Econometrica Supplementary Material

SUPPLEMENT TO "CONSISTENT PSEUDO-MAXIMUM LIKELIHOOD ESTIMATORS AND GROUPS OF TRANSFORMATIONS"

(*Econometrica*, Vol. 87, No. 1, January 2019, 327–345)

C. GOURIÉROUX
University of Toronto, TSE, and PSL, Dauphine

A. MONFORT
CREST, ENSAE

J-M. ZAKOÏAN
CREST, ENSAE and University of Lille

THIS DOCUMENT consists of three sections of additional results: (i) regularity conditions for Proposition 3; (ii) sketch of proof of Proposition 3; (iii) derivatives of functions based on exponential of matrices.

8. REGULARITY CONDITIONS FOR PROPOSITION 3

Let $\boldsymbol{\vartheta} = (\boldsymbol{\theta}', \boldsymbol{\lambda}')' = (\vartheta_j)_{1\leq j\leq K+J}$ and $l_t(\boldsymbol{\vartheta}) = l[a(\boldsymbol{x}_t; \boldsymbol{\theta}) + \boldsymbol{\lambda}, \boldsymbol{y}_t]$.

Together with Assumptions A.1–A.7, Proposition 3 requires the following assumptions:

ASSUMPTION A.9: $\boldsymbol{\vartheta}_0^* = (\boldsymbol{\theta}_0, \boldsymbol{\lambda}_0^*)$ belongs to the interior of $\boldsymbol{\theta}(\mathcal{B}) \times \mathcal{A}$.

ASSUMPTION A.10: For any $\boldsymbol{x}$, the function $\boldsymbol{\theta} \to a(\boldsymbol{x}; \boldsymbol{\theta})$ has continuous third-order derivatives. The pseudo-density function g is three times continuously differentiable.

ASSUMPTION A.11: The matrices $\boldsymbol{K}$ and $\boldsymbol{L}$ defined in the proof are positive definite.

ASSUMPTION A.12: For at least one $j \in \{1, \dots, J\}$, the matrix $\mathrm{V}_0(\frac{\partial a_j}{\partial \boldsymbol{\theta}}(x_t, \boldsymbol{\theta}_0))$ is positive definite.

ASSUMPTION A.13: There exists a neighborhood $V(\boldsymbol{\vartheta}_0^*)$ of $\boldsymbol{\vartheta}_0^*$ such that, for $i, j = 1, \dots, r$, for all $\boldsymbol{\vartheta} \in V(\boldsymbol{\vartheta}_0^*)$, the process $\{\frac{\partial}{\partial \boldsymbol{\vartheta}'}(\frac{\partial^2}{\partial\vartheta_i\partial\vartheta_j}\ell_t(\boldsymbol{\vartheta}))\}$ is strictly stationary and ergodic, and,

$$E_0 \sup_{\boldsymbol{\vartheta}\in V(\boldsymbol{\vartheta}_0^*)} \left\| \frac{\partial}{\partial \boldsymbol{\vartheta}'}\left(\frac{\partial^2}{\partial\vartheta_i\partial\vartheta_j}\ell_t(\boldsymbol{\vartheta})\right)\right\| < \infty.$$

9. PROOF OF PROPOSITION 3

In this section, we will explain how to use an appropriate central limit theorem (CLT), and we will derive the asymptotic covariance matrix.

C. Gouriéroux: Christian.Gourieroux@ensae.fr
A. Monfort: Alain.Monfort@ensae.fr
J-M. Zakoïan: Jean-Michel.Zakoian@ensae.fr

 https://doi.org/10.3982/ECTA14727

9.1. *The Pseudo-Score*

For model (6.1), the pseudo log-likelihood for one observation takes the form

$$l_t(\boldsymbol{\vartheta}) = l_t(\boldsymbol{\theta}, \boldsymbol{\lambda}) = \log g\left[\exp\left\{\sum_{j=1}^{J}\left[a_j(\boldsymbol{x}_t;\boldsymbol{\theta}) + \lambda_j\right]\boldsymbol{C}_j\right\}\boldsymbol{y}_t\right] + \sum_{j=1}^{J}\left[a_j(\boldsymbol{x}_t;\boldsymbol{\theta}) + \lambda_j\right]\mathrm{Tr}(\boldsymbol{C}_j).$$

