Web-based Supplementary Materials for
Selection Consistency of EBIC for GLIM with Non-canonical Links and Diverging Number of Parameters

by Shan Luo and Zehua Chen

Verification of condition C6 for GLIM with non-canonical link functions

In this supplementary document, we verify condition C6 for some common GLIMs with non-canonical link functions while assuming that $\sigma_i^2$ are bounded away from 0 and from above. These GLIMs were considered in Wedderburn (1976). In particular, we consider the following exponential families and their corresponding link functions:

1. Poisson Distribution: $\eta = \ln(\mu)$, $\mu^\gamma$ where $0 < \gamma < 1$;
2. Binomial Distribution: $\eta = \mu$, $\arcsin(\mu)$, $\ln(\frac{\mu}{1-\mu})$, $\ln(-\ln(1-\mu))$, $\Phi^{-1}(\mu)$.
3. Gamma Distribution ($G(1, \mu)$): $\eta = \ln \mu$, $\mu^\gamma$ where $-1 \leq \gamma < 0$.

The corresponding function $\theta = h(\eta)$ for the models above are as follows:

1. Poisson Distribution: $\theta = \eta$, $\frac{1}{\gamma} \ln \eta$ where $0 < \gamma < 1$;
2. Binomial Distribution: $\theta = \ln \frac{\eta}{1-\eta}$, $\ln \frac{\sin(\eta)}{1-\sin(\eta)}$, $\eta$, $\ln(\exp(e^\eta) - 1)$, $\ln \left( \frac{\Phi(\eta)}{1-\Phi(\eta)} \right)$.
3. Gamma Distribution: $\theta = -e^{-\eta}$, $-\eta^{-\frac{\gamma}{\gamma+1}}$.

1 Poisson Distribution

The link $\eta = \mu^\gamma$ where $0 < \gamma < 1$ and $\mu \in [a, b]$. In this situation,

$$h'(\eta) = \frac{1}{\gamma \eta}, \quad h''(\eta) = -\frac{1}{\gamma \eta^2}, \quad \sigma^2 = \eta^{\frac{2}{\gamma+1}}.$$
Hence under the assumption, \( \forall 1 \leq i \leq n, \)

\[
|h'(x_i^\tau \beta_0)| \in \left[ \frac{1}{\gamma b^{2\gamma}}, \frac{1}{\gamma a^{2\gamma}} \right], \sigma_i^2 \in [a, b], |h''(x_i^\tau \beta_0)| \in \left[ \frac{1}{\gamma b^{2\gamma}}, \frac{1}{\gamma a^{2\gamma}} \right],
\]

\[
\frac{x_{i,j}^2 (h'(x_i^\tau \beta_0))^2}{\sum_{i=1}^{n} \sigma_i^2 x_{i,j}^2 (h'(x_i^\tau \beta_0))^2} = O \left( \frac{x_{i,j}^2}{\sum_{i=1}^{n} x_{i,j}^2} \right)
\]

\[
\frac{(h''(x_i^\tau \beta_0))^2}{\sum_{i=1}^{n} \sigma_i^2 (h''(x_i^\tau \beta_0))^2} = O \left( \frac{(h''(x_i^\tau \beta_0))^2}{\sum_{i=1}^{n} (h''(x_i^\tau \beta_0))^2} \right).
\]

when \( 0 < a < b < +\infty, \) (A6) is true when \( \max_{1 \leq i \leq n} \max_{1 \leq j \leq p} \frac{x_{i,j}^2}{\sum_{i=1}^{n} x_{i,j}^2} = o(n^{-1/3}). \)

2 \ Binomial Distribution

For binomial distribution, \( \sigma_i^2 = \mu_i(1 - \mu_i) = \frac{e^{\theta_i}}{(1 + e^{\theta_i})^2}. \) Here we assume

\[
\min_{1 \leq i \leq n} (\mu_i \wedge (1 - \mu_i)) \geq c \text{ where } 0 < c \leq 1/2.
\]

This implies, \( c^2 \leq \min_{1 \leq i \leq n} \sigma_i^2 \leq \max_{1 \leq i \leq n} \sigma_i^2 \leq 1/4. \) Therefore,

\[
\frac{x_{i,j}^2 (h'(x_i^\tau \beta_0))^2}{\sum_{i=1}^{n} \sigma_i^2 x_{i,j}^2 (h'(x_i^\tau \beta_0))^2} = O \left( \frac{x_{i,j}^2 (h'(x_i^\tau \beta_0))^2}{\sum_{i=1}^{n} x_{i,j}^2 (h'(x_i^\tau \beta_0))^2} \right)
\]

\[
\frac{(h''(x_i^\tau \beta_0))^2}{\sum_{i=1}^{n} \sigma_i^2 (h''(x_i^\tau \beta_0))^2} = O \left( \frac{(h''(x_i^\tau \beta_0))^2}{\sum_{i=1}^{n} (h''(x_i^\tau \beta_0))^2} \right).
\]

