a})', \ldots, b_{k+1}^{(p)}(\mathbf{b}_k \otimes \cdots \otimes \mathbf{b}_1 \otimes \mathbf{a})'\right] \begin{bmatrix} \frac{\partial}{\partial x_1}\left(\frac{\mathrm{d}}{\mathrm{d}\mathbf{x}}\right)^{<k>} A(\mathbf{x}) \\ \vdots \\ \frac{\partial}{\partial x_p}\left(\frac{\mathrm{d}}{\mathrm{d}\mathbf{x}}\right)^{<k>} A(\mathbf{x}) \end{bmatrix} \\
&\quad = \sum_{j=1}^{p} b_{k+1}^{(j)} \frac{\partial}{\partial x_j}(\mathbf{b}_k \otimes \cdots \otimes \mathbf{b}_1 \otimes \mathbf{a})' \left(\frac{\mathrm{d}}{\mathrm{d}\mathbf{x}}\right)^{<k>} A(\mathbf{x}) \\
&\quad = \sum_{j=1}^{p} b_{k+1}^{(j)} \frac{\partial}{\partial x_j} \left(\mathbf{b}_k' \frac{\mathrm{d}}{\mathrm{d}\mathbf{x}}\right) \cdots \left(\mathbf{b}_1' \frac{\mathrm{d}}{\mathrm{d}\mathbf{x}}\right) [\mathbf{a}' A(\mathbf{x})] \\
&\qquad\qquad \text{(by the assumption due to induction)} \\
&\quad = \left(\mathbf{b}_{k+1}' \frac{\mathrm{d}}{\mathrm{d}\mathbf{x}}\right) \left(\mathbf{b}_k' \frac{\mathrm{d}}{\mathrm{d}\mathbf{x}}\right) \cdots \left(\mathbf{b}_1' \frac{\mathrm{d}}{\mathrm{d}\mathbf{x}}\right) [\mathbf{a}' A(\mathbf{x})].
\end{aligned}
$$

$N_{p<k>}$ and higher-order derivatives. Let $A(\mathbf{x})$ be an $a_1 \times a_2$-matrix valued function. Then

i) $\left[\left(\frac{\mathrm{d}}{\mathrm{d}\mathbf{x}}\right)^{<k>} A(\mathbf{x})\right]' = \left(\frac{\mathrm{d}}{\mathrm{d}\mathbf{x}'}\right)^{<k>} A(\mathbf{x})'$;

ii) $\left(N_{p<k>} \otimes I_{a_1}\right) \left[\left(\frac{\mathrm{d}}{\mathrm{d}\mathbf{x}}\right)^{<k>} A(\mathbf{x})\right] = \left(\frac{\mathrm{d}}{\mathrm{d}\mathbf{x}}\right)^{<k>} A(\mathbf{x})$;

iii) $\left[\left(\frac{\mathrm{d}}{\mathrm{d}\mathbf{x}'}\right)^{<k>} A(\mathbf{x})\right] \left(N_{p<k>} \otimes I_{a_2}\right) = \left(\frac{\mathrm{d}}{\mathrm{d}\mathbf{x}'}\right)^{<k>} A(\mathbf{x})$.

Remark. The equality iii) reduces to (2.1) in Kano (1993) when $a_2 = 1$.

Proof of i). The proof is due to induction. We have

$$
\begin{aligned}
\left[\left(\frac{\mathrm{d}}{\mathrm{d}\mathbf{x}}\right)^{<k+1>} A(\mathbf{x})\right]' &= \begin{bmatrix} \frac{\partial}{\partial x_1}\left(\frac{\mathrm{d}}{\mathrm{d}\mathbf{x}}\right)^{<k>} A(\mathbf{x}) \\ \vdots \\ \frac{\partial}{\partial x_p}\left(\frac{\mathrm{d}}{\mathrm{d}\mathbf{x}}\right)^{<k>} A(\mathbf{x}) \end{bmatrix}' \\
&= \left[\frac{\partial}{\partial x_1}\left(\frac{\mathrm{d}}{\mathrm{d}\mathbf{x}'}\right)^{<k>} A(\mathbf{x})', \ldots, \frac{\partial}{\partial x_p}\left(\frac{\mathrm{d}}{\mathrm{d}\mathbf{x}'}\right)^{<k>} A(\mathbf{x})'\right] \\
&\qquad\qquad \text{(by the assumption due to induction)} \\
&= \left(\frac{\mathrm{d}}{\mathrm{d}\mathbf{x}'}\right)^{<k+1>} A(\mathbf{x})'.
\end{aligned}
$$

q.e.d.

Proof of ii). We write $I_{a_1} = \sum_{j=1}^{a_1} \widetilde{\mathbf{e}}_j \widetilde{\mathbf{e}}_j'$ with $\widetilde{\mathbf{e}}_j$ the j-th canonical base of $\mathbf{R}^{a_1}$. Using (2), we have

$$
\begin{aligned}
&\left(N_{p^{<k>}} \otimes I_{a_1}\right)\left[\left(\frac{\mathrm{d}}{\mathrm{d}\mathbf{x}}\right)^{<k>} A(\mathbf{x})\right] \\
&\quad = \left[\frac{1}{k!}\sum_{i_1=1}^{p}\cdots\sum_{i_k=1}^{p}\sum_{\sigma\in S_k}\Big\{(\mathbf{e}_{i_1}\otimes\cdots\otimes\mathbf{e}_{i_k}) \right. \\
&\qquad\qquad \left.\cdot(\mathbf{e}_{i_{\sigma(1)}}\otimes\cdots\otimes\mathbf{e}_{i_{\sigma(k)}})'\Big\}\otimes\left(\sum_{j=1}^{a_1}\widetilde{\mathbf{e}}_j\widetilde{\mathbf{e}}_j'\right)\right]\left(\frac{\mathrm{d}}{\mathrm{d}\mathbf{x}}\right)^{<k>} A(\mathbf{x}) \\
&\quad = \frac{1}{k!}\sum_{i_1=1}^{p}\cdots\sum_{i_k=1}^{p}\sum_{\sigma\in S_k}\sum_{j=1}^{a_1}(\mathbf{e}_{i_1}\otimes\cdots\otimes\mathbf{e}_{i_k}\otimes\widetilde{\mathbf{e}}_j) \\
&\qquad\qquad \cdot(\mathbf{e}_{i_{\sigma(1)}}\otimes\cdots\otimes\mathbf{e}_{i_{\sigma(k)}}\otimes\widetilde{\mathbf{e}}_j)'\left(\frac{\mathrm{d}}{\mathrm{d}\mathbf{x}}\right)^{<k>} A(\mathbf{x}) \\
&\quad = \frac{1}{k!}\sum_{i_1=1}^{p}\cdots\sum_{i_k=1}^{p}\sum_{\sigma\in S_k}\sum_{j=1}^{a_1}(\mathbf{e}_{i_1}\otimes\cdots\otimes\mathbf{e}_{i_k}\otimes\widetilde{\mathbf{e}}_j) \\
&\qquad\qquad \cdot\frac{\partial}{\partial x_{i_{\sigma(1)}}}\cdots\frac{\partial}{\partial x_{i_{\sigma(k)}}}\left[\widetilde{\mathbf{e}}_j' A(\mathbf{x})\right] \qquad \text{(by (3))} \\
&\quad = \frac{1}{k!}\sum_{i_1=1}^{p}\cdots\sum_{i_k=1}^{p}\sum_{\sigma\in S_k}\sum_{j=1}^{a_1}(\mathbf{e}_{i_1}\otimes\cdots\otimes\mathbf{e}_{i_k}\otimes\widetilde{\mathbf{e}}_j)\frac{\partial}{\partial x_{i_1}}\cdots\frac{\partial}{\partial x_{i_k}}\left[\widetilde{\mathbf{e}}_j' A(\mathbf{x})\right] \\
&\quad = \sum_{i_1=1}^{p}\cdots\sum_{i_k=1}^{p}\sum_{j=1}^{a_1}(\mathbf{e}_{i_1}\otimes\cdots\otimes\mathbf{e}_{i_k}\otimes\widetilde{\mathbf{e}}_j) \\
&\qquad\qquad \cdot(\mathbf{e}_{i_1}\otimes\cdots\otimes\mathbf{e}_{i_k}\otimes\widetilde{\mathbf{e}}_j)'\left(\frac{\mathrm{d}}{\mathrm{d}\mathbf{x}}\right)^{<k>} A(\mathbf{x}) \\
&\quad = \left(\frac{\mathrm{d}}{\mathrm{d}\mathbf{x}}\right)^{<k>} A(\mathbf{x}).
\end{aligned}
$$

q.e.d.

