The Extended Bayesian Information Criterion for Gaussian Graphical Models

Rina Foygel
University of Chicago
rina@uchicago.edu

Mathias Drton
University of Chicago
drton@uchicago.edu

A Supplementary materials

This document gives the proof for Lemma 1 in the main paper, and fills in the details of the proofs of Theorem 3 and Theorem 4 in the main paper.

Lemma 1. *For any $\lambda \in (0, 1)$, for any $n \geq 4\lambda^{-2} + 1$,*

$$P\{\chi_n^2 < n(1-\lambda)\} \leq \frac{1}{\lambda\sqrt{\pi(n-1)}} e^{-\frac{n-1}{2}(\lambda - \log(1+\lambda))} \quad .$$

Remark 1. We note that some lower bound on n is intuitively necessary in order to be able to bound the 'left tail', because the mode of the χ_n^2 distribution is at $x = n - 2$ (for $n \geq 2$). If λ is very close to zero, then the 'left tail' ($\chi_n^2 \in [0, n(1-\lambda)]$) actually includes the mode $x = n - 2 \leq n(1-\lambda)$; therefore, we could not hope to get an exponentially small probability for being in the tail. However, this intuitive explanation suggests that we should have $n \geq \mathbf{O}(\lambda^{-1})$; perhaps the bound in this lemma could be tightened.

We first prove a preliminary lemma:

Lemma A.1. *For any $\lambda > 0$, for any $n \geq 4\lambda^{-2} + 1$,*

$$P\{\chi_{n+1}^2 < (n+1)(1-\lambda)\} \leq P\{\chi_n^2 \leq n(1-\lambda)\} \quad .$$

Proof. Let f_n denote the density function for χ_n^2, and let $\tilde{f}_n$ denote the density function for $\frac{1}{n}\chi_n^2$. Then, using $y = x/n$, we get:

$$f_n(x) = \frac{1}{2^{n/2}\Gamma(n/2)} x^{n/2-1} e^{-x/2} \;\Rightarrow\; \tilde{f}_n(y) = \frac{1}{2^{n/2}\Gamma(n/2)} y^{n/2-1} e^{-ny/2} n^{n/2} \quad .$$

So,

$$\tilde{f}_{n+1}(y) = \tilde{f}_n(y) \times \left[\sqrt{\frac{n+1}{2}} \frac{\Gamma(n/2)}{\Gamma((n+1)/2)} \left(\frac{n+1}{n}\right)^{n/2} \sqrt{ye^{-y}}\right] \quad .$$

First, note that ye^{-y} is an increasing function for $y < 1$, and therefore

$$y \in [0, 1-\lambda] \;\Rightarrow\; ye^{-y} \leq (1-\lambda)e^{-(1-\lambda)} \leq e^{-1}\left(1 - \frac{\lambda^2}{2}\right) \quad .$$

(Here the last inequality is from the Taylor series). Next, since $\log\Gamma(x)$ is a convex function (where $x > 0$), and since $\Gamma((n+1)/2) = \Gamma((n-1)/2) \times \frac{n-1}{2}$, we see that

$$\frac{\Gamma((n+1)/2)}{\Gamma(n/2)} \geq \sqrt{\frac{n-1}{2}} \quad .$$

Finally, it is a fact that $(1 + \frac{1}{n})^n \leq e$. Putting the above bounds together, and assuming that $y \in [0, 1-\lambda]$, we obtain

$$\tilde{f}_{n+1}(y) \leq \tilde{f}_n(y) \times \left[\sqrt{\frac{n+1}{2}} \sqrt{\frac{2}{n-1}} \sqrt{e} \sqrt{e^{-1}\left(1 - \frac{\lambda^2}{2}\right)}\right]$$

$$= \tilde{f}_n(y) \times \left[\sqrt{\frac{n+1}{n-1}} \sqrt{1 - \frac{\lambda^2}{2}} \right] \quad .$$

Since we require $n \geq 4\lambda^{-2} + 1$, the quantity in the brackets is at most 1, and so

$$\tilde{f}_{n+1}(y) \leq \tilde{f}_n(y) \, \forall \, y \in [0, 1-\lambda] \quad .$$

Therefore,

$$P\left\{ \frac{1}{n+1} \chi^2_{n+1} < (1-\lambda) \right\} \leq P\left\{ \frac{1}{n} \chi^2_n < (1-\lambda) \right\} \quad .$$

□

Now we prove Lemma 1.