Let

$$z_t(\boldsymbol{\theta}, \boldsymbol{\lambda}) = \exp\left\{\sum_{j=1}^{J}\left[a_j(\boldsymbol{x}_t;\boldsymbol{\theta}) + \lambda_j\right]\boldsymbol{C}_j\right\}\boldsymbol{y}_t.$$

For $\boldsymbol{\gamma} = (\mathrm{Tr}(\boldsymbol{C}_1), \dots, \mathrm{Tr}(\boldsymbol{C}_J))'$, we have

$$l_t(\boldsymbol{\theta}, \boldsymbol{\lambda}) = \log g\left\{z_t(\boldsymbol{\theta}, \boldsymbol{\lambda})\right\} + \boldsymbol{\gamma}'\left\{\boldsymbol{a}(\boldsymbol{x}_t;\boldsymbol{\theta}) + \boldsymbol{\lambda}\right\}.$$

Using the computations of Section 10.3, it follows that

$$\begin{aligned}\frac{\partial}{\partial\boldsymbol{\theta}'}l_t(\boldsymbol{\theta}, \boldsymbol{\lambda}) &= \frac{\partial \log g}{\partial \boldsymbol{u}'}\left\{z_t(\boldsymbol{\theta}, \boldsymbol{\lambda})\right\}\frac{\partial z_t(\boldsymbol{\theta}, \boldsymbol{\lambda})}{\partial\boldsymbol{\theta}'} + \boldsymbol{\gamma}'\left\{\boldsymbol{a}(\boldsymbol{x}_t;\boldsymbol{\theta}) + \boldsymbol{\lambda}\right\} \\ &= \left(\frac{\partial \log g}{\partial \boldsymbol{u}'}\left\{z_t(\boldsymbol{\theta}, \boldsymbol{\lambda})\right\}\boldsymbol{C}\left[\boldsymbol{I}_J \otimes z_t(\boldsymbol{\theta}, \boldsymbol{\lambda})\right] + \boldsymbol{\gamma}'\right)\frac{\partial \boldsymbol{a}(\boldsymbol{x}_t;\boldsymbol{\theta})}{\partial\boldsymbol{\theta}'} \\ &:= h_t'(\boldsymbol{\theta}, \boldsymbol{\lambda})\frac{\partial \boldsymbol{a}(\boldsymbol{x}_t;\boldsymbol{\theta})}{\partial\boldsymbol{\theta}'}.\end{aligned}$$

Proceeding similarly with parameter $\boldsymbol{\lambda}$, we find that

$$\frac{\partial}{\partial\boldsymbol{\vartheta}}l_t(\boldsymbol{\theta}, \boldsymbol{\lambda}) = \begin{pmatrix}\frac{\partial}{\partial\boldsymbol{\theta}}l_t(\boldsymbol{\theta}, \boldsymbol{\lambda}) \\ \frac{\partial}{\partial\boldsymbol{\lambda}}l_t(\boldsymbol{\theta}, \boldsymbol{\lambda})\end{pmatrix} = \begin{pmatrix}\frac{\partial \boldsymbol{a}'(\boldsymbol{x}_t;\boldsymbol{\theta})}{\partial\boldsymbol{\theta}} \\ \boldsymbol{I}_J\end{pmatrix} h_t(\boldsymbol{\theta}, \boldsymbol{\lambda}).$$

9.2. *The Martingale Difference Property*

Replacing $\boldsymbol{y}_t$ by $\exp\{-\sum_{j=1}^{J}[a_j(\boldsymbol{x}_t;\boldsymbol{\theta}_0) + \lambda_{0j}]\boldsymbol{C}_j\}\boldsymbol{u}_t$, we find that

$$z_t\left(\boldsymbol{\theta}_0, \boldsymbol{\lambda}_0^*\right) = \exp\left\{\sum_{j=1}^{J}\left(\lambda_{0j}^* - \lambda_{0j}\right)\boldsymbol{C}_j\right\}\boldsymbol{u}_t := \boldsymbol{\Gamma}\left(\boldsymbol{\lambda}_0^* - \boldsymbol{\lambda}_0\right)\boldsymbol{u}_t.$$

