*Statistica Sinica* **19**(2009): Supplement, S1–S7

# AN INFORMATION MATRIX PRIOR FOR BAYESIAN ANALYSIS IN GENERALIZED LINEAR MODELS WITH HIGH DIMENSIONAL DATA

Mayetri Gupta and Joseph G. Ibrahim

*Boston University and University of North Carolina at Chapel Hill*

## Supplementary Material

This note contains proofs for Theorems 1-6, and their corollaries.

**S.1. Proof of Theorem 1.** Without loss of generality, let $\boldsymbol{\mu}_0 = 0$.We need to show that the prior MGF: $\int e^{\boldsymbol{t}'\boldsymbol{\beta}}\pi_{IMR}(\boldsymbol{\beta})d\boldsymbol{\beta}$

$$\propto \int |X'\Omega(\boldsymbol{\beta})X+\lambda I|^{1/2}\exp\left\{-\frac{1}{2c_0}\boldsymbol{\beta}'(X'\Omega(\boldsymbol{\beta})X+\lambda I)\boldsymbol{\beta}+\boldsymbol{t}'\boldsymbol{\beta}\right\}d\boldsymbol{\beta} \qquad \text{(S.1.1)}$$

exists for some $\boldsymbol{t}$ in a neighborhood including zero. As previously, we have $I(\boldsymbol{\beta}) = X'\Omega(\boldsymbol{\beta})X$. By Corollary 13.7.4 in Harville (1997), we have,

$$|I(\boldsymbol{\beta}) + \lambda I| = \sum_{s=0}^{p} \lambda^s \sum_T |I(\boldsymbol{\beta})^{(i_1,\ldots,i_s)}|, \qquad \text{(S.1.2)}$$

where $T = \{i_1, \ldots, i_s\}$ is an $s$-dimensional subset of the $p$ positive integers $\{1, \ldots, p\}$ and the summation is over all such $\binom{p}{s}$ subsets; and $|I(\boldsymbol{\beta})^{(i_1,\ldots,i_s)}|$ is the determinant of the $(p-s)\times(p-s)$ submatrix of $I(\boldsymbol{\beta})$ obtained by leaving out the $(i_1, \ldots, i_s)$-th rows and columns of $I(\boldsymbol{\beta})$. Then, using (S.1.2), we have

$$\begin{aligned}
\text{(S.1.1)} \quad &= \int \left[\sum_{s=0}^{p}\lambda^s \sum_T |I(\boldsymbol{\beta})^{(i_1,\ldots,i_s)}|\right]^{\frac{1}{2}} \exp\left\{-\frac{1}{2c_0}\boldsymbol{\beta}'(I(\boldsymbol{\beta})+\lambda I)\boldsymbol{\beta}+t'\boldsymbol{\beta}\right\}d\boldsymbol{\beta}\\
&\le \int \sum_{s=0}^{p}\lambda^{s/2} \sum_T |I(\boldsymbol{\beta})^{(i_1,\ldots,i_s)}|^{\frac{1}{2}} \exp\left\{-\frac{1}{2c_0}\boldsymbol{\beta}'(I(\boldsymbol{\beta})+\lambda I)\boldsymbol{\beta}+\boldsymbol{t}'\boldsymbol{\beta}\right\}d\boldsymbol{\beta}.
\end{aligned} \qquad \text{(S.1.3)}$$

Now by applying the Cauchy-Binet formula, we have

$$|I(\boldsymbol{\beta})^{(i_1,\ldots,i_s)}| = \sum_{V(T)} c(a_{i_1}, \ldots, a_{i_{p-s}}) \prod_{j=1}^{p-s} \omega_{i_j},$$

where $V(T) = \{i_1, \ldots, i_{p-s}\}$, and $c(a_{i_1}, \ldots, a_{i_{p-s}}) = |X_*^{(p-s)}|^2$, with $X_*^{(p-s)'} = (x_{i_1}, \ldots, x_{i_{p-s}})$ being a $(p-s)\times(p-s)$ matrix with $j$-th column $x_{i_j}$ $(j = 1, \ldots, p-s)$; so that

