Quasi-maximum likelihood estimation and penalized estimation under non-standard conditions

Junichiro Yoshida · Nakahiro Yoshida

Graduate School of Mathematical Sciences, University of Tokyo: 3-8-1 Komaba, Meguro-ku, Tokyo 153-8914, Japan.

Supplementary material

6 Appendix: proof of (22)

Let

\[ K(\lambda, a, b) = \frac{b^\lambda}{2} \int_0^\infty t^{\lambda-1} \exp \left[ -\frac{1}{2} \left( bt + \frac{a}{t} \right) \right] dt \]

Then \( K(\lambda, \delta^2, \gamma^2) = (\gamma \delta)^\lambda K(\gamma \delta) \).

Simply denoted by \( p(x; \lambda, \delta, \gamma) \), the density \( p_{GIG}(x; \lambda, \delta, \gamma) \) is expressed as

\[ p(x; \lambda, \delta, \gamma) = \frac{\gamma^2 \lambda}{2K(\lambda, \delta^2, \gamma^2)} x^{\lambda-1} \exp \left[ -\frac{1}{2} \left( \delta^2 x + \gamma^2 x \right) \right] \quad (x > 0) \]

This model is a curved exponential family:

\[ p(x; \lambda, \delta, \gamma) = \exp \left[ (\lambda - 1) \log x - \frac{\delta^2}{2x} - \frac{\gamma^2}{2} x - \Psi(\lambda, \delta^2, \gamma^2) \right] \quad (x > 0) \]

with the potential \( \Psi(\lambda, a, b) = -\log \frac{b^\lambda}{2K(\lambda, a, b)} \).

We have

\[ \Psi(\lambda, \delta^2, \gamma^2) - \Psi(\lambda^*, 0, (\gamma^*)^2) = \log \int_0^\infty \exp \left[ (\lambda - 1) \log x - \frac{\delta^2}{2x} - \frac{\gamma^2}{2} x \right] dx \]

\[ - \log \int_0^\infty \exp \left[ (\lambda^* - 1) \log x - \frac{0}{2x} - \frac{(\gamma^*)^2}{2} x \right] dx \]

\[ = D[\Delta(\lambda, \delta^2, \gamma^2)] + \int_0^1 (1-s)H(s, \lambda, \delta^2, \gamma^2)[(\Delta(\lambda, \delta^2, \gamma^2)^2)] ds. \]

Here

\[ \Delta(\lambda, \delta^2, \gamma^2) = \left( \frac{\lambda - \lambda^*}{\delta^2}, \frac{\gamma^2}{(\gamma^*)^2} \right), \]

\[ D = (\partial_{(\lambda, a, b)} \Psi) (\lambda^*, 0, (\gamma^*)^2) = \begin{pmatrix} E[\log \xi_0] \\ -2^{-1}E[\xi_0^{-1}] \end{pmatrix} \]

\[ -2^{-1}E[\xi_0] \]
and

\[ H(s, \lambda, \delta^2, \gamma^2) = (s^2 \lambda, s^4, s^3 + (1-s)\gamma^2) \]

\[ = \left( \frac{\vartheta}{\vartheta^3} \right) (s^2 \lambda + (1-s)\lambda^*, s^4 \delta^2, s^3 + (1-s)(\gamma^*)^2) \]

\[ = \begin{pmatrix}
-2^{-1} \text{Cov}[\log \xi] & -2^{-1} \text{Cov}[\log \xi, \xi^{-1}] & -2^{-1} \text{Cov}[\log \xi, \xi_0]
-2^{-1} \text{Cov}[\log \xi, \xi_0] & 4^{-1} \text{Var}[\xi^{-1}] & 4^{-1} \text{Cov}[\xi^{-1}, \xi_0]
-2^{-1} \text{Cov}[\log \xi, \xi_0] & 4^{-1} \text{Cov}[\xi^{-1}, \xi_0] & 4^{-1} \text{Var}[\xi_0]
\end{pmatrix}, \]

with \( \xi = \xi(s\lambda + (1-s)\lambda^*, s\delta^2, s\gamma^2 + (1-s)(\gamma^*)^2) \), where \( \xi(s\lambda, \delta, \gamma) \) denotes a random variable such that \( \xi(s\lambda, \delta, \gamma) \sim \text{GIG}(\lambda, \delta, \gamma) \).