(1) \( \mu = \eta, 0 < \eta < 1: \)

\[
h'(\eta) = \frac{1}{\eta (1 - \eta)}, \quad h''(\eta) = \frac{2\eta - 1}{\eta^2 (1 - \eta)^2}, \quad \sigma^2 = \eta (1 - \eta).
\]

2 Binomial Distribution

For binomial distribution, \( \sigma_i^2 = \mu_i(1 - \mu_i) = \frac{e^{\theta_i}}{(1 + e^{\theta_i})^2}. \) Here we assume

\[
\min_{1 \leq i \leq n} (\mu_i \wedge (1 - \mu_i)) \geq c \text{ where } 0 < c \leq 1/2.
\]

This implies, \( c^2 \leq \min_{1 \leq i \leq n} \sigma_i^2 \leq \max_{1 \leq i \leq n} \sigma_i^2 \leq 1/4. \) Therefore,

\[
\frac{x_{i,j}^2 (h'(x_i^\tau \beta_0))^2}{\sum_{i=1}^{n} \sigma_i^2 x_{i,j}^2 (h'(x_i^\tau \beta_0))^2} = O \left( \frac{x_{i,j}^2 (h'(x_i^\tau \beta_0))^2}{\sum_{i=1}^{n} x_{i,j}^2 (h'(x_i^\tau \beta_0))^2} \right)
\]

\[
\frac{(h''(x_i^\tau \beta_0))^2}{\sum_{i=1}^{n} \sigma_i^2 (h''(x_i^\tau \beta_0))^2} = O \left( \frac{(h''(x_i^\tau \beta_0))^2}{\sum_{i=1}^{n} (h''(x_i^\tau \beta_0))^2} \right).
\]

(1) \( \mu = \eta, 0 < \eta < 1: \)

\[
h'(\eta) = \frac{1}{\eta (1 - \eta)}, \quad h''(\eta) = \frac{2\eta - 1}{\eta^2 (1 - \eta)^2}, \quad \sigma^2 = \eta (1 - \eta).
\]
Under assumption (1),

\[ 4 \leq h'(x_i^\top \beta_0) \leq \frac{1}{c^2}; \quad |h''(x_i^\top \beta_0)| \leq \frac{1-2c}{c^4} \]

for all \(1 \leq i \leq n\). (A.6) holds when

\[
\max_{1 \leq j \leq p} \max_{1 \leq i \leq n} \left\{ \frac{x_{i,j}^2}{n \sum_{i=1}^{n} x_{i,j}^2} \right\} = o(n^{-1/3}).
\]

(2) \( \eta = \arcsin \mu : \)

\[
h'(\eta) = \frac{\cos \eta}{\sin \eta (1 - \sin \eta)}, \quad h''(\eta) = \frac{\sin \eta}{1 - \sin \eta} - \frac{\cos^2 \eta}{\sin^2 \eta}, \quad \sigma^2 = \sin(1 - \sin \eta).
\]

Under assumption (1),

\[
4\sqrt{2c - c^2} \leq \left| h'(x_i^\top \beta_0) \right| \leq \frac{\sqrt{1-c^2}}{c^2};
\]

\[
\frac{3c - c^2 - 1}{(1-c)^2c} \leq \left| h''(x_i^\top \beta_0) \right| \leq \frac{1-c^2-c}{c^2(1-c)}
\]

for all \(1 \leq i \leq n\). (A.6) holds when

\[
\max_{1 \leq j \leq p} \max_{1 \leq i \leq n} \left\{ \frac{x_{i,j}^2}{n \sum_{i=1}^{n} x_{i,j}^2} \right\} = o(n^{-1/3}).
\]

(3) \( \eta = g(\mu) = \ln \{ -\ln(1-\mu) \} \) or \( \eta = g(\mu) = \ln \{ -\ln(\mu) \} \).

For the first link function, complementary log-log link, we have

\[
\theta = \ln\left( \frac{\mu}{1-\mu} \right) = h(\eta) = \ln \{ \exp(\eta) - 1 \}, \quad \sigma^2 = \frac{\exp(\eta) - 1}{\exp(2\eta)}.
\]

(2)

Therefore, the first and second order derivatives of \( h(\cdot) \) are

\[
h'(\eta) = \frac{e^{\eta} + e^{-\eta}}{e^{\eta} - 1}; \quad h''(\eta) = \frac{e^{\eta} + e^{-\eta}[e^{2\eta} - e^{2\eta} - 1]}{(e^{\eta} - 1)^2}.
\]

(3)
It is easy to see that \( e^n \leq h'(\eta) \leq e^n \). Now let us look at \( h''(\eta) \). It is straightforward that \(|h''(\eta)| \leq |h'(\eta)| \leq e^n \). Consider the function \( f(x) = \frac{e^x(x - 1)}{(e^x - 1)^2} \) on \((0, +\infty)\). Because

\[
\lim_{x \to 0} f(x) = \lim_{x \to 1} \frac{x^2/2}{x^2} = \frac{1}{2}; \quad \lim_{x \to +\infty} f(x) = \lim_{x \to +\infty} \frac{1 - x}{e^x} = 1.
\]

