Economic Review (Otaru University of Commerce),
Vol. 67, No. 2 & 3, 9-44, December, 2016.

Supplement I to the paper "Asymptotic cumulants of some information criteria" —Proofs and technical results

Haruhiko Ogasawara

This article is to supplement Ogasawara (2016) with proofs and technical results.

Appendix. Proofs and technical results

A. Proofs and associated expressions

A1. Proof of Theorem 1

We obtain an expression of b_2 which is different from that of Konishi and Kitagawa (2003) with b_1 being well known. For the expression, we use the formula of the expansion of $\hat{\boldsymbol{\theta}}_{\mathrm{W}} = \boldsymbol{\theta}_{\mathrm{W}}(\mathbf{X}^*)$ given by Ogasawara (2015a, Equation (2.1) (see also 2015b for correction); 2014, Equation (2.4)):

$$
\begin{aligned}
\hat{\boldsymbol{\theta}}_{\mathrm{W}} - \boldsymbol{\theta}_0 &= -n^{-1}\boldsymbol{\Lambda}^{-1}\mathbf{q}_0^* + \sum_{j=1}^{3}\boldsymbol{\Lambda}^{(j)}\mathbf{l}_0^{(j)} - n^{-1}(\hat{\mathbf{L}}_{\mathrm{W}}^{-1}\hat{\mathbf{q}}_{\mathrm{W}}^* - \boldsymbol{\Lambda}^{-1}\mathbf{q}_0^*)_{O_p(n^{-1/2})} \\
&\quad + O_p(n^{-2}) \\
&= -n^{-1}\boldsymbol{\Lambda}^{-1}\mathbf{q}_0^* + \sum_{j=1}^{3}\boldsymbol{\Lambda}^{(j)}\mathbf{l}_0^{(j)} + n^{-1}\left[\boldsymbol{\Lambda}^{-1}\mathbf{M}\boldsymbol{\Lambda}^{-1}\mathbf{q}_0^* - \boldsymbol{\Lambda}^{-1}\frac{\partial \mathbf{q}^*}{\partial \boldsymbol{\theta}_0{}'}\boldsymbol{\Lambda}^{(1)}\mathbf{l}_0^{(1)}\right. \\
&\quad \left. - \boldsymbol{\Lambda}^{-1}\mathrm{E}_g(\mathbf{J}_0^{(3)})\{(\boldsymbol{\Lambda}^{-1}\mathbf{q}_0^*)\otimes\boldsymbol{\Lambda}^{-1}\}\mathbf{l}_0^{(1)}\right] + O_p(n^{-2}) \qquad \text{(A1.1)} \\
&\equiv -n^{-1}\boldsymbol{\Lambda}^{-1}\mathbf{q}_0^* + \sum_{j=1}^{3}\boldsymbol{\Lambda}^{(j)}\mathbf{l}_0^{(j)} + n^{-1}(\mathbf{l}_0^{(\mathrm{W})})_{O_p(n^{-1/2})} + O_p(n^{-2}),
\end{aligned}
$$

where

$$
\boldsymbol{\Lambda} = \mathrm{E}_g(\partial^2\bar{l}/\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}) \equiv \mathrm{E}_g(\partial^2\bar{l}/\partial\boldsymbol{\theta}_0\partial\boldsymbol{\theta}_0{}') = O(1),\ \ \mathbf{q}_0^* = \mathbf{q}^*(\boldsymbol{\theta}_0),
$$

$$\boldsymbol{\Lambda}^{(j)} = O(1),\ \mathbf{l}_0^{(j)} = O_p(n^{-j/2})\ (j=1,\ 2,\ 3),$$

$$\hat{\mathbf{L}}_{\mathrm{W}} = \frac{\partial^2 \bar{l}}{\partial\boldsymbol{\theta}\,\partial\boldsymbol{\theta}'}\Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_{\mathrm{W}}} \equiv \frac{\partial^2 \bar{l}}{\partial\hat{\boldsymbol{\theta}}_{\mathrm{W}}\,\partial\hat{\boldsymbol{\theta}}_{\mathrm{W}}'},$$

$$\mathbf{q}_{\mathrm{W}}^{*} = \mathbf{q}^{*}(\hat{\boldsymbol{\theta}}_{\mathrm{W}}),\ \mathbf{M} = \frac{\partial^2 \bar{l}}{\partial\boldsymbol{\theta}_0\,\partial\boldsymbol{\theta}_0'} - \boldsymbol{\Lambda} = O_p(n^{-1/2}),$$

$$\frac{\partial\mathbf{q}^{*}}{\partial\boldsymbol{\theta}_0'} = \frac{\partial\mathbf{q}^{*}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}'}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0},$$

$\mathbf{J}_0^{(3)} = \dfrac{\partial^3 \bar{l}}{\partial\boldsymbol{\theta}_0\,(\partial\boldsymbol{\theta}_0')^{<2>}}$, $\mathbf{x}^{<k>} = \mathbf{x}\otimes\cdots\otimes\mathbf{x}$ (k times of $\mathbf{x}$), $\otimes$ denotes the Kronecker product, and $(\cdot)_{O_p(n^{-1/2})}$ indicates that $(\cdot)$ is of order $O_p(n^{-1/2})$ with other similar expressions.

The term $\sum_{j=1}^{3}\boldsymbol{\Lambda}^{(j)}\mathbf{l}_0^{(j)}$ in (A1.1) (Ogasawara, 2010, Equation (2.4)) is given from the following expansion:

$$\hat{\boldsymbol{\theta}}_{\mathrm{ML}} - \boldsymbol{\theta}_0 = \sum_{j=1}^{3}\boldsymbol{\Lambda}^{(j)}\mathbf{l}_0^{(j)} + O_p(n^{-2}), \tag{A1.2}$$

$$\boldsymbol{\Lambda}^{(1)}\mathbf{l}_0^{(1)} = -\boldsymbol{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0},$$

$$\boldsymbol{\Lambda}^{(2)}\mathbf{l}_0^{(2)} = \boldsymbol{\Lambda}^{-1}\mathbf{M}\boldsymbol{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0} - \frac{1}{2}\boldsymbol{\Lambda}^{-1}\mathrm{E}_g(\mathbf{J}_0^{(3)})\left(\boldsymbol{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0}\right)^{<2>}$$

$$\boldsymbol{\Lambda}^{(3)}\mathbf{l}_0^{(3)} = -\boldsymbol{\Lambda}^{-1}\mathbf{M}\boldsymbol{\Lambda}^{-1}\mathbf{M}\boldsymbol{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0} + \frac{1}{2}\boldsymbol{\Lambda}^{-1}\mathbf{M}\boldsymbol{\Lambda}^{-1}\mathrm{E}_g(\mathbf{J}_0^{(3)})\left(\boldsymbol{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0}\right)^{<2>}$$

$$+\boldsymbol{\Lambda}^{-1}\mathrm{E}_g(\mathbf{J}_0^{(3)})\left\{\left(\boldsymbol{\Lambda}^{-1}\mathbf{M}\boldsymbol{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0}\right)\otimes\left(\boldsymbol{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0}\right)\right\}$$

$$-\frac{1}{2}\boldsymbol{\Lambda}^{-1}\{\mathbf{J}_0^{(3)}-\mathrm{E}_g(\mathbf{J}_0^{(3)})\}\left(\boldsymbol{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0}\right)^{<2>}$$

$$-\frac{1}{2}\boldsymbol{\Lambda}^{-1}\mathrm{E}_g(\mathbf{J}_0^{(3)})\left[\left(\boldsymbol{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0}\right)\otimes\left\{\boldsymbol{\Lambda}^{-1}\mathrm{E}_g(\mathbf{J}_0^{(3)})\left(\boldsymbol{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0}\right)^{<2>}\right\}\right]$$

$$+\frac{1}{6}\boldsymbol{\Lambda}^{-1}\mathrm{E}_g(\mathbf{J}_0^{(4)})\left(\boldsymbol{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0}\right)^{<3>},$$

$$\mathbf{J}_0^{(4)}\equiv\frac{\partial^4\bar{l}}{\partial\boldsymbol{\theta}_0(\partial\boldsymbol{\theta}_0{}')^{<3>}},\ \mathbf{l}_0^{(1)}=\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0},$$

$$\mathbf{l}_0^{(2)}=\left\{\mathrm{v}'(\mathbf{M})\otimes\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0{}'},\left(\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0{}'}\right)^{<2>}\right\}'\equiv(\mathbf{l}_0^{(2-1)}{}',\mathbf{l}_0^{(2-2)}{}')'=O_p(n^{-1}),$$

$$\mathbf{l}_0^{(3)}=\left[\mathrm{v}'(\mathbf{M})^{<2>}\otimes\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0{}'},\mathrm{v}'(\mathbf{M})\otimes\left(\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0{}'}\right)^{<2>},\right.$$

$$\left.\mathrm{vec}'\{\mathbf{J}_0^{(3)}-\mathrm{E}_g(\mathbf{J}_0^{(3)})\}\otimes\left(\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0{}'}\right)^{<2>},\left(\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0{}'}\right)^{<3>}\right]'$$

$$\equiv(\mathbf{l}_0^{(3-1)}{}',\mathbf{l}_0^{(3-2)}{}',\mathbf{l}_0^{(3-3)}{}',\mathbf{l}_0^{(3-4)}{}')'=O_p(n^{-3/2}),$$

where $\boldsymbol{\Lambda}^{(2-j)}=O(1)\ (j=1,2)$ and $\boldsymbol{\Lambda}^{(3-j)}=O(1)\ (j=1,...,4)$ are defined implicitly by $\boldsymbol{\Lambda}^{(2)}\mathbf{l}_0^{(2)}=\sum_{j=1}^{2}\boldsymbol{\Lambda}^{(2-j)}\mathbf{l}_0^{(2-j)}$ and $\boldsymbol{\Lambda}^{(3)}\mathbf{l}_0^{(3)}=\sum_{j=1}^{4}\boldsymbol{\Lambda}^{(3-j)}\mathbf{l}_0^{(3-j)}$; $\mathrm{v}'(\mathbf{M})^{<2>}=[\{\mathrm{v}(\mathbf{M})\}']^{<2>}$; $\mathrm{v}(\cdot)$ is the vectorizing operator taking the non-duplicated elements of a symmetric matrix in parentheses; and $\mathrm{vec}(\cdot)$ is the vectorizing operator stacking the columns of a matrix sequentially.

Expand $-2\hat{\bar{l}}_{\mathrm{W}}$ and $-2\hat{\bar{l}}_{\mathrm{W}}^{*}$ as

$$-2\hat{\bar{l}}_{\mathrm{W}} = -2(\bar{l}_0)_{O_p(1)} - 2\sum_{j=1}^{4}\frac{1}{j!}\left\{\frac{\partial^j \bar{l}}{(\partial\boldsymbol{\theta}_0{}')^{<j>}}\right\}_{O_p(1)} \{(\hat{\boldsymbol{\theta}}_{\mathrm{W}} - \boldsymbol{\theta}_0)^{<j>}\}_{O_p(n^{-j/2})} + O_p(n^{-5/2}) \qquad \text{(A1.3)}$$

and

$$-2\hat{\bar{l}}_{\mathrm{W}}^{*} = -2(\bar{l}_0^{*})_{O(1)} - 2\sum_{j=1}^{4}\frac{1}{j!}\left[\mathrm{E}_g\left\{\frac{\partial^j \bar{l}}{(\partial\boldsymbol{\theta}_0{}')^{<j>}}\right\}\right]_{O(1)} \{(\hat{\boldsymbol{\theta}}_{\mathrm{W}} - \boldsymbol{\theta}_0)^{<j>}\}_{O_p(n^{-j/2})} + O_p(n^{-5/2}),$$

respectively. Then, recalling $\mathrm{E}_g(\bar{l}_0) = \bar{l}_0^{*}$, we have

$$\begin{aligned}
&-2\mathrm{E}_g(\hat{\bar{l}}_{\mathrm{W}} - \hat{\bar{l}}_{\mathrm{W}}^{*}) \\
&= -2\mathrm{E}_g\left[\sum_{j=1}^{3}\frac{1}{j!}\left[\frac{\partial^j \bar{l}}{(\partial\boldsymbol{\theta}_0{}')^{<j>}} - \mathrm{E}_g\left\{\frac{\partial^j \bar{l}}{(\partial\boldsymbol{\theta}_0{}')^{<j>}}\right\}\right]_{O_p(n^{-1/2})} (\hat{\boldsymbol{\theta}}_{\mathrm{W}} - \boldsymbol{\theta}_0)^{<j>}\right] \qquad \text{(A1.4)} \\
&\quad + O(n^{-3}) \\
&= -2\mathrm{E}_g\left\{\frac{\partial \bar{l}}{\partial\boldsymbol{\theta}_0{}'}(\hat{\boldsymbol{\theta}}_{\mathrm{W}} - \boldsymbol{\theta}_0)\right\}_{\to O(n^{-2})} - \mathrm{E}_g\{\mathrm{vec}'(\mathbf{M})(\hat{\boldsymbol{\theta}}_{\mathrm{W}} - \boldsymbol{\theta}_0)^{<2>}\}_{\to O(n^{-2})} \\
&\quad - \frac{1}{3}\mathrm{E}_g\{\mathrm{vec}'\{\mathbf{J}_0^{(3)} - \mathrm{E}_g(\mathbf{J}_0^{(3)})\}(\hat{\boldsymbol{\theta}}_{\mathrm{W}} - \boldsymbol{\theta}_0)^{<3>}\}_{\to O(n^{-2})} + O(n^{-3}),
\end{aligned}$$

where the terms of $j = 4$ in $\sum_{j=1}^{4}(\cdot)$ of (A1.3), when the expectation is taken, are absorbed in the remainder term of order $O(n^{-3})$; and $\mathrm{E}_g(\cdot)_{\to O(n^{-2})}$ indicates that the expectation is taken up to order $O(n^{-2})$.

Let $\boldsymbol{\Gamma} = n\mathrm{E}_g\left(\frac{\partial \bar{l}}{\partial\boldsymbol{\theta}_0}\frac{\partial \bar{l}}{\partial\boldsymbol{\theta}_0{}'}\right)$. When the model is true, $\boldsymbol{\Gamma} = -\boldsymbol{\Lambda} = \mathbf{I}_0$, where $\mathbf{I}_0$ is the population Fisher information matrix per observation. Under possible model misspecification, the last three expectations in (A1.4) are given as

$$-2\mathrm{E}_g\left\{\frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0{}'}(\hat{\boldsymbol{\theta}}_{\mathrm{W}}-\boldsymbol{\theta}_0)\right\}$$

$$=-2\mathrm{E}_g\left[\frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0{}'}\{-n^{-1}\boldsymbol{\Lambda}^{-1}\mathbf{q}_0^* + \sum_{j=1}^{3}\boldsymbol{\Lambda}^{(j)}\mathbf{l}_0^{(j)} + n^{-1}\mathbf{l}_0^{(\mathrm{W})} + O_p(n^{-2})\}\right]$$

$$=\left\{2\mathrm{E}_g\left(\frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0{}'}\boldsymbol{\Lambda}^{-1}\frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0}\right)\right\}_{O(n^{-1})} - \left\{2\mathrm{E}_g\left(\frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0{}'}\boldsymbol{\Lambda}^{(2)}\mathbf{l}_0^{(2)}\right)\right\}_{O(n^{-2})}$$

$$-\left\{2\mathrm{E}_g\left(\frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0{}'}\boldsymbol{\Lambda}^{(3)}\mathbf{l}_0^{(3)}\right)\right\}_{O(n^{-2})} - \left\{2\mathrm{E}_g\left(n^{-1}\frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0{}'}\mathbf{l}_0^{(\mathrm{W})}\right)\right\}_{O(n^{-2})} + O(n^{-3})$$

$$=2n^{-1}\mathrm{tr}(\boldsymbol{\Lambda}^{-1}\boldsymbol{\Gamma}) - 2n^{-2}\Bigg[\underset{(\mathrm{A})}{\sum_{a\ge b}\sum_{c,d=1}^{q}}(\boldsymbol{\Lambda}^{(2-1)})_{(d:ab,\,c)}n^2\mathrm{E}_g\left(m_{ab}\frac{\partial \bar{l}}{\partial \theta_{0c}}\frac{\partial \bar{l}}{\partial \theta_{0d}}\right) \quad \text{(A1.5)}$$

$$+\sum_{a,b,c=1}^{q}(\boldsymbol{\Lambda}^{(2-2)})_{(c:a,b)}n^2\mathrm{E}_g\left(\frac{\partial \bar{l}}{\partial \theta_{0a}}\frac{\partial \bar{l}}{\partial \theta_{0b}}\frac{\partial \bar{l}}{\partial \theta_{0c}}\right) + \sum_{a\ge b}\sum_{c\ge d}\sum_{e,f=1}^{q}(\boldsymbol{\Lambda}^{(3-1)})_{(f:ab,\,cd,e)}$$

$$\times\left\{n\,\mathrm{cov}_g(m_{ab},m_{cd})\gamma_{ef} + \sum_{(e,f)}^{2} n\,\mathrm{cov}_g\left(m_{ab},\frac{\partial \bar{l}}{\partial \theta_{0e}}\right)n\,\mathrm{cov}_g\left(m_{cd},\frac{\partial \bar{l}}{\partial \theta_{0f}}\right)\right\}$$

$$+\sum_{a\ge b}\sum_{c,d,e=1}^{q}(\boldsymbol{\Lambda}^{(3-2)})_{(e:ab,\,c,d)}\sum_{(c,d,e)}^{3} n\,\mathrm{cov}_g\left(m_{ab},\frac{\partial \bar{l}}{\partial \theta_{0c}}\right)\gamma_{de}$$

$$+\sum_{a,b,c,d,e,f=1}^{q}(\boldsymbol{\Lambda}^{(3-3)})_{(f:abc,\,d,e)}\sum_{(d,e,f)}^{3} n\,\mathrm{cov}_g\left\{(\mathbf{J}_0^{(3)})_{(a,b,c)},\frac{\partial \bar{l}}{\partial \theta_{0d}}\right\}\gamma_{ef}$$

$$+\sum_{a,b,c,d=1}^{q}(\boldsymbol{\Lambda}^{(3-4)})_{(d:a,b,c)}(\gamma_{ab}\gamma_{cd}+\gamma_{ac}\gamma_{bd}+\gamma_{ad}\gamma_{bc})$$

$$+\sum_{a,b,c=1}^{q}\lambda^{ab}(\boldsymbol{\Lambda}^{-1}\mathbf{q}_0^*)_c n\,\mathrm{cov}_g\left(\frac{\partial \bar{l}}{\partial \theta_{0a}},m_{bc}\right) + \mathrm{tr}\left(\frac{\partial \mathbf{q}^*}{\partial \boldsymbol{\theta}_0{}'}\boldsymbol{\Lambda}^{-1}\boldsymbol{\Gamma}\boldsymbol{\Lambda}^{-1}\right)$$

$$-\mathrm{tr}[\mathrm{E}_g(\mathbf{J}_0^{(3)})\{(\mathbf{\Lambda}^{-1}\mathbf{q}_0^*)\otimes(\mathbf{\Lambda}^{-1}\mathbf{\Gamma}\mathbf{\Lambda}^{-1})\}]\underset{\text{(A)}}{\Big]}+O(n^{-3})$$

$$\equiv n^{-1}b_1+n^{-2}c_1+O(n^{-3})\ (b_1=2\mathrm{tr}(\mathbf{\Lambda}^{-1}\mathbf{\Gamma}),\ c_1=-2\underset{\text{(A)}}{[}\cdot\underset{\text{(A)}}{]}),$$

where $(\mathbf{\Lambda}^{(2-1)})_{(d:ab,\,c)}$ indicates the element of the d-th row and the column corresponding to $(\mathbf{M})_{ab}\equiv m_{ab}$ (the $(a,\ b)$th element of $\mathbf{M}$) and $\partial\bar{l}/\partial(\mathbf{\theta}_0)_c\equiv\partial\bar{l}/\partial\theta_{0c}$ of $\mathbf{\Lambda}^{(2-1)}$ with $(\cdot)_c$ being the c-th element of a vector with other expressions defined similarly;