Thus,

$$\frac{\partial}{\partial\boldsymbol{\vartheta}}l_t\left(\boldsymbol{\theta}_0, \boldsymbol{\lambda}_0^*\right) = \begin{pmatrix}\frac{\partial}{\partial\boldsymbol{\theta}}l_t\left(\boldsymbol{\theta}_0, \boldsymbol{\lambda}_0^*\right) \\ \frac{\partial}{\partial\boldsymbol{\lambda}}l_t\left(\boldsymbol{\theta}_0, \boldsymbol{\lambda}_0^*\right)\end{pmatrix} = \begin{pmatrix}\frac{\partial a'(\boldsymbol{x}_t;\boldsymbol{\theta}_0)}{\partial\boldsymbol{\theta}} \\ \boldsymbol{I}_J\end{pmatrix}\boldsymbol{k}(\boldsymbol{u}_t),$$

where

$$\boldsymbol{k}(\boldsymbol{u}_t) = \left[\boldsymbol{I}_J \otimes \left(\boldsymbol{\Gamma}\left(\boldsymbol{\lambda}_0^* - \boldsymbol{\lambda}_0\right)\boldsymbol{u}_t\right)'\right]\boldsymbol{C}'\frac{\partial \log g}{\partial \boldsymbol{u}}\left\{\boldsymbol{\Gamma}\left(\boldsymbol{\lambda}_0^* - \boldsymbol{\lambda}_0\right)\boldsymbol{u}_t\right\} + \boldsymbol{\gamma}.$$

Noting that

$$\boldsymbol{k}(\boldsymbol{u}_t) = \big[\boldsymbol{I}_J \otimes \big(\boldsymbol{\Gamma}\big(\boldsymbol{a}_0^*\big)\boldsymbol{u}_t\big)'\big]\boldsymbol{C}'\frac{\partial \log g}{\partial \boldsymbol{u}}\big\{\boldsymbol{\Gamma}\big(\boldsymbol{a}_0^*\big)\boldsymbol{u}_t\big\} + \boldsymbol{\gamma},$$

where $\boldsymbol{a}_0^*$ is defined in Assumption A.5, we have $E\{\boldsymbol{k}(\boldsymbol{u}_t)\} = 0$ from the first-order conditions in the generic model. Thus, using Assumption A.8, $(\frac{\partial}{\partial \boldsymbol{\vartheta}} l_t(\boldsymbol{\theta}_0, \boldsymbol{\lambda}_0^*), \mathcal{F}_t)$ is a martingale difference sequence, where $\mathcal{F}_t = \sigma\{(\boldsymbol{u}_i, \boldsymbol{x}_i), i \le t\}$. The asymptotic normality follows from applying a CLT for the square integrable, ergodic, and stationary martingale difference (see Billingsley (1961)). We get

$$V_{\text{as}}\left[\sqrt{T}\begin{pmatrix}\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}_0\\ \hat{\boldsymbol{\lambda}}_T - \boldsymbol{\lambda}_0^*\end{pmatrix}\right] = \boldsymbol{A}^{-1}\boldsymbol{B}\boldsymbol{A}^{-1},$$

where

$$\boldsymbol{A} = E_0\left[-\frac{\partial^2}{\partial \boldsymbol{\vartheta}\partial \boldsymbol{\vartheta}'} l_t\big(\boldsymbol{\theta}_0, \boldsymbol{\lambda}_0^*\big)\right], \qquad \boldsymbol{B} = V_0\left[\frac{\partial}{\partial \boldsymbol{\vartheta}} l_t\big(\boldsymbol{\theta}_0, \boldsymbol{\lambda}_0^*\big)\right].$$