$$\text{(S.1.3)} \quad = \quad \int \sum_{s=0}^{p} \lambda^{s/2} \sum_{T} \Bigg[ \sum_{V(T)} c(a_{i_1}, \ldots, a_{i_{p-s}}) \prod_{j=1}^{p-s} \omega_{i_j} \Bigg]^{\frac{1}{2}} \exp\Big\{ -\frac{1}{2c_0}\boldsymbol{\beta}'(I(\boldsymbol{\beta}) + \lambda I)\boldsymbol{\beta} + t'\boldsymbol{\beta}\Big\} d\boldsymbol{\beta}$$

$$\leq \quad \int \sum_{s=0}^{p} \lambda^{s/2} \sum_{T} \sum_{V(T)} c^{\frac{1}{2}}(a_{i_1}, \ldots, a_{i_{p-s}}) \prod_{j=1}^{p-s} \omega_{i_j}^{\frac{1}{2}} e^{-\frac{1}{2c_0}\boldsymbol{\beta}'(I(\boldsymbol{\beta})+\lambda I)\boldsymbol{\beta}+\boldsymbol{t}'\boldsymbol{\beta}} d\boldsymbol{\beta}. \qquad \text{(S.1.4)}$$

Now, for any $s$ such that $p - s > n$, $c(a_{i_1}, \ldots, a_{i_{p-s}}) = 0$. So

$$\text{(S.1.4)} \quad \leq \quad \int \sum_{s=p-n}^{p} \lambda^{s/2} \sum_{T} \sum_{V(T)} c^{\frac{1}{2}}(a_{i_1}, \ldots, a_{i_{p-s}}) \prod_{j=1}^{p-s} \omega_{i_j}^{\frac{1}{2}} e^{-\frac{1}{2c_0}\boldsymbol{\beta}'(I(\boldsymbol{\beta})+\lambda I)\boldsymbol{\beta}+\boldsymbol{t}'\boldsymbol{\beta}} d\boldsymbol{\beta}. \qquad \text{(S.1.5)}$$

Now, without loss of generality, taking $(i_1, \ldots, i_{p-s}) = (1, \ldots, p-s)$, the finiteness of (S.1.5) is equivalent to the finiteness of

$$\int \sum_{s=p-n}^{p} \prod_{j=1}^{p-s} \omega_j^{\frac{1}{2}} \exp\Big\{ -\frac{1}{2c_0}\boldsymbol{\beta}'(I(\boldsymbol{\beta}) + \lambda I)\boldsymbol{\beta} + \boldsymbol{t}'\boldsymbol{\beta}\Big\} d\boldsymbol{\beta}, \qquad \text{(S.1.6)}$$

as $\lambda$, $c(a_{i_1}, \ldots, a_{i_{p-s}})$ are constants that do not depend on $\boldsymbol{\beta}$. For any positive semi-definite matrices $A$ and $B$, $|A+B| \geq |A| \Rightarrow x'(A+B)x \geq x'Ax$. So we can simplify (S.1.6) as

$$\text{(S.1.6)} \quad \leq \quad \int \sum_{s=p-n}^{p} \prod_{j=1}^{p-s} \Big[ v_j(\boldsymbol{\beta})\delta_j^2(\boldsymbol{\beta}) \Big]^{\frac{1}{2}} \exp\Big\{ -\frac{\lambda}{2c_0}\boldsymbol{\beta}'\boldsymbol{\beta} + \boldsymbol{t}'\boldsymbol{\beta}\Big\} d\boldsymbol{\beta}. \qquad \text{(S.1.7)}$$

Now let $v_j(\boldsymbol{\beta}) = \delta_j(\boldsymbol{\beta}) = 1$ when $j = n+1, \ldots, p$. Let us construct a $p \times p$ matrix $X^* = \begin{bmatrix} X \\ \boldsymbol{x}_0 \end{bmatrix}$, where $\boldsymbol{x}_0$ is a $(p-n) \times p$ matrix such that $X^*$ is positive definite (p.d.). If $X$ is of rank $n$, this can always be done. Then, make the substitution $\boldsymbol{u} = X^*\boldsymbol{\beta}$, so that $\boldsymbol{\beta} = (X^*)^{-1}\boldsymbol{u} \Rightarrow |J| = |(X^*)^{-1}|$. Denoting $Q = (X^*)^{-1}$, we have $\boldsymbol{\beta}'\boldsymbol{\beta} = \boldsymbol{u}'Q'Q\boldsymbol{u}$. So, finiteness of (S.1.7) is equivalent to finiteness of

$$= \quad \int \Bigg[ \prod_{j=1}^{p} v(\theta(u_j))^{\frac{1}{2}} \delta^2(\theta(u_j)) \Bigg] \exp\Big\{ -\frac{\lambda}{2c_0}\boldsymbol{u}'Q'Q\boldsymbol{u} + \boldsymbol{t}'Q\boldsymbol{u}\Big\} d\boldsymbol{u}. \qquad \text{(S.1.8)}$$