A proof of iii) is obtained by transposing both sides of ii) and using i). *q.e.d.*

Proposition 2.3. Let $\mathbf{x}$ be a p-vector and let $A(\mathbf{x})$ and $B(\mathbf{x})$ be of order $a_1 \times a_2$ and $b_1 \times b_2$. Suppose that the following matrix products in the LHS are permissible.

i) $\frac{\mathrm{d}}{\mathrm{d}\mathbf{x}'}[A(\mathbf{x})B(\mathbf{x})] = \frac{\mathrm{d}A(\mathbf{x})}{\mathrm{d}\mathbf{x}'}(I_p \otimes B(\mathbf{x})) + A(\mathbf{x})\frac{\mathrm{d}B(\mathbf{x})}{\mathrm{d}\mathbf{x}'}$;

ii) $\frac{\mathrm{d}}{\mathrm{d}\mathbf{x}'}[A(\mathbf{x})\otimes B(\mathbf{x})] = \left(\frac{\mathrm{d}A(\mathbf{x})}{\mathrm{d}\mathbf{x}'}\otimes B(\mathbf{x})\right) + \left(A(\mathbf{x})\otimes\frac{\mathrm{d}B(\mathbf{x})}{\mathrm{d}\mathbf{x}'}\right)(K_{a_2 p}\otimes I_{b_2})$;

iii) $\frac{\mathrm{d}}{\mathrm{d}\mathbf{t}'}[A(\mathbf{x}(\mathbf{t}))B] = \frac{\mathrm{d}A(\mathbf{x})}{\mathrm{d}\mathbf{x}'}\left(\frac{\mathrm{d}\mathbf{x}(\mathbf{t})}{\mathrm{d}\mathbf{t}'}\otimes B\right)$, where $\mathbf{x}(\mathbf{t})$ is a p-vector valued function of $\mathbf{t}$ and B is a constant matrix.

Proof of i). It follows from the definition of $\frac{\mathrm{d}}{\mathrm{d}\mathbf{x}'}$ that

$$
\frac{\mathrm{d}}{\mathrm{d}\mathbf{x}'}[AB(\mathbf{x})C] = \left[A\frac{\partial B(\mathbf{x})}{\partial x_1}C,\ldots,A\frac{\partial B(\mathbf{x})}{\partial x_1}C\right] = A\frac{\mathrm{d}B(\mathbf{x})}{\mathrm{d}\mathbf{x}'}(I_p\otimes C). \tag{4}
$$

We can apply (4) to show i). *q.e.d.*

Proof of ii). Let A, B and C be matrices of order $a_1 \times a_2$, $b_1 \times b_2$, $c_1 \times c_2$, respectively. Then

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\mathbf{x}'}[A \otimes B(\mathbf{x}) \otimes C] &= \frac{\mathrm{d}}{\mathrm{d}\mathbf{x}'}[A \otimes (B(\mathbf{x}) \otimes C)] \\
&= \frac{\mathrm{d}}{\mathrm{d}\mathbf{x}'}\left[K_{a_1,b_1c_1}\{(B(\mathbf{x}) \otimes C) \otimes A\}K_{b_2c_2,a_2}\right] \qquad \text{(by Prop.2.1 ii))} \\
&= K_{a_1,b_1c_1}\frac{\mathrm{d}}{\mathrm{d}\mathbf{x}'}[B(\mathbf{x}) \otimes C \otimes A](I_p \otimes K_{b_2c_2,a_2}) \qquad \text{(by (4))} \\
&= K_{a_1,b_1c_1}\left[\frac{\partial B(\mathbf{x})}{\partial x_1} \otimes C \otimes A, \ldots, \frac{\partial B(\mathbf{x})}{\partial x_p} \otimes C \otimes A\right](I_p \otimes K_{b_2c_2,a_2}) \\
&= K_{a_1,b_1c_1}\left(\frac{\mathrm{d}B(\mathbf{x})}{\mathrm{d}\mathbf{x}'} \otimes C \otimes A\right)(I_p \otimes K_{b_2c_2,a_2}) \\
&= \left(A \otimes \frac{\mathrm{d}B(\mathbf{x})}{\mathrm{d}\mathbf{x}'} \otimes C\right)K_{a_2,pb_2c_2}(I_p \otimes K_{b_2c_2,a_2}) \\
&= \left(A \otimes \frac{\mathrm{d}B(\mathbf{x})}{\mathrm{d}\mathbf{x}'} \otimes C\right)(K_{a_2p} \otimes I_{b_2c_2}) \qquad \text{(by Prop.2.1 iii)).}
\end{aligned}$$

The expression is available to prove ii). *q.e.d.*

Proof of iii). We first note that

$$\begin{aligned}
\frac{\partial}{\partial t_j}[A(\mathbf{x}(\mathbf{t}))] &= \sum_{i=1}^{p}\frac{\partial A(\mathbf{x})}{\partial x_i}\frac{\partial x_i(\mathbf{t})}{\partial t_j} = \sum_{i=1}^{p}\frac{\partial A(\mathbf{x})}{\partial x_i}\left(\frac{\partial x_i(\mathbf{t})}{\partial t_j}I_{a_2}\right) \\
&= \left[\frac{\partial A(\mathbf{x})}{\partial x_1}, \cdots, \frac{\partial A(\mathbf{x})}{\partial x_p}\right]\begin{bmatrix} \frac{\partial x_1(\mathbf{t})}{\partial t_j}I_{a_2} \\ \vdots \\ \frac{\partial x_p(\mathbf{t})}{\partial t_j}I_{a_2} \end{bmatrix} \\
&= \frac{\mathrm{d}A(\mathbf{x})}{\mathrm{d}\mathbf{x}'}\left(\frac{\partial \mathbf{x}(\mathbf{t})}{\partial t_j} \otimes I_{a_2}\right),
\end{aligned}$$

which implies that $\frac{\partial}{\partial t_j}[A(\mathbf{x}(\mathbf{t}))B] = \frac{\mathrm{d}A(\mathbf{x})}{\mathrm{d}\mathbf{x}'}\left(\frac{\partial \mathbf{x}(\mathbf{t})}{\partial t_j} \otimes B\right)$. Hence,

$$\begin{aligned}
\frac{\partial}{\partial \mathbf{t}'}[A(\mathbf{x}(\mathbf{t}))B] &= \left[\frac{\partial}{\partial t_1}[A(\mathbf{x}(\mathbf{t}))B], \ldots, \frac{\partial}{\partial t_{p'}}[A(\mathbf{x}(\mathbf{t}))B]\right] \\
&= \frac{\mathrm{d}A(\mathbf{x})}{\mathrm{d}\mathbf{x}'}\left(\frac{\partial \mathbf{x}(\mathbf{t})}{\partial \mathbf{t}'} \otimes B\right).
\end{aligned}$$

q.e.d.

Moments.

$$
\begin{aligned}
\mathrm{E}[\mathbf{z}_n\mathbf{z}_n'] =& \Psi_{11}, \qquad \mathrm{E}[\mathbf{z}_n^{<2>}] = \Psi_{20}, \qquad \mathrm{E}[\mathbf{z}_n^{<3>}] = \frac{1}{\sqrt{n}}\Psi_{30};\\
\mathrm{E}[\mathbf{z}_n^{<2>}\mathbf{z}_n^{<2>\prime}] =& 2N_{p^{<2>}}\Psi_{11}^{<2>}N_{p^{<2>}} + \Psi_{20}\Psi_{02} + O\left(\frac{1}{n}\right);\\
\mathrm{E}[\mathbf{z}_n^{<3>}\mathbf{z}_n'] =& 3N_{p^{<3>}}(\Psi_{20}\otimes\Psi_{11}) + O\left(\frac{1}{n}\right) = 3N_{p^{<3>}}(\Psi_{11}\otimes\Psi_{20}) + O\left(\frac{1}{n}\right);\\
\mathrm{E}[\mathbf{z}_n^{<4>}] =& 3N_{p^{<4>}}\Psi_{20}^{<2>} + O\left(\frac{1}{n}\right);\\
\mathrm{E}[\mathbf{z}_n^{<3>}\mathbf{z}_n^{<3>\prime}] =& 6N_{p^{<3>}}\Psi_{11}^{<3>}N_{p^{<3>}}\\
&+ 9N_{p^{<3>}}(\Psi_{20}\otimes\Psi_{11}\otimes\Psi_{02})N_{p^{<3>}} + O\left(\frac{1}{n}\right).
\end{aligned}
$$

$$
N_{p^{<3>}}(\Psi_{02}\otimes\Psi_{11}\otimes\Psi_{20})N_{p^{<3>}} = N_{p^{<3>}}(\Psi_{20}\otimes\Psi_{11}\otimes\Psi_{02})N_{p^{<3>}}. \tag{2.2}
$$