Let \( a_n = \text{diag}[n^{-1/2}, n^{-1/4}, n^{-1/2}] \), and define \( U_n \) and \( U \) as (5) and (7), respectively, for \( \Theta := [\lambda, \delta, \gamma] \times [0, \delta, \gamma] \) and \( \theta^* := (\lambda^*, 0, \gamma^*) \). For \( H_n(\theta) = \sum_{j=1}^n \log p(X_j; \theta) \) (\( \theta = (\lambda, \delta, \gamma) \in \Theta \)) and \( u = (u_1, u_2, u_3) \in U_n \), we obtain

\[ \log Z_n(u) = H_n(\theta^* + a_n u) - H_n(\theta^*) \]

\[ = u_1 n^{-1/2} \sum_{j=1}^n \log \hat{X}_j - u_2 n^{-1/2} \sum_{j=1}^n \log \hat{X}_j - u_3 \gamma^* n^{-1/2} \sum_{j=1}^n \hat{X}_j 
\]

\[ -n \int_0^1 (1-s)H(s, \lambda^* + n^{-1/2}u_1, n^{-1/2}u_2, (\gamma^* + n^{-1/2}u_3)^2) \]

\[ \times \left[ (n^{-1/2}u_1, n^{-1/2}u_2, (\gamma^* + n^{-1/2}u_3)^2 - (\gamma^*)^2) \right] ds \]

\[ -u_2^2 - n^{-1} \sum_{j=1}^n \hat{X}_j, \]

where we are writing \( \hat{F}(X_j) = F(X_j) - E[F(X_j)] \) for a function \( F(X_j) \) of a random variable \( X_j \) satisfying \( X_j \sim \text{GIG}(\lambda^*, 0, \gamma^*) = I(\lambda^*, (\gamma^*)^2/2) \). Then it is possible to write it as

\[ \log Z_n(u) = u_1 n^{-1/2} \sum_{j=1}^n \log \hat{X}_j - u_2 n^{-1/2} \sum_{j=1}^n \log \hat{X}_j - u_3 \gamma^* n^{-1/2} \sum_{j=1}^n \hat{X}_j 
\]

\[ -\frac{1}{2} \text{C}[u_1^2, u_2^2, u_3^2] + r_n(u) \]

with the positive-definite covariance matrix

\[ C = \begin{pmatrix}
\text{Var}[\log \xi_0] & -2^{-1} \text{Cov}[\log \xi_0, \xi_0^{-1}] & -\gamma^* \text{Cov}[\log \xi_0, \xi_0]
-2^{-1} \text{Cov}[\log \xi_0, \xi_0^{-1}] & 4^{-1} \text{Var}[\xi_0^{-1}] & 2^{-1} \gamma^* \text{Cov}[\xi_0, \xi_0^{-1}]
-\gamma^* \text{Cov}[\log \xi_0, \xi_0] & 2^{-1} \gamma^* \text{Cov}[\xi_0^{-1}, \xi_0] & (\gamma^*)^2 \text{Var}[\xi_0]
\end{pmatrix}, \]

and the term \( r_n(u) \) satisfying

\[ \sup_{u \in U_n} \frac{|r_n(u)|}{1 + |u_1|^2 + |u_2|^2 + |u_3|^2} = O_p(n^{-1/2}). \]
Then, Condition [A1] is verified by the estimate

$$\lim_{R \to \infty} \limsup_{n \to \infty} P \left[ M_n \geq \left( 2^{-1} \lambda_{\text{min}}[C] + O_p(n^{-1/2}) \right) R \right] = 0$$

for $M_n = \left| n^{-1/2} \sum_{j=1}^n \log \bar{X}_j \right| + \left| n^{-1/2} \sum_{j=1}^n 2^{-1} \bar{X}_j^{-1} \right| + \left| n^{-1/2} \gamma^{*} \sum_{j=1}^n \bar{X}_j \right|$. Condition [A2] is satisfied with

$$V_n(u) = \exp \left( u_1 n^{-1/2} \sum_{j=1}^n \log \bar{X}_j - u_2 n^{-1/2} \sum_{j=1}^n 2^{-1} \bar{X}_j^{-1} ight. \left. - u_3 \gamma^{*} n^{-1/2} \sum_{j=1}^n \bar{X}_j - \frac{1}{2} C[(u_1, u_2, u_3) \otimes 2] \right),$$

$$Z(u) = \exp \left( \Delta \cdot (u_1, u_2, u_3) - \frac{1}{2} C[(u_1, u_2, u_3) \otimes 2] \right)$$

with a three-dimensional random vector $\Delta \sim N_3(0, C)$. Now from Example 1, [A3] holds, and we obtain $U = \mathbb{R} \times [0, \infty) \times \mathbb{R}$. Condition [A4] obviously holds. Therefore, Theorem 1 concludes that the MLE $\left( \hat{\lambda}_n, \hat{\delta}_n, \hat{\gamma}_n \right)$ admits (22). \qed