Now, we need to find a scalar constant $M_1 > 0$ such that $\boldsymbol{u}'Q'Q\boldsymbol{u} \geq M_1\boldsymbol{u}'\boldsymbol{u}$ for all $\boldsymbol{u}$. Since $Q$ is p.d., a necessary and sufficient condition for this to hold is $|Q'Q| \geq M_1^p \Rightarrow |X^*| \leq M_1^{-p/2}$. So the required $M_1 > 0$ exists, as $|X^*|$ is bounded above. Next, we have,

$$\text{(S.1.8)} \quad \leq \quad \int \prod_{j=1}^{p} \Big[ v(\theta(u_j))^{1/2} \delta(\theta(u_j)) \Big] \exp\Big\{ -\frac{M_1\lambda}{2c_0}\boldsymbol{u}'\boldsymbol{u} + \boldsymbol{t}'Q\boldsymbol{u}\Big\} d\boldsymbol{u}$$

$$= \quad \prod_{j=1}^{p} \int \Big[ v(\theta(u_j))^{1/2} \delta(\theta(u_j)) \Big] e^{-\frac{M_1\lambda}{2c_0}u_j^2 + \tau_j u_j} du_j, \qquad \text{(S.1.9)}$$

where $\boldsymbol{\tau}' = \boldsymbol{t}'Q$. Finally, make the transformation $r_j = \theta(u_j) \Rightarrow u_j = \theta^{-1}(r_j)$. Then (S.1.9) becomes a product of $p$ one dimensional integrals

$$\prod_{j=1}^{p} \int \Big[v(r_j)^{1/2}\delta(r_j)\Big] e^{-\frac{M_1\lambda}{2c_0}[\theta^{-1}(r_j)]^2 + \tau_j\theta^{-1}(r_j)} \left|\frac{du_j}{dr_j}\right| dr_j$$
$$= \prod_{j=1}^{p} \int e^{-\frac{M_1\lambda}{2c_0}[\theta^{-1}(r_j)]^2 + \tau_j\theta^{-1}(r_j)} \left[\frac{d^2 b(r_j)}{dr_j^2}\right]^{1/2} dr_j, \qquad \text{(S.1.10)}$$

as $\frac{d}{dr_j}\theta^{-1}(r_j) = \delta(r_j) = \frac{dr_j}{du_j}$. So if each integral in (S.1.10) is finite for some $\tau_j$ in a open interval containing zero, the prior MGF exists. The corollaries immediately follow.

*Proof of Corollary 1.2.* To see that Corollary 1.2 holds, take $\lambda = 0$ in (S.1.3), then the expression reduces to only the term with $s = 0$. Continue the proof until Eqn (S.1.7), where, instead of augmenting the matrix, delete the last $n - p$ rows to get a square $p \times p$ matrix. The condition then follows.

**S.2. Proof of Theorem 2.** Starting with (S.1.1) in the proof of Theorem 1, we have

$$\text{(S.1.1)} \quad = \quad \int \left[\sum_{s=0}^{p} \lambda^s \sum_{T} |I(\boldsymbol{\beta})^{(i_1,\ldots,i_s)}|\right]^{\frac{1}{2}} \exp\left\{-\frac{1}{2c_0}\boldsymbol{\beta}'(I(\boldsymbol{\beta}) + \lambda I)\boldsymbol{\beta} + \boldsymbol{t}'\boldsymbol{\beta}\right\} d\boldsymbol{\beta}$$
$$\geq \quad \int \left[\lambda^s |I(\boldsymbol{\beta})^{(i_1,\ldots,i_s)}|\right]^{\frac{1}{2}} \exp\left\{-\frac{1}{2c_0}\boldsymbol{\beta}'(I(\boldsymbol{\beta}) + \lambda I)\boldsymbol{\beta} + \boldsymbol{t}'\boldsymbol{\beta}\right\} d\boldsymbol{\beta}, \quad \text{(S.2.1)}$$