Proof. The first three formulas are obvious. For simplicity we assume $\mathrm{E}[\mathbf{x}] = 0$. From the definition of the commutation matrix K_{pp}, we have $\mathbf{x}_1\otimes\mathbf{x}_2 = K_{pp}(\mathbf{x}_2\otimes\mathbf{x}_1)$, from which it follows that

$$
\begin{aligned}
&\mathrm{E}[\mathbf{z}_n^{<2>}\mathbf{z}_n^{<2>\prime}]\\
&=\mathrm{E}[(\mathbf{x}_1\otimes\mathbf{x}_2)(\mathbf{x}_2\otimes\mathbf{x}_1)'] + \mathrm{E}[(\mathbf{x}_1\otimes\mathbf{x}_2)(\mathbf{x}_1\otimes\mathbf{x}_2)'] + \mathrm{E}[(\mathbf{x}_1\otimes\mathbf{x}_1)(\mathbf{x}_2\otimes\mathbf{x}_2)'] + O(\frac{1}{n})\\
&=K_{pp}\mathrm{E}[(\mathbf{x}_2\mathbf{x}_2')\otimes(\mathbf{x}_1\mathbf{x}_1')] + \mathrm{E}[(\mathbf{x}_1\mathbf{x}_1')\otimes(\mathbf{x}_2\mathbf{x}_2')] + \mathrm{E}[(\mathbf{x}_1\otimes\mathbf{x}_1)(\mathbf{x}_2\otimes\mathbf{x}_2)'] + O(\frac{1}{n})\\
&=(K_{pp}+I_{p^2})(\Psi_{11}\otimes\Psi_{11}) + \Psi_{20}\Psi_{02} + O(\frac{1}{n}),
\end{aligned}
$$

from which, along with Prop.2.2 iii) and $N_{p^{<2>}} = (K_{pp}+I_{p^2})/2$, the result follows.

Using equalities such as $(\mathbf{x}_1\otimes\mathbf{x}_2)\otimes\mathbf{x}_1 = K_{p^2,p}\{\mathbf{x}_1\otimes(\mathbf{x}_1\otimes\mathbf{x}_2)\}$, we get

$$
\begin{aligned}
\mathrm{E}[\mathbf{z}_n^{<3>}\mathbf{z}_n'] =& \mathrm{E}[(\mathbf{x}_1\otimes\mathbf{x}_1\otimes\mathbf{x}_2)\mathbf{x}_2'] + \mathrm{E}[(\mathbf{x}_1\otimes\mathbf{x}_2\otimes\mathbf{x}_1)\mathbf{x}_2'] + \mathrm{E}[(\mathbf{x}_2\otimes\mathbf{x}_1\otimes\mathbf{x}_1)\mathbf{x}_2'] + O(\frac{1}{n})\\
=&(I_{p^3}+K_{p^2,p}+K_{p,p^2})\mathrm{E}[(\mathbf{x}_1\otimes\mathbf{x}_1\otimes\mathbf{x}_2)\mathbf{x}_2']\\
=&(I_{p^3}+K_{p^2,p}+K_{p,p^2})(\Psi_{20}\otimes\Psi_{11})
\end{aligned}
$$

because $(\mathbf{x}_1\otimes\mathbf{x}_1\otimes\mathbf{x}_2)\mathbf{x}_2' = (\mathbf{x}_1\otimes\mathbf{x}_1\otimes\mathbf{x}_2)(1\otimes1\otimes\mathbf{x}_2') = \mathbf{x}_1\otimes\mathbf{x}_1\otimes(\mathbf{x}_2\mathbf{x}_2')$. Premultiplying both sides by $N_{p^{<3>}}$ and using $N_{p^{<3>}}\mathbf{z}_n^{<3>} = \mathbf{z}_n^{<3>}$ and $N_{p^{<3>}}K_{p^2,p} = N_{p^{<3>}}K_{p,p^2} = N_{p^{<3>}}$ [see Prop.2.2 ii), iv)] show that

$$
\mathrm{E}[\mathbf{z}_n^{<3>}\mathbf{z}_n'] = 3N_{p^{<3>}}(\Psi_{20}\otimes\Psi_{11}),
$$

and furthermore

$$
\mathrm{E}[\mathbf{z}_n^{<3>}\mathbf{z}_n'] = 3N_{p^{<3>}}K_{p^2,p}(\Psi_{11}\otimes\Psi_{20}) = 3N_{p^{<3>}}(\Psi_{11}\otimes\Psi_{20})
$$

by virtue of Prop.2.1 ii) and Prop.2.2 iv). The same idea is available to show

$$\begin{aligned}
\mathrm{E}[\mathbf{z}_n^{<4>}] =& \mathrm{E}[\mathbf{x}_1 \otimes \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_2] + \mathrm{E}[\mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_1 \otimes \mathbf{x}_2] + \mathrm{E}[\mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_2 \otimes \mathbf{x}_1] + O(\frac{1}{n}) \\
=& \{I_{p^4} + (I_p \otimes K_{pp} \otimes I_p) + (I_p \otimes K_{p^2,p})\}\mathrm{E}[\mathbf{x}_1 \otimes \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_2] + O(\frac{1}{n}) \\
=& \{I_{p^4} + (I_p \otimes K_{pp} \otimes I_p) + (I_p \otimes K_{p^2,p})\}\Psi_{20}^{<2>} + O(\frac{1}{n}).
\end{aligned}$$

Premultiplication with $N_{p^{<4>}}$ results in $\mathrm{E}[\mathbf{z}_n^{<4>}] = 3N_{p^{<4>}}\Psi_{20}^{<2>}$. More generally, we can show

$$\mathrm{E}[\mathbf{z}_n^{<2k>}] = \frac{(2k)!}{2^k k!} N_{p^{<k>}} \Psi_{20}^{<k>} + O(\frac{1}{n})$$

[cf. Holmquist (1988)].

The last formula is a bit complicated, but can be shown on the same lines. The following fifteen expectations are divided into two groups, first six and last nine. Each expectation of the first group has a Kronecker product of permutations of $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ in each of the two parentheses involved, so that the two parentheses have $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ in common; in the second group the two parentheses in each expectation have only one variate $\mathbf{x}_3$ in common. The commutation matrix is very useful in showing

$$\begin{aligned}
\mathrm{E}[\mathbf{z}_n^{<3>}\mathbf{z}_n^{<3>\prime}] =& \mathrm{E}[(\mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3)(\mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3)'] + \mathrm{E}[(\mathbf{x}_1 \otimes \mathbf{x}_3 \otimes \mathbf{x}_2)(\mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3)'] \\
&+ \mathrm{E}[(\mathbf{x}_2 \otimes \mathbf{x}_1 \otimes \mathbf{x}_3)(\mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3)'] + \mathrm{E}[(\mathbf{x}_2 \otimes \mathbf{x}_3 \otimes \mathbf{x}_1)(\mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3)'] \\
&+ \mathrm{E}[(\mathbf{x}_3 \otimes \mathbf{x}_1 \otimes \mathbf{x}_2)(\mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3)'] + \mathrm{E}[(\mathbf{x}_3 \otimes \mathbf{x}_2 \otimes \mathbf{x}_1)(\mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3)'] \\
(5) \\
&+ \mathrm{E}[(\mathbf{x}_1 \otimes \mathbf{x}_1 \otimes \mathbf{x}_3)(\mathbf{x}_3 \otimes \mathbf{x}_2 \otimes \mathbf{x}_2)'] + \mathrm{E}[(\mathbf{x}_1 \otimes \mathbf{x}_1 \otimes \mathbf{x}_3)(\mathbf{x}_2 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3)'] \\
&+ \mathrm{E}[(\mathbf{x}_1 \otimes \mathbf{x}_1 \otimes \mathbf{x}_3)(\mathbf{x}_2 \otimes \mathbf{x}_3 \otimes \mathbf{x}_2)'] \\
&+ \mathrm{E}[(\mathbf{x}_1 \otimes \mathbf{x}_3 \otimes \mathbf{x}_1)(\mathbf{x}_3 \otimes \mathbf{x}_2 \otimes \mathbf{x}_2)'] + \mathrm{E}[(\mathbf{x}_1 \otimes \mathbf{x}_3 \otimes \mathbf{x}_1)(\mathbf{x}_2 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3)'] \\
&+ \mathrm{E}[(\mathbf{x}_1 \otimes \mathbf{x}_3 \otimes \mathbf{x}_1)(\mathbf{x}_2 \otimes \mathbf{x}_3 \otimes \mathbf{x}_2)'] \\
&+ \mathrm{E}[(\mathbf{x}_3 \otimes \mathbf{x}_1 \otimes \mathbf{x}_1)(\mathbf{x}_3 \otimes \mathbf{x}_2 \otimes \mathbf{x}_2)'] + \mathrm{E}[(\mathbf{x}_3 \otimes \mathbf{x}_1 \otimes \mathbf{x}_1)(\mathbf{x}_2 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3)'] \\
&+ \mathrm{E}[(\mathbf{x}_3 \otimes \mathbf{x}_1 \otimes \mathbf{x}_1)(\mathbf{x}_2 \otimes \mathbf{x}_3 \otimes \mathbf{x}_2)'] + O(\frac{1}{n}) \\
=& \left\{I_{p^3} + (I_p \otimes K_{pp}) + (K_{pp} \otimes I_p) + K_{p^2,p} + K_{p,p^2} + K_{p,p^2}(I_p \otimes K_{pp})\right\} \\
&\quad \cdot \mathrm{E}[(\mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3)(\mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3)'] \\
&+ (I_{p^3} + K_{p,p^2} + K_{p^2,p})\mathrm{E}[(\mathbf{x}_1 \otimes \mathbf{x}_1 \otimes \mathbf{x}_3)(\mathbf{x}_3 \otimes \mathbf{x}_2 \otimes \mathbf{x}_2)'] \\
&\quad \cdot (I_{p^3} + K_{p,p^2} + K_{p^2,p}) + O(\frac{1}{n}).
\end{aligned}$$