9.3. *Computation of the Asymptotic Covariance Matrix*

We have

$$\boldsymbol{B} = V_0\left[\frac{\partial}{\partial \boldsymbol{\vartheta}} l_t\big(\boldsymbol{\theta}_0, \boldsymbol{\lambda}_0^*\big)\right] = E_0\left[\begin{pmatrix}\dfrac{\partial a'(\boldsymbol{x}_t;\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}}\\ \boldsymbol{I}_J\end{pmatrix}\boldsymbol{K}\begin{pmatrix}\dfrac{\partial a'(\boldsymbol{x}_t;\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}}\\ \boldsymbol{I}_J\end{pmatrix}'\right], \quad \boldsymbol{K} = V_0\big[\boldsymbol{k}(\boldsymbol{u}_t)\big].$$

Now,

$$\frac{\partial^2}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}'} l_t(\boldsymbol{\theta}, \boldsymbol{\lambda}) = \frac{\partial h_t'(\boldsymbol{\theta}, \boldsymbol{\lambda})}{\partial \boldsymbol{\theta}}\frac{\partial \boldsymbol{a}(\boldsymbol{x}_t;\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} + \big\{h_t'(\boldsymbol{\theta}, \boldsymbol{\lambda}) \otimes \boldsymbol{I}_p\big\}A(\boldsymbol{x}_t;\boldsymbol{\theta}),$$

where $A(\boldsymbol{x}_t;\boldsymbol{\theta})$ is the $Jp \times p$ matrix:

$$A(\boldsymbol{x}_t;\boldsymbol{\theta}) = \begin{pmatrix}\dfrac{\partial^2 a_1(\boldsymbol{x}_t;\boldsymbol{\theta})}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}'}\\ \vdots \\ \dfrac{\partial^2 a_J(\boldsymbol{x}_t;\boldsymbol{\theta})}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}'}\end{pmatrix}.$$

Noting that

$$h_t'(\boldsymbol{\theta}, \boldsymbol{\lambda}) = \left[\frac{\partial \log g}{\partial \boldsymbol{u}'}\big\{z_t(\boldsymbol{\theta}, \boldsymbol{\lambda})\big\}\boldsymbol{C}_1 z_t(\boldsymbol{\theta}, \boldsymbol{\lambda}), \dots, \frac{\partial \log g}{\partial \boldsymbol{u}'}\big\{z_t(\boldsymbol{\theta}, \boldsymbol{\lambda})\big\}\boldsymbol{C}_J z_t(\boldsymbol{\theta}, \boldsymbol{\lambda})\right] + \boldsymbol{\gamma}',$$

we compute

$$\begin{aligned}&\frac{\partial}{\partial \boldsymbol{\theta}}\left\{\frac{\partial \log g}{\partial \boldsymbol{u}'}\big\{z_t(\boldsymbol{\theta}, \boldsymbol{\lambda})\big\}\boldsymbol{C}_j z_t(\boldsymbol{\theta}, \boldsymbol{\lambda})\right\}\\ &\quad = \frac{\partial}{\partial \boldsymbol{\theta}}\left\{\frac{\partial \log g}{\partial \boldsymbol{u}'}\big\{z_t(\boldsymbol{\theta}, \boldsymbol{\lambda})\big\}\right\}\boldsymbol{C}_j z_t(\boldsymbol{\theta}, \boldsymbol{\lambda}) + \left[\frac{\partial}{\partial \boldsymbol{\theta}}\big\{\boldsymbol{C}_j z_t(\boldsymbol{\theta}, \boldsymbol{\lambda})\big\}'\right]\left\{\frac{\partial \log g}{\partial \boldsymbol{u}'}\big\{z_t(\boldsymbol{\theta}, \boldsymbol{\lambda})\big\}\right\}'\end{aligned}$$