$\sum_{a\geq b}(\cdot)\equiv\sum_{b=1}^{a}\sum_{a=1}^{q}(\cdot),\ \sum_{e,f=1}^{q}(\cdot)=\sum_{e=1}^{q}\sum_{f=1}^{q}(\cdot);\ \mathrm{cov}_g(\cdot)$ is the covariance using the distribution $g(\mathbf{X}^*\,|\,\zeta_0)$; $\sum_{(e,f)}^{2}(\cdot)$ is the sum of two symmetric terms with respect to e and f with $\sum_{(c,d,e)}^{3}(\cdot)$ defined similarly; and $\underset{\text{(A)}}{[}\cdot\underset{\text{(A)}}{]}$ is for ease of finding correspondence;

$$-\mathrm{E}_g\{\mathrm{vec}'(\mathbf{M})(\hat{\mathbf{\theta}}_{\mathrm{W}}-\mathbf{\theta}_0)^{<2>}\}$$

$$=-\mathrm{E}_g\left[\mathrm{vec}'(\mathbf{M})\left\{2(-n^{-1}\mathbf{\Lambda}^{-1}\mathbf{q}_0^*)\otimes\left(-\mathbf{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\mathbf{\theta}_0}\right)+\left(\mathbf{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\mathbf{\theta}_0}\right)^{<2>}\right.\right.$$

$$\left.\left.+2\left(-\mathbf{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\mathbf{\theta}_0}\right)\otimes(\mathbf{\Lambda}^{(2)}\mathbf{l}_0^{(2)})\right\}\right]$$

$$=-n^{-2}\underset{\text{(A)}}{\Big[}\ 2\sum_{a,b,c=1}^{q}(\mathbf{\Lambda}^{-1}\mathbf{q}_0^*)_a\lambda^{bc}n\,\mathrm{cov}_g\left(m_{ab},\frac{\partial\bar{l}}{\partial\theta_{0c}}\right)$$

$$+\sum_{a,b,c,d=1}^{q}\lambda^{ac}\lambda^{bd}n^2\mathrm{E}_g\left(m_{ab}\frac{\partial\bar{l}}{\partial\theta_{0c}}\frac{\partial\bar{l}}{\partial\theta_{0d}}\right)$$

$$-2\sum_{a,b,c=1}^{q}\sum_{d\geq e}\sum_{f=1}^{q}(\mathbf{\Lambda}^{(2-1)})_{(b:de,\,f)}$$

$$\times\lambda^{ac}\left\{\sum_{(c,f)}^{2} n\,\mathrm{cov}_g\left(m_{ab},\frac{\partial\bar{l}}{\partial\theta_{0c}}\right)n\,\mathrm{cov}_g\left(m_{de},\frac{\partial\bar{l}}{\partial\theta_{0f}}\right)\right.$$

$$\left.+n\,\mathrm{cov}_g(m_{ab},m_{de})\gamma_{cf}\right\}$$

$$\left.-2\sum_{a,b,c,d,e=1}^{q}(\mathbf{\Lambda}^{(2-2)})_{(b:d,e)}\lambda^{ac}\sum_{(c,d,e)}^{3}n\,\mathrm{cov}_g\left(m_{ab},\frac{\partial\bar{l}}{\partial\theta_{0c}}\right)\gamma_{de}\right]_{(\mathrm{A})}+O(n^{-3}) \qquad \text{(A1.6)}$$

$$\equiv n^{-2}c_2+O(n^{-3}),$$

$$-\frac{1}{3}\mathrm{E}_g[\mathrm{vec}'\{\mathbf{J}_0^{(3)}-\mathrm{E}_g(\mathbf{J}_0^{(3)})\}(\hat{\boldsymbol{\theta}}_{\mathrm{W}}-\boldsymbol{\theta}_0)^{<3>}]$$

$$=-\frac{1}{3}\mathrm{E}_g\left[\mathrm{vec}'\{\mathbf{J}_0^{(3)}-\mathrm{E}_g(\mathbf{J}_0^{(3)})\}\left(-\mathbf{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\theta_{0d}}\right)^{<3>}\right]+O(n^{-3})$$

$$=n^{-2}\sum_{a,b,c,d,e,f=1}^{q}\lambda^{ad}\lambda^{be}\lambda^{cf}n\,\mathrm{cov}_g\left\{(\mathbf{J}_0^{(3)})_{(a,b,c)},\frac{\partial\bar{l}}{\partial\theta_{0d}}\right\}\gamma_{ef}+O(n^{-3}) \qquad \text{(A1.7)}$$

$$\equiv n^{-2}c_3+O(n^{-3}),$$

where $\lambda^{bc}=(\mathbf{\Lambda}^{-1})_{bc}$. Then, from (A1.5) to (A1.7) we have (2.9).

A2. Proof of Corollary 1

Under canonical parametrization in the exponential family, it is known that $\frac{\partial^j\bar{l}}{(\partial\boldsymbol{\theta}_0)^{<j>}}=\mathrm{E}_g\left\{\frac{\partial^j\bar{l}}{(\partial\boldsymbol{\theta}_0)^{<j>}}\right\}(j=2,3,...)$, which gives c_1 of (2.11) from (A1.5) with $\mathbf{M}=\mathbf{O}$ (a zero matrix of an appropriate size) and $\mathbf{J}_0^{(3)}-\mathrm{E}_g(\mathbf{J}_0^{(3)})=\mathbf{O}$. The results of $c_2=c_3=0$ are derived similarly from (A1.6) and (A1.7) with $\mathbf{M}=\mathbf{O}$ and $\mathbf{J}_0^{(3)}-\mathrm{E}_g(\mathbf{J}_0^{(3)})=\mathbf{O}$, respectively.

A3. Proof of Corollary 2

Recalling (A1.2) for $\mathbf{\Lambda}^{(2-2)}$ and $\mathbf{\Lambda}^{(3-4)}$ in c_1 of (2.11), we have

$$
\begin{aligned}
c_1 = -2\Bigg[& \sum_{a,b,c=1}^{q} (\mathbf{\Lambda}^{(2-2)})_{(c:a,b)} n^2 \mathrm{E}_g\left(\frac{\partial \bar{l}}{\partial \theta_{0a}}\frac{\partial \bar{l}}{\partial \theta_{0b}}\frac{\partial \bar{l}}{\partial \theta_{0c}}\right) \\
& + \sum_{a,b,c,d=1}^{q} (\mathbf{\Lambda}^{(3-4)})_{(d:a,b,c)} (\gamma_{ab}\gamma_{cd} + \gamma_{ac}\gamma_{bd} + \gamma_{ad}\gamma_{bc}) \\
& + \mathrm{tr}\left(\frac{\partial \mathbf{q}^*}{\partial \boldsymbol{\theta}_0{}'}\mathbf{\Lambda}^{-1}\mathbf{\Gamma}\mathbf{\Lambda}^{-1}\right) - \mathrm{tr}[\mathbf{J}_0^{(3)}\{(\mathbf{\Lambda}^{-1}\mathbf{q}_0^*)\otimes(\mathbf{\Lambda}^{-1}\mathbf{\Gamma}\mathbf{\Lambda}^{-1})\}] \Bigg] \\
= -2\Bigg[& \sum_{a,b,c=1}^{q} \left\{-\frac{1}{2}\mathbf{\Lambda}^{-1}\mathbf{J}_0^{(3)}(\mathbf{\Lambda}^{-1})^{<2>}\right\}_{(c:a,b)} n^2 \mathrm{E}_g\left(\frac{\partial \bar{l}}{\partial \theta_{0a}}\frac{\partial \bar{l}}{\partial \theta_{0b}}\frac{\partial \bar{l}}{\partial \theta_{0c}}\right) \\
& - \sum_{a,b,c,d=1}^{q} \frac{1}{2}(\mathbf{\Lambda}^{-1})_{d\cdot}\mathbf{J}_0^{(3)}\left[(\mathbf{\Lambda}^{-1})_{\cdot a}\otimes[\mathbf{\Lambda}^{-1}\mathbf{J}_0^{(3)}\{(\mathbf{\Lambda}^{-1})_{\cdot b}\otimes(\mathbf{\Lambda}^{-1})_{\cdot c}\}]\right] \\
& \quad \times(\gamma_{ab}\gamma_{cd} + \gamma_{ac}\gamma_{bd} + \gamma_{ad}\gamma_{bc}) \\
& + \sum_{a,b,c,d=1}^{q} \frac{1}{6}\{\mathbf{\Lambda}^{-1}\mathbf{J}_0^{(4)}(\mathbf{\Lambda}^{-1})^{<3>}\}_{(d:a,b,c)} 3\gamma_{ab}\gamma_{cd} \\
& + \mathrm{tr}\left(\frac{\partial \mathbf{q}^*}{\partial \boldsymbol{\theta}_0{}'}\mathbf{\Lambda}^{-1}\mathbf{\Gamma}\mathbf{\Lambda}^{-1}\right) - \mathrm{tr}[\mathbf{J}_0^{(3)}\{(\mathbf{\Lambda}^{-1}\mathbf{q}_0^*)\otimes(\mathbf{\Lambda}^{-1}\mathbf{\Gamma}\mathbf{\Lambda}^{-1})\}] \Bigg] \\
= -\mathrm{vec}'(\mathbf{J}_0^{(3)}) & n^2 \mathrm{E}_g\left\{\left(-\mathbf{\Lambda}^{-1}\frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0}\right)^{<3>}\right\} \\
& + \mathrm{vec}'(\mathbf{\Lambda}^{-1}\mathbf{\Gamma}\mathbf{\Lambda}^{-1})\mathbf{J}_0^{(3)}{}'\mathbf{\Lambda}^{-1}\mathbf{J}_0^{(3)}\mathrm{vec}(\mathbf{\Lambda}^{-1}\mathbf{\Gamma}\mathbf{\Lambda}^{-1}) \\
& + 2\mathrm{vec}'(\mathbf{J}_0^{(3)})\{\mathbf{\Lambda}^{-1}\otimes(\mathbf{\Lambda}^{-1}\mathbf{\Gamma}\mathbf{\Lambda}^{-1})^{<2>}\}\mathrm{vec}(\mathbf{J}_0^{(3)}) \\
& - \mathrm{vec}'(\mathbf{J}_0^{(4)})\mathrm{vec}\{(\mathbf{\Lambda}^{-1}\mathbf{\Gamma}\mathbf{\Lambda}^{-1})^{<2>}\} \\
& - 2\mathrm{tr}\left(\frac{\partial \mathbf{q}^*}{\partial \boldsymbol{\theta}_0{}'}\mathbf{\Lambda}^{-1}\mathbf{\Gamma}\mathbf{\Lambda}^{-1}\right) + 2\mathrm{tr}[\mathbf{J}_0^{(3)}\{(\mathbf{\Lambda}^{-1}\mathbf{q}_0^*)\otimes(\mathbf{\Lambda}^{-1}\mathbf{\Gamma}\mathbf{\Lambda}^{-1})\}],
\end{aligned}
\tag{A3.1}
$$

where $(\cdot)_{d\cdot}$ is the d-th row of a matrix and $(\cdot)_{\cdot a}$ is the a-th column of a matrix.

Under correct model specification and canonical parametrization, since

$\partial l_j / \partial \boldsymbol{\theta}_0 = \mathbf{x}^* - \mathrm{E}_f(\mathbf{x}^*)$ and $-\boldsymbol{\Lambda} = \boldsymbol{\Gamma} = \mathbf{I}_0$, (A3.1) becomes

$$
\begin{aligned}
c_1 &= \boldsymbol{\kappa}_{f3}{}'(\mathbf{x}^*)\boldsymbol{\kappa}_{f3}\left(\mathbf{I}_0^{-1}\frac{\partial l_j}{\partial \boldsymbol{\theta}_0}\right) - \mathrm{vec}'(\mathbf{I}_0^{-1})\mathbf{J}_0^{(3)}{}'\mathbf{I}_0^{-1}\mathbf{J}_0^{(3)}\mathrm{vec}(\mathbf{I}_0^{-1}) \\
&\quad - 2\mathrm{vec}'(\mathbf{J}_0^{(3)})(\mathbf{I}_0^{-1})^{<3>}\mathrm{vec}(\mathbf{J}_0^{(3)}) + \boldsymbol{\kappa}_{f4}{}'(\mathbf{x}^*)\mathrm{vec}\{(\mathbf{I}_0^{-1})^{<2>}\} \\
&\quad - 2\mathrm{tr}\left(\frac{\partial \mathbf{q}^*}{\partial \boldsymbol{\theta}_0{}'}\mathbf{I}_0^{-1}\right) - 2\mathrm{tr}[\mathbf{J}_0^{(3)}\{(\mathbf{I}_0^{-1}\mathbf{q}_0^*)\otimes \mathbf{I}_0^{-1}\}] \\
&= \boldsymbol{\kappa}_{f3}{}'(\mathbf{I}_0^{-1/2}\mathbf{x}^*)\boldsymbol{\kappa}_{f3}\left(\mathbf{I}_0^{-1/2}\frac{\partial l_j}{\partial \boldsymbol{\theta}_0}\right) \\
&\quad - \boldsymbol{\kappa}_{f3}{}'(\mathbf{I}_0^{-1/2}\mathbf{x}^*)[\mathbf{I}_{(q)}\otimes\{\mathrm{vec}(\mathbf{I}_{(q)})\mathrm{vec}'(\mathbf{I}_{(q)})\}]\boldsymbol{\kappa}_{f3}(\mathbf{I}_0^{-1/2}\mathbf{x}^*) \qquad \text{(A3.2)} \\
&\quad - 2\boldsymbol{\kappa}_{f3}{}'(\mathbf{I}_0^{-1/2}\mathbf{x}^*)\boldsymbol{\kappa}_{f3}(\mathbf{I}_0^{-1/2}\mathbf{x}^*) + \boldsymbol{\kappa}_{f4}{}'(\mathbf{I}_0^{-1/2}\mathbf{x}^*)\mathrm{vec}\{(\mathbf{I}_{(q)})^{<2>}\} \\
&\quad - 2\mathrm{tr}\left(\frac{\partial \mathbf{q}^*}{\partial \boldsymbol{\theta}_0{}'}\mathbf{I}_0^{-1}\right) + 2\boldsymbol{\kappa}_{f3}{}'(\mathbf{I}_0^{-1/2}\mathbf{x}^*)\{(\mathbf{I}_0^{-1/2}\mathbf{q}_0^*)\otimes \mathrm{vec}(\mathbf{I}_{(q)})\} \\
&= \boldsymbol{\kappa}_{f3}{}'(\tilde{\mathbf{x}}^*)\boldsymbol{\kappa}_{f3}(\tilde{\mathbf{x}}^*) - \boldsymbol{\kappa}_{f3}{}'(\tilde{\mathbf{x}}^*)[\mathbf{I}_{(q)}\otimes\{\mathrm{vec}(\mathbf{I}_{(q)})\mathrm{vec}'(\mathbf{I}_{(q)})\}]\boldsymbol{\kappa}_{f3}(\tilde{\mathbf{x}}^*) \\
&\quad - 2\boldsymbol{\kappa}_{f3}{}'(\tilde{\mathbf{x}}^*)\boldsymbol{\kappa}_{f3}(\tilde{\mathbf{x}}^*) + \boldsymbol{\kappa}_{f4}{}'(\tilde{\mathbf{x}}^*)\mathrm{vec}(\mathbf{I}_{(q^2)}) \\
&\quad - 2\mathrm{tr}\left(\frac{\partial \mathbf{q}^*}{\partial \boldsymbol{\theta}_0{}'}\mathbf{I}_0^{-1}\right) + 2\boldsymbol{\kappa}_{f3}{}'(\tilde{\mathbf{x}}^*)\{(\mathbf{I}_0^{-1/2}\mathbf{q}_0^*)\otimes \mathrm{vec}(\mathbf{I}_{(q)})\},
\end{aligned}
$$

which gives (2.12).