Pre- and post-multiplying both sides by $N_{p^{<3>}}$ and using

$$\begin{aligned}
&(\mathbf{x}_1 \otimes \mathbf{x}_1 \otimes \mathbf{x}_3)(\mathbf{x}_3 \otimes \mathbf{x}_2 \otimes \mathbf{x}_2)' \\
&\qquad = (\mathbf{x}_1 \otimes \mathbf{x}_1 \otimes \mathbf{x}_3 \otimes 1 \otimes 1)(1 \otimes 1 \otimes \mathbf{x}_3 \otimes \mathbf{x}_2 \otimes \mathbf{x}_2)' \\
&\qquad = (\mathbf{x}_1 \otimes \mathbf{x}_1) \otimes (\mathbf{x}_3\mathbf{x}_3') \otimes (\mathbf{x}_2 \otimes \mathbf{x}_2)',
\end{aligned}$$

we establish

$$\begin{aligned}
\mathrm{E}[\mathbf{z}_n^{<3>}\mathbf{z}_n^{<3>\prime}] =& 6N_{p^{<3>}}\Psi_{11}^{<3>}N_{p^{<3>}} \\
&+ 9N_{p^{<3>}}(\Psi_{20} \otimes \Psi_{11} \otimes \Psi_{02})N_{p^{<3>}} + O(\frac{1}{n}).
\end{aligned}$$

q.e.d.

The claim (2.2) can be proved in a similar manner. In fact,

$$\begin{aligned}\Psi_{02}\otimes\Psi_{11}\otimes\Psi_{20} =&\Psi_{02}\otimes\{K_{p,p^2}(\Psi_{20}\otimes\Psi_{11})\}\\ =&(1\otimes K_{p,p^2})(\Psi_{02}\otimes\Psi_{20}\otimes\Psi_{11})\\ =&K_{p,p^2}\{\Psi_{02}\otimes(\Psi_{20}\otimes\Psi_{11})\}\\ =&K_{p,p^2}\{(\Psi_{20}\otimes\Psi_{11})\otimes\Psi_{02}\}K_{p,p^2}.\end{aligned}$$

q.e.d.

Results in Section 2.4.

We have assumed that the exponential family is defined on any θ in an open subset of $\mathbf{R}^p$. The moment generating function exists and can be expressed as

$$\mathrm{E}[e^{\mathbf{t}'\mathbf{x}}] = \int e^{\mathbf{t}'x}e^{\theta'x-\psi(\theta)}\mu(\mathrm{d}x) = e^{\psi(\mathbf{t}+\theta)-\psi(\theta)},$$

and hence $\mathrm{E}[\mathbf{x}] = \Psi_{10}(\theta)$. Since

$$\mathrm{E}[e^{\mathbf{t}'(\mathbf{x}-\Psi_{10}(\theta))}] = e^{\psi(\mathbf{t}+\theta)-\psi(\theta)-\mathbf{t}'\Psi_{10}(\theta)},$$

the cumulant generating function $\kappa(\mathbf{t})$ about mean is of the form:

$$\kappa(\mathbf{t}) = \psi(\mathbf{t}+\theta) - \psi(\theta) - \mathbf{t}'\Psi_{10}(\theta).$$

Thus, we have

$$\frac{\mathrm{d}}{\mathrm{d}\mathbf{t}}\kappa(\mathbf{t}) = \Psi_{10}(\mathbf{t}+\theta) - \Psi_{10}(\theta),$$

and hence

$$\left(\frac{\mathrm{d}}{\mathrm{d}\mathbf{t}}\right)^{<k>}\left(\frac{\mathrm{d}}{\mathrm{d}\mathbf{t}'}\right)^{<\ell>}\kappa(\mathbf{t})\Big|_{\mathbf{t}=0} = \Psi_{k\ell}(\theta) \qquad \text{for} \quad k+\ell\geq 2.$$

This shows

$$\mathrm{E}[(\mathbf{x}_1-\Psi_{10}(\theta))^{<k>}(\mathbf{x}_1-\Psi_{10}(\theta))^{<\ell>\prime}] = \Psi_{k\ell}(\theta) \qquad \text{for} \quad k+\ell = 2,\,3.$$

q.e.d.

We have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}u'}\log L(u|\bar{\mathbf{x}}) =&\frac{\mathrm{d}}{\mathrm{d}u'}\left[n\{\theta(u)'\bar{\mathbf{x}}-\psi(\theta(u))\}\right] = \frac{\mathrm{d}}{\mathrm{d}u'}\left[n\{\bar{\mathbf{x}}'\theta(u)-\psi(\theta(u))\}\right]\\ =&n(\bar{\mathbf{x}}-\Psi_{10}(\theta(u))'\Theta_1(u).\end{aligned}$$

Recall that $\mathrm{E}[\mathbf{x}] = \Psi_{10}(\theta(u)) = \eta(u)$. We have

$$\frac{\mathrm{d}}{\mathrm{d}u}\log L(u|\bar{\mathbf{x}}) = n\Theta_1(u)'\left\{\bar{\mathbf{x}}-\eta(u)\right\}.$$

q.e.d.

The Fisher information matrix is expressed as

$$\begin{aligned} i_u =& \frac{1}{n}\mathrm{E}\left[\left\{\frac{\mathrm{d}}{\mathrm{d}u}\log L(u|\bar{\mathbf{x}})\right\}\left\{\frac{\mathrm{d}}{\mathrm{d}u}\log L(u|\bar{\mathbf{x}})\right\}'\right] \\ =& n\Theta_1(u)'\mathrm{E}[(\bar{\mathbf{x}}-\eta(u))(\bar{\mathbf{x}}-\eta(u))']\Theta_1(u) \\ =& \Theta_1(u)'\Psi_{11}(\theta(u))\Theta_1(u). \end{aligned}$$

q.e.d.

Suppose that the MLE $\hat{u}(\bar{\mathbf{x}})$ satisfies

$$\Theta_1(\hat{u}(\bar{\mathbf{x}}))'\left\{\bar{\mathbf{x}}-\Psi_{10}(\theta(\hat{u}(\bar{\mathbf{x}})))\right\}=0.$$

Then, differentiation by $\bar{\mathbf{x}}'$ and evaluation at $\bar{\mathbf{x}}=\Psi_{10}(\theta(u))$ lead to

$$\Theta_1(u)'\left\{I_p-\Psi_{11}(\theta(u))\Theta_1(u)\left.\frac{\mathrm{d}\hat{u}(\bar{\mathbf{x}})}{\mathrm{d}\bar{\mathbf{x}}'}\right|_{\bar{\mathbf{x}}=\Psi_{10}}\right\}=0$$

because $\bar{\mathbf{x}}-\Psi_{10}(\theta(\hat{u}(\bar{\mathbf{x}})))|_{\bar{\mathbf{x}}=\Psi_{10}}=0$. The equation leads to

$$\left.\frac{\mathrm{d}\hat{u}(\bar{\mathbf{x}})}{\mathrm{d}\bar{\mathbf{x}}'}\right|_{\bar{\mathbf{x}}=\Psi_{10}}=i_u^{-1}\Theta_1(u)',$$

provided that i_u is positive definite. *q.e.d.*

Remark 4.2.