for every $s = 0, \ldots, p$. So finiteness of (S.2.1) necessitates finiteness of each integral

$$\int \lambda^{s/2} |I(\boldsymbol{\beta})^{(i_1,\ldots,i_s)}|^{\frac{1}{2}} \exp\left\{-\frac{1}{2c_0}\boldsymbol{\beta}'(I(\boldsymbol{\beta}) + \lambda I)\boldsymbol{\beta} + \boldsymbol{t}'\boldsymbol{\beta}\right\} d\boldsymbol{\beta}. \qquad \text{(S.2.2)}$$

In practice, it is sufficient to check the non-existence of (S.2.2) for any $s = 0, \ldots, p$, to disprove existence of the prior MGF.

*Proof of Corollary 2.1.* To see that Corollary 2.1 holds, take $\lambda = 0$ in (S.2.1), and thus the only non-zero term is the $s = 0$ term. Continuing the proof the same way, the necessary condition is the finiteness of the same integral, with $s = \lambda = 0$.

**S.3. Proof of Theorem 3.** This proof follows along the same lines as the proof of Theorem 1, with the likelihood term added, and setting $w_i = 0$ (for $i = n+1, \ldots, p$), so that $\sum_{i=n+1}^{p} \phi^{-1} w_i [y_i\theta_i - b(\theta_i)] = 0$.

**S.4. Proof of Theorem 4.** Existence of MGFs for specific models are shown through application of Theorems 1-3 below.

- **Binomial with canonical link.** Here $b(\theta) = \log(1 + e^\theta)$, so $b''(\theta) = \frac{e^\theta}{(1+e^\theta)^2}$. Thus the sufficient condition (3.3) is satisfied.

- **Binomial with probit link.** For the probit link, the link function is given by $\theta(\eta) = \log\left[\frac{\Phi(\eta)}{1-\Phi(\eta)}\right]$, so that $\theta^{-1}(\eta) = \Phi^{-1}\left(\frac{e^\eta}{1+e^\eta}\right)$. To check if the sufficient condition (3.3) holds, we need to check finiteness of

$$\int_{-\infty}^{\infty} \exp\left[-\frac{\lambda M}{2c_0}[\Phi^{-1}\left(\frac{e^r}{1+e^r}\right)]^2 + \tau\Phi^{-1}\left(\frac{e^r}{1+e^r}\right)\right]\frac{e^{r/2}}{1+e^r}dr$$
$$= \int \exp\left[-\frac{\lambda M}{2c_0}z^2 + \tau z\right]\frac{\phi(z)}{\sqrt{\Phi(z)[1-\Phi(z)]}}dz,$$

which is finite, since $\frac{\phi(z)}{\sqrt{\Phi(z)[1-\Phi(z)]}}$ is bounded.
- **Poisson with canonical link.** $b(\theta) = e^\theta \Rightarrow b''(\theta) = e^\theta$. $E(e^{(1/2+\tau)\theta})$ exists, so the sufficient condition (3.3) is satisfied.
- **Poisson with identity link.** Here $\theta(\eta) = \log(\eta) \Rightarrow \theta^{-1}(\eta) = e^\eta \Rightarrow \frac{d}{dr}\theta^{-1}(r) = e^r$.

$$\text{Now,} \qquad \int_{-\infty}^{\infty} e^{-\frac{\lambda M}{2c_0}e^{2r}+\tau e^r}e^{r/2}dr = \int_0^\infty e^{-\frac{\lambda M}{2c_0}u+\tau\sqrt{u}}u^{-\frac{3}{4}}du. \tag{S.4.1}$$