A4. Proof of Corollary 4

Since $\mathbf{J}_0^{(j)} \equiv \dfrac{\partial^j \bar{l}}{\partial \boldsymbol{\theta}_0(\partial \boldsymbol{\theta}_0{}')^{<j-1>}} = \mathrm{E}_g\left\{\dfrac{\partial^j \bar{l}}{\partial \boldsymbol{\theta}_0(\partial \boldsymbol{\theta}_0{}')^{<j-1>}}\right\}(j = 2,3,\ldots)$ under canonical parametrization, the asymptotic expansion using the MLE corresponding to (A1.4) higher than (A1.4) is given only by the first term $-2\mathrm{E}_g\left\{\dfrac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0{}'}(\hat{\boldsymbol{\theta}}_{\mathrm{ML}} - \boldsymbol{\theta}_0)\right\}$, which is also given only by $-2\mathrm{E}_g\left\{\dfrac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0{}'}\left(-\boldsymbol{\Lambda}^{-1}\dfrac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0}\right)\right\}$ and $-2\mathrm{E}_g\{h(\mathbf{J}_0^{(3)},\mathbf{J}_0^{(4)},\ldots)\}$, where $h(\cdot)$ is the sum of multiplicative functions of the powers of the arguments.

In the only non-vanishing term $-2\mathrm{E}_g\left\{\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0{}'}(\hat{\boldsymbol{\theta}}_{\mathrm{ML}}-\boldsymbol{\theta}_0)\right\}$ for the expansion of the left-hand side of (2.15),

$$-2\mathrm{E}_g\left\{\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0{}'}\left(-\boldsymbol{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0}\right)\right\}=-2\mathrm{tr}\left\{\boldsymbol{\Sigma}\mathrm{E}_g\left(\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0}\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0{}'}\right)\right\}$$

$=-n^{-1}2\mathrm{tr}(\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{-1})=-n^{-1}2q$ under arbitrary distributions as long as $\boldsymbol{\Sigma}$ and $\boldsymbol{\Sigma}^{-1}$ exist. The remaining terms $-2\mathrm{E}_g\{h(\mathbf{J}_0^{(3)},\mathbf{J}_0^{(4)},\ldots)\}$ vanish when we use the normal distribution even under non-normality since $\mathbf{J}_0^{(j)}=\mathbf{O}\,(j=3,4,\ldots)$ in this case.

An alternative direct proof is given as follows. Let $\mathbf{z}_j\,(j=1,\ldots,n)$ be independent copies of $\mathbf{x}^*$ and $\mathrm{E}_{\mathbf{Z}^*}(\cdot)$ denote an expectation over the distribution of $\mathbf{Z}^*$ or $\mathbf{z}_j\,(j=1,\ldots,n)$. Then, by definition,

$$\begin{aligned}-2\hat{\bar{l}}^*_{\mathrm{ML}}&=-2\mathrm{E}_{\mathbf{Z}^*}\left[-\frac{n^{-1}}{2}\sum_{j=1}^{n}(\mathbf{z}_j-\bar{\mathbf{x}})'\boldsymbol{\Sigma}^{-1}(\mathbf{z}_j-\bar{\mathbf{x}})-\frac{1}{2}\log\{(2\pi)^q\,|\,\boldsymbol{\Sigma}\,|\}\right]\\&=\mathrm{tr}(\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma})+(\boldsymbol{\mu}_0-\bar{\mathbf{x}})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_0-\bar{\mathbf{x}})'+\log\{(2\pi)^q\,|\,\boldsymbol{\Sigma}\,|\}\\&=q+(\boldsymbol{\mu}_0-\bar{\mathbf{x}})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_0-\bar{\mathbf{x}})'+\log\{(2\pi)^q\,|\,\boldsymbol{\Sigma}\,|\},\end{aligned}\tag{A4.1}$$

which gives $-2\mathrm{E}_g(\hat{\bar{l}}^*_{\mathrm{ML}})=(1+n^{-1})q+\log\{(2\pi)^q\,|\,\boldsymbol{\Sigma}\,|\}$. On the other hand,

$$\begin{aligned}-2\mathrm{E}_g(\hat{\bar{l}}_{\mathrm{ML}})&=-2\mathrm{E}_g\left[-\frac{n^{-1}}{2}\sum_{j=1}^{n}(\mathbf{x}_j-\bar{\mathbf{x}})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}_j-\bar{\mathbf{x}})-\frac{1}{2}\log\{(2\pi)^q\,|\,\boldsymbol{\Sigma}\,|\}\right]\\&=(1-n^{-1})\mathrm{tr}(\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma})+\log\{(2\pi)^q\,|\,\boldsymbol{\Sigma}\,|\}\\&=(1-n^{-1})q+\log\{(2\pi)^q\,|\,\boldsymbol{\Sigma}\,|\}.\end{aligned}\tag{A4.2}$$

Consequently, (A4.1) and (A4.2) yield $-2\mathrm{E}_g(\hat{\bar{l}}_{\mathrm{ML}}-\hat{\bar{l}}^*_{\mathrm{ML}})=-n^{-1}2q$.

A4a. Proof of Corollary 5

Let c_1^* be the sum of the first three terms on the right-hand side of (2.12). Then,

$$c_1 = c_1^* - 2\mathrm{tr}\left(\frac{\partial \mathbf{q}^*}{\partial \boldsymbol{\theta}_0{}'}\mathbf{I}_0^{-1}\right) + 2\boldsymbol{\kappa}_{f3}{}'(\tilde{\mathbf{x}}^*)\{(\mathbf{I}_0^{-1/2}\mathbf{q}_0^*)\otimes \mathrm{vec}(\mathbf{I}_{(q)})\}. \tag{A4.3}$$

Since

$$\begin{aligned}\mathbf{q}_0^* &= \frac{1}{2}\frac{\partial \log|\mathbf{I}|}{\partial \boldsymbol{\theta}_0} = \frac{1}{2}\frac{\partial \mathrm{vec}'(\mathbf{I})}{\partial \boldsymbol{\theta}_0}\mathrm{vec}(\mathbf{I}_0^{-1}) \\ &= \frac{1}{2}\{\mathbf{I}_0^{1/2}\otimes \mathrm{vec}'(\mathbf{I}_{(q)})\}\boldsymbol{\kappa}_{f3}(\tilde{\mathbf{x}}^*),\end{aligned} \tag{A4.4}$$

the last term in (A4.3) becomes

$$\begin{aligned}&2\boldsymbol{\kappa}_{f3}{}'(\tilde{\mathbf{x}}^*)\{(\mathbf{I}_0^{-1/2}\mathbf{q}_0^*)\otimes \mathrm{vec}(\mathbf{I}_{(q)})\} \\ &= \boldsymbol{\kappa}_{f3}{}'(\tilde{\mathbf{x}}^*)[\mathbf{I}_{(q)}\otimes \mathrm{vec}(\mathbf{I}_{(q)})\mathrm{vec}'(\mathbf{I}_{(q)})]\boldsymbol{\kappa}_{f3}(\tilde{\mathbf{x}}^*).\end{aligned} \tag{A4.5}$$

On the other hand, the second term on the right-hand side of (A4.3) is

$$\begin{aligned}&-2\mathrm{tr}\left(\frac{\partial \mathbf{q}^*}{\partial \boldsymbol{\theta}_0{}'}\mathbf{I}_0^{-1}\right) \\ &= -\mathrm{tr}\left[\left[\left\{\frac{\partial^2}{\partial \boldsymbol{\theta}_0 \partial \boldsymbol{\theta}_0{}'}\mathrm{vec}'(\mathbf{I})\right\}\mathrm{vec}(\mathbf{I}_0^{-1})\right]\mathbf{I}_0^{-1}\right]_{(\mathrm{A})} \\ &+ \mathrm{tr}\left[\left[\frac{\partial \mathrm{vec}'(\mathbf{I})}{\partial \boldsymbol{\theta}_0}\left\{\mathrm{vec}\left(\mathbf{I}_0^{-1}\frac{\partial \mathbf{I}}{\partial \theta_{01}}\mathbf{I}_0^{-1}\right),\ldots,\mathrm{vec}\left(\mathbf{I}_0^{-1}\frac{\partial \mathbf{I}}{\partial \theta_{0q}}\mathbf{I}_0^{-1}\right)\right\}\right]\mathbf{I}_0^{-1}\right]_{(\mathrm{B})} \\ &= -\boldsymbol{\kappa}_{f4}{}'(\tilde{\mathbf{x}}^*)\mathrm{vec}(\mathbf{I}_{(q^2)}) + \boldsymbol{\kappa}_{f3}{}'(\tilde{\mathbf{x}}^*)\boldsymbol{\kappa}_{f3}{}'(\tilde{\mathbf{x}}^*),\end{aligned} \tag{A4.6}$$

where $\left\{\frac{\partial^2}{\partial \boldsymbol{\theta}_0 \partial \boldsymbol{\theta}_0{}'}\mathrm{vec}'(\mathbf{A})\right\}\mathrm{vec}(\mathbf{B}_0) = \sum_{a,b=1}^{q}\frac{\partial^2 (\mathbf{A})_{ab}}{\partial \boldsymbol{\theta}_0 \partial \boldsymbol{\theta}_0{}'}(\mathbf{B}_0)_{ab}$ and $(\cdot)_{ab}$ is the (a, b)th element of a matrix. From the definition of c_1^* in (A4.3) (see (2.12)), we find that the sum of (A4.5) and (A4.6) is $-c_1^*$, which gives $c_1 = c_1^* - c_1^* = 0$.

A5. Expressions of $-\boldsymbol{\Lambda}_{\mathbf{M}}^{-1(\Delta)}$, $-\boldsymbol{\Lambda}_{\mathbf{M}}^{-1(\Delta\Delta)}$, $\boldsymbol{\Gamma}_{\mathbf{M}}^{(\Delta)}$ and $\boldsymbol{\Gamma}_{\mathbf{M}}^{(\Delta\Delta)}$

Let $\mathbf{L}_0 = \left(\dfrac{\partial^2 \bar{l}}{\partial \boldsymbol{\theta}_0 \partial \boldsymbol{\theta}_0{}'}\right)_{O_p(1)}$, then

$$
\begin{aligned}
-\hat{\mathbf{L}}_{\mathrm{W}}^{-1} &= -\mathbf{L}_0^{-1} + \sum_{j=1}^{q} \mathbf{L}_0^{-1} \frac{\partial \mathbf{L}_0}{\partial (\boldsymbol{\theta}_0)_j} \mathbf{L}_0^{-1} (\hat{\boldsymbol{\theta}}_{\mathrm{W}} - \boldsymbol{\theta}_0)_j \\
&\quad + \sum_{j,k=1}^{q} \left\{ -\mathbf{L}_0^{-1} \frac{\partial \mathbf{L}_0}{\partial (\boldsymbol{\theta}_0)_j} \mathbf{L}_0^{-1} \frac{\partial \mathbf{L}_0}{\partial (\boldsymbol{\theta}_0)_k} \mathbf{L}_0^{-1} + \frac{1}{2} \mathbf{L}_0^{-1} \frac{\partial^2 \mathbf{L}_0}{\partial (\boldsymbol{\theta}_0)_j \partial (\boldsymbol{\theta}_0)_k} \mathbf{L}_0^{-1} \right\} \\
&\qquad \times (\hat{\boldsymbol{\theta}}_{\mathrm{W}} - \boldsymbol{\theta}_0)_j (\hat{\boldsymbol{\theta}}_{\mathrm{W}} - \boldsymbol{\theta}_0)_k + O_p(n^{-3/2}) \\
&= -\boldsymbol{\Lambda}^{-1} + \boldsymbol{\Lambda}^{-1}\mathbf{M}\boldsymbol{\Lambda}^{-1} - \boldsymbol{\Lambda}^{-1}\mathbf{M}\boldsymbol{\Lambda}^{-1}\mathbf{M}\boldsymbol{\Lambda}^{-1} \\
&\quad + (\boldsymbol{\Lambda}^{-1} - \boldsymbol{\Lambda}^{-1}\mathbf{M}\boldsymbol{\Lambda}^{-1}) \sum_{j=1}^{q} \left[\mathrm{E}_g\left(\frac{\partial \mathbf{L}_0}{\partial (\boldsymbol{\theta}_0)_j}\right) + \left\{ \frac{\partial \mathbf{L}_0}{\partial (\boldsymbol{\theta}_0)_j} - \mathrm{E}_g\left(\frac{\partial \mathbf{L}_0}{\partial (\boldsymbol{\theta}_0)_j}\right) \right\} \right] \\
&\quad \times (\boldsymbol{\Lambda}^{-1} - \boldsymbol{\Lambda}^{-1}\mathbf{M}\boldsymbol{\Lambda}^{-1}) \left(-\boldsymbol{\Lambda}^{-1} \frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0} - n^{-1}\boldsymbol{\Lambda}^{-1}\mathbf{q}_0^{*} + \boldsymbol{\Lambda}^{(2)}\mathbf{l}_0^{(2)} \right)_j \\
&\quad + \sum_{j,k=1}^{q} \left\{ -\boldsymbol{\Lambda}^{-1} \mathrm{E}_g\left(\frac{\partial \mathbf{L}_0}{\partial (\boldsymbol{\theta}_0)_j}\right) \boldsymbol{\Lambda}^{-1} \mathrm{E}_g\left(\frac{\partial \mathbf{L}_0}{\partial (\boldsymbol{\theta}_0)_k}\right) \boldsymbol{\Lambda}^{-1} + \frac{1}{2} \boldsymbol{\Lambda}^{-1} \mathrm{E}_g\left(\frac{\partial^2 \mathbf{L}_0}{\partial (\boldsymbol{\theta}_0)_j \partial (\boldsymbol{\theta}_0)_k}\right) \boldsymbol{\Lambda}^{-1} \right\} \\
&\qquad \times \left(\boldsymbol{\Lambda}^{-1} \frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0} \right)_j \left(\boldsymbol{\Lambda}^{-1} \frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0} \right)_k + O_p(n^{-3/2})
\end{aligned}
$$

$$= -\boldsymbol{\Lambda}^{-1} + \left[\boldsymbol{\Lambda}^{-1}\mathbf{M}\boldsymbol{\Lambda}^{-1} - \boldsymbol{\Lambda}^{-1}\mathrm{E}_g(\mathbf{J}_0^{(3)})\left\{ \boldsymbol{\Lambda}^{-1} \otimes \left(\boldsymbol{\Lambda}^{-1}\frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0} \right) \right\} \right]_{O_p(n^{-1/2})}$$

$$+ \Bigg[-\boldsymbol{\Lambda}^{-1}\mathbf{M}\boldsymbol{\Lambda}^{-1}\mathbf{M}\boldsymbol{\Lambda}^{-1} + \boldsymbol{\Lambda}^{-1}\mathbf{M}\boldsymbol{\Lambda}^{-1}\mathrm{E}_g(\mathbf{J}_0^{(3)})\left\{ \boldsymbol{\Lambda}^{-1} \otimes \left(\boldsymbol{\Lambda}^{-1}\frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0} \right) \right\}$$
$$_{(\mathrm{A})}$$

$$+ \boldsymbol{\Lambda}^{-1}\mathrm{E}_g(\mathbf{J}_0^{(3)})\left\{ (\boldsymbol{\Lambda}^{-1}\mathbf{M}\boldsymbol{\Lambda}^{-1}) \otimes \left(\boldsymbol{\Lambda}^{-1}\frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0} \right) \right\}$$

$$- \boldsymbol{\Lambda}^{-1}\{\mathbf{J}_0^{(3)} - \mathrm{E}_g(\mathbf{J}_0^{(3)})\}\left\{ \boldsymbol{\Lambda}^{-1} \otimes \left(\boldsymbol{\Lambda}^{-1}\frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0} \right) \right\}$$

$$+ \boldsymbol{\Lambda}^{-1}\mathrm{E}_g(\mathbf{J}_0^{(3)})\left\{ \boldsymbol{\Lambda}^{-1} \otimes (-n^{-1}\boldsymbol{\Lambda}^{-1}\mathbf{q}_0^* + \boldsymbol{\Lambda}^{(2)}\mathbf{l}_0^{(2)}) \right\}$$

$$- \boldsymbol{\Lambda}^{-1}\left[\mathrm{E}_g(\mathbf{J}_0^{(3)})\left\{ \boldsymbol{\Lambda}^{-1} \otimes \left(\boldsymbol{\Lambda}^{-1}\frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0} \right) \right\} \right]^2$$

$$+ \frac{1}{2}\boldsymbol{\Lambda}^{-1}\mathrm{E}_g(\mathbf{J}_0^{(4)})\left\{ \boldsymbol{\Lambda}^{-1} \otimes \left(\boldsymbol{\Lambda}^{-1}\frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0} \right)^{<2>} \right\} \Bigg]_{(\mathrm{A})O_p(n^{-1})} + O_p(n^{-3/2})$$

$$\equiv -\boldsymbol{\Lambda}^{-1} + (-\boldsymbol{\Lambda}_{\mathbf{M}}^{-1(\Delta)})_{O_p(n^{-1/2})} + (-\boldsymbol{\Lambda}_{\mathbf{M}}^{-1(\Delta\Delta)})_{O_p(n^{-1})} + O_p(n^{-3/2}). \qquad \text{(A5.1)}$$