It is obvious to show $\mathrm{E}[(\hat{u}(\bar{\mathbf{x}})-u)(\hat{u}(\bar{\mathbf{x}})-u)']=\frac{1}{n}\bar{G}_1\Psi_{11}\bar{G}_1'+o\left(\frac{1}{n}\right)$. Applying Lemma 4.3, we have

$$\begin{aligned} &\mathrm{E}\left[\left(\hat{u}^*(\bar{\mathbf{x}})-u-\frac{1}{\sqrt{n}}\bar{G}_1\mathbf{z}_n\right)\left(\frac{1}{\sqrt{n}}\bar{G}_1\mathbf{z}_n\right)'\right] \\ &\quad=\frac{1}{\sqrt{n}}\mathrm{E}[(\hat{u}^*(\bar{\mathbf{x}})-u)\mathbf{z}_n'\Theta_1]i_u^{-1}-\frac{1}{n}\bar{G}_1\Psi_{11}\bar{G}_1' \\ &\quad=\frac{1}{n}\left\{\frac{\mathrm{d}}{\mathrm{d}u'}\mathrm{E}[\hat{u}^*(\bar{\mathbf{x}})-u]-\mathrm{E}\left[\frac{\mathrm{d}}{\mathrm{d}u'}(\hat{u}^*(\bar{\mathbf{x}})-u)\right]\right\}i_u^{-1}-i_u^{-1} \\ &\quad=\frac{1}{n}\frac{\mathrm{d}}{\mathrm{d}u'}\mathrm{E}[\hat{u}^*(\bar{\mathbf{x}})-u]i_u^{-1}. \end{aligned}$$

Therefore we obtain an explicit form of the quadratic loss as

$$\begin{aligned} &\mathrm{E}[(\hat{u}^*(\bar{\mathbf{x}})-u)(\hat{u}^*(\bar{\mathbf{x}})-u)'] \\ &\quad=\mathrm{E}\left[\left(\hat{u}^*(\bar{\mathbf{x}})-u-\frac{1}{\sqrt{n}}\bar{G}_1\mathbf{z}_n+\frac{1}{\sqrt{n}}\bar{G}_1\mathbf{z}_n\right)\left(\hat{u}^*(\bar{\mathbf{x}})-u-\frac{1}{\sqrt{n}}\bar{G}_1\mathbf{z}_n+\frac{1}{\sqrt{n}}\bar{G}_1\mathbf{z}_n\right)'\right] \\ &\quad=\mathrm{E}\left[\left(\hat{u}^*(\bar{\mathbf{x}})-u-\frac{1}{\sqrt{n}}\bar{G}_1\mathbf{z}_n\right)\left(\hat{u}^*(\bar{\mathbf{x}})-u-\frac{1}{\sqrt{n}}\bar{G}_1\mathbf{z}_n\right)'\right]+\frac{1}{n}\bar{G}_1\Psi_{11}\bar{G}_1' \\ &\qquad+\frac{1}{n}\left\{\frac{\mathrm{d}}{\mathrm{d}u'}\mathrm{E}[\hat{u}^*(\bar{\mathbf{x}})-u]i_u^{-1}+i_u^{-1}\frac{\mathrm{d}}{\mathrm{d}u}\mathrm{E}[\hat{u}^*(\bar{\mathbf{x}})-u]'\right\}. \end{aligned}$$

As a result of the bias correction, we know that $\mathrm{E}[\hat{u}^*(\bar{\mathbf{x}}) - u] = o\left(\frac{1}{n}\right)$, and from (4.8) we know that

$$\hat{u}^*(\bar{\mathbf{x}}) - u - \frac{1}{\sqrt{n}}\bar{G}_1\mathbf{z}_n = \frac{1}{2n}\bar{G}_2(\mathbf{z}_n^{<2>} - \Psi_{20}) + o_p\left(\frac{1}{n}\right).$$

From these facts, we establish

$$\mathrm{E}[(\hat{u}^*(\bar{\mathbf{x}}) - u)(\hat{u}^*(\bar{\mathbf{x}}) - u)'] = \frac{1}{n}\bar{G}_1\Psi_{11}\bar{G}_1' + \frac{1}{2n^2}\bar{G}_2\Psi_{11}^{<2>}\bar{G}_2' + o\left(\frac{1}{n^2}\right).$$

q.e.d.

We additionally notice the following. In the same manner we can show

$$\mathrm{E}[(g^*(\bar{\mathbf{x}}) - u)(g^*(\bar{\mathbf{x}}) - u)'] = \frac{1}{n}\bar{G}_1\Psi_{11}\bar{G}_1' + \frac{1}{2n^2}G_2\Psi_{11}^{<2>}G_2' + o\left(\frac{1}{n^2}\right)$$

for any first-order efficient estimator $g^*(\bar{\mathbf{x}})$ after bias correction.

References

Holmquist, B. (1988). Moments and cumulants of the multivariate normal distribution. *Stoch. Anal. App.* **6** 273-278.

Kano, Y. (1993). Beyond third-order efficiency. (submitted).

Magnus, J. R. and Neudecker, H. (1988). *Matrix Differential Calculus with Applications in Statistics and Econometrics.* John Wiley, Chichester.

Magnus, J. R. and Wansbeek, T. (1983). Some results on commutation matrices, with statistical applications. *Canadian J. Statist.* **11** 221-231.

(Technical Report, University of Osaka Prefecture, Nov. 1993)

Curvatures, and relationship with Efron's result

Here we study how the expression developed in Kano [(1993), formula (4.4b)] relates to Efron's result [(1975), formula (10.1)]. Efron introduces two geometrical curvatures, namely the statistical curvature γ_u and the naming curvature Γ_u, to evaluate the quadratic loss of Fisher consistent estimators in a curved exponential family with a uni-dimensional structural parameter u [$\dim(u) = q = 1$ in our notation]. We also give a *matrix-form* representation [see (15) and (16)] for each of the two curvatures in the *multi*-dimensional case [cf. Amari (1985)].

Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ be an i.i.d. sample drawn from $f = \exp\{\theta(u)'x - \psi(\theta(u))\}$, where u is a q-vector of structural parameters. We write $\ell(u) = \log f = \theta(u)'\mathbf{x} - \psi(\theta(u))$. Recall that $g(\bar{\mathbf{x}})$ is an arbitrary differentiable Fisher consistent estimator and that i_u is the Fisher information matrix. Efron [(1975), formula (10.1)] shows that under $q = 1$, if $g(\bar{\mathbf{x}})$ is first-order efficient,

$$\begin{aligned}
\mathrm{Var}[g(\bar{\mathbf{x}})] &= \mathrm{E}[(g(\bar{\mathbf{x}}) - u)^2] - \mathrm{E}^2[g(\bar{\mathbf{x}}) - u] \\
&= \frac{1}{ni_u} + \frac{1}{n^2 i_u}\left\{\gamma_u^2 + 4\frac{\Gamma_u^2}{i_u} + \Delta_u^{g(\bar{\mathbf{x}})}\right\} + \frac{2}{ni_u}\frac{\mathrm{d}}{\mathrm{d}u}\mathrm{E}[g(\bar{\mathbf{x}}) - u] + o\left(\frac{1}{n^2}\right) \\
&= \frac{1}{ni_u} + \frac{1}{n^2}\left\{\widetilde{\gamma}_u^2 + \widetilde{\Gamma}_u^2 + \Delta_u^{g(\bar{\mathbf{x}})}\right\} + \frac{2}{ni_u}\frac{\mathrm{d}}{\mathrm{d}u}\mathrm{E}[g(\bar{\mathbf{x}}) - u] + o\left(\frac{1}{n^2}\right), \qquad (6)
\end{aligned}$$

where $\widetilde{\gamma}_u^2 = \gamma_u^2/i_u$ and $\widetilde{\Gamma}_u^2 = 4\Gamma_u^2/i_u^2$, and γ_u^2 and Γ_u^2 are (the square of) the statistical curvature and the naming curvature, respectively.