The existence of (S.4.1) is equivalent to the existence of $E\left[e^{\tau\sqrt{u}}u^{-1}\right]$ when $u \sim \text{Gamma}(p, \alpha)$, where the shape parameter $p = \frac{5}{4}$ and the scale parameter $\alpha = \frac{\lambda M}{2c_0}$. If $\tau < 0$, $E\left[e^{\tau\sqrt{u}}u^{-1}\right] < E(u^{-1})$, which exists for any $p > 1$ (by Corollary 2.1 of Piegorsch and Casella (1985)). If $\tau > 0$, the existence of (S.4.1) is equivalent to the existence of $E(u^{-1})$ where $u \sim \text{Gamma}(\frac{5}{4}, \frac{\lambda M}{2c_0} - \tau)$, which exists as long as $\tau < \frac{\lambda M}{2c_0}$. So (S.4.1) exists for any $\tau \in (-\infty, \frac{\lambda M}{2c_0})$, so the sufficient condition (3.3) is satisfied.
- **Gamma with canonical link.** Here, $b(\theta) = -\log(-\theta)$, so $b''(\theta) = \frac{1}{\theta^2}$. The sufficient condition would be satisfied by the existence of $\int_0^\infty e^{-\frac{\lambda M}{2c_0}r^2+\tau r}r^{-1}dr = \int_0^\infty e^{-\frac{\lambda M}{2c_0}u+\tau\sqrt{u}}u^{-1}du$, which does not exist for $\tau > 0$, since in that case, with $u \sim \text{Exponential}(\frac{\lambda M}{2c_0})$, $E(u^{-1}e^{\tau\sqrt{u}}) > E(u^{-1}) = \infty$. Since the sufficient condition (3.3) fails, we then check if the minimum necessary condition holds, by taking $p = 1$ and checking the $(p-1)$-th term in the integral. Since $a_{11}(\beta) = x^2vx = \frac{x^2}{x^2\beta^2} = \frac{1}{\beta^2}$, the integral in the necessity condition (3.4) for the prior MGF reduces to

$$\int_{-\infty}^{\infty} \beta^{-2}e^{-\frac{\beta^2}{2c_0}\left(\frac{x^2}{\beta^2}+\lambda\right)}d\beta = e^{-\frac{x^2}{2c_0}}\int_{-\infty}^{\infty}\beta^{-2}e^{-\frac{\lambda}{2c_0}\beta^2}d\beta = \infty,$$

since the second negative moment of a Gaussian distribution is infinite. Hence the necessary condition fails.
- **Gamma with log link.** With the log link, $\eta = \log(\mu)$, or $\log(-1/\theta) = \eta$,

which implies that $\theta^{-1}(\eta) = -\log(-\eta)$. So the integral in (3.3) reduces to

$$\int_0^\infty \exp\left\{-\frac{\lambda M}{2c_0}(\log r)^2 + \tau(-\log r)\right\}\left(\frac{1}{r}\right)^{1+1/2} dr = \int_0^\infty \left[\frac{1}{r^{\log r}}\right]^{\frac{\lambda M}{2c_0}}\left(\frac{1}{r}\right)^{\tau+\frac{3}{2}} dr$$

which is infinite, and hence the sufficient condition does not hold. It is straightforward to check that the necessity condition (3.4) holds. Hence from these conditions we cannot determine whether the prior MGF exists. We will return to this model in Section 3.2 and show that the prior MGF does exist here due to a special result.

- **Inverse Gaussian with canonical link.** Here $b(\theta) = -(-2\theta)^{1/2} \Rightarrow b''(\theta) = -(-2\theta)^{-3/2}$, and $\theta^{-1}(r) = r$. So the prior MGF would exist if the integral

$$\int_0^\infty e^{-\frac{\lambda M}{2c_0}r^2 + \tau r} r^{-\frac{3}{4}} dr \quad = \quad \int_0^\infty e^{-\frac{\lambda M}{2c_0}u + \tau\sqrt{u}} u^{-\frac{3}{4}-\frac{1}{2}} du \qquad \text{(S.4.2)}$$

were finite. Now (S.4.2) is greater than $E(u^{-5/4})$ for an exponential distribution with mean $\frac{2c_0}{\lambda M}$ when $\tau < 0$, and hence is infinite. Next we check the necessity condition for $p = 1$. $a_{11}(\beta) = x^{1/2}\beta^{-3/2}$, and the integral to check necessity is