$$
\begin{aligned}
&= \frac{\partial z_t'(\boldsymbol{\theta},\boldsymbol{\lambda})}{\partial \boldsymbol{\theta}} \frac{\partial^2 \log g}{\partial \boldsymbol{u}\partial \boldsymbol{u}'}\{z_t(\boldsymbol{\theta},\boldsymbol{\lambda})\}\boldsymbol{C}_j z_t(\boldsymbol{\theta},\boldsymbol{\lambda}) \\
&\quad + \frac{\partial \boldsymbol{a}'(\boldsymbol{x}_t;\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\big[\boldsymbol{I}_J \otimes z_t(\boldsymbol{\theta},\boldsymbol{\lambda})\big]'\boldsymbol{C}'\boldsymbol{C}_j' \frac{\partial \log g}{\partial \boldsymbol{u}}\{z_t(\boldsymbol{\theta},\boldsymbol{\lambda})\} \\
&= \frac{\partial \boldsymbol{a}'(\boldsymbol{x}_t;\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\big[\boldsymbol{I}_J \otimes z_t(\boldsymbol{\theta},\boldsymbol{\lambda})\big]'\boldsymbol{C}' \\
&\quad \times \left\{\frac{\partial^2 \log g}{\partial \boldsymbol{u}\partial \boldsymbol{u}'}\{z_t(\boldsymbol{\theta},\boldsymbol{\lambda})\}\boldsymbol{C}_j z_t(\boldsymbol{\theta},\boldsymbol{\lambda}) + \boldsymbol{C}_j' \frac{\partial \log g}{\partial \boldsymbol{u}}\{z_t(\boldsymbol{\theta},\boldsymbol{\lambda})\}\right\}.
\end{aligned}
$$

It follows that

$$
\begin{aligned}
&\frac{\partial^2}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}'} l_t(\boldsymbol{\theta},\boldsymbol{\lambda}) \\
&= \frac{\partial \boldsymbol{a}'(\boldsymbol{x}_t;\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\big[\boldsymbol{I}_J \otimes z_t(\boldsymbol{\theta},\boldsymbol{\lambda})\big]'\boldsymbol{C}' \\
&\quad \times \sum_{j=1}^{J}\left\{\frac{\partial^2 \log g}{\partial \boldsymbol{u}\partial \boldsymbol{u}'}\{z_t(\boldsymbol{\theta},\boldsymbol{\lambda})\}\boldsymbol{C}_j z_t(\boldsymbol{\theta},\boldsymbol{\lambda}) + \boldsymbol{C}_j' \frac{\partial \log g}{\partial \boldsymbol{u}}\{z_t(\boldsymbol{\theta},\boldsymbol{\lambda})\}\right\}\frac{\partial a_j(\boldsymbol{x}_t;\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \\
&\quad + \{h_t'(\boldsymbol{\theta},\boldsymbol{\lambda}) \otimes \boldsymbol{I}_p\} A(\boldsymbol{x}_t;\boldsymbol{\theta}) \\
&= \frac{\partial \boldsymbol{a}'(\boldsymbol{x}_t;\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\big[\boldsymbol{I}_J \otimes z_t(\boldsymbol{\theta},\boldsymbol{\lambda})\big]'\boldsymbol{C}' \frac{\partial^2 \log g}{\partial \boldsymbol{u}\partial \boldsymbol{u}'}\{z_t(\boldsymbol{\theta},\boldsymbol{\lambda})\}\boldsymbol{C}\big[\boldsymbol{I}_J \otimes z_t(\boldsymbol{\theta},\boldsymbol{\lambda})\big]\frac{\partial \boldsymbol{a}(\boldsymbol{x}_t;\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \\
&\quad + \frac{\partial \boldsymbol{a}'(\boldsymbol{x}_t;\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\big[\boldsymbol{I}_J \otimes z_t(\boldsymbol{\theta},\boldsymbol{\lambda})\big]'\boldsymbol{C}'\boldsymbol{G}\big[z_t(\boldsymbol{\theta},\boldsymbol{\lambda})\big]\frac{\partial \boldsymbol{a}(\boldsymbol{x}_t;\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} + \{h_t'(\boldsymbol{\theta},\boldsymbol{\lambda}) \otimes \boldsymbol{I}_p\} A(\boldsymbol{x}_t;\boldsymbol{\theta}),
\end{aligned}
$$

where $G(\boldsymbol{u})$ is the $n \times J$ matrix:

$$
\boldsymbol{G}(\boldsymbol{u}) = \left[\boldsymbol{C}_1' \frac{\partial \log g}{\partial \boldsymbol{u}}(\boldsymbol{u}) \quad \boldsymbol{C}_2' \frac{\partial \log g}{\partial \boldsymbol{u}}(u) \quad \dots \quad \boldsymbol{C}_J' \frac{\partial \log g}{\partial \boldsymbol{u}}(\boldsymbol{u})\right].
$$