Our expression for the quadratic loss, expanded up to order n^{-2}, of a bias-corrected first-order efficient estimator $g^*(\bar{\mathbf{x}})$ is

$$\mathrm{E}[(g^*(\bar{\mathbf{x}}) - u)(g^*(\bar{\mathbf{x}}) - u)'] = \frac{1}{n}\bar{G}_1\Psi_{11}\bar{G}_1' + \frac{1}{2n^2}G_2\Psi_{11}^{<2>}G_2' + o\left(\frac{1}{n^2}\right)$$

[see (4.4b) and Rem.4.2. in Kano (1993)]. In a similar manner, we can show for a first-order efficient $g(\bar{\mathbf{x}})$ *without* bias correction,

$$\begin{aligned}
\mathrm{E}[(g(\bar{\mathbf{x}}) - u)(g(\bar{\mathbf{x}}) - u)'] = &\frac{1}{n}\bar{G}_1\Psi_{11}\bar{G}_1' + \frac{1}{4n^2}G_2(2\Psi_{11}^{<2>} + \Psi_{20}\Psi_{02})G_2' \\
&+ \frac{1}{n^2}\left\{\frac{\mathrm{d}}{\mathrm{d}u'}\mathrm{E}[g(\bar{\mathbf{x}}) - u]i_u^{-1} + i_u^{-1}\frac{\mathrm{d}}{\mathrm{d}u}\mathrm{E}[g(\bar{\mathbf{x}}) - u]'\right\} + o\left(\frac{1}{n^2}\right),
\end{aligned}$$

or equivalently,

$$\begin{aligned}
&\mathrm{E}[(g(\bar{\mathbf{x}})-u)(g(\bar{\mathbf{x}})-u)'] - \left(\frac{1}{2n}G_2\Psi_{20}\right)\left(\frac{1}{2n}G_2\Psi_{20}\right)' \\
&\quad = \frac{1}{n}i_u^{-1} + \frac{1}{2n^2}G_2\Psi_{11}^{<2>}G_2' \\
&\qquad + \frac{1}{n^2}\left\{\frac{\mathrm{d}}{\mathrm{d}u'}\mathrm{E}[g(\bar{\mathbf{x}})-u]i_u^{-1} + i_u^{-1}\frac{\mathrm{d}}{\mathrm{d}u}\mathrm{E}[g(\bar{\mathbf{x}})-u]'\right\} + o\left(\tfrac{1}{n^2}\right) \\
&\quad = \frac{1}{n}i_u^{-1} + \frac{1}{n^2}\left\{\frac{1}{2}\bar{G}_2\Psi_{11}^{<2>}\bar{G}_2' + \frac{1}{2}(G_2-\bar{G}_2)\Psi_{11}^{<2>}(G_2-\bar{G}_2)'\right\} \\
&\qquad + \frac{1}{n^2}\left\{\frac{\mathrm{d}}{\mathrm{d}u'}\mathrm{E}[g(\bar{\mathbf{x}})-u]i_u^{-1} + i_u^{-1}\frac{\mathrm{d}}{\mathrm{d}u}\mathrm{E}[g(\bar{\mathbf{x}})-u]'\right\} + o\left(\tfrac{1}{n^2}\right) \qquad (7)
\end{aligned}$$

because $\bar{G}_1\Psi_{11}\bar{G}_1' = i_u^{-1}$ and $(G_2-\bar{G}_2)\Psi_{11}^{<2>}\bar{G}_2' = 0$ as in the proof of Theorem 3.1 in Kano (1993). Since $\mathrm{E}[g(\bar{\mathbf{x}})-u] = \frac{1}{2n}G_2\Psi_{20} + o\left(\frac{1}{n}\right)$, comparison between (6) and (7) gives

$$\frac{1}{2}\bar{G}_2\Psi_{11}^{<2>}\bar{G}_2' + \frac{1}{2}(G_2-\bar{G}_2)\Psi_{11}^{<2>}(G_2-\bar{G}_2)' = \widetilde{\gamma}_u^2 + \widetilde{\Gamma}_u^2 + \Delta_u^{g(\bar{\mathbf{x}})}$$

for $q=1$. In particular, we will show

$$\frac{1}{2}\bar{G}_2\Psi_{11}^{<2>}\bar{G}_2' = \widetilde{\gamma}_u^2 + \widetilde{\Gamma}_u^2 \qquad (8)$$

for a general $q \in \mathbf{N}$.

First, we shall generalize $\widetilde{\gamma}_u^2$ and $\widetilde{\Gamma}_u^2$ to those for the multi-parameter case. Before this, we describe notation. Recall that $\Psi_{k\ell} = \left(\frac{\mathrm{d}}{\mathrm{d}\theta}\right)^{<k>}\left(\frac{\mathrm{d}}{\mathrm{d}\theta'}\right)^{<\ell>}\psi(\theta)$ and we define

$$\Theta_{k\ell} = \left(\frac{\mathrm{d}}{\mathrm{d}u}\right)^{<k>}\left(\frac{\mathrm{d}}{\mathrm{d}u'}\right)^{<\ell>}\theta(u).$$

The old notation Θ_ℓ equals $\Theta_{0\ell}$. The Fisher information matrix is $i_u = \Theta_{01}'\Psi_{11}\Theta_{01}$. Using the formula $\left[\left(\frac{\mathrm{d}}{\mathrm{d}\mathbf{x}}\right)^{<k>}A(\mathbf{x})\right]' = \left(\frac{\mathrm{d}}{\mathrm{d}\mathbf{x}'}\right)^{<k>}A(\mathbf{x})'$ [see page 4 in this manuscript)], we have

$$\frac{\mathrm{d}}{\mathrm{d}u'}\Theta_{01}' = \left(\frac{\mathrm{d}}{\mathrm{d}u}\Theta_{01}\right)' = \Theta_{11}', \quad \frac{\mathrm{d}}{\mathrm{d}u'}\Psi_{10}' = \left(\frac{\mathrm{d}}{\mathrm{d}u}\Psi_{10}\right)' = \Psi_{20}'. \qquad (9)$$

For the univariate case [i.e. $q=1$],

$$\Theta_{01} = \Theta_{10} = \frac{\partial\theta(u)}{\partial u}, \quad \Theta_{02} = \Theta_{11} = \Theta_{20} = \frac{\partial^2\theta(u)}{\partial u^2}. \qquad (10)$$

The statistical curvature $\widetilde{\gamma}_u^2$, divided by i_u, of Efron [(1975), formula (5.3) and (5.4)] is of the form:

$$\begin{aligned}
\widetilde{\gamma}_u^2 = \frac{\gamma_u^2}{i_u} &= \frac{1}{i_u}\left(\frac{\mathrm{E}\left[\left(\frac{\mathrm{d}^2\ell(u)}{\mathrm{d}u^2}\right)^2\right] - i_u^2}{i_u^2} - \frac{\mathrm{E}^2\left[\frac{\mathrm{d}\ell(u)}{\mathrm{d}u}\frac{\mathrm{d}^2\ell(u)}{\mathrm{d}u^2}\right]}{i_u^3}\right) \\
&= i_u^{-1}\mathrm{E}\left[\left(\frac{\mathrm{d}^2\ell(u)}{\mathrm{d}u^2}+i_u\right)i_u^{-1}\left(\frac{\mathrm{d}^2\ell(u)}{\mathrm{d}u^2}+i_u\right)\right]i_u^{-1} \\
&\qquad - i_u^{-1}\mathrm{E}\left[\frac{\mathrm{d}^2\ell(u)}{\mathrm{d}u^2}\left(i_u^{-1}\frac{\mathrm{d}\ell(u)}{\mathrm{d}u}\right)\right](i_u i_u^{-1})\,\mathrm{E}\left[\frac{\mathrm{d}^2\ell(u)}{\mathrm{d}u^2}\left(i_u^{-1}\frac{\mathrm{d}\ell(u)}{\mathrm{d}u}\right)\right]i_u^{-1}
\end{aligned}$$

because $\mathrm{E}[\frac{\mathrm{d}^2\ell(u)}{\mathrm{d}u^2}] = -i_u$. Thus, we define the statistical curvature $\widetilde{\gamma}_u^2$ for the multi-dimensional case as

$$\widetilde{\gamma}_u^2 = i_u^{-1}\mathrm{E}\left[\left(\frac{\mathrm{d}^2\ell(u)}{\mathrm{d}u\mathrm{d}u'} + i_u\right) i_u^{-1} \left(\frac{\mathrm{d}^2\ell(u)}{\mathrm{d}u\mathrm{d}u'} + i_u\right)'\right] i_u^{-1}$$

(11)

$$- i_u^{-1}\mathrm{E}\left[\frac{\mathrm{d}^2\ell(u)}{\mathrm{d}u\mathrm{d}u'}\left(i_u^{-1}\otimes\frac{\mathrm{d}\ell(u)}{\mathrm{d}u'}\right)\right](i_u\otimes i_u^{-1})\,\mathrm{E}\left[\frac{\mathrm{d}^2\ell(u)}{\mathrm{d}u\mathrm{d}u'}\left(i_u^{-1}\otimes\frac{\mathrm{d}\ell(u)}{\mathrm{d}u'}\right)\right]' i_u^{-1}.$$

Note that $\widetilde{\gamma}_u^2$ is the minimum value of the following minimization problem:

$$\min_{A:\,q\times q,\ B:\,q\times q^2} i_u^{-1}\mathrm{E}\left[\left\{\frac{\mathrm{d}^2\ell(u)}{\mathrm{d}u\mathrm{d}u'} - A - B\left(I_q\otimes\frac{\mathrm{d}\ell(u)}{\mathrm{d}u}\right)\right\} i_u^{-1} \cdot\left\{\frac{\mathrm{d}^2\ell(u)}{\mathrm{d}u\mathrm{d}u'} - A - B\left(I_q\otimes\frac{\mathrm{d}\ell(u)}{\mathrm{d}u}\right)\right\}'\right] i_u^{-1}.$$

The fact that $\widetilde{\gamma}_u^2$ is a regression residual was noted by Reeds [(1975), page 1237].