Let

$$\mathbf{G} \equiv \mathbf{G}(\boldsymbol{\theta}) \equiv \mathbf{G}(\boldsymbol{\theta}, \mathbf{X}^*) \equiv \left(n^{-1}\sum_{j=1}^{n}\frac{\partial l_j}{\partial \boldsymbol{\theta}}\frac{\partial l_j}{\partial \boldsymbol{\theta}'} \right)_{O_p(1)},$$

$$\mathbf{G}_0 \equiv \mathbf{G}(\boldsymbol{\theta}_0) \equiv \mathbf{G}(\boldsymbol{\theta}_0, \mathbf{X}^*),$$

$$\mathbf{G}_0 = \boldsymbol{\Gamma} + (\mathbf{M}_{\mathbf{G}})_{O_p(n^{-1/2})},\ \mathrm{E}_g(\mathbf{G}_0) = \boldsymbol{\Gamma},\ \mathbf{G}_{0(j)}^{(3)} = \partial \mathbf{G}_0 / \partial(\boldsymbol{\theta}_0)_j, \qquad \text{(A5.2)}$$

$$\mathbf{G}_{0(j,k)}^{(4)} = \partial^2 \mathbf{G}_0 / \partial(\boldsymbol{\theta}_0)_j \partial(\boldsymbol{\theta}_0)_j\ (j,k = 1,\ldots,q),$$

then

$$
\begin{aligned}
\hat{\boldsymbol{\Gamma}}_{\mathrm{W}} &= \mathbf{G}(\hat{\boldsymbol{\theta}}_{\mathrm{W}}, \mathbf{X}^{*}) = (\mathbf{G}_0)_{O_p(1)} + (\hat{\boldsymbol{\Gamma}}_{\mathrm{W}} - \mathbf{G}_0)_{O_p(n^{-1/2})} \\
&= \boldsymbol{\Gamma} + \mathbf{M}_{\mathbf{G}} + \sum_{j=1}^{q} \mathbf{G}_{0(j)}^{(3)} (\hat{\boldsymbol{\theta}}_{\mathrm{W}} - \boldsymbol{\theta}_0)_j + \frac{1}{2} \sum_{j,k=1}^{q} \mathbf{G}_{0(j,k)}^{(4)} (\hat{\boldsymbol{\theta}}_{\mathrm{W}} - \boldsymbol{\theta}_0)_j (\hat{\boldsymbol{\theta}}_{\mathrm{W}} - \boldsymbol{\theta}_0)_k \\
&\quad + O_p(n^{-3/2}) \\
&= \boldsymbol{\Gamma} + \mathbf{M}_{\mathbf{G}} + \sum_{j=1}^{q} \mathbf{G}_{0(j)}^{(3)} \left(-n^{-1} \boldsymbol{\Lambda}^{-1} \mathbf{q}_0^{*} + \sum_{k=1}^{2} \boldsymbol{\Lambda}^{(k)} \mathbf{l}_0^{(k)} \right)_j \\
&\quad + \frac{1}{2} \sum_{j,k=1}^{q} \mathbf{G}_{0(j,k)}^{(4)} (\boldsymbol{\Lambda}^{(1)} \mathbf{l}_0^{(1)})_j (\boldsymbol{\Lambda}^{(1)} \mathbf{l}_0^{(1)})_k + O_p(n^{-3/2}) \\
&= \boldsymbol{\Gamma} + \left\{ \mathbf{M}_{\mathbf{G}} - \sum_{j=1}^{q} \mathrm{E}_g(\mathbf{G}_{0(j)}^{(3)}) \left(\boldsymbol{\Lambda}^{-1} \frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0} \right)_j \right\}_{O_p(n^{-1/2})} \\
&\quad + \Bigg[_{(\mathrm{A})} - \sum_{j=1}^{q} \{ \mathbf{G}_{0(j)}^{(3)} - \mathrm{E}_g(\mathbf{G}_{0(j)}^{(3)}) \} \left(\boldsymbol{\Lambda}^{-1} \frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0} \right)_j \\
&\quad + \sum_{j=1}^{q} \mathrm{E}_g(\mathbf{G}_{0(j)}^{(3)}) (-n^{-1} \boldsymbol{\Lambda}^{-1} \mathbf{q}_0^{*} + \boldsymbol{\Lambda}^{(2)} \mathbf{l}_0^{(2)})_j \\
&\quad + \frac{1}{2} \sum_{j,k=1}^{q} \mathrm{E}_g(\mathbf{G}_{0(j,k)}^{(4)}) \left(\boldsymbol{\Lambda}^{-1} \frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0} \right)_j \left(\boldsymbol{\Lambda}^{-1} \frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0} \right)_k \Bigg]_{(\mathrm{A}) O_p(n^{-1})} + O_p(n^{-3/2}) \\
&\equiv \boldsymbol{\Gamma} + (\boldsymbol{\Gamma}_{\mathbf{M}}^{(\Delta)})_{O_p(n^{-1/2})} + (\boldsymbol{\Gamma}_{\mathbf{M}}^{(\Delta\Delta)})_{O_p(n^{-1})} + O_p(n^{-3/2}).
\end{aligned}
\tag{A5.3}
$$

A6. Actual expressions of $-\boldsymbol{\Lambda}_{\mathbf{I}}^{-1(\Delta)}$, $-\boldsymbol{\Lambda}_{\mathbf{I}}^{-1(\Delta\Delta)}$, $\boldsymbol{\Gamma}_{\mathbf{I}}^{(\Delta)}$ and $\boldsymbol{\Gamma}_{\mathbf{I}}^{(\Delta\Delta)}$

Omitting terms with $\mathbf{M}$, $\mathbf{J}_0^{(3)} - \mathrm{E}_g(\mathbf{J}_0^{(3)})$, $\mathbf{M}_{\mathbf{G}}$ and $\mathbf{G}_{0(j)}^{(3)} - \mathrm{E}_g(\mathbf{G}_{0(j)}^{(3)})$ in (A5.1) and (A5.3), we obtain

$$\hat{\mathbf{I}}_{\mathrm{W}}^{(-\boldsymbol{\Lambda})-1} = -\boldsymbol{\Lambda}^{-1} - \left[\boldsymbol{\Lambda}^{-1}\mathrm{E}_g(\mathbf{J}_0^{(3)})\left\{ \boldsymbol{\Lambda}^{-1} \otimes \left(\boldsymbol{\Lambda}^{-1} \frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0} \right) \right\} \right]_{O_p(n^{-1/2})}$$

$$+ \Big[\boldsymbol{\Lambda}^{-1}\mathrm{E}_g(\mathbf{J}_0^{(3)})\{ \boldsymbol{\Lambda}^{-1} \otimes (-n^{-1}\boldsymbol{\Lambda}^{-1}\mathbf{q}_0^* + \boldsymbol{\Lambda}^{(2)}\mathbf{l}_0^{(2)}) \}$$

$$- \boldsymbol{\Lambda}^{-1} \left[\mathrm{E}_g(\mathbf{J}_0^{(3)}) \left\{ \boldsymbol{\Lambda}^{-1} \otimes \left(\boldsymbol{\Lambda}^{-1} \frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0} \right) \right\} \right]^2$$

$$+ \frac{1}{2}\boldsymbol{\Lambda}^{-1}\mathrm{E}_g(\mathbf{J}_0^{(4)}) \left\{ \boldsymbol{\Lambda}^{-1} \otimes \left(\boldsymbol{\Lambda}^{-1} \frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0} \right)^{<2>} \right\} \Big]_{O_p(n^{-1})}^{(\mathrm{A})} + O_p(n^{-3/2}) \qquad \text{(A6.1)}$$

$$\equiv -\boldsymbol{\Lambda}^{-1} + (-\boldsymbol{\Lambda}_{\mathbf{I}}^{-1(\Delta)})_{O_p(n^{-1/2})} + (-\boldsymbol{\Lambda}_{\mathbf{I}}^{-1(\Delta\Delta)})_{O_p(n^{-1})} + O_p(n^{-3/2}),$$

$$\hat{\mathbf{I}}_{\mathrm{W}}^{(\boldsymbol{\Gamma})} = \boldsymbol{\Gamma} - \left\{ \sum_{j=1}^{q} \mathrm{E}_g(\mathbf{G}_{0(j)}^{(3)}) \left(\boldsymbol{\Lambda}^{-1} \frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0} \right)_j \right\}_{O_p(n^{-1/2})}$$

$$+ \Big[\sum_{j=1}^{n} \mathrm{E}_g(\mathbf{G}_{0(j)}^{(3)})(-n^{-1}\boldsymbol{\Lambda}^{-1}\mathbf{q}_0^* + \boldsymbol{\Lambda}^{(2)}\mathbf{l}_0^{(2)})_j$$

$$+ \frac{1}{2}\sum_{j,k=1}^{q} \mathrm{E}_g(\mathbf{G}_{0(j,k)}^{(4)}) \left(\boldsymbol{\Lambda}^{-1} \frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0} \right)_j \left(\boldsymbol{\Lambda}^{-1} \frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0} \right)_k \Big]_{O_p(n^{-1})}^{(\mathrm{A})} + O_p(n^{-3/2}) \qquad \text{(A6.2)}$$

$$\equiv \boldsymbol{\Gamma} + (\boldsymbol{\Gamma}_{\mathbf{I}}^{(\Delta)})_{O_p(n^{-1/2})} + (\boldsymbol{\Gamma}_{\mathbf{I}}^{(\Delta\Delta)})_{O_p(n^{-1})} + O_p(n^{-3/2}).$$

A7. Actual expressions of $d^{(\mathrm{T1})}$ in $\mathrm{E}_g\{2(\mathrm{tr}_{\Delta\Delta}^{(\mathrm{T1})})\}$ and $d^{(\mathrm{T2})}$ in $\mathrm{E}_g\{2(\mathrm{tr}_{\Delta\Delta}^{(\mathrm{T2})})\}$

$$
\begin{aligned}
&\mathrm{E}_g\{2(\mathrm{tr}_{\Delta\Delta}^{(\mathrm{T1})})\} \\
&= 2\mathrm{E}_g\{\mathrm{tr}(-\boldsymbol{\Lambda}_{\mathbf{M}}^{-1(\Delta)}\boldsymbol{\Gamma}_{\mathbf{M}}^{(\Delta)} - \boldsymbol{\Lambda}_{\mathbf{M}}^{-1(\Delta\Delta)}\boldsymbol{\Gamma} - \boldsymbol{\Lambda}^{-1}\boldsymbol{\Gamma}_{\mathbf{M}}^{(\Delta\Delta)})\} \\
&= n^{-1}2\mathrm{E}_g\mathrm{tr}\Bigg[_{(\mathrm{A})}\; n\left[\boldsymbol{\Lambda}^{-1}\mathbf{M}\boldsymbol{\Lambda}^{-1} - \boldsymbol{\Lambda}^{-1}\mathrm{E}_g(\mathbf{J}_0^{(3)})\left\{\boldsymbol{\Lambda}^{-1}\otimes\left(\boldsymbol{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0}\right)\right\}\right] \\
&\qquad\qquad \times\left\{\mathbf{M}_{\mathbf{G}} - \sum_{j=1}^{q}\mathrm{E}_g(\mathbf{G}_{0(j)}^{(3)})\left(\boldsymbol{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0}\right)_j\right\} \\
&\quad + n\Bigg[_{(\mathrm{B})} -\boldsymbol{\Lambda}^{-1}\mathbf{M}\boldsymbol{\Lambda}^{-1}\mathbf{M}\boldsymbol{\Lambda}^{-1} + \boldsymbol{\Lambda}^{-1}\mathbf{M}\boldsymbol{\Lambda}^{-1}\mathrm{E}_g(\mathbf{J}_0^{(3)})\left\{\boldsymbol{\Lambda}^{-1}\otimes\left(\boldsymbol{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0}\right)\right\} \\
&\qquad + \boldsymbol{\Lambda}^{-1}\mathrm{E}_g(\mathbf{J}_0^{(3)})\left\{(\boldsymbol{\Lambda}^{-1}\mathbf{M}\boldsymbol{\Lambda}^{-1})\otimes\left(\boldsymbol{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0}\right)\right\} \\
&\qquad - \boldsymbol{\Lambda}^{-1}\{\mathbf{J}_0^{(3)} - \mathrm{E}_g(\mathbf{J}_0^{(3)})\}\left\{\boldsymbol{\Lambda}^{-1}\otimes\left(\boldsymbol{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0}\right)\right\} \\
&\qquad + \boldsymbol{\Lambda}^{-1}\mathrm{E}_g(\mathbf{J}_0^{(3)})\left\{\boldsymbol{\Lambda}^{-1}\otimes(-n^{-1}\boldsymbol{\Lambda}^{-1}\mathbf{q}_0^{*} + \boldsymbol{\Lambda}^{(2)}\mathbf{l}_0^{(2)})\right\} \\
&\quad - \boldsymbol{\Lambda}^{-1}\left[\mathrm{E}_g(\mathbf{J}_0^{(3)})\left\{\boldsymbol{\Lambda}^{-1}\otimes\left(\boldsymbol{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0}\right)\right\}\right]^2 \\
&\qquad + \frac{1}{2}\boldsymbol{\Lambda}^{-1}\mathrm{E}_g(\mathbf{J}_0^{(4)})\left\{\boldsymbol{\Lambda}^{-1}\otimes\left(\boldsymbol{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0}\right)^{<2>}\right\}\Bigg]_{(\mathrm{B})}\boldsymbol{\Gamma}
\end{aligned}
\tag{A7.1}
$$

$$
\begin{aligned}
&-n\mathbf{\Lambda}^{-1}\Bigg[_{(\mathrm{C})}\; -\sum_{j=1}^{q}\{\mathbf{G}_{0(j)}^{(3)}-\mathrm{E}_g(\mathbf{G}_{0(j)}^{(3)})\}\left(\mathbf{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0}\right)_j \\
&\qquad +\sum_{j=1}^{q}\mathrm{E}_g(\mathbf{G}_{0(j)}^{(3)})(-n^{-1}\mathbf{\Lambda}^{-1}\mathbf{q}_0^{*}+\mathbf{\Lambda}^{(2)}\mathbf{l}_0^{(2)})_j \\
&\qquad +\frac{1}{2}\sum_{j,k=1}^{q}\mathrm{E}_g(\mathbf{G}_{0(j,k)}^{(4)})\left(\mathbf{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0}\right)_j\left(\mathbf{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0}\right)_k\;\Bigg]_{(\mathrm{C})}\;\Bigg]_{(\mathrm{A})} \\
&= n^{-1}2\Bigg[_{(\mathrm{A})}\; \mathrm{vec}'(\mathbf{\Lambda}^{-1})n\mathrm{E}_g(\mathbf{M}_{\mathbf{G}}\otimes\mathbf{M})\mathrm{vec}(\mathbf{\Lambda}^{-1}) \\
&\qquad -\sum_{a,b,c=1}^{q}(\mathbf{\Lambda}^{-1})_{a\cdot}\mathrm{E}_g(\mathbf{J}_0^{(3)})\{(\mathbf{\Lambda}^{-1})_{\cdot b}\otimes(\mathbf{\Lambda}^{-1})_{\cdot c}\}n\mathrm{E}_g\left\{\frac{\partial\bar{l}}{\partial\theta_{0c}}(\mathbf{M}_{\mathbf{G}})_{ba}\right\} \\
&\qquad -\sum_{a,b,c,d,e=1}^{q}\lambda^{ac}\lambda^{db}\sum_{j=1}^{q}\mathrm{E}_g(\mathbf{G}_{0(j)}^{(3)})_{ab}\lambda^{je}n\mathrm{E}_g\left(m_{cd}\frac{\partial\bar{l}}{\partial\theta_{0e}}\right) \\
&\qquad +\sum_{a,b,c,d=1}^{q}(\mathbf{\Lambda}^{-1})_{a\cdot}\mathrm{E}_g(\mathbf{J}_0^{(3)})\{(\mathbf{\Lambda}^{-1})_{\cdot b}\otimes(\mathbf{\Lambda}^{-1})_{\cdot c}\}\mathrm{E}_g(\mathbf{G}_{0(j)}^{(3)})_{ab}\lambda^{jd}\gamma_{cd} \\
&+\Bigg[_{(\mathrm{B})}\; -\mathrm{vec}'(\mathbf{\Lambda}^{-1})n\mathrm{E}_g(\mathbf{M}^{<2>})\mathrm{vec}(\mathbf{\Lambda}^{-1}\mathbf{\Gamma}\mathbf{\Lambda}^{-1}) \\
&\qquad +2\sum_{a,b,c=1}^{q}(\mathbf{\Lambda}^{-1})_{a\cdot}\mathrm{E}_g(\mathbf{J}_0^{(3)})\{(\mathbf{\Lambda}^{-1}\mathbf{\Gamma}\mathbf{\Lambda}^{-1})_{\cdot b}\otimes(\mathbf{\Lambda}^{-1})_{\cdot c}\}n\mathrm{E}_g\left(m_{ab}\frac{\partial\bar{l}}{\partial\theta_{0c}}\right) \\
&\qquad -\sum_{a=1}^{q}\mathrm{tr}\left[n\mathrm{E}_g\left\{\{\mathbf{J}_0^{(3)}-\mathrm{E}_g(\mathbf{J}_0^{(3)})\}\frac{\partial\bar{l}}{\partial\theta_{0a}}\right\}\{(\mathbf{\Lambda}^{-1}\mathbf{\Gamma}\mathbf{\Lambda}^{-1})\otimes(\mathbf{\Lambda}^{-1})_{\cdot a}\}\right] \\
&\qquad +\mathrm{tr}[\mathrm{E}_g(\mathbf{J}_0^{(3)})\{(\mathbf{\Lambda}^{-1}\mathbf{\Gamma}\mathbf{\Lambda}^{-1})\otimes\boldsymbol{\alpha}_{\mathrm{W1}}\}]
\end{aligned}
$$