Note that the first-order conditions imply $Eh_t'(\boldsymbol{\theta}_0, \boldsymbol{\lambda}_0^*) = 0$. Therefore,

$$
\boldsymbol{A} = E_0\left[-\frac{\partial^2}{\partial \boldsymbol{\vartheta}\partial \boldsymbol{\vartheta}'} l_t(\boldsymbol{\theta}_0, \boldsymbol{\lambda}_0^*)\right] = E_0\left[\begin{pmatrix} \frac{\partial a'(\boldsymbol{x}_t;\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}} \\ \boldsymbol{I}_J \end{pmatrix} \boldsymbol{L} \begin{pmatrix} \frac{\partial a'(\boldsymbol{x}_t;\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}} \\ \boldsymbol{I}_J \end{pmatrix}'\right],
$$

where

$$
\begin{aligned}
\boldsymbol{L} = -E_0\Big\{&\big[\boldsymbol{I}_J \otimes \boldsymbol{\Gamma}(\boldsymbol{\lambda}_0^* - \boldsymbol{\lambda}_0)\boldsymbol{u}_t\big]'\boldsymbol{C}'\frac{\partial^2 \log g}{\partial \boldsymbol{u}\partial \boldsymbol{u}'}\{\boldsymbol{\Gamma}(\boldsymbol{\lambda}_0^* - \boldsymbol{\lambda}_0)\boldsymbol{u}_t\}\boldsymbol{C}\big[\boldsymbol{I}_J \otimes \boldsymbol{\Gamma}(\boldsymbol{\lambda}_0^* - \boldsymbol{\lambda}_0)\boldsymbol{u}_t\big] \\
&+ \big[\boldsymbol{I}_J \otimes \boldsymbol{\Gamma}(\boldsymbol{\lambda}_0^* - \boldsymbol{\lambda}_0)\boldsymbol{u}_t\big]'\boldsymbol{C}'\boldsymbol{G}\big(\boldsymbol{\Gamma}(\boldsymbol{\lambda}_0^* - \boldsymbol{\lambda}_0)\boldsymbol{u}_t\big)\Big\}.
\end{aligned}
$$

10. DERIVATIVES OF FUNCTIONS BASED ON EXPONENTIAL OF MATRICES

10.1. *Derivatives of* $a \to \log g(e^{a\boldsymbol{C}}\boldsymbol{y})$

For $a \in \mathbb{R}$, $\boldsymbol{y} \in \mathbb{R}^n$, $\boldsymbol{C}$ an $n \times n$ matrix, $g : \mathbb{R}^n \to \mathbb{R}^+$ a function,

$$\frac{\partial}{\partial a}(e^{a\boldsymbol{C}}\boldsymbol{y}) = \boldsymbol{C}e^{a\boldsymbol{C}}\boldsymbol{y}, \qquad \frac{\partial^2}{\partial a^2}(e^{a\boldsymbol{C}}\boldsymbol{y}) = \boldsymbol{C}^2 e^{a\boldsymbol{C}}\boldsymbol{y},$$

$$\frac{\partial}{\partial a}\log g(e^{a\boldsymbol{C}}\boldsymbol{y}) = \left[\frac{\partial \log g}{\partial \boldsymbol{u}'}(e^{a\boldsymbol{C}}\boldsymbol{y})\right]\boldsymbol{C}e^{a\boldsymbol{C}}\boldsymbol{y},$$

$$\frac{\partial^2}{\partial a^2}\log g(e^{a\boldsymbol{C}}\boldsymbol{y}) = (\boldsymbol{C}e^{a\boldsymbol{C}}\boldsymbol{y})'\left[\frac{\partial^2 \log g}{\partial \boldsymbol{u}\partial \boldsymbol{u}'}(e^{a\boldsymbol{C}}\boldsymbol{y})\right]\boldsymbol{C}e^{a\boldsymbol{C}}\boldsymbol{y} + (\boldsymbol{C}^2 e^{a\boldsymbol{C}}\boldsymbol{y})'\left[\frac{\partial \log g}{\partial \boldsymbol{u}}(e^{a\boldsymbol{C}}\boldsymbol{y})\right].$$