The naming curvature $\widetilde{\Gamma}_u^2$, multiplied by $4/i_u^2$, of Efron [(1975), formula (10.23) and the discussion just after it] is of the form:

$$\begin{aligned}\widetilde{\Gamma}_u^2 =&\Gamma_u^2\times\frac{4}{i_u^2} = \frac{\mathrm{E}^2\left[\left(\frac{\mathrm{d}f}{\mathrm{d}u}\cdot\frac{1}{f}\right)\left(\frac{\mathrm{d}^2f}{\mathrm{d}u^2}\cdot\frac{1}{f}\right)\right]}{8i_u^2}\times\frac{4}{i_u^2}\\ =&\mathrm{E}^2\left[\left(\frac{\mathrm{d}f}{\mathrm{d}u}\cdot\frac{1}{f}\right)\left(\frac{\mathrm{d}^2f}{\mathrm{d}u^2}\cdot\frac{1}{f}\right)\right]/2i_u^4\\ =&\frac{1}{2}i_u^{-1}\mathrm{E}\left[\left(\frac{\mathrm{d}f}{\mathrm{d}u}\cdot\frac{1}{f}\right)\left(\frac{\mathrm{d}^2f}{\mathrm{d}u^2}\cdot\frac{1}{f}\right)\right] i_u^{-1} i_u^{-1}\mathrm{E}\left[\left(\frac{\mathrm{d}f}{\mathrm{d}u}\cdot\frac{1}{f}\right)\left(\frac{\mathrm{d}^2f}{\mathrm{d}u^2}\cdot\frac{1}{f}\right)\right] i_u^{-1}.\end{aligned}$$

Thus, we define the naming curvature $\widetilde{\Gamma}_u^2$ for the multi-parameter case as

(12)

$$\widetilde{\Gamma}_u^2 = \frac{1}{2}i_u^{-1}\mathrm{E}\left[\left(\frac{\mathrm{d}f}{\mathrm{d}u}\cdot\frac{1}{f}\right)\left(\frac{\mathrm{d}^{<2>}f}{\mathrm{d}u^{<2>}}\cdot\frac{1}{f}\right)'\right](i_u^{-1}\otimes i_u^{-1})\,\mathrm{E}\left[\left(\frac{\mathrm{d}f}{\mathrm{d}u}\cdot\frac{1}{f}\right)\left(\frac{\mathrm{d}^{<2>}f}{\mathrm{d}u^{<2>}}\cdot\frac{1}{f}\right)'\right]' i_u^{-1},$$

where $\left(\frac{\mathrm{d}^{<2>}f}{\mathrm{d}u^{<2>}}\right)' = \left(\frac{\mathrm{d}}{\mathrm{d}u'}\right)^{<2>} f$.

Apparently, these definitions (11) and (12) of the curvatures for the multi-parameter case reduces to those of Efron when $\dim(u) = 1$.

Let us derive the expressions of $\widetilde{\gamma}_u^2$ and $\widetilde{\Gamma}_u^2$ in terms of $\Theta_{k\ell}$ and $\Psi_{k\ell}$. Since $\ell(u) = \theta(u)'\mathbf{x} - \psi(\theta(u))$, we get

$$\frac{\mathrm{d}\ell(u)}{\mathrm{d}u} = \Theta_{01}'(\mathbf{x}-\Psi_{10}), \tag{13a}$$

$$\frac{\mathrm{d}^2\ell(u)}{\mathrm{d}u\mathrm{d}u'} = \frac{\mathrm{d}}{\mathrm{d}u'}\left[\Theta_{01}'(\mathbf{x}-\Psi_{10})\right] = \Theta_{11}'(I_q\otimes(\mathbf{x}-\Psi_{10})) - i_u, \tag{13b}$$

(13c)

$$\left(\frac{\mathrm{d}}{\mathrm{d}u'}\right)^{<2>}\ell(u) = \frac{\mathrm{d}}{\mathrm{d}u'}\left[(\mathbf{x}-\Psi_{10})'\Theta_{01}\right] = (\mathbf{x}-\Psi_{10})'\Theta_{02} - \Psi_{20}'(\Theta_{01}\otimes\Theta_{01})$$

in view of (9) and the rules for differentiations in Section 2 in Kano (1993). Note that

(13b) is symmetric. Using (13), we calculate

$$\begin{aligned}
&\mathrm{E}\left[\left(\frac{\mathrm{d}^2\ell(u)}{\mathrm{d}u\mathrm{d}u'}+i_u\right)i_u^{-1}\left(\frac{\mathrm{d}^2\ell(u)}{\mathrm{d}u\mathrm{d}u'}+i_u\right)'\right]\\
&\qquad=\mathrm{E}\left[\{\Theta_{11}'(I_q\otimes(\mathbf{x}-\Psi_{10}))\}\,i_u^{-1}\,\{\Theta_{11}'(I_q\otimes(\mathbf{x}-\Psi_{10}))\}'\right]\\
&\qquad=\Theta_{11}'(i_u^{-1}\otimes\Psi_{11})\Theta_{11}, \qquad\text{(14a)}
\end{aligned}$$

$$\begin{aligned}
\mathrm{E}\left[\frac{\mathrm{d}^2\ell(u)}{\mathrm{d}u\mathrm{d}u'}\left(i_u^{-1}\otimes\frac{\mathrm{d}\ell(u)}{\mathrm{d}u'}\right)\right]&=\mathrm{E}\left[\{\Theta_{11}'(I_q\otimes(\mathbf{x}-\Psi_{10}))-i_u\}\left\{i_u^{-1}\otimes((\mathbf{x}-\Psi_{10})'\Theta_{01})\right\}\right]\\
&=\Theta_{11}'(i_u^{-1}\otimes\Psi_{11}\Theta_{01}). \qquad\text{(14b)}
\end{aligned}$$

Here we have used $\mathrm{E}[\mathbf{x}-\Psi_{10}]=0$. Substitution of (14) into (11) gives an explicit form of $\widetilde{\gamma}_u^2$ as

$$\begin{aligned}
\widetilde{\gamma}_u^2=&i_u^{-1}\Theta_{11}'(i_u^{-1}\otimes\Psi_{11})\Theta_{11}i_u^{-1}\\
&-i_u^{-1}\Theta_{11}'(i_u^{-1}\otimes\Psi_{11}\Theta_{01})(i_u\otimes i_u^{-1})(i_u^{-1}\otimes\Theta_{01}'\Psi_{11})\Theta_{11}i_u^{-1}\\
=&i_u^{-1}\Theta_{11}'\left\{i_u^{-1}\otimes(\Psi_{11}-\Psi_{11}\Theta_{01}i_u^{-1}\Theta_{01}'\Psi_{11})\right\}\Theta_{11}i_u^{-1}. \qquad\text{(15)}
\end{aligned}$$

Since

$$\begin{aligned}
\left(\frac{\mathrm{d}}{\mathrm{d}u'}\right)^{<2>}\ell(u)=&\frac{\mathrm{d}}{\mathrm{d}u'}\left[\frac{1}{f}\cdot\frac{\mathrm{d}f}{\mathrm{d}u'}\right]=-\frac{1}{f^2}\left(\frac{\mathrm{d}f}{\mathrm{d}u'}\right)^{<2>}+\frac{1}{f}\left(\frac{\mathrm{d}}{\mathrm{d}u'}\right)^{<2>}f\\
&=-\left(\frac{\mathrm{d}\ell(u)}{\mathrm{d}u'}\right)^{<2>}+\frac{1}{f}\left(\frac{\mathrm{d}}{\mathrm{d}u'}\right)^{<2>}f,
\end{aligned}$$

we have from (13) again

$$\begin{aligned}
\mathrm{E}\left[\left(\frac{\mathrm{d}f}{\mathrm{d}u}\cdot\frac{1}{f}\right)\left(\frac{\mathrm{d}^{<2>}f}{\mathrm{d}u^{<2>}}\cdot\frac{1}{f}\right)'\right]&=\mathrm{E}\left[\frac{\mathrm{d}\ell(u)}{\mathrm{d}u}\left\{\left(\frac{\mathrm{d}}{\mathrm{d}u'}\right)^{<2>}\ell(u)+\left(\frac{\mathrm{d}\ell(u)}{\mathrm{d}u'}\right)^{<2>}\right\}\right]\\
&=\mathrm{E}\left[\Theta_{01}'(\mathbf{x}-\Psi_{10})\left\{(\mathbf{x}-\Psi_{10})'\Theta_{02}-\Psi_{20}'(\Theta_{01}\otimes\Theta_{01})+((\mathbf{x}-\Psi_{10})'\Theta_{01})^{<2>}\right\}\right]\\
&=\Theta_{01}'\Psi_{11}\Theta_{02}+\Theta_{01}'\mathrm{E}\left[(\mathbf{x}-\Psi_{10})(\mathbf{x}-\Psi_{10})^{<2>\prime}\right]\Theta_{01}^{<2>}\\
&=\Theta_{01}'\Psi_{11}\Theta_{02}+\Theta_{01}'\Psi_{12}\Theta_{01}^{<2>}.
\end{aligned}$$