$$\begin{aligned}\int_0^\infty \beta^{-\frac{3}{4}} e^{-\frac{1}{2c_0}(\sqrt{x}\beta^{1/2} + \lambda\beta^2) + t\beta} d\beta \quad &= \quad \int_0^\infty u^{-\frac{3}{8}-\frac{1}{2}} e^{-\frac{\sqrt{x}}{2c_0}u^{1/4} + tu^{1/2} - \frac{\lambda}{2c_0}u} du \\ < \int_0^\infty u^{-\frac{7}{8}} e^{-(\frac{\sqrt{x}}{2c_0} + t - \frac{\lambda}{2c_0})u^{1/4}} du \quad &= \quad \int_0^\infty v^{-\frac{7}{2}+3} e^{-(\frac{\sqrt{x}}{2c_0} + t - \frac{\lambda}{2c_0})v} dv < \infty.\end{aligned}$$

So the sufficiency condition (3.3) is not satisfied, but the necessity condition (3.4) is satisfied, so that we cannot directly determine whether the prior MGF of $\boldsymbol{\beta}$ exists.

- **Posterior MGF existence.** The existence of the posterior MGFs can be checked in the same way as the prior MGFs, and are shown below.
  - *Binomial with canonical link.* The posterior MGF exists, as the sufficient condition (3.5) is satisfied: $\int \exp\left[-\frac{\lambda M}{2c_0}r^2 + (\tau + \phi^{-1}wy)r\right]\frac{e^r}{(1+e^r)^{\phi^{-1}w+1}} < \infty$.
  - *Gamma with log link.* Prior MGF existence is shown in Section 3.2. The posterior MGF also exists, as

$$\int_0^\infty \left[\frac{1}{r^{\log r}}\right]^{\frac{\lambda M}{2c_0}}\left(\frac{1}{r}\right)^{\tau+\frac{3}{2}-\phi^{-1}w} e^{-\phi^{-1}wry} dr < \int_0^\infty r^{-\tau-\frac{3}{2}+\alpha} e^{-\alpha ry} dr < \infty \qquad \text{(S.4.3)}$$

for $\tau + \frac{3}{2} - \alpha < 1$, by Corollary 2.1 of Piegorsch and Casella (1985). (For a gamma model, $\phi^{-1} = \alpha, w = 1$.) A special case is the exponential model, with $\alpha = 1$. In this case, it can be seen that the posterior MGF exists, as the integral (S.4.3) is finite for $\tau < \frac{1}{2}$.

– *Poisson with canonical link.* The sufficient condition (3.5) would be satisfied by the finiteness of the integral $\int_{-\infty}^{\infty} e^{-\frac{\lambda M}{2c_0}r^2+\tau r+yr-e^r}dr$, that is, if $E(e^{(\tau+y)r-e^r})$ exists for a Gaussian distribution with variance $\frac{c_0}{\lambda M}$. Since $0 < e^{e^r} < 1$, $E(e^{(\tau+y)r-e^r}) < E(e^{(\tau+y)r}) < \infty$, and the posterior MGF exists.

### S.5. Proof of Theorem 5.

**Upper bound**

From (3.7), $\Sigma^{-1} = (1+1/c_0)X'X + \lambda/c_0 I_p$ and since both are positive semi-definite,

$$\Rightarrow |\Sigma^{-1}| \geq |(1+1/c_0)X'X| + |\lambda/c_0 I_p|. \tag{S.5.1}$$

When $p > n$, $|(1+1/c_0)X'X| = 0$. Also, from (S.5.1), $|\Sigma| \leq \left(\frac{c_0}{\lambda}\right)^p$. So as $p \longrightarrow \infty$, $|\Sigma| \longrightarrow 0$ if $c_0 < \lambda$.

**Lower bound**

$$|\Sigma^{-1}| = |(1+1/c_0)X'X + \lambda/c_0 I_p| = \left(\frac{\lambda}{c_0}\right)^p \left|I_n + \frac{c_0+1}{\lambda}XX'\right|,$$

since $|A+BC| = |A||I_k + CA^{-1}B|$. By Hadamard's inequality,

$$|\Sigma^{-1}|^2 \leq \left(\frac{\lambda}{c_0}\right)^p \prod_{i=1}^n \left[\sum_{k=1}^n \left(1 + \frac{c_0+1}{\lambda}\boldsymbol{x}_i'\boldsymbol{x}_k\right)^2\right]. \tag{S.5.2}$$