10.2. *Derivatives of* $\boldsymbol{\theta} \to e^{a(\boldsymbol{\theta})\boldsymbol{C}}\boldsymbol{y}$ *and* $\boldsymbol{\theta} \to \log g(e^{a(\boldsymbol{\theta})\boldsymbol{C}}\boldsymbol{y})$

For $a : \mathbb{R}^p \to \mathbb{R}$ and $\boldsymbol{\theta} \in \mathbb{R}^p$,

$$\frac{\partial}{\partial \boldsymbol{\theta}'}\{e^{a(\boldsymbol{\theta})\boldsymbol{C}}\boldsymbol{y}\} = \boldsymbol{C}e^{a(\boldsymbol{\theta})\boldsymbol{C}}\boldsymbol{y}\frac{\partial a(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'},$$

$$\begin{aligned}\frac{\partial}{\partial \boldsymbol{\theta}}\log g\{e^{a(\boldsymbol{\theta})\boldsymbol{C}}\boldsymbol{y}\} &= \frac{\partial}{\partial a}\log g\{e^{a(\boldsymbol{\theta})\boldsymbol{C}}\boldsymbol{y}\} \cdot \frac{\partial a(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \\ &= \left\{\left[\frac{\partial \log g}{\partial \boldsymbol{u}'}\{e^{a(\boldsymbol{\theta})\boldsymbol{C}}\boldsymbol{y}\}\right]\boldsymbol{C}e^{a(\boldsymbol{\theta})\boldsymbol{C}}\boldsymbol{y}\right\} \cdot \frac{\partial a(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}},\end{aligned}$$

$$\begin{aligned}\frac{\partial^2}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}'}\log g\{e^{a(\boldsymbol{\theta})\boldsymbol{C}}\boldsymbol{y}\} &= \frac{\partial}{\partial a}\log g\{e^{a(\boldsymbol{\theta})\boldsymbol{C}}\boldsymbol{y}\} \cdot \frac{\partial^2 a(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}'} + \frac{\partial^2}{\partial a^2}\log g(e^{a\boldsymbol{C}}\boldsymbol{y}) \cdot \frac{\partial a(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\frac{\partial a(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \\ &= \left\{\left[\frac{\partial \log g}{\partial \boldsymbol{u}'}\{e^{a(\boldsymbol{\theta})\boldsymbol{C}}\boldsymbol{y}\}\right]\boldsymbol{C}e^{a(\boldsymbol{\theta})\boldsymbol{C}}\boldsymbol{y}\right\} \cdot \frac{\partial^2 a(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}'} \\ &\quad + \left\{(\boldsymbol{C}e^{a(\boldsymbol{\theta})\boldsymbol{C}}\boldsymbol{y})'\left[\frac{\partial^2 \log g}{\partial \boldsymbol{u}\partial \boldsymbol{u}'}(e^{a(\boldsymbol{\theta})\boldsymbol{C}}\boldsymbol{y})\right]\boldsymbol{C}e^{a(\boldsymbol{\theta})\boldsymbol{C}}\boldsymbol{y}\right. \\ &\quad \left. + (\boldsymbol{C}^2 e^{a(\boldsymbol{\theta})\boldsymbol{C}}\boldsymbol{y})'\left[\frac{\partial \log g}{\partial \boldsymbol{u}}(e^{a(\boldsymbol{\theta})\boldsymbol{C}}\boldsymbol{y})\right]\right\} \cdot \frac{\partial a(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\frac{\partial a(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}.\end{aligned}$$

In these equalities, "·" indicates the multiplication of a matrix by a scalar.