$$-\sum_{a,b=1}^{q}\mathrm{tr}[\boldsymbol{\Gamma}\boldsymbol{\Lambda}^{-1}\mathrm{E}_g(\mathbf{J}_0^{(3)})\{\boldsymbol{\Lambda}^{-1}\otimes(\boldsymbol{\Lambda}^{-1})_{\bullet a}\}\mathrm{E}_g(\mathbf{J}_0^{(3)})\{\boldsymbol{\Lambda}^{-1}\otimes(\boldsymbol{\Lambda}^{-1})_{\bullet b}\}]\gamma_{ab}$$

$$+\frac{1}{2}\mathrm{vec}'\mathrm{E}_g(\mathbf{J}_0^{(4)})\{\mathrm{vec}(\boldsymbol{\Lambda}^{-1}\boldsymbol{\Gamma}\boldsymbol{\Lambda}^{-1})\}^{<2>}\Bigg]_{(\mathrm{B})}$$

$$-\Bigg[_{(\mathrm{C})}-\sum_{a,b=1}^{q}\mathrm{tr}\,n\mathrm{E}_g\left[\boldsymbol{\Lambda}^{-1}\{\mathbf{G}_{0(a)}^{(3)}-\mathrm{E}_g(\mathbf{G}_{0(a)}^{(3)})\}\frac{\partial\bar{l}}{\partial\theta_{0b}}\right]\lambda^{ab}$$

$$+\sum_{j=1}^{q}\mathrm{tr}\{\boldsymbol{\Lambda}^{-1}\mathrm{E}_g(\mathbf{G}_{0(j)}^{(3)})(\boldsymbol{\alpha}_{\mathrm{W1}})_j\}+\frac{1}{2}\sum_{j,k=1}^{q}\mathrm{tr}\{\boldsymbol{\Lambda}^{-1}\mathrm{E}_g(\mathbf{G}_{0(j,k)}^{(4)})(\boldsymbol{\Lambda}^{-1}\boldsymbol{\Gamma}\boldsymbol{\Lambda}^{-1})_{jk}\}\Bigg]_{(\mathrm{C})}\Bigg]_{(\mathrm{A})}$$

$$\equiv n^{-1}d^{(\mathrm{T1})},$$

where $n^{-1}\boldsymbol{\alpha}_{\mathrm{W1}}\equiv -n^{-1}\boldsymbol{\Lambda}^{-1}\mathbf{q}_0^{*}+\mathrm{E}_g(\boldsymbol{\Lambda}^{(2)}\mathbf{l}_0^{(2)})$ is the vector of the asymptotic biases of $\hat{\boldsymbol{\theta}}_{\mathrm{W}}$ up to order $O(n^{-1})$ under possible model misspecification.

On the other hand,

$$\begin{aligned}
&\mathrm{E}_g\{2(\mathrm{tr}_{\Delta\Delta}^{(\mathrm{T2})})\}\\
&=2\mathrm{E}_g\{\mathrm{tr}(-\boldsymbol{\Lambda}_{\mathbf{I}}^{-1(\Delta)}\boldsymbol{\Gamma}_{\mathbf{I}}^{(\Delta)}-\boldsymbol{\Lambda}_{\mathbf{I}}^{-1(\Delta\Delta)}\boldsymbol{\Gamma}-\boldsymbol{\Lambda}^{-1}\boldsymbol{\Gamma}_{\mathbf{I}}^{(\Delta\Delta)})\}\\
&=n^{-1}2\mathrm{E}_g\mathrm{tr}\Bigg[_{(\mathrm{A})}\;n\Bigg[-\boldsymbol{\Lambda}^{-1}\mathrm{E}_g(\mathbf{J}_0^{(3)})\left\{\boldsymbol{\Lambda}^{-1}\otimes\left(\boldsymbol{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0}\right)\right\}\\
&\qquad\times\left\{-\sum_{j=1}^{q}\mathrm{E}_g(\mathbf{G}_{0(j)}^{(3)})\left(\boldsymbol{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0}\right)_j\right\}\Bigg]\\
&\qquad+n\Bigg[_{(\mathrm{B})}\;\boldsymbol{\Lambda}^{-1}\mathrm{E}_g(\mathbf{J}_0^{(3)})\left\{\boldsymbol{\Lambda}^{-1}\otimes(-n^{-1}\boldsymbol{\Lambda}^{-1}\mathbf{q}_0^{*}+\boldsymbol{\Lambda}^{(2)}\mathbf{l}_0^{(2)})\right\}\\
&\qquad-\boldsymbol{\Lambda}^{-1}\left[\mathrm{E}_g(\mathbf{J}_0^{(3)})\left\{\boldsymbol{\Lambda}^{-1}\otimes\left(\boldsymbol{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0}\right)\right\}\right]^2
\end{aligned}\tag{A7.2}$$

$$+\frac{1}{2}\mathbf{\Lambda}^{-1}\mathrm{E}_g(\mathbf{J}_0^{(4)})\left\{\mathbf{\Lambda}^{-1}\otimes\left(\mathbf{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\mathbf{\theta}_0}\right)^{<2>}\right\}\Bigg]_{(\mathrm{B})}\mathbf{\Gamma}\Bigg]_{(\mathrm{A})}$$

$$-n^{-1}2\mathrm{E}_g\{\mathrm{tr}(\mathbf{\Lambda}^{-1}\mathbf{\Gamma}_{\mathbf{I}}^{(\Delta\Delta)})\}$$

$$=n^{-1}2\Bigg[_{(\mathrm{A})}\sum_{a,b,c,d=1}^{q}(\mathbf{\Lambda}^{-1})_{a\bullet}\mathrm{E}_g(\mathbf{J}_0^{(3)})\{(\mathbf{\Lambda}^{-1})_{\bullet b}\otimes(\mathbf{\Lambda}^{-1})_{\bullet c}\}\sum_{j=1}^{q}\mathrm{E}_g(\mathbf{G}_{0(j)}^{(3)})_{ab}\lambda^{jd}\gamma_{cd}$$

$$+\Bigg[_{(\mathrm{B})}\ \mathrm{tr}[\mathrm{E}_g(\mathbf{J}_0^{(3)})\{(\mathbf{\Lambda}^{-1}\mathbf{\Gamma}\mathbf{\Lambda}^{-1})\otimes\boldsymbol{\alpha}_{\mathrm{W}1}\}]$$

$$-\sum_{a,b=1}^{q}\mathrm{tr}[\mathbf{\Gamma}\mathbf{\Lambda}^{-1}\mathrm{E}_g(\mathbf{J}_0^{(3)})\{\mathbf{\Lambda}^{-1}\otimes(\mathbf{\Lambda}^{-1})_{\bullet a}\}\mathrm{E}_g(\mathbf{J}_0^{(3)})\{\mathbf{\Lambda}^{-1}\otimes(\mathbf{\Lambda}^{-1})_{\bullet b}\}\gamma_{ab}$$

$$+\frac{1}{2}\mathrm{vec}'\mathrm{E}_g(\mathbf{J}_0^{(4)})\{\mathrm{vec}(\mathbf{\Lambda}^{-1}\mathbf{\Gamma}\mathbf{\Lambda}^{-1})\}^{<2>}\ \Bigg]_{(\mathrm{B})}$$

$$-\Bigg[_{(\mathrm{C})}\sum_{j=1}^{q}\mathrm{tr}\{\mathbf{\Lambda}^{-1}\mathrm{E}_g(\mathbf{G}_{0(j)}^{(3)})(\boldsymbol{\alpha}_{\mathrm{W}1})_j\}+\frac{1}{2}\sum_{j,k=1}^{q}\mathrm{tr}\{\mathbf{\Lambda}^{-1}\mathrm{E}_g(\mathbf{G}_{0(j,k)}^{(4)})\}(\mathbf{\Lambda}^{-1}\mathbf{\Gamma}\mathbf{\Lambda}^{-1})_{jk}\ \Bigg]_{(\mathrm{C})}\Bigg]_{(\mathrm{A})}$$

$$\equiv n^{-1}d^{(\mathrm{T2})}.$$

A8. The derivation and actual expressions of $(\bar{l}_{\mathrm{ML}}^{(j)})_{O_p(n^{-j/2})}$ $(j=1,\ldots,4)$

The five terms up to order $O_p(n^{-2})$ in the last expression of (4.3) are further expanded one by one as follows:

(i)

$$-2(\bar{l}_0)_{O_p(1)}=-2\mathrm{E}_g(\bar{l}_0)-2\{\bar{l}_0-\mathrm{E}_g(\bar{l}_0)\}$$
$$\equiv-2(\bar{l}_0^*)_{O(1)}-2(\bar{l}_0-\bar{l}_0^*)_{O_p(n^{-1/2})},$$

(ii)

$$
\begin{aligned}
&-2\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0{}'}\left(-n^{-1}\boldsymbol{\Lambda}^{-1}\mathbf{q}_0^* + \sum_{k=1}^{3}\boldsymbol{\Lambda}^{(k)}\mathbf{l}_0^{(k)} + n^{-1}\mathbf{l}_0^{(\mathrm{W})}\right)\\
&=2\left(n^{-1}\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0{}'}\boldsymbol{\Lambda}^{-1}\mathbf{q}_0^*\right)_{O_p(n^{-3/2})}+2\left(\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0{}'}\boldsymbol{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0}\right)_{O_p(n^{-1})}\\
&\quad-2\left(\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0{}'}\boldsymbol{\Lambda}^{(2)}\mathbf{l}_0^{(2)}\right)_{O_p(n^{-3/2})}-2\left(\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0{}'}\boldsymbol{\Lambda}^{(3)}\mathbf{l}_0^{(3)}\right)_{O_p(n^{-2})}\\
&\quad-2\left(n^{-1}\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0{}'}\mathbf{l}_0^{(\mathrm{W})}\right)_{O_p(n^{-2})},
\end{aligned}
\tag{A8.1}
$$

(iii)

$$
\begin{aligned}
&-\left\{\frac{\partial^2\bar{l}}{(\partial\boldsymbol{\theta}_0{}')^{<2>}}\right\}_{O_p(1)}\left(-n^{-1}\boldsymbol{\Lambda}^{-1}\mathbf{q}_0^* + \sum_{k=1}^{3}\boldsymbol{\Lambda}^{(k)}\mathbf{l}_0^{(k)} + n^{-1}\mathbf{l}_0^{(\mathrm{W})}\right)^{<2>}\\
&=-\mathrm{vec}'\{\boldsymbol{\Lambda}+(\mathbf{M})_{O_p(n^{-1/2})}\}\left(-n^{-1}\boldsymbol{\Lambda}^{-1}\mathbf{q}_0^* + \sum_{k=1}^{3}\boldsymbol{\Lambda}^{(k)}\mathbf{l}_0^{(k)} + n^{-1}(\mathbf{l}_0^{(\mathrm{W})})\right)^{<2>}\\
&=-\{n^{-2}\mathrm{vec}'(\boldsymbol{\Lambda})(\boldsymbol{\Lambda}^{-1}\mathbf{q}_0^*)^{<2>}\}_{O(n^{-2})}\\
&\quad-2\left[n^{-1}\mathrm{vec}'(\boldsymbol{\Lambda})\left\{(\boldsymbol{\Lambda}^{-1}\mathbf{q}_0^*)\otimes\left(\boldsymbol{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0}\right)\right\}\right]_{O_p(n^{-3/2})}\\
&+2[n^{-1}\mathrm{vec}'(\boldsymbol{\Lambda})\{(\boldsymbol{\Lambda}^{-1}\mathbf{q}_0^*)\otimes(\boldsymbol{\Lambda}^{(2)}\mathbf{l}_0^{(2)})\}]_{O_p(n^{-2})}-\left\{\mathrm{vec}'(\boldsymbol{\Lambda})\left(\boldsymbol{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0}\right)^{<2>}\right\}_{O_p(n^{-1})}\\
&+2\left[\mathrm{vec}'(\boldsymbol{\Lambda})\left\{\left(\boldsymbol{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0}\right)\otimes(\boldsymbol{\Lambda}^{(2)}\mathbf{l}_0^{(2)})\right\}\right]_{O_p(n^{-3/2})}\\
&+2\left[\mathrm{vec}'(\boldsymbol{\Lambda})\left\{\left(\boldsymbol{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0}\right)\otimes(\boldsymbol{\Lambda}^{(3)}\mathbf{l}_0^{(3)})\right\}\right]_{O_p(n^{-2})}\\
&-\{\mathrm{vec}'(\boldsymbol{\Lambda})(\boldsymbol{\Lambda}^{(2)}\mathbf{l}_0^{(2)})^{<2>}\}_{O_p(n^{-2})}+2\left[n^{-1}\mathrm{vec}'(\boldsymbol{\Lambda})\left\{\left(\boldsymbol{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0}\right)\otimes\mathbf{l}_0^{(\mathrm{W})}\right\}\right]_{O_p(n^{-2})}
\end{aligned}
$$

$$-2\left[n^{-1}\mathrm{vec}'(\mathbf{M})\left\{(\boldsymbol{\Lambda}^{-1}\mathbf{q}_0^*)\otimes\left(\boldsymbol{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0}\right)\right\}\right]_{O_p(n^{-2})}$$

$$-\left\{\mathrm{vec}'(\mathbf{M})\left(\boldsymbol{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0}\right)^{<2>}\right\}_{O_p(n^{-3/2})}+2\left[\mathrm{vec}'(\mathbf{M})\left\{\left(\boldsymbol{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0}\right)\otimes(\boldsymbol{\Lambda}^{(2)}\mathbf{l}_0^{(2)})\right\}\right]_{O_p(n^{-2})},$$

(iv)

$$-\frac{1}{3}\frac{\partial^3\bar{l}}{(\partial\boldsymbol{\theta}_0')^{<3>}}\left(-n^{-1}\boldsymbol{\Lambda}^{-1}\mathbf{q}_0^*+\sum_{k=1}^{2}\boldsymbol{\Lambda}^{(k)}\mathbf{l}_0^{(k)}\right)^{<3>}$$

$$=-\frac{1}{3}\mathrm{vec}'[\mathrm{E}_g(\mathbf{J}_0^{(3)})+\{\mathbf{J}_0^{(3)}-\mathrm{E}_g(\mathbf{J}_0^{(3)})\}_{O_p(n^{-1/2})}]\left(-n^{-1}\boldsymbol{\Lambda}^{-1}\mathbf{q}_0^*+\sum_{k=1}^{2}\boldsymbol{\Lambda}^{(k)}\mathbf{l}_0^{(k)}\right)^{<3>}$$

$$=\left[n^{-1}\mathrm{vec}'\{\mathrm{E}_g(\mathbf{J}_0^{(3)})\}\left\{(\boldsymbol{\Lambda}^{-1}\mathbf{q}_0^*)\otimes\left(\boldsymbol{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0}\right)^{<2>}\right\}\right]_{O_p(n^{-2})}$$

$$+\frac{1}{3}\left[\mathrm{vec}'\{\mathrm{E}_g(\mathbf{J}_0^{(3)})\}\left(\boldsymbol{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0}\right)^{<3>}\right]_{O_p(n^{-3/2})}$$

$$-\left[\mathrm{vec}'\{\mathrm{E}_g(\mathbf{J}_0^{(3)})\}\left\{\left(\boldsymbol{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0}\right)^{<2>}\otimes(\boldsymbol{\Lambda}^{(2)}\mathbf{l}_0^{(2)})\right\}\right]_{O_p(n^{-2})}$$

$$+\frac{1}{3}\left[\mathrm{vec}'\{\mathbf{J}_0^{(3)}-\mathrm{E}_g(\mathbf{J}_0^{(3)})\}\left(\boldsymbol{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0}\right)^{<3>}\right]_{O_p(n^{-2})},$$

(v)

$$-\frac{1}{12}\mathrm{E}_g\left\{\frac{\partial^4\bar{l}}{(\partial\boldsymbol{\theta}_0')^{<4>}}\right\}(\boldsymbol{\Lambda}^{(1)}\mathbf{l}_0^{(1)})^{<4>}=-\frac{1}{12}\left[\mathrm{vec}'\{\mathrm{E}_g(\mathbf{J}_0^{(4)})\}\left(\boldsymbol{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0}\right)^{<4>}\right]_{O_p(n^{-2})}.$$