Proof. First suppose that n is even. Let f_n denote the density function of the χ^2_n distribution. From Cai [10] (as cited in the paper), if $n > 2$,

$$P\{\chi^2_n < x\} = 1 - 2f_n(x) - P\{\chi^2_{n-2} > x\} = -2f_n(x) + P\{\chi^2_{n-2} < x\} \quad .$$

Iterating this identity, we get

$$\begin{aligned}
P\{\chi^2_n < x\} &= P\{\chi^2_2 < x\} - 2f_n(x) - 2f_{n-2}(x) - \cdots - 2f_4(x) \\
&= 1 - e^{-\frac{x}{2}} - 2 \sum_{k=1}^{n/2-1} f_{2k+2}(x) \\
&= 1 - e^{-\frac{x}{2}} - 2 \sum_{k=1}^{n/2-1} \frac{1}{2^{k+1}\Gamma(k+1)} x^k e^{-\frac{x}{2}} \\
&= 1 - e^{-\frac{x}{2}} \left(\sum_{k=0}^{n/2-1} \frac{1}{2^k k!} x^k \right) \\
&= 1 - e^{-\frac{x}{2}} \left(\sum_{k=0}^{\infty} \frac{(x/2)^k}{k!} - \sum_{k=n/2}^{\infty} \frac{(x/2)^k}{k!} \right) \\
&= 1 - e^{-\frac{x}{2}} \left(e^{\frac{x}{2}} - \sum_{k=n/2}^{\infty} \frac{(x/2)^k}{k!} \right) \\
&= e^{-\frac{x}{2}} \sum_{k=n/2}^{\infty} \frac{(x/2)^k}{k!} \quad .
\end{aligned}$$

Now set $x = n(1-\lambda)$ for $\lambda \in (0,1)$. We obtain

$$\begin{aligned}
P\{\chi^2_n < x\} &= e^{-\frac{n(1-\lambda)}{2}} \sum_{k=n/2}^{\infty} \frac{(n(1-\lambda)/2)^k}{k!} \\
&\leq e^{-\frac{n(1-\lambda)}{2}} \frac{(n/2)^{n/2}}{(n/2)!} \sum_{k=n/2}^{\infty} (1-\lambda)^k \\
&= e^{-\frac{n(1-\lambda)}{2}} \frac{(n/2)^{n/2}}{(n/2)!} \frac{(1-\lambda)^{n/2}}{\lambda} \quad .
\end{aligned}$$

By Stirling's formula,

$$\frac{(n/2)^{n/2}}{(n/2)!} \leq \frac{e^{\frac{n}{2}}}{\sqrt{\pi n}} \quad ,$$

and so,

$$P\{\chi_n^2 < n(1-\lambda)\} \le e^{-\frac{n(1-\lambda)}{2}} \frac{e^{\frac{n}{2}}}{\sqrt{\pi n}} \frac{(1-\lambda)^{n/2}}{\lambda} = \frac{1}{\lambda\sqrt{\pi n}} e^{\frac{n}{2}(\lambda+\log(1-\lambda))} \quad .$$

This is clearly sufficient to prove the desired bound in the case that n is even. Next we turn to the odd case; let n be odd. First observe that if $\lambda > 1$, the statement is trivial, while if $\lambda \le 1$, then $n \ge 4\lambda^{-2} + 1 \ge 5$, therefore $n-1$ is positive. By Lemma A.1 and the expression above,

$$P\{\chi_n^2 < n(1-\lambda)\} \le P\{\chi_{n-1}^2 \le (n-1)(1-\lambda)\} \le \frac{1}{\lambda\sqrt{\pi(n-1)}} e^{\frac{n-1}{2}(\lambda+\log(1-\lambda))} \quad .$$

□

Next we turn to the theorems in the paper. Recall the assumptions made in the paper: we assume the following, where $\epsilon_0, \epsilon_1 > 0$, $C \ge \sigma_{max}^2 \lambda_{max}$, $\kappa = \log_n p$, and $\gamma_0 = \gamma - (1 - \frac{1}{4\kappa})$:

$$\frac{(p+2q)\log p}{n} \times \frac{\lambda_{max}^2}{\theta_0^2} \le \frac{1}{3200 \max\{1+\gamma_0, \left(1+\frac{\epsilon_1}{2}\right) C^2\}} \quad , \tag{1}$$

$$2(\sqrt{1+\gamma_0} - 1) - \frac{\log\log p + \log(4\sqrt{1+\gamma_0}) + 1}{2\log p} \ge \epsilon_0 \quad . \tag{2}$$

Lemmas A.2 and A.3 below are sufficient to fill in the details of Theorem 3 in the paper.