10.3. *Derivatives of* $\boldsymbol{\theta} \to \exp\{\sum_{j=1}^{J}[a_j(\boldsymbol{x}_t;\boldsymbol{\theta}) + \lambda_j]\boldsymbol{C}_j\}\boldsymbol{y}_t$

Let $z_t(\lambda, \boldsymbol{\theta}) = \exp\{\sum_{j=1}^{J}[a_j(\boldsymbol{x}_t;\boldsymbol{\theta}) + \lambda_j]\boldsymbol{C}_j\}\boldsymbol{y}_t$, where $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_J)' \in \mathbb{R}^J$, $a_j(\cdot)$ are real valued functions with $\boldsymbol{\theta} \in \mathbb{R}^p$, $\boldsymbol{y}_t \in \mathbb{R}^n$, $\boldsymbol{C}_j$ are $n \times n$ matrices. Let $a(\boldsymbol{x}_t;\boldsymbol{\theta}) = (a_1(\boldsymbol{x}_t;\boldsymbol{\theta}), \ldots, a_J(\boldsymbol{x}_t;\boldsymbol{\theta}))'$. For $i = 1, \ldots, J$, let the $n \times 1$ vectors

$$\begin{aligned}z_t^{(i)}(\boldsymbol{\theta}, \boldsymbol{\lambda}) = \exp&\left\{\sum_{j=1}^{i-1}[a_j(\boldsymbol{x}_t;\boldsymbol{\theta}) + \lambda_j]\boldsymbol{C}_j\right\}\boldsymbol{C}_i \exp\{[a_i(\boldsymbol{x}_t;\boldsymbol{\theta}) + \lambda_i]\boldsymbol{C}_i\} \\ &\times \exp\left\{\sum_{j=i+1}^{J}[a_j(\boldsymbol{x}_t;\boldsymbol{\theta}) + \lambda_j]\boldsymbol{C}_j\right\}\boldsymbol{y}_t,\end{aligned}$$

where the first and last sums are replaced by 0 when $i = 1$ and $i = J$, respectively. Let the $n \times J$ block-matrix

$$\boldsymbol{Z}_t(\boldsymbol{\theta}, \boldsymbol{\lambda}) = \big[z_t^{(1)}(\boldsymbol{\theta}, \boldsymbol{\lambda})| \dots |z_t^{(J)}(\boldsymbol{\theta}, \boldsymbol{\lambda})\big].$$

We have

$$\frac{\partial z_t(\boldsymbol{\theta}, \boldsymbol{\lambda})}{\partial \boldsymbol{\theta}'} = \boldsymbol{Z}_t(\boldsymbol{\theta}, \boldsymbol{\lambda}) \frac{\partial a(\boldsymbol{x}_t; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}.$$

When the matrices $\boldsymbol{C}_j$ commute, we have $z_t^{(i)}(\boldsymbol{\theta}, \boldsymbol{\lambda}) = \boldsymbol{C}_i z_t(\boldsymbol{\theta}, \boldsymbol{\lambda})$ and

$$\boldsymbol{Z}_t(\boldsymbol{\theta}, \boldsymbol{\lambda}) = \big[\boldsymbol{C}_1 z_t(\boldsymbol{\theta}, \boldsymbol{\lambda})| \dots |\boldsymbol{C}_J z_t(\boldsymbol{\theta}, \boldsymbol{\lambda})\big] = \boldsymbol{C}\big[\boldsymbol{I}_J \otimes z_t(\boldsymbol{\theta}, \boldsymbol{\lambda})\big],$$

where $\boldsymbol{C} = [\boldsymbol{C}_1| \dots |\boldsymbol{C}_J]$. Thus, when the $\boldsymbol{C}_j$ commute,

$$\frac{\partial z_t(\boldsymbol{\theta}, \boldsymbol{\lambda})}{\partial \boldsymbol{\theta}'} = \boldsymbol{C}\big[\boldsymbol{I}_J \otimes z_t(\boldsymbol{\theta}, \boldsymbol{\lambda})\big] \frac{\partial a(\boldsymbol{x}_t; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}.$$

REFERENCES

BILLINGSLEY, P. (1961): “The Lindeberg–Lévy Theorem for Martingales,” *Proceedings of the American Mathematical Society*, 12, 788–792.[3]

Co-editor Ulrich K. Müller handled this manuscript.

Manuscript received 23 September, 2016; final version accepted 9 July, 2018; available online 10 August, 2018.