Using $\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0'}\boldsymbol{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0}=\mathrm{vec}'(\boldsymbol{\Lambda})\left(\boldsymbol{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0}\right)^{<2>}$ and similar results in (A8.1), (4.3) becomes

$$
\begin{aligned}
-2\hat{\bar{l}}_{\mathrm{W}} &= -2(\bar{l}_0^{\,*})_{O(1)} - 2(\bar{l}_0 - \bar{l}_0^{\,*})_{O_p(n^{-1/2})} + \left(\frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0{}'}\boldsymbol{\Lambda}^{-1}\frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0}\right)_{O_p(n^{-1})} \\
&\quad + \underset{(\mathrm{A})}{\Big[}\; 2n^{-1}\frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0{}'}\boldsymbol{\Lambda}^{-1}\mathbf{q}_0^{*} - 2\frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0{}'}\boldsymbol{\Lambda}^{(2)}\mathbf{l}_0^{(2)} \\
&\quad - 2n^{-1}\mathrm{vec}'(\boldsymbol{\Lambda})\left\{(\boldsymbol{\Lambda}^{-1}\mathbf{q}_0^{*}) \otimes \left(\boldsymbol{\Lambda}^{-1}\frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0}\right)\right\} \\
&\quad + 2\mathrm{vec}'(\boldsymbol{\Lambda})\left\{\left(\boldsymbol{\Lambda}^{-1}\frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0}\right) \otimes (\boldsymbol{\Lambda}^{(2)}\mathbf{l}_0^{(2)})\right\} - \mathrm{vec}'(\mathbf{M})\left(\boldsymbol{\Lambda}^{-1}\frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0}\right)^{<2>} \qquad (\mathrm{A8.2}) \\
&\quad + \frac{1}{3}\mathrm{vec}'\{\mathrm{E}_g(\mathbf{J}_0^{(3)})\}\left(\boldsymbol{\Lambda}^{-1}\frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0}\right)^{<3>}\; \Big]_{(\mathrm{A})\,O_p(n^{-3/2})} \\
&\quad - \{n^{-2}\mathrm{vec}'(\boldsymbol{\Lambda})(\boldsymbol{\Lambda}^{-1}\mathbf{q}_0^{*})^{<2>}\}_{O(n^{-2})} \\
&\quad + \underset{(\mathrm{B})}{\Big[}\; \underbrace{-2\frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0{}'}\boldsymbol{\Lambda}^{(3)}\mathbf{l}_0^{(3)}}_{(1)} \underbrace{- 2n^{-1}\frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0{}'}\mathbf{l}_0^{(\mathrm{W})}}_{(2)} + \underbrace{2n^{-1}\mathrm{vec}'(\boldsymbol{\Lambda})\{(\boldsymbol{\Lambda}^{-1}\mathbf{q}_0^{*}) \otimes (\boldsymbol{\Lambda}^{(2)}\mathbf{l}_0^{(2)})\}}_{(3)} \\
&\quad + \underbrace{2\mathrm{vec}'(\boldsymbol{\Lambda})\left\{\left(\boldsymbol{\Lambda}^{-1}\frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0}\right) \otimes (\boldsymbol{\Lambda}^{(3)}\mathbf{l}_0^{(3)})\right\}}_{-(1)} \underbrace{-\mathrm{vec}'(\boldsymbol{\Lambda})(\boldsymbol{\Lambda}^{(2)}\mathbf{l}_0^{(2)})^{<2>}}_{(4)} \\
&\quad + \underbrace{2n^{-1}\mathrm{vec}'(\boldsymbol{\Lambda})\left\{\left(\boldsymbol{\Lambda}^{-1}\frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0}\right) \otimes \mathbf{l}_0^{(\mathrm{W})}\right\}}_{-(2)} \underbrace{-2n^{-1}\mathrm{vec}'(\mathbf{M})\left\{(\boldsymbol{\Lambda}^{-1}\mathbf{q}_0^{*}) \otimes \left(\boldsymbol{\Lambda}^{-1}\frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0}\right)\right\}}_{-(3)}
\end{aligned}
$$

$$+2\mathrm{vec}'(\mathbf{M})\left\{\left(\mathbf{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\mathbf{\theta}_0}\right)\otimes(\mathbf{\Lambda}^{(2)}\mathbf{l}_0^{(2)})\right\}$$

−2×(4)..

$$+n^{-1}\mathrm{vec}'\{\mathrm{E}_g(\mathbf{J}_0^{(3)})\}\left\{(\mathbf{\Lambda}^{-1}\mathbf{q}_0^*)\otimes\left(\mathbf{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\mathbf{\theta}_0}\right)^{<2>}\right\}$$

−(3)..

$$-\mathrm{vec}'\{\mathrm{E}_g(\mathbf{J}_0^{(3)})\}\left\{\left(\mathbf{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\mathbf{\theta}_0}\right)^{<2>}\otimes(\mathbf{\Lambda}^{(2)}\mathbf{l}_0^{(2)})\right\}$$

−2×(4)..

$$+\frac{1}{3}\mathrm{vec}'\{\mathbf{J}_0^{(3)}-\mathrm{E}_g(\mathbf{J}_0^{(3)})\}\left(\mathbf{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\mathbf{\theta}_0}\right)^{<3>}$$

$$\left.-\frac{1}{12}\mathrm{vec}'\{\mathrm{E}_g(\mathbf{J}_0^{(4)})\}\left(\mathbf{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\mathbf{\theta}_0}\right)^{<4>}\right]_{(\mathrm{B})O_p(n^{-2})}+O_p(n^{-5/2})$$

$$=-2(\bar{l}_0^*)_{O(1)}-2(\bar{l}_0-\bar{l}_0^*)_{O_p(n^{-1/2})}+\left(\frac{\partial\bar{l}}{\partial\mathbf{\theta}_0{}'}\mathbf{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\mathbf{\theta}_0}\right)_{O_p(n^{-1})}$$

$$+\left[-\mathrm{vec}'(\mathbf{M})\left(\mathbf{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\mathbf{\theta}_0}\right)^{<2>}+\frac{1}{3}\mathrm{vec}'\{\mathrm{E}_g(\mathbf{J}_0^{(3)})\}\left(\mathbf{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\mathbf{\theta}_0}\right)^{<3>}\right]_{O_p(n^{-3/2})}$$

$$-(n^{-2}\mathbf{q}_0^*{}'\mathbf{\Lambda}^{-1}\mathbf{q}_0^*)_{O(n^{-2})}$$

$$+\left[\mathrm{vec}'(\mathbf{\Lambda})(\mathbf{\Lambda}^{(2)}\mathbf{l}_0^{(2)})^{<2>}+\frac{1}{3}\mathrm{vec}'\{\mathbf{J}_0^{(3)}-\mathrm{E}_g(\mathbf{J}_0^{(3)})\}\left(\mathbf{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\mathbf{\theta}_0}\right)^{<3>}\right.$$

$$\left.-\frac{1}{12}\mathrm{vec}'\{\mathrm{E}_g(\mathbf{J}_0^{(4)})\}\left(\mathbf{\Lambda}^{-1}\frac{\partial\bar{l}}{\partial\mathbf{\theta}_0}\right)^{<4>}\right]_{O_p(n^{-2})}+O_p(n^{-5/2})$$

$$\equiv-2(\bar{l}_0^*)_{O(1)}+\sum_{j=1}^{4}(\bar{l}_{\mathrm{ML}}^{(j)})_{O_p(n^{-j/2})}-(n^{-2}\mathbf{q}_0^*{}'\mathbf{\Lambda}^{-1}\mathbf{q}_0^*)_{O(n^{-2})}+O_p(n^{-5/2})$$

$$(\bar{l}_{\mathrm{W}}^{(j)}=\bar{l}_{\mathrm{ML}}^{(j)},\ j=1,\ldots,4),$$

where the underline with a number in parentheses indicates a quantity and the negative number e.g., “$-a\times(4)\ldots$” indicates $-a$ times the quantity which has the symbol “(4)…” when the quantities with “$-a\times(4)\ldots$” are summed.

In the last result of (A8.2), the first term for $\overline{l}_{\mathrm{W}}^{(4)}$ can also be written as

$$
\begin{aligned}
&\mathrm{vec}'(\mathbf{\Lambda})(\mathbf{\Lambda}^{(2)}\mathbf{l}_0^{(2)})^{<2>}\\
&=\mathrm{vec}'(\mathbf{M})\left\{\left(\mathbf{\Lambda}^{-1}\frac{\partial\overline{l}}{\partial\boldsymbol{\theta}_0}\right)\otimes\left(\mathbf{\Lambda}^{-1}\mathbf{M}\mathbf{\Lambda}^{-1}\frac{\partial\overline{l}}{\partial\boldsymbol{\theta}_0}\right)\right\}\\
&\quad-\mathrm{vec}'\{\mathrm{E}_g(\mathbf{J}_0^{(3)})\}\left\{\left(\mathbf{\Lambda}^{-1}\frac{\partial\overline{l}}{\partial\boldsymbol{\theta}_0}\right)^{<2>}\otimes\left(\mathbf{\Lambda}^{-1}\mathbf{M}\mathbf{\Lambda}^{-1}\frac{\partial\overline{l}}{\partial\boldsymbol{\theta}_0}\right)\right\}\\
&\quad+\frac{1}{4}\mathrm{vec}'\{\mathrm{E}_g(\mathbf{J}_0^{(3)})\}\left[\left(\mathbf{\Lambda}^{-1}\frac{\partial\overline{l}}{\partial\boldsymbol{\theta}_0}\right)^{<2>}\otimes\left\{\mathbf{\Lambda}^{-1}\mathrm{E}_g(\mathbf{J}_0^{(3)})\left(\mathbf{\Lambda}^{-1}\frac{\partial\overline{l}}{\partial\boldsymbol{\theta}_0}\right)^{<2>}\right\}\right]
\end{aligned}
\tag{A8.3}
$$

(recall (A1.2)).

A9. The derivation and actual expressions of Theorem 4

$$
\begin{aligned}
&\kappa_{g1}(n^{-1}\mathrm{AIC}_{\mathrm{W}})\\
&=-2(\overline{l}_0^{*})_{O(1)}+n^{-1}\{n\mathrm{E}_g(\overline{l}_{\mathrm{ML}}^{(2)})+2q\}_{O(1)}\\
&\quad+n^{-2}\{n^2\mathrm{E}_g(\overline{l}_{\mathrm{ML}}^{(3)}+\overline{l}_{\mathrm{ML}}^{(4)})-\mathbf{q}_0^{*}{}'\mathbf{\Lambda}^{-1}\mathbf{q}_0^{*}\}_{O(1)}+O(n^{-3})\\
&=-2\overline{l}_0^{*}+n^{-1}\{\mathrm{tr}(\mathbf{\Lambda}^{-1}\mathbf{\Gamma})+2q\}\\
&\quad+n^{-2}\{n^2\mathrm{E}_g(\overline{l}_{\mathrm{ML}}^{(3)}+\overline{l}_{\mathrm{ML}}^{(4)})-\mathbf{q}_0^{*}{}'\mathbf{\Lambda}^{-1}\mathbf{q}_0^{*}\}+O(n^{-3})\qquad\qquad(\mathrm{A9.1})\\
&\equiv-2\overline{l}_0^{*}+n^{-1}\alpha_{\mathrm{ML1}}^{(\mathrm{A})}+n^{-2}\alpha_{\mathrm{W\Delta1}}^{(\mathrm{A})}+O(n^{-3})\\
&(\alpha_{\mathrm{W1}}^{(\mathrm{A})}=\alpha_{\mathrm{ML1}}^{(\mathrm{A})}=\mathrm{tr}(\mathbf{\Lambda}^{-1}\mathbf{\Gamma})+2q),
\end{aligned}
$$

$$
\begin{aligned}
\kappa_{g2}(n^{-1}\mathrm{AIC}_{\mathrm{W}})&=n^{-1}[n\mathrm{E}_g\{(\overline{l}_{\mathrm{ML}}^{(1)})^2\}]_{O(1)}\\
&\quad+n^{-2}[2n^2\mathrm{E}_g(\overline{l}_{\mathrm{ML}}^{(1)}\overline{l}_{\mathrm{ML}}^{(2)})+2n^2\mathrm{E}_g(\overline{l}_{\mathrm{ML}}^{(1)}\overline{l}_{\mathrm{ML}}^{(3)})\\
&\qquad+n^2\mathrm{E}_g\{(\overline{l}_{\mathrm{ML}}^{(2)})^2\}-(\alpha_{\mathrm{ML1}}^{(\mathrm{A})}-2q)^2]+O(n^{-3})
\end{aligned}
$$

$$\equiv n^{-1}\alpha_{\mathrm{ML2}}^{(\mathrm{A})}+n^{-2}\alpha_{\mathrm{ML\Delta2}}^{(\mathrm{A})}+O(n^{-3})$$
$$(n\mathrm{E}_g(\bar{l}_{\mathrm{ML}}^{(2)})=\alpha_{\mathrm{ML1}}^{(\mathrm{A})}-2q=\mathrm{tr}(\mathbf{\Lambda}^{-1}\mathbf{\Gamma}))$$
$$(\alpha_{\mathrm{W2}}^{(\mathrm{A})}=\alpha_{\mathrm{ML2}}^{(\mathrm{A})}=n\mathrm{E}_g\{(\bar{l}_{\mathrm{ML}}^{(1)})^2\}=4\mathrm{E}_g\{(l_j-\bar{l}_0^*)^2\}$$
$$=4\,\mathrm{var}_g(l_j),\ \alpha_{\mathrm{W\Delta2}}^{(\mathrm{A})}=\alpha_{\mathrm{ML\Delta2}}^{(\mathrm{A})}),$$

$$\kappa_{g3}(n^{-1}\mathrm{AIC_W})=n^{-2}[\,n^2\mathrm{E}_g\{(\bar{l}_{\mathrm{ML}}^{(1)})^3\}+3n^2\mathrm{E}_g\{(\bar{l}_{\mathrm{ML}}^{(1)})^2\bar{l}_{\mathrm{ML}}^{(2)}\}$$
$$-3n\mathrm{E}_g(\bar{l}_{\mathrm{ML}}^{(2)})\alpha_{\mathrm{ML2}}^{(\mathrm{A})}\,]+O(n^{-3})$$
$$\equiv n^{-2}\alpha_{\mathrm{ML3}}^{(\mathrm{A})}+O(n^{-3})\ \ (\alpha_{\mathrm{W3}}^{(\mathrm{A})}=\alpha_{\mathrm{ML3}}^{(\mathrm{A})}),$$

$$\kappa_{g4}(n^{-1}\mathrm{AIC_W})$$
$$=\mathrm{E}_g[\{n^{-1}\mathrm{AIC_W}-\mathrm{E}_g(n^{-1}\mathrm{AIC_W})\}^4]$$
$$-3\{n^{-1}\alpha_{\mathrm{ML2}}^{(\mathrm{A})}+n^{-2}\alpha_{\mathrm{ML\Delta2}}^{(\mathrm{A})}\}^2+O(n^{-4})$$
$$=\mathrm{E}_g\{(n^{-1}\mathrm{AIC_W}+2\bar{l}_0^*)^4\}$$
$$+n^{-3}[-4(\alpha_{\mathrm{ML1}}^{(\mathrm{A})}-2q)\{\alpha_{\mathrm{ML3}}^{(\mathrm{A})}+3(\alpha_{\mathrm{ML1}}^{(\mathrm{A})}-2q)\alpha_{\mathrm{ML2}}^{(\mathrm{A})}\}$$
$$+6(\alpha_{\mathrm{ML1}}^{(\mathrm{A})}-2q)^2\alpha_{\mathrm{ML2}}^{(\mathrm{A})}]-3n^{-2}(\alpha_{\mathrm{ML2}}^{(\mathrm{A})})^2$$
$$-6n^{-3}\alpha_{\mathrm{ML2}}^{(\mathrm{A})}\alpha_{\mathrm{ML\Delta2}}^{(\mathrm{A})}+O(n^{-4})$$
$$=\mathrm{E}_g\{(n^{-1}\mathrm{AIC_W}+2\bar{l}_0^*)^4\}-3n^{-2}(\alpha_{\mathrm{ML2}}^{(\mathrm{A})})^2$$
$$-n^{-3}\{4(\alpha_{\mathrm{ML1}}^{(\mathrm{A})}-2q)\alpha_{\mathrm{ML3}}^{(\mathrm{A})}+6\alpha_{\mathrm{ML2}}^{(\mathrm{A})}\alpha_{\mathrm{ML\Delta2}}^{(\mathrm{A})}$$
$$+6\alpha_{\mathrm{ML2}}^{(\mathrm{A})}(\alpha_{\mathrm{ML1}}^{(\mathrm{A})}-2q)^2\}+O(n^{-4})$$
$$=n^{-3}[\,n^3\{\kappa_{g4}(\bar{l}_{\mathrm{ML}}^{(1)})\}_{O(n^{-3})}+4n^3\mathrm{E}_g\{(\bar{l}_{\mathrm{ML}}^{(1)})^3\bar{l}_{\mathrm{ML}}^{(2)}\}$$
$$+6n^3\mathrm{E}_g\{(\bar{l}_{\mathrm{ML}}^{(1)})^2(\bar{l}_{\mathrm{ML}}^{(2)})^2\}$$
$$+4n^3\mathrm{E}_g\{(\bar{l}_{\mathrm{ML}}^{(1)})^3\bar{l}_{\mathrm{ML}}^{(3)}\}-4(\alpha_{\mathrm{ML1}}^{(\mathrm{A})}-2q)\alpha_{\mathrm{ML3}}^{(\mathrm{A})}$$
$$-6\alpha_{\mathrm{ML2}}^{(\mathrm{A})}\alpha_{\mathrm{ML\Delta2}}^{(\mathrm{A})}-6\alpha_{\mathrm{ML2}}^{(\mathrm{A})}(\alpha_{\mathrm{ML1}}^{(\mathrm{A})}-2q)^2\,]+O(n^{-4})$$
$$\equiv n^{-3}\alpha_{\mathrm{ML4}}^{(\mathrm{A})}+O(n^{-4})\ \ (\alpha_{\mathrm{W4}}^{(\mathrm{A})}=\alpha_{\mathrm{ML4}}^{(\mathrm{A})}).$$

B. Expository notes in regression models

For n^{-1}CAIC under canonical parametrization, from Corollary 2,

$$c_1 = -\mathrm{vec}'(\mathbf{I}_0^{-1})\mathbf{J}_0^{(3)}{}'\mathbf{I}_0^{-1}\mathbf{J}_0^{(3)}\mathrm{vec}'(\mathbf{I}_0^{-1}) - \mathrm{vec}'(\mathbf{J}_0^{(3)})(\mathbf{I}_0^{-1})^{<3>}\mathrm{vec}(\mathbf{J}_0^{(3)})$$
$$-\mathrm{vec}'(\mathbf{J}_0^{(4)})\mathrm{vec}\{(\mathbf{I}_0^{-1})^{<2>}\}$$
$$= -\sum_{\substack{a,b,c,\\ d,e,f=1}}^{p} i_0^{ab}(\mathbf{J}_0^{(3)})_{(a,b,c)} i_0^{cd}(\mathbf{J}_0^{(3)})_{(d,e,f)} i_0^{ef} - \sum_{\substack{a,b,c,\\ d,e,f=1}}^{p} (\mathbf{J}_0^{(3)})_{(a,b,c)} i_0^{ad} i_0^{be} i_0^{cf} (\mathbf{J}_0^{(3)})_{(d,e,f)}$$
$$-\sum_{a,b,c,d=1}^{p} (\mathbf{J}_0^{(4)})_{(a,b,c,d)} i_0^{ab} i_0^{cd},$$
$$(\mathbf{I}_0^{-1})_{ab} = i_0^{ab}\,(a,b=1,\ldots,p).$$

When the Jeffreys prior is used,

$$\mathbf{q}_0^* = \frac{1}{2}\frac{\partial \mathrm{vec}'(\mathbf{I})}{\partial \boldsymbol{\theta}_0}\mathrm{vec}(\mathbf{I}_0^{-1}),\ \frac{\partial(\mathbf{I})_{ij}}{\partial \theta_{0k}} = -(\mathbf{J}_0^{(3)})_{(i,j,k)},$$
$$\frac{\partial(\mathbf{q}_0^*)_k}{\partial \theta_{0l*}} = \frac{1}{2}\frac{\partial^2 \mathrm{vec}'(\mathbf{I})}{\partial \theta_{0k}\partial \theta_{0l*}}\mathrm{vec}(\mathbf{I}_0^{-1}) + \frac{1}{2}\frac{\partial \mathrm{vec}'(\mathbf{I})}{\partial \theta_{0k}}\mathrm{vec}\left(\frac{\partial \mathbf{I}^{-1}}{\partial \theta_{0l*}}\right),$$
$$\frac{\partial^2(\mathbf{I})_{ij}}{\partial \theta_{0k}\partial \theta_{0l*}} = -(\mathbf{J}_0^{(4)})_{(i,j,k,l*)},\ \frac{\partial \mathbf{I}^{-1}}{\partial \theta_{0l*}} = -\mathbf{I}^{-1}\frac{\partial \mathbf{I}}{\partial \theta_{0l*}}\mathbf{I}^{-1}$$
$$(i,j,k,l^* = 1,\ldots,p).$$

When the shape parameter in a regression model is unknown, $p+1$ in place of p in c_1 and $\mathbf{q}_0^*$ should be used.