Now let $x_0$ be chosen such that $\boldsymbol{x}_i'\boldsymbol{x}_k \leq x_0^2$ (for $1 \leq i, k \leq n$). Then,

$$\begin{aligned}
(\text{S.5.2}) &\leq \left(\frac{\lambda}{c_0}\right)^p n^n \left(1 + \frac{c_0+1}{\lambda}x_0^2\right)^{2n} \\
\Rightarrow |\Sigma| &\geq \left(\frac{c_0}{\lambda}\right)^{p/2} \frac{\left(1 + \frac{c_0+1}{\lambda}x_0^2\right)^{-n}}{n^{n/2}}.
\end{aligned} \tag{S.5.3}$$

From (S.5.2) and (S.5.3) it is easy to see that $\lim_{p\longrightarrow\infty} |\Sigma| = 0$ if $c_0 < \lambda$. However, if $c_0 > \lambda$, and $n$ is fixed, then $\lim_{p\longrightarrow\infty} |\Sigma| = \infty$, which proves the corollary.

**S.6. Proof of Theorem 6.** Bias $= E\Big[E(\boldsymbol{\beta}|Y) - \boldsymbol{\beta}\Big] = (\Sigma X'X - I)\boldsymbol{\beta}$, where $\Sigma = \left[\frac{c_0+1}{c_0}X'X + \frac{\lambda}{c_0}I_p\right]^{-1}$.

**Part 1: Average bias.** This is given by

$$\begin{aligned}
\frac{1}{p}\sum_{i=1}^p bias(\beta_p) &= \boldsymbol{J}'(\Sigma X'X - I)\boldsymbol{\beta} \\
&= -\frac{1}{p(c_0+1)}\left[\boldsymbol{J}' + \frac{\lambda c_0}{c_0+1}\boldsymbol{J}'\left(X'X + \frac{\lambda}{c_0+1}I_p\right)^{-1}\right]\boldsymbol{\beta}.
\end{aligned}$$

Now, let $\boldsymbol{\beta} = B\boldsymbol{J}$, where $B = \text{diag}(|\beta_1|, \ldots, |\beta_p|)$, and $\boldsymbol{J}$ denotes a $p$-dimensional vector of ones. Now,

$$\left|\left(X'X + \frac{\lambda}{c_0+1}I_p\right)^{-1} B\right| = \left|X'X + \frac{\lambda}{c_0+1}I_p\right|^{-1}|B| \leq \left|\frac{c_0+1}{\lambda}B\right|, \quad \text{(S.6.1)}$$

since $|X'X + \frac{\lambda}{c_0+1}I_p| \geq |\frac{\lambda}{c_0+1}I_p|$. So (S.6.1) implies that

$$\boldsymbol{J}'\left(X'X + \frac{\lambda}{c_0+1}I_p\right)^{-1} B\boldsymbol{J} \leq \frac{c_0+1}{\lambda}\boldsymbol{J}'B\boldsymbol{J}, \quad \text{so finally,}$$

$$\| \textit{ average bias } \| \leq \frac{1}{p(c_0+1)}\left(1 + \frac{c_0+1}{\lambda}\frac{c_0\lambda}{c_0+1}\right)\sum_{i=1}^{p}|\beta_i| = \frac{1}{p}\sum_{i=1}^{p}|\beta_i|.$$

**Part 2: Bound on determinant of bias matrix.** Let $D = \Sigma X'X - I$. Then,

$$D = [(c_0X'X + X'X + \lambda I]^{-1}c_0(X'X) - I = -[c_0X'X + X'X + \lambda I]^{-1}(X'X + \lambda I).$$

$$\text{So,} \quad |D| = \frac{|X'X + \lambda I|}{|c_0X'X + X'X + \lambda I|} \leq 1,$$

which again shows that $\| D\boldsymbol{\beta} \|^2 = \boldsymbol{\beta}'D'D\boldsymbol{\beta} \leq \boldsymbol{\beta}'\boldsymbol{\beta}$.

## References

Harville, D. A. (1997). *Matrix Algebra from a Statistician's Perspective.* Springer-Verlag.

Piegorsch, W. W. and Casella, G. (1985). The existence of the first negative moment. *Amer. Statist.*, 39(1):60–62.