Lemma A.2. *With probability at least* $1 - \frac{1}{\sqrt{\pi \log p}} e^{-\epsilon_1 \log p}$*, the following holds for all edges* $\mathbf{e}$ *in the complete graph:*

$$(s_n(\Theta_0))_{\mathbf{e}}^2 \le 6\sigma_{max}^4 (2+\epsilon_1) n \log p \quad .$$

Proof. Fix some edge $\mathbf{e} = \{j, k\}$. Then

$$(s_n(\Theta_0))_{(j,k)} = \frac{n}{2}(\Sigma_0)_{jk} - \frac{1}{2} X_j^T X_k = -\frac{1}{2} \sum_{i=1}^{n} ((X_j)_i (X_k)_i - (\Sigma_0)_{jk}) \quad .$$

Write $Y_j = ((\Sigma_0)_{jj})^{-1} X_j$, $Y_k = ((\Sigma_0)_{kk})^{-1} X_k$, $\rho = ((\Sigma_0)_{jj}(\Sigma_0)_{kk})^{-1} (\Sigma_0)_{jk} = \mathrm{corr}(Y_j, Y_k)$. Then

$$(s_n(\Theta_0))_{(j,k)} = -\frac{1}{2} (\Sigma_0)_{jj} (\Sigma_0)_{kk} \sum_{i=1}^{n} ((Y_j)_i (Y_k)_i - \rho) \quad .$$

By Lemma 2 in the paper, there are some independent $A, B \sim \chi_n^2$ such that

$$(s_n(\Theta_0))_{(j,k)} = -\frac{1}{2} (\Sigma_0)_{jj} (\Sigma_0)_{kk} \left[\left(\frac{1+\rho}{2}\right)(A-n) - \left(\frac{1-\rho}{2}\right)(B-n) \right] \quad .$$

There are $\binom{p}{2} \le \frac{1}{2}p^2$ edges in the complete graph. Therefore, by the union bound, it will suffice to show that, with probability at least $1 - (\frac{1}{2}p^2)^{-1} \frac{1}{\sqrt{\pi \log p}} e^{-\epsilon_1 \log p}$,

$$\frac{1}{4} \sigma_{max}^4 \left[\left(\frac{1+\rho}{2}\right)(A-n) - \left(\frac{1-\rho}{2}\right)(B-n) \right]^2 \le 6\sigma_{max}^4 (2+\epsilon_1) n \log p \quad .$$

Suppose this bound does not hold. Then

$$\left| \left(\frac{1+\rho}{2}\right)(A-n) \right| > \sqrt{6(2+\epsilon_1) n \log p} \ \text{ or } \ \left| \left(\frac{1-\rho}{2}\right)(B-n) \right| > \sqrt{6(2+\epsilon_1) n \log p} \quad .$$

Since $\rho \in [-1, 1]$, this implies that

$$|A-n| > \sqrt{6(2+\epsilon_1) n \log p} \ \text{ or } \ |B-n| > \sqrt{6(2+\epsilon_1) n \log p} \quad .$$

Since $A \overset{\mathcal{D}}{=} B$, it will suffice to show that with probability at least $1 - p^{-2} \frac{1}{\sqrt{\pi \log p}} e^{-\epsilon_1 \log p}$,

$$|A-n| \le \sqrt{6(2+\epsilon_1) n \log p} \quad .$$

Write $\lambda = \sqrt{6(2+\epsilon_1)\frac{\log p}{n}}$. Observe that, by assumption (1), $\lambda \le \frac{1}{2}$ and $n \ge 3$; therefore (by Taylor series),

$$\frac{n}{2}(\lambda - \log(1+\lambda)) \ge \frac{n}{2}\left(\frac{\lambda^2}{2} - \frac{\lambda^3}{3}\right) \ge \frac{n}{2}\cdot\frac{\lambda^2}{3} = (2+\epsilon_1)\log p \ \text{, and}$$

$$-\frac{n-1}{2}(\lambda + \log(1-\lambda)) \ge \frac{n-1}{2}\left(\frac{\lambda^2}{2}\right) \ge \frac{n}{2}\cdot\frac{\lambda^2}{3} = (2+\epsilon_1)\log p \ .$$