B1. Logistic regression

$$\Pr(y_i^* = y_i \mid \pi_{0i}) = \pi_{0i}^{y_i}(1-\pi_{0i})^{1-y_i}\ (y_i = 0,1),$$
$$\pi_{0i} = \frac{1}{1+\exp(-\mathbf{x}_i'\boldsymbol{\beta}_0)},\ \boldsymbol{\theta}_0 = \boldsymbol{\beta}_0, \mathbf{X} = (\mathbf{x}_1,\ldots,\mathbf{x}_n)', \mathbf{y} = (y_1,\ldots,y_n)',$$
$$\pi_i = \frac{1}{1+\exp(-\mathbf{x}_i'\boldsymbol{\beta})},\ \boldsymbol{\theta} = \boldsymbol{\beta}\ (i=1,\ldots,n),$$

$$\bar{l}(\boldsymbol{\theta}\,|\,\mathbf{X},\mathbf{y}) = \bar{l}(\boldsymbol{\theta}) \equiv \bar{l} = n^{-1}\sum_{i=1}^{n}\{y_i \log\pi_i + (1-y_i)\log(1-\pi_i)\}$$

$$= n^{-1}\sum_{i=1}^{n}\{y_i\mathbf{x}_i'\boldsymbol{\beta} + \log(1-\pi_i)\},$$

$$\frac{\partial\bar{l}}{\partial\boldsymbol{\beta}} = n^{-1}\sum_{i=1}^{n}(y_i-\pi_i)\,\mathbf{x}_i,\quad \hat{\boldsymbol{\beta}}_{\mathrm{ML}} \text{ is given by solving } \frac{\partial\bar{l}}{\partial\boldsymbol{\beta}}\Big|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}_{\mathrm{ML}}} = \mathbf{0},$$

$$\frac{\partial\bar{l}}{\partial\boldsymbol{\beta}_0} = n^{-1}\sum_{i=1}^{n}(y_i-\pi_{0i})\,\mathbf{x}_i,$$

$$\boldsymbol{\Lambda} = \frac{\partial^2\bar{l}}{\partial\boldsymbol{\beta}_0\partial\boldsymbol{\beta}_0'} = -n^{-1}\sum_{i=1}^{n}\pi_{0i}(1-\pi_{0i})\,\mathbf{x}_i\mathbf{x}_i' = -n^{-1}\sum_{i=1}^{n}\kappa_2(y_i^*)\,\mathbf{x}_i\mathbf{x}_i' = -\mathbf{I}_0,$$

$$\boldsymbol{\Gamma} = n\mathrm{E}_f\left(\frac{\partial\bar{l}}{\partial\boldsymbol{\beta}_0}\frac{\partial\bar{l}}{\partial\boldsymbol{\beta}_0'}\right) = n^{-1}\sum_{i=1}^{n}\mathrm{E}_f\{(y_i^*-\pi_{0i})^2\}\mathbf{x}_i\mathbf{x}_i'$$

$$= n^{-1}\sum_{i=1}^{n}\pi_{0i}(1-\pi_{0i})\mathbf{x}_i\mathbf{x}_i' = \mathbf{I}_0\ \ (\mathrm{var}(y_i^*) = \pi_{0i}(1-\pi_{0i})),$$

$$(\mathbf{J}_0^{(3)})_{(a,b,c)} = \frac{\partial^3\bar{l}}{\partial\beta_{0a}\partial\beta_{0b}\partial\beta_{0c}} = -n^{-1}\sum_{i=1}^{n}(1-2\pi_{0i})\pi_{0i}(1-\pi_{0i})\,x_{ia}x_{ib}x_{ic}$$

$$= -n^{-1}\sum_{i=1}^{n}\kappa_3(y_i^*)\,x_{ia}x_{ib}x_{ic},$$

$$(\mathbf{J}_0^{(4)})_{(a,b,c,d)} = \frac{\partial^4\bar{l}}{\partial\beta_{0a}\partial\beta_{0b}\partial\beta_{0c}\partial\beta_{0d}}$$

$$= -n^{-1}\sum_{i=1}^{n}(1-6\pi_{0i}+6\pi_{0i}^2)\pi_{0i}(1-\pi_{0i})\,x_{ia}x_{ib}x_{ic}x_{id}$$

$$= -n^{-1}\sum_{i=1}^{n}\kappa_4(y_i^*)\,x_{ia}x_{ib}x_{ic}x_{id}$$

$$(a,b,c,d = 1,\ldots,n),$$

$$-2\hat{\bar{l}}_{\mathrm{ML}} = -2n^{-1}\sum_{i=1}^{n}\{y_i \log\hat{\pi}_{\mathrm{ML}i} + (1-y_i)\log(1-\hat{\pi}_{\mathrm{ML}i})\}$$

$$= -2n^{-1}\sum_{i=1}^{n}\{y_i\mathbf{x}_i{}'\hat{\boldsymbol{\beta}}_{\mathrm{ML}} + \log(1-\hat{\pi}_{\mathrm{ML}i})\},$$

$$\hat{\pi}_{\mathrm{ML}i} = \frac{1}{1+\exp(-\mathbf{x}_i{}'\hat{\boldsymbol{\beta}}_{\mathrm{ML}})},$$

$$n^{-1}\mathrm{AIC} = -2\hat{\bar{l}}_{\mathrm{ML}} + n^{-1}2p,$$

$$n^{-1}\mathrm{CAIC} = -2\hat{\bar{l}}_{\mathrm{ML}} + n^{-1}2p - n^{-2}\hat{c}_1.$$

B2. Poisson regression

$$\Pr(y_i^* = y_i \mid \lambda_{0i}) = \lambda_{0i}^{y_i}\exp(-\lambda_{0i})/y_i!\ (y_i = 0,1,2,\ldots),$$

$$\lambda_{0i} = \exp(\mathbf{x}_i{}'\boldsymbol{\beta}_0),\ \boldsymbol{\theta}_0 = \boldsymbol{\beta}_0,\ \lambda_i = \exp(\mathbf{x}_i{}'\boldsymbol{\beta}),\ \boldsymbol{\theta} = \boldsymbol{\beta}\ (i=1,\ldots,n),$$

$$\bar{l} = n^{-1}\sum_{i=1}^{n}\{y_i\mathbf{x}_i{}'\boldsymbol{\beta} - \lambda_i - \log(y_i!)\},$$

$$\frac{\partial\bar{l}}{\partial\boldsymbol{\beta}} = n^{-1}\sum_{i=1}^{n}(y_i-\lambda_i)\mathbf{x}_i,\quad \hat{\boldsymbol{\beta}}_{\mathrm{ML}}\ \text{ is given by solving }\ \frac{\partial\bar{l}}{\partial\boldsymbol{\beta}}\Big|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}_{\mathrm{ML}}} = \mathbf{0},$$

$$\frac{\partial\bar{l}}{\partial\boldsymbol{\beta}_0} = n^{-1}\sum_{i=1}^{n}(y_i-\lambda_{0i})\mathbf{x}_i\ \ (\mathrm{E}_f(y_i^*) = \lambda_{0i}),$$

$$\boldsymbol{\Lambda} = \frac{\partial^2\bar{l}}{\partial\boldsymbol{\beta}_0\partial\boldsymbol{\beta}_0{}'} = -n^{-1}\sum_{i=1}^{n}\lambda_{0i}\mathbf{x}_i\mathbf{x}_i{}' = -n^{-1}\sum_{i=1}^{n}\kappa_2(y_i^*)\mathbf{x}_i\mathbf{x}_i{}' = -\mathbf{I}_0,$$

$$\boldsymbol{\Gamma} = n\mathrm{E}_f\left(\frac{\partial\bar{l}}{\partial\boldsymbol{\beta}_0}\frac{\partial\bar{l}}{\partial\boldsymbol{\beta}_0{}'}\right) = n^{-1}\sum_{i=1}^{n}\mathrm{E}_f\{(y_i^*-\lambda_{0i})^2\}\mathbf{x}_i\mathbf{x}_i{}' = n^{-1}\sum_{i=1}^{n}\lambda_{0i}\mathbf{x}_i\mathbf{x}_i{}' = \mathbf{I}_0$$

$$(\mathrm{var}(y_i^*) = \lambda_{0i}),$$

$$(\mathbf{J}_0^{(3)})_{(a,b,c)} = \frac{\partial^3 \bar{l}}{\partial\beta_{0a}\partial\beta_{0b}\partial\beta_{0c}} = -n^{-1}\sum_{i=1}^{n}\lambda_{0i}\, x_{ia}x_{ib}x_{ic}$$

$$= -n^{-1}\sum_{i=1}^{n}\kappa_3(y_i^*)x_{ia}x_{ib}x_{ic},$$

$$(\mathbf{J}_0^{(4)})_{(a,b,c,d)} = \frac{\partial^4 \bar{l}}{\partial\beta_{0a}\partial\beta_{0b}\partial\beta_{0c}\partial\beta_{0d}} = -n^{-1}\sum_{i=1}^{n}\lambda_{0i}\, x_{ia}x_{ib}x_{ic}x_{id}$$

$$= -n^{-1}\sum_{i=1}^{n}\kappa_4(y_i^*)x_{ia}x_{ib}x_{ic}x_{id}$$

$$(a,b,c,d = 1,\ldots,n),$$

$$-2\hat{\bar{l}}_{\mathrm{ML}} = -2n^{-1}\sum_{i=1}^{n}\{y_i\mathbf{x}_i{}'\hat{\boldsymbol{\beta}}_{\mathrm{ML}} - \hat{\lambda}_{\mathrm{ML}i} - \log(y_i!)\},\ \hat{\lambda}_{\mathrm{ML}i} = \exp(\mathbf{x}_i{}'\hat{\boldsymbol{\beta}}_{\mathrm{ML}}),$$

$$n^{-1}\mathrm{AIC} = -2\hat{\bar{l}}_{\mathrm{ML}} + n^{-1}2p,$$

$$n^{-1}\mathrm{CAIC} = -2\hat{\bar{l}}_{\mathrm{ML}} + n^{-1}2p - n^{-2}\hat{c}_1.$$

B3. Negative binomial (NB) regression

B3.1 NB regression when the shape parameter *r* is known

$$\Pr(y_i^* = y_i \mid \pi_{0i}, r) = \binom{y_i + r - 1}{y_i}\pi_{0i}^{y_i}(1-\pi_{0i})^r$$

$$= \frac{\Gamma(y_i + r)}{y_i!\Gamma(r)}\pi_{0i}^{y_i}(1-\pi_{0i})^r\ (y_i = 0,1,2,\ldots),$$

$$\pi_{0i} = \exp(\mathbf{x}_i{}'\boldsymbol{\beta}_0),\ \boldsymbol{\theta}_0 = \boldsymbol{\beta}_0,\ \pi_i = \exp(\mathbf{x}_i{}'\boldsymbol{\beta}),\ \boldsymbol{\theta} = \boldsymbol{\beta}\ (i = 1,\ldots,n),$$

$$\bar{l} = n^{-1}\sum_{i=1}^{n}\{\log\Gamma(y_i + r) - \log(y_i!) - \log\Gamma(r) + y_i\mathbf{x}_i{}'\boldsymbol{\beta} + r\log(1-\pi_i)\},$$

$$\frac{\partial\bar{l}}{\partial\boldsymbol{\beta}} = n^{-1}\sum_{i=1}^{n}\left(y_i - \frac{r\pi_i}{1-\pi_i}\right)\mathbf{x}_i,\ \ \hat{\boldsymbol{\beta}}_{\mathrm{ML}}\ \text{ is given by solving }\ \frac{\partial\bar{l}}{\partial\boldsymbol{\beta}}\Big|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}_{\mathrm{ML}}} = \mathbf{0},$$

$$\frac{\partial\bar{l}}{\partial\boldsymbol{\beta}_0}=n^{-1}\sum_{i=1}^{n}\left(y_i-\frac{r\pi_{0i}}{1-\pi_{0i}}\right)\mathbf{x}_i\ \ (\mathrm{E}_f(y_i^*)=r\pi_{0i}/(1-\pi_{0i})),$$

$$\boldsymbol{\Lambda}=\frac{\partial^2\bar{l}}{\partial\boldsymbol{\beta}_0\partial\boldsymbol{\beta}_0'}=-n^{-1}\sum_{i=1}^{n}\frac{r\pi_{0i}}{(1-\pi_{0i})^2}\mathbf{x}_i\mathbf{x}_i'=-n^{-1}\sum_{i=1}^{n}\kappa_2(y_i^*)\mathbf{x}_i\mathbf{x}_i'=-\mathbf{I}_0,$$

$$\boldsymbol{\Gamma}=n\mathrm{E}_f\left(\frac{\partial\bar{l}}{\partial\boldsymbol{\beta}_0}\frac{\partial\bar{l}}{\partial\boldsymbol{\beta}_0'}\right)=n^{-1}\sum_{i=1}^{n}\mathrm{E}_f\left\{\left(y_i^*-\frac{r\pi_{0i}}{1-\pi_{0i}}\right)^2\right\}\mathbf{x}_i\mathbf{x}_i'$$

$$=n^{-1}\sum_{i=1}^{n}\frac{r\pi_{0i}}{(1-\pi_{0i})^2}\mathbf{x}_i\mathbf{x}_i'=\mathbf{I}_0\quad(\mathrm{var}(y_i^*)=r\pi_{0i}/(1-\pi_{0i})^2),$$

$$(\mathbf{J}_0^{(3)})_{(a,b,c)}=\frac{\partial^3\bar{l}}{\partial\beta_{0a}\partial\beta_{0b}\partial\beta_{0c}}$$

$$=-n^{-1}\sum_{i=1}^{n}r\pi_{0i}\left\{\frac{1}{(1-\pi_{0i})^2}+\frac{2\pi_{0i}}{(1-\pi_{0i})^3}\right\}x_{ia}x_{ib}x_{ic}$$

$$=-n^{-1}\sum_{i=1}^{n}\kappa_3(y_i^*)x_{ia}x_{ib}x_{ic},$$

$$(\mathbf{J}_0^{(4)})_{(a,b,c,d)}=\frac{\partial^4\bar{l}}{\partial\beta_{0a}\partial\beta_{0b}\partial\beta_{0c}\partial\beta_{0d}}$$

$$=-n^{-1}\sum_{i=1}^{n}r\pi_{0i}\left\{\frac{1}{(1-\pi_{0i})^2}+\frac{6\pi_{0i}}{(1-\pi_{0i})^3}+\frac{6\pi_{0i}^2}{(1-\pi_{0i})^4}\right\}x_{ia}x_{ib}x_{ic}x_{id}$$

$$=-n^{-1}\sum_{i=1}^{n}\kappa_4(y_i^*)x_{ia}x_{ib}x_{ic}x_{id}$$

$$(a,b,c,d=1,\ldots,n),$$

$$-2\hat{\bar{l}}_{\mathrm{ML}}=-2n^{-1}\sum_{i=1}^{n}\{\log\Gamma(y_i+r)-\log(y_i!)-\log\Gamma(r)$$

$$+y_i\mathbf{x}_i'\hat{\boldsymbol{\beta}}_{\mathrm{ML}}+r\log(1-\hat{\pi}_{\mathrm{ML}i}),$$

$$\hat{\pi}_{\mathrm{ML}i} = \exp(\mathbf{x}_i{}'\hat{\boldsymbol{\beta}}_{\mathrm{ML}}),$$
$$n^{-1}\mathrm{AIC} = -2\hat{\bar{l}}_{\mathrm{ML}} + n^{-1}2p,$$
$$n^{-1}\mathrm{CAIC} = -2\hat{\bar{l}}_{\mathrm{ML}} + n^{-1}2p - n^{-2}\hat{c}_1.$$