(4)

Therefore, there exists a positive constant \( C_1, C_2 \) independent of \( x \) such that \( C_1 \leq f(x) \leq C_2 \). That is, \( C_1 e^n \leq h''(\eta) \leq C_2 e^n \). When \( \sigma^2_i \in [a, b] \) for some \( 0 < a \leq b \leq 1/4 \), for \( 1 \leq i \leq n \), we have

\[
\frac{1 + \sqrt{1 - 4b}}{2b} \leq \exp(e^n) \leq \frac{1 + \sqrt{1 - 4a}}{2a} \quad \text{or} \quad \frac{1 - \sqrt{1 - 4a}}{2a} \leq \exp(e^n) \leq \frac{1 - \sqrt{1 - 4b}}{2b}.
\]

That is, \(|h'(\eta)|\) and \(|h''(\eta)|\) are both bounded away from 0 and finite. (A.6) holds when

\[
\max_{1 \leq j \leq p_n} \max_{1 \leq i \leq n} \left\{ \frac{x_{i,j}^2}{\sum_{i=1}^{n} x_{i,j}^2} \right\} = o(n^{-1/3}).
\]

The same argument applies to the second link function by changing \( \eta \) to \(-\eta\).

(4) \( \eta = \Phi^{-1}(\mu) : \)

\[
h'(\eta) = \frac{f(\eta)}{\Phi(\eta)(1 - \Phi(\eta))}, \quad h''(\eta) = \frac{f'(\eta)}{\Phi(\eta)(1 - \Phi(\eta))} + f^2(\eta)\left[\frac{1}{(1 - \Phi(\eta))^2} - \frac{1}{\Phi^2(\eta)}\right]
\]

\[
\sigma^2 = \Phi(\eta)(1 - \Phi(\eta))
\]

Under assumption (1), \( \Phi^{-1}(c) \leq |x^T \beta| \leq \Phi^{-1}(1 - c) \). Note that for

\[
1 - \Phi(t) \leq \frac{f(t)}{t}, \quad \forall t > 0
\]
therefore, we have

\[ 4c\Phi^{-1}(c) \leq 4f(x_i^*\beta_0) \leq \left\lfloor \frac{1}{c^2} f(x_i^*\beta_0) \right\rfloor \leq \frac{1}{\sqrt{2\pi c^2}}; \]

\[ 4f'(x_i^*\beta_0) \leq \left| \frac{f'(x_i^*\beta_0)}{\Phi(x_i^*\beta_0)(1 - \Phi(x_i^*\beta_0))} \right| \leq \frac{1}{c^2} f'(x_i^*\beta_0) \leq \frac{\Phi^{-1}(1-c)}{\sqrt{2\pi c^2}}; \]

\[ \left| f^2(x_i^*\beta_0) \frac{1}{(1 - \Phi(x_i^*\beta_0))^2} - \frac{1}{\Phi^2(x_i^*\beta_0)} \right| \leq \frac{|2c - 1|}{c^2(1-c)^2} f^2(x_i^*\beta_0) \leq \frac{|2c - 1|}{2\pi c^2(1-c)^2}. \]

for all \( 1 \leq i \leq n \). (A.6) holds when

\[ \max_{1 \leq j \leq p_n} \max_{1 \leq i \leq n} \left\lfloor \frac{n^2 x_{i,j}^2}{\sum_{i=1}^n x_{i,j}^2} \right\rfloor = o(n^{-1/3}). \]

3 Gamma Distribution

(1) \( \eta = \ln(\mu) : h'(\eta) = e^{-\eta}, h''(\eta) = -e^{-\eta}, \sigma^2 = e^{2\eta} \). When \( \sigma_i^2 \) is away from 0 and finite, \( |h'|, |h''| \) are bounded. (A.6) holds when

\[ \max_{1 \leq j \leq p_n} \max_{1 \leq i \leq n} \left\lfloor \frac{n^2 x_{i,j}^2}{\sum_{i=1}^n x_{i,j}^2} \right\rfloor = o(n^{-1/3}). \]

(2) \( \eta = \mu^\gamma \) where \(-1 \leq \gamma < 0 \). Let \( \hat{\gamma} = -\frac{1}{\gamma} \), then \( 0 < \hat{\gamma} \leq 1 \). Then

\[ h'(\eta) = -\hat{\gamma}\eta^{\hat{\gamma}-1}, h''(\eta) = \hat{\gamma}(1 - \hat{\gamma})\eta^{\hat{\gamma}-2}, \sigma^2 = \eta^{2\hat{\gamma}}. \]

When \( \sigma_i^2 \) is away from 0 and finite, \( |h'|, |h''| \) are bounded. (A.6) holds when

\[ \max_{1 \leq j \leq p_n} \max_{1 \leq i \leq n} \left\lfloor \frac{n^2 x_{i,j}^2}{\sum_{i=1}^n x_{i,j}^2} \right\rfloor = o(n^{-1/3}). \]