Furthermore,

$$\lambda\sqrt{n-1} = \sqrt{6(2+\epsilon_1)\log p \times \frac{n-1}{n}} \ge \sqrt{\log p} \ .$$

By (CSB) from the paper,

$$P\{A - n > \sqrt{6(2+\epsilon_1)n\log p}\} = P\{A > n(1+\lambda)\} \le \frac{1}{\lambda\sqrt{\pi n}}e^{-\frac{n}{2}(\lambda - \log(1+\lambda))}$$

$$\le \frac{1}{\lambda\sqrt{\pi(n-1)}}e^{-(2+\epsilon_1)\log p} \le \frac{1}{\sqrt{\pi\log p}}e^{-(2+\epsilon_1)\log p} \ ,$$

and also,

$$P\{A - n < -\sqrt{6(2+\epsilon_1)n\log p}\} = P\{A < n(1-\lambda)\} \le \frac{1}{\lambda\sqrt{\pi(n-1)}}e^{\frac{n-1}{2}(\lambda + \log(1-\lambda))}$$

$$\le \frac{1}{\lambda\sqrt{\pi(n-1)}}e^{-(2+\epsilon_1)\log p} \le \frac{1}{\sqrt{\pi\log p}}e^{-(2+\epsilon_1)\log p} \ .$$

This gives the desired result. $\square$

Lemma A.3. *Recall that, in the proof of Theorem 3 in the paper, we showed that*

$$l_n(\Theta) - l_n(\Theta_0) \le \sqrt{\theta_0^2(p+2q)\times 6\sigma_{max}^4(2+\epsilon_1)n\log p} - \frac{1}{2}\theta_0^2 \times \frac{n}{2}(2\lambda_{max})^{-2} \ .$$

Then this implies that

$$l_n(\Theta) - l_n(\Theta_0) \le -2q(\log p)(1+\gamma_0) \ .$$

Proof. It is sufficient to show that

$$\sqrt{\theta_0^2(p+2q)\times 6\sigma_{max}^4(2+\epsilon_1)n\log p} - \frac{1}{2}\theta_0^2 \times \frac{n}{2}(2\lambda_{max})^{-2} \le -(p+2q)(\log p)(1+\gamma_0) \ .$$

We rewrite this as

$$\sqrt{A \times n^2\theta_0^4\lambda_{max}^{-2}6\sigma_{max}^4(2+\epsilon_1)} - \frac{1}{2}\theta_0^2 \times \frac{n}{2}(2\lambda_{max})^{-2} \le -A \times n \times \theta_0^2\lambda_{max}^{-2}(1+\gamma_0) \ ,$$

where

$$A = \frac{(p+2q)\log p}{n} \times \frac{\lambda_{max}^2}{\theta_0^2} \ .$$

Using $C \ge \sigma_{max}^2\lambda_{max}$, it's sufficient to show that

$$\sqrt{A \times n^2\theta_0^4\lambda_{max}^{-4}6C^2(2+\epsilon_1)} - \frac{1}{2}\theta_0^2 \times \frac{n}{2}(2\lambda_{max})^{-2} \le -A \times n \times \theta_0^2\lambda_{max}^{-2}(1+\gamma_0) \ .$$

Dividing out common factors, the above is equivalent to showing that

$$\sqrt{A \times 6C^2(2+\epsilon_1)} - \frac{1}{16} \le -A \times (1+\gamma_0) \ .$$

By assumption (1), we know:

$$A \times (1+\gamma_0) \le \frac{1}{3200} \ ,$$

and also,

$$A \times 6C^2(2+\epsilon_1) \le 12 \times \frac{1}{3200} \ .$$

Therefore,

$$A \times (1+\gamma_0) + \sqrt{A \times 6C^2(2+\epsilon_1)} \le \frac{1}{3200} + \sqrt{\frac{12}{3200}} < \frac{1}{16} \ ,$$

as desired. $\square$

Lemma A.4 below is sufficient to fill in the details of Theorem 4 in the paper.