B3.2 NB regression when the shape parameter r_0 is unknown

$$\boldsymbol{\theta}_0 = (\boldsymbol{\beta}_0{}', r_0)', \quad \boldsymbol{\theta} = (\boldsymbol{\beta}', r)',$$

$$\bar{l} = n^{-1}\sum_{i=1}^{n}\{\log\Gamma(y_i + r) - \log(y_i!) - \log\Gamma(r) + y_i\mathbf{x}_i{}'\boldsymbol{\beta} + r\log(1-\pi_i)\},$$

$$\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}} = \begin{pmatrix} n^{-1}\sum_{i=1}^{n}\left(y_i - \dfrac{r\pi_i}{1-\pi_i}\right)\mathbf{x}_i \\ n^{-1}\sum_{i=1}^{n}\{\psi(y_i + r) - \psi(r) + \log(1-\pi_i)\} \end{pmatrix},$$

$\hat{\boldsymbol{\beta}}_{\mathrm{ML}}$ is given by solving $\dfrac{\partial\bar{l}}{\partial\boldsymbol{\beta}}\Big|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}_{\mathrm{ML}}} = \mathbf{0}$,

$$\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0} = \begin{pmatrix} n^{-1}\sum_{i=1}^{n}\left(y_i - \dfrac{r_0\pi_{0i}}{1-\pi_{0i}}\right)\mathbf{x}_i \\ n^{-1}\sum_{i=1}^{n}\{\psi(y_i + r_0) - \psi(r_0) + \log(1-\pi_{0i})\} \end{pmatrix}$$

$$(\mathrm{E}_f(y_i^*) = r_0\pi_{0i}/(1-\pi_{0i})),$$

$$\boldsymbol{\Lambda} = \frac{\partial^2\bar{l}}{\partial\boldsymbol{\theta}_0\partial\boldsymbol{\theta}_0{}'} = -\begin{pmatrix} n^{-1}\sum_{i=1}^{n}\dfrac{r_0\pi_{0i}}{(1-\pi_{0i})^2}\mathbf{x}_i\mathbf{x}_i{}' & n^{-1}\sum_{i=1}^{n}\dfrac{\pi_{0i}}{1-\pi_{0i}}\mathbf{x}_i \\ n^{-1}\sum_{i=1}^{n}\dfrac{\pi_{0i}}{1-\pi_{0i}}\mathbf{x}_i{}' & -n^{-1}\sum_{i=1}^{n}\{\psi'(y_i + r_0) - \psi'(r_0)\} \end{pmatrix}$$

$$\neq -\mathbf{I}_0$$

(note that the $(p+1, p+1)$th element of $\boldsymbol{\Lambda}$ is stochastic).

Non-zero elements of $\mathbf{J}_0^{(3)}$ and $\mathbf{J}_0^{(4)}$ are

$$(\mathbf{J}_0^{(3)})_{(a,b,c)} = \frac{\partial^3 \bar{l}}{\partial\beta_{0a}\partial\beta_{0b}\partial\beta_{0c}}$$

$$= -n^{-1}\sum_{i=1}^{n} r_0 \pi_{0i} \left\{ \frac{1}{(1-\pi_{0i})^2} + \frac{2\pi_{0i}}{(1-\pi_{0i})^3} \right\} x_{ia} x_{ib} x_{ic},$$

$$(\mathbf{J}_0^{(3)})_{(a,b,p+1)} = \frac{\partial^3 \bar{l}}{\partial\beta_{0a}\partial\beta_{0b}\partial r_0} = -n^{-1}\sum_{i=1}^{n} \frac{\pi_{0i}}{(1-\pi_{0i})^2} x_{ia} x_{ib}$$

$(a,b,c=1,\ldots,p)$,

$$(\mathbf{J}_0^{(3)})_{(p+1,p+1,p+1)} = \frac{\partial^3 \bar{l}}{\partial r_0^3} = n^{-1}\sum_{i=1}^{n} \{\psi''(y_i + r_0) - \psi''(r_0)\}$$

(note that the last element is stochastic),

$$(\mathbf{J}_0^{(4)})_{(a,b,c,d)} = \frac{\partial^4 \bar{l}}{\partial\beta_{0a}\partial\beta_{0b}\partial\beta_{0c}\partial\beta_{0d}}$$

$$= -n^{-1}\sum_{i=1}^{n} r_0 \pi_{0i} \left\{ \frac{1}{(1-\pi_{0i})^2} + \frac{6\pi_{0i}}{(1-\pi_{0i})^3} + \frac{6\pi_{0i}^2}{(1-\pi_{0i})^4} \right\} x_{ia} x_{ib} x_{ic} x_{id}$$

$$(\mathbf{J}_0^{(4)})_{(a,b,c,p+1)} = \frac{\partial^4 \bar{l}}{\partial\beta_{0a}\partial\beta_{0b}\partial\beta_{0c}\partial r_0}$$

$$= -n^{-1}\sum_{i=1}^{n} \pi_{0i} \left\{ \frac{1}{(1-\pi_{0i})^2} + \frac{2\pi_{0i}}{(1-\pi_{0i})^3} \right\} x_{ia} x_{ib}$$

$(a,b,c,d=1,\ldots,p)$,

$$(\mathbf{J}_0^{(4)})_{(p+1,p+1,p+1,p+1)} = \frac{\partial^4 \bar{l}}{\partial r_0^4} = n^{-1}\sum_{i=1}^{n} \{\psi'''(y_i + r_0) - \psi'''(r_0)\}$$

(note that the last element is stochastic),

$$n^{-1}\mathrm{AIC} = -2\hat{\bar{l}}_{\mathrm{ML}} + n^{-1}2(p+1),$$

$$n^{-1}\mathrm{CAIC}^* = -2\hat{\bar{l}}_{\mathrm{ML}} + n^{-1}2(p+1) - n^{-2}\hat{c}_1 \quad (n^{-1}\mathrm{CAIC}^* \doteq n^{-1}\mathrm{CAIC}).$$

B4. Gamma regression

B4.1 Gamma regression when the shape parameter α is known

$$f(y_i^* = y_i \mid \lambda_{0i}, \alpha) = y_i^{\alpha-1}\lambda_{0i}^{\alpha}\exp(-\lambda_{0i}y_i)/\Gamma(\alpha)\ (y_i > 0),$$

$$\lambda_{0i} = \mathbf{x}_i{}'\boldsymbol{\beta}_0,\ \boldsymbol{\theta}_0 = \boldsymbol{\beta}_0,\ \lambda_i = \mathbf{x}_i{}'\boldsymbol{\beta},\ \boldsymbol{\theta} = \boldsymbol{\beta}\ (i = 1,\ldots,n),$$

$$\kappa_j(y_i^*) = (j-1)!\alpha/\lambda_{0i}^j\ (j = 1,2,\ldots),$$

$$\bar{l} = n^{-1}\sum_{i=1}^{n}\{(\alpha-1)\log(y_i) + \alpha\log(\lambda_i) - \lambda_i y_i\} - \log\Gamma(\alpha),$$

$$\frac{\partial\bar{l}}{\partial\boldsymbol{\beta}} = n^{-1}\sum_{i=1}^{n}\left(\frac{\alpha}{\lambda_i} - y_i\right)\mathbf{x}_i,\ \hat{\boldsymbol{\beta}}_{\mathrm{ML}}\ \text{is given by solving}\ \frac{\partial\bar{l}}{\partial\boldsymbol{\beta}}\Big|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}_{\mathrm{ML}}} = \mathbf{0},$$

$$\frac{\partial\bar{l}}{\partial\boldsymbol{\beta}_0} = n^{-1}\sum_{i=1}^{n}\left(\frac{\alpha}{\lambda_{0i}} - y_i\right)\mathbf{x}_i\ (\mathrm{E}_f(y_i^*) = \alpha/\lambda_{0i}),$$

$$\boldsymbol{\Lambda} = \frac{\partial^2\bar{l}}{\partial\boldsymbol{\beta}_0\partial\boldsymbol{\beta}_0{}'} = -n^{-1}\alpha\sum_{i=1}^{n}\frac{\mathbf{x}_i\mathbf{x}_i{}'}{\lambda_{0i}^2} = -n^{-1}\sum_{i=1}^{n}\kappa_2(-y_i^*)\mathbf{x}_i\mathbf{x}_i{}'$$

$$= -n^{-1}\sum_{i=1}^{n}\kappa_2(y_i^*)\mathbf{x}_i\mathbf{x}_i{}' = -\mathbf{I}_0,$$

$$\boldsymbol{\Gamma} = n\mathrm{E}_f\left(\frac{\partial\bar{l}}{\partial\boldsymbol{\beta}_0}\frac{\partial\bar{l}}{\partial\boldsymbol{\beta}_0{}'}\right) = n^{-1}\sum_{i=1}^{n}\mathrm{E}_f\left\{\left(y_i^* - \frac{\alpha}{\lambda_{0i}}\right)^2\right\}\mathbf{x}_i\mathbf{x}_i{}'$$

$$= n^{-1}\alpha\sum_{i=1}^{n}\frac{\mathbf{x}_i\mathbf{x}_i{}'}{\lambda_{0i}^2} = \mathbf{I}_0\quad(\mathrm{var}(y_i^*) = \alpha/\lambda_{0i}^2),$$

$$(\mathbf{J}_0^{(3)})_{(a,b,c)} = \frac{\partial^3\bar{l}}{\partial\beta_{0a}\partial\beta_{0b}\partial\beta_{0c}} = n^{-1}2\alpha\sum_{i=1}^{n}\frac{x_{ia}x_{ib}x_{ic}}{\lambda_{0i}^3},$$

$$= -n^{-1}\sum_{i=1}^{n}\kappa_3(-y_i^*)x_{ia}x_{ib}x_{ic} = n^{-1}\sum_{i=1}^{n}\kappa_3(y_i^*)x_{ia}x_{ib}x_{ic},$$

$$(\mathbf{J}_0^{(4)})_{(a,b,c,d)} = \frac{\partial^4 \bar{l}}{\partial\beta_{0a}\partial\beta_{0b}\partial\beta_{0c}\partial\beta_{0d}} = -n^{-1}6\alpha\sum_{i=1}^{n}\frac{x_{ia}x_{ib}x_{ic}x_{id}}{\lambda_{0i}^3}$$

$$= -n^{-1}\sum_{i=1}^{n}\kappa_4(-y_i^*)x_{ia}x_{ib}x_{ic}x_{id} = -n^{-1}\sum_{i=1}^{n}\kappa_4(y_i^*)x_{ia}x_{ib}x_{ic}x_{id}$$

$(a,b,c,d = 1,\ldots,n)$,

$$-2\hat{\bar{l}}_{\mathrm{ML}} = -2n^{-1}\sum_{i=1}^{n}\{(\alpha-1)\log(y_i) + \alpha\log(\hat{\lambda}_{\mathrm{ML}i}) - \hat{\lambda}_{\mathrm{ML}i}y_i\} + 2\log\Gamma(\alpha)$$

$\hat{\lambda}_{\mathrm{ML}i} = \mathbf{x}_i{}'\hat{\boldsymbol{\beta}}_{\mathrm{ML}}$,

$n^{-1}\mathrm{AIC} = -2\hat{\bar{l}}_{\mathrm{ML}} + n^{-1}2p$,

$n^{-1}\mathrm{CAIC} = -2\hat{\bar{l}}_{\mathrm{ML}} + n^{-1}2p - n^{-2}\hat{c}_1$.

B4.2 Gamma regression when the shape parameter α_0 is unknown

$\boldsymbol{\theta}_0 = (\boldsymbol{\beta}_0{}', \alpha_0)'$,　$\boldsymbol{\theta} = (\boldsymbol{\beta}', \alpha)'$,

$$\bar{l} = n^{-1}\sum_{i=1}^{n}\{(\alpha-1)\log(y_i) + \alpha\log(\lambda_i) - \lambda_i y_i\} - \log\Gamma(\alpha),$$

$$\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}} = \begin{pmatrix} n^{-1}\sum_{i=1}^{n}\left(\frac{\alpha}{\lambda_i} - y_i\right)\mathbf{x}_i \\ n^{-1}\sum_{i=1}^{n}\log(\lambda_i y_i) - \psi(\alpha) \end{pmatrix},$$

$\hat{\boldsymbol{\beta}}_{\mathrm{ML}}$ is given by solving $\frac{\partial\bar{l}}{\partial\boldsymbol{\beta}}\Big|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}_{\mathrm{ML}}} = \mathbf{0}$,

$$\frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0} = \begin{pmatrix} n^{-1}\sum_{i=1}^{n}\left(\frac{\alpha_0}{\lambda_{0i}} - y_i\right)\mathbf{x}_i \\ n^{-1}\sum_{i=1}^{n}\log(\lambda_i y_i) - \psi(\alpha_0) \end{pmatrix} (\mathrm{E}_f(y_i^*) = \alpha_0 / \lambda_{0i}),$$

$$\boldsymbol{\Lambda} = \frac{\partial^2 \bar{l}}{\partial \boldsymbol{\theta}_0 \partial \boldsymbol{\theta}_0'} = -\begin{pmatrix} n^{-1}\alpha_0 \sum_{i=1}^{n} \frac{\mathbf{x}_i \mathbf{x}_i'}{\lambda_{0i}^2} & -n^{-1}\sum_{i=1}^{n}\frac{\mathbf{x}_i}{\lambda_{0i}} \\ -n^{-1}\sum_{i=1}^{n}\frac{\mathbf{x}_i'}{\lambda_{0i}} & \psi'(\alpha_0) \end{pmatrix} = -\mathbf{I}_0.$$

Non-zero elements of $\mathbf{J}_0^{(3)}$ and $\mathbf{J}_0^{(4)}$ are

$$(\mathbf{J}_0^{(3)})_{(a,b,c)} = \frac{\partial^3 \bar{l}}{\partial \beta_{0a} \partial \beta_{0b} \partial \beta_{0c}} = n^{-1} 2\alpha_0 \sum_{i=1}^{n} \frac{x_{ia} x_{ib} x_{ic}}{\lambda_{0i}^3},$$

$$(\mathbf{J}_0^{(3)})_{(a,b,p+1)} = \frac{\partial^3 \bar{l}}{\partial \beta_{0a} \partial \beta_{0b} \partial \alpha_0} = -n^{-1} \sum_{i=1}^{n} \frac{x_{ia} x_{ib}}{\lambda_{0i}^2}$$

$(a,b,c = 1,\ldots,p)$,

$$(\mathbf{J}_0^{(3)})_{(p+1,p+1,p+1)} = \frac{\partial^3 \bar{l}}{\partial \alpha_0^3} = -\psi''(\alpha_0),$$

$$(\mathbf{J}_0^{(4)})_{(a,b,c,d)} = \frac{\partial^4 \bar{l}}{\partial \beta_{0a} \partial \beta_{0b} \partial \beta_{0c} \partial \beta_{0d}} = -n^{-1} 6\alpha_0 \sum_{i=1}^{n} \frac{x_{ia} x_{ib} x_{ic} x_{id}}{\lambda_{0i}^3}$$

$$(\mathbf{J}_0^{(4)})_{(a,b,c,p+1)} = \frac{\partial^3 \bar{l}}{\partial \beta_{0a} \partial \beta_{0b} \partial \beta_{0c} \partial \alpha_0} = n^{-1} 2 \sum_{i=1}^{n} \frac{x_{ia} x_{ib} x_{ic}}{\lambda_{0i}^3}$$

$(a,b,c,d = 1,\ldots,n)$,

$$(\mathbf{J}_0^{(4)})_{(p+1,p+1,p+1,p+1)} = \frac{\partial^4 \bar{l}}{\partial \alpha_0^4} = -\psi'''(\alpha_0),$$

$$n^{-1}\mathrm{AIC} = -2\hat{\bar{l}}_{\mathrm{ML}} + n^{-1}2(p+1),$$

$$n^{-1}\mathrm{CAIC} = -2\hat{\bar{l}}_{\mathrm{ML}} + n^{-1}2(p+1) - n^{-2}\hat{c}_1.$$

References

Konishi, S., & Kitagawa, G. (2003). Asymptotic theory for information criteria in model selection – functional approach. *Journal of Statistical Planning and Inference, 114*, 45-61.

Ogasawara, H. (2010). Asymptotic expansions for the pivots using log-likelihood derivatives with an application in item response theory. *Journal of Multivariate Analysis, 101*, 2149-2167.

Ogasawara, H. (2014). Optimization of the Gaussian and Jeffreys power priors with emphasis on the canonical parameters in the exponential family. *Behaviormetrika, 41,* 195-223.

Ogasawara, H. (2015a). Bias adjustment minimizing the asymptotic mean square error. *Communications in Statistics – Theory and Methods, 44,* 3503-3522.

Ogasawara, H. (2015b). An expository supplement to the paper "Bias adjustment minimizing the asymptotic mean square error" with errata. *Economic Review* (*Otaru University of Commerce), 65* (4), 121-130. http://www.res.otaru-uc.ac.jp/~hogasa/, Permalink: http://hdl.handle.net/10252/5427.

Ogasawara, H. (2016). Asymptotic cumulants of some information criteria. To appear in *Journal of the Japanese Society of Computational Statistics*.