Thus, the naming curvature $\widetilde{\Gamma}_u^2$ in (12) can be expressed in the form:

$$\begin{aligned}
\widetilde{\Gamma}_u^2=&\frac{1}{2}i_u^{-1}\left(\Theta_{01}'\Psi_{11}\Theta_{02}+\Theta_{01}'\Psi_{12}\Theta_{01}^{<2>}\right)(i_u^{-1}\otimes i_u^{-1})\left(\Theta_{01}'\Psi_{11}\Theta_{02}+\Theta_{01}'\Psi_{12}\Theta_{01}^{<2>}\right)'i_u^{-1}\\
&\qquad\text{(16)}\\
=&\frac{1}{2}i_u^{-1}\Theta_{01}'\left(\Psi_{11}\Theta_{02}+\Psi_{12}\Theta_{01}^{<2>}\right)\left(i_u^{-1}\right)^{<2>}\left(\Psi_{11}\Theta_{02}+\Psi_{12}\Theta_{01}^{<2>}\right)'\Theta_{01}i_u^{-1}.
\end{aligned}$$

We are now in a position to show the equality (8). For the purpose, we need to get an explicit form of $\bar{G}_2$. Recall that

$$\Theta_{11}'(\bar{\mathbf{x}})\left\{\bar{G}_1(\bar{\mathbf{x}})\otimes(\bar{\mathbf{x}}-\Psi_{10}(\bar{\mathbf{x}}))\right\}+\Theta_{01}(\bar{\mathbf{x}})'\left\{I_p-\Psi_{11}(\bar{\mathbf{x}})\Theta_{01}(\bar{\mathbf{x}})\bar{G}_1(\mathbf{x})\right\}=0,$$

which appeared at the bottom on page 19 in Kano (1993). Here we have used (9). Oper-

ating $\frac{\mathrm{d}}{\mathrm{d}\bar{\mathbf{x}}'}$ and evaluating $\bar{\mathbf{x}} = \Psi_{10}$, the LHS above becomes

$$\begin{aligned}
&\Theta'_{11}\left(\bar{G}_1 \otimes (I_p - \Psi_{11}\Theta_{01}\bar{G}_1)\right) K_{pp} + \Theta'_{11}\left(\bar{G}_1 \otimes (I_p - \Psi_{11}\Theta_{01}\bar{G}_1)\right) \\
&\qquad - \Theta'_{01}\left[\Psi_{12}(\Theta_{01}\bar{G}_1 \otimes \Theta_{01}\bar{G}_1) + \Psi_{11}\Theta_{02}(\bar{G}_1 \otimes \bar{G}_1) + \Psi_{11}\Theta_{01}\bar{G}_2\right] \\
&= 2\Theta'_{11}\left(\bar{G}_1 \otimes (I_p - \Psi_{11}\Theta_{01}\bar{G}_1)\right) N_{p^{<2>}} \\
&\qquad - \Theta'_{01}\Psi_{12}(\Theta_{01}\bar{G}_1)^{<2>} - \Theta'_{01}\Psi_{11}\Theta_{02}\bar{G}_1^{<2>} - \Theta'_{01}\Psi_{11}\Theta_{01}\bar{G}_2
\end{aligned}$$

in view of Prop. 2.3. in Kano (1993). Thus, the explicit form of $\bar{G}_2$ is obtained as

$$\begin{aligned}
\bar{G}_2 = 2i_u^{-1}\Theta'_{11}\left(\bar{G}_1 \otimes (I_p - \Psi_{11}\Theta_{01}\bar{G}_1)\right) N_{p^{<2>}} \\
- i_u^{-1}\Theta'_{01}\left[\Psi_{12}\Theta_{01}^{<2>} + \Psi_{11}\Theta_{02}\right]\bar{G}_1^{<2>}.
\end{aligned} \tag{17}$$

Recall that $\bar{G}_1 = i_u^{-1}\Theta'_{01}$. Since

$$(I_p - \Psi_{11}\Theta_{01}\bar{G}_1)\Psi_{11}\bar{G}'_1 = 0, \tag{18}$$

we have

$$\begin{aligned}
&\left\{\bar{G}_1 \otimes (I_p - \Psi_{11}\Theta_{01}\bar{G}_1)\right\} N_{p^{<2>}} \Psi_{11}^{<2>}\bar{G}_1^{<2>\prime} \\
&\qquad = \left\{\bar{G}_1 \otimes (I_p - \Psi_{11}\Theta_{01}\bar{G}_1)\right\}\left(\Psi_{11}\bar{G}'_1\right)^{<2>} N_{q^{<2>}} = 0,
\end{aligned}$$

which, along with (17), implies that

$$\begin{aligned}
&\frac{1}{2}\bar{G}_2\Psi_{11}^{<2>}\bar{G}'_2 \\
&= 2i_u^{-1}\Theta'_{11}\left\{\bar{G}_1 \otimes (I_p - \Psi_{11}\Theta_{01}\bar{G}_1)\right\} N_{p^{<2>}}\Psi_{11}^{<2>}N_{p^{<2>}}\left\{\bar{G}_1 \otimes (I_p - \Psi_{11}\Theta_{01}\bar{G}_1)\right\}'\Theta_{11}i_u^{-1} \\
&\quad + \frac{1}{2}i_u^{-1}\Theta'_{01}\left\{\Psi_{12}\Theta_{01}^{<2>} + \Psi_{11}\Theta_{02}\right\}\bar{G}_1^{<2>}\Psi_{11}^{<2>}\bar{G}_1^{<2>\prime}\left\{\Psi_{12}\Theta_{01}^{<2>} + \Psi_{11}\Theta_{02}\right\}'\Theta_{01}i_u^{-1}.
\end{aligned}$$

The second term is obviously identical with $\widetilde{\Gamma}_u^2$ in (16) because $\bar{G}_1\Psi_{11}\bar{G}'_1 = i_u^{-1}$.

Finally we shall show the first term is equal to $\widetilde{\gamma}_u^2$ in (15). Note that

$$2N_{p^{<2>}}\Psi_{11}^{<2>}N_{p^{<2>}} = 2N_{p^{<2>}}\Psi_{11}^{<2>} = (I_p + K_{pp})\Psi_{11}^{<2>}$$

[see Prop. 2.2. and the first line on page 5 in Kano (1993)]. Using (18), we can evaluate the first term as

$$\begin{aligned}
&i_u^{-1}\Theta'_{11}\left[\bar{G}_1 \otimes (I_p - \Psi_{11}\Theta_{01}\bar{G}_1) + K_{qp}\left\{(I_p - \Psi_{11}\Theta_{01}\bar{G}_1) \otimes \bar{G}_1\right\}\right] \\
&\qquad \cdot \Psi_{11}^{<2>}\left\{\bar{G}_1 \otimes (I_p - \Psi_{11}\Theta_{01}\bar{G}_1)\right\}'\Theta_{11}i_u^{-1} \\
&\quad = i_u^{-1}\Theta'_{11}\left\{\bar{G}_1 \otimes (I_p - \Psi_{11}\Theta_{01}\bar{G}_1)\right\}\Psi_{11}^{<2>}\left\{\bar{G}_1 \otimes (I_p - \Psi_{11}\Theta_{01}\bar{G}_1)\right\}'\Theta_{11}i_u^{-1} \\
&\quad = i_u^{-1}\Theta'_{11}\left[\left(\bar{G}_1\Psi_{11}\bar{G}'_1\right) \otimes \left\{(I_p - \Psi_{11}\Theta_{01}\bar{G}_1)\Psi_{11}(I_p - \Psi_{11}\Theta_{01}\bar{G}_1)'\right\}\right]\Theta_{11}i_u^{-1} \\
&\quad = i_u^{-1}\Theta'_{11}\left\{i_u^{-1} \otimes (I_p - \Psi_{11}\Theta_{01}i_u^{-1}\Theta'_{01}\Psi_{11})\right\}\Theta_{11}i_u^{-1} \\
&\quad = \widetilde{\gamma}_u^2.
\end{aligned}$$

q.e.d.

Additional Reference

Reeds, J. (1975). Discussion of the Efron paper. *Ann. Statist.* **3** 1234-1238.

(Technical Report, University of Osaka Prefecture, Dec. 1993)