Lemma A.4. *Recall that, in the proof of Theorem 4 in the paper, we showed that, stochastically,*

$$l_n(\hat{\Theta}(\mathbf{s})) - l_n(\hat{\Theta}(\mathbf{s_0})) \leq \frac{n}{2} \times \frac{1}{n - \sqrt{2q} - 1}\chi^2_m \quad .$$

Then this implies that

$$P\{l_n(\hat{\Theta}(\mathbf{s})) - l_n(\hat{\Theta}(\mathbf{s_0})) \geq 2(1+\gamma_0)m\log(p)\} \leq \frac{1}{4\sqrt{\pi}\log p} e^{-\frac{m}{2}(4(1+\frac{\epsilon_0}{2})\log p)} \quad .$$

Proof. First, we show that $\frac{n-\sqrt{2q}-1}{n} \geq (1+\gamma_0)^{-\frac{1}{2}}$. By assumption (2) we see that:

$$\frac{1}{\log p} \leq 4(\sqrt{1+\gamma_0} - 1) \quad .$$

Now turn to assumption (1). We see that the right-hand side of (1) is $\leq \frac{1}{4\sqrt{1+\gamma_0}}$. On the left-hand side of (1), by definition, $\lambda^2_{max} \geq \theta^2_0$. Therefore,

$$\frac{(p+2q)\log p}{n} \leq \frac{1}{4\sqrt{1+\gamma_0}} \quad .$$

Therefore,

$$\frac{\sqrt{2q}+1}{n} \leq \frac{p+2q}{n} \leq \frac{4(\sqrt{1+\gamma_0}-1)}{4\sqrt{1+\gamma_0}} = 1 - \frac{1}{\sqrt{1+\gamma_0}} \quad ,$$

and so,

$$\frac{n-\sqrt{2q}-1}{n} \geq (1+\gamma_0)^{-\frac{1}{2}} \quad .$$

Therefore, using the stochastic inequality in the statement in the lemma,

$$\begin{aligned}
P\{l_n(\hat{\Theta}(\mathbf{s})) - l_n(\hat{\Theta}(\mathbf{s_0})) &\geq 2(1+\gamma_0)m\log(p)\} \\
&\leq P\{\chi^2_m \geq 4(1+\gamma_0)m\log p \times \frac{n-\sqrt{2q}-1}{n}\} \\
&\leq P\{\chi^2_m \geq 4\sqrt{1+\gamma_0}m\log p\} \quad .
\end{aligned}$$

Now we apply Cai's [10] (CSB) as cited in the paper, and obtain that

$$P\{\chi^2_m \geq 4\sqrt{1+\gamma_0}m\log p\} \leq \frac{1}{(4\sqrt{1+\gamma_0}\log p - 1)\sqrt{\pi m}} e^{-\frac{m}{2}(4\sqrt{1+\gamma_0}\log p - 1 - \log(4\sqrt{1+\gamma_0}\log p))} \quad .$$

Since $m \geq 1$ and $\frac{1}{\log p} \leq 4(\sqrt{1+\gamma_0} - 1)$, we obtain that the upper bound is at most

$$\begin{aligned}
&\frac{1}{4\sqrt{\pi}\log p} e^{-\frac{m}{2}(4\sqrt{1+\gamma_0}\log p - 1 - \log(4\sqrt{1+\gamma_0}\log p))} \\
= \; &\frac{1}{4\sqrt{\pi}\log p} e^{-\frac{m}{2}(4\sqrt{1+\gamma_0}\log p - (\log\log p + \log(4\sqrt{1+\gamma_0}) + 1))} \\
= \; &\frac{1}{4\sqrt{\pi}\log p} e^{-\frac{m}{2}(2\log p)(2\sqrt{1+\gamma_0} - (\log\log p + \log(4\sqrt{1+\gamma_0}) + 1)/(2\log p))} \quad .
\end{aligned}$$

By assumption (2), we may further bound this expression from above as

$$\frac{1}{4\sqrt{\pi}\log p} e^{-\frac{m}{2}(2\log p)(2+\epsilon_0)} = \frac{1}{4\sqrt{\pi}\log p} e^{-\frac{m}{2}4(1+\frac{\epsilon_0}{2})\log p} \quad .$$

□