Covariance Estimation via Fiducial Inference

Wen Jenny Shi[1]
joint work with Jan Hannig[2], Thomas Lee[3], and Randy Lai[3]

1. University of Colorado Anschutz Medical Campus
2. University of North Carolina at Chapel Hill
3. University of California at Davis

August 11, 2015

Brief history of fiducial inference

(1930)

Sir Ronald Aylmer Fisher
Proposed *fiducial inference*

(1989-2001)

- Generalized *p*-value (Tsui & Weerahndi)
- Generalized CI (Weerahndi)
- Surrogate variable method (Chiang)

Not exact!
Not unique!

Major controversy (1930-1960)

- Generalized fiducial inference
- Established asymptotic exactness
- A variety of applications to both continuous and discrete models

(2004-present)

Various applications of the generalized fiducial approach

- bioequivalence [Hannig, Abdel-Karim & Iyer (2006)]
- mixture of normal and Cauchy distributions [Glagovskiy (2006)]
- metrology [e.g. Hannig, Iyer & Wang (2007)]
- logistic regression and LD_{50} [Hannig & Iyer (2009)]
- multiple comparisons [Wandler & Hannig (2012)]
- extreme value estimation [Wandler & Hannig (2012b)]
- variance components [e.g. Cisewski & Hannig (2012)]
- wavelet regression [Hannig & Lee (2009)]
- free-knot splines [Sonderegger & Hannig (2013)]
- volatility estimation [Katsoridas & Hannig (2015+)]
- survey analysis [Liu & Hannig (2015+)]
- ...

Generalized fiducial inference (GFI) recipe

- Data generating equation (DGE)

$$\mathbf{X} = \mathbf{G}(\mathbf{U}, \xi)$$

- Given observations & independent copies of $\mathbf{U}$,

$$\mathbf{x} = \mathbf{G}(\mathbf{U}^{\star}, \xi)$$

- A functional inverse of $\mathbf{G}$

$$\lim_{\varepsilon \downarrow 0} \left[\operatorname*{argmin}_{\xi} \|\mathbf{x} - \boldsymbol{G}(\mathbf{U}^{\star}, \xi)\| \mid \{\min_{\xi} \|\mathbf{x} - \boldsymbol{G}(\mathbf{U}^{\star}, \xi)\| \leq \varepsilon\} \right] \qquad (1)$$

- Generalized fiducial distribution (GFD) := distribution of (1)

Generalized fiducial inference (GFI) recipe

- Data generating equation (DGE)

$$\mathbf{X} = \mathbf{G}(\mathbf{U}, \xi)$$

- Given observations & independent copies of $\mathbf{U}$,

$$\mathbf{x} = \mathbf{G}(\mathbf{U}^\star, \xi)$$

- A functional inverse of $\mathbf{G}$

$$\lim_{\varepsilon \downarrow 0} \left[\operatorname*{argmin}_{\xi} \|\mathbf{x} - \boldsymbol{G}(\mathbf{U}^\star, \xi)\| \mid \{\min_{\xi} \|\mathbf{x} - \boldsymbol{G}(\mathbf{U}^\star, \xi)\| \le \varepsilon\} \right] \qquad (1)$$

- Generalized fiducial distribution (GFD) := distribution of (1)

Generalized fiducial inference (GFI) recipe

- Data generating equation (DGE)

$$\mathbf{X} = \mathbf{G}(\mathbf{U}, \xi)$$

- Given observations & independent copies of $\mathbf{U}$,

$$\mathbf{x} = \mathbf{G}(\mathbf{U}^\star, \xi)$$

- A functional inverse of $\mathbf{G}$

$$\lim_{\varepsilon \downarrow 0} \left[\operatorname*{argmin}_{\xi} \|\mathbf{x} - \boldsymbol{G}(\mathbf{U}^\star, \xi)\| \mid \{\min_{\xi} \|\mathbf{x} - \boldsymbol{G}(\mathbf{U}^\star, \xi)\| \le \varepsilon\} \right] \qquad (1)$$

- Generalized fiducial distribution (GFD) := distribution of (1)

Generalized fiducial inference (GFI) recipe

- Data generating equation (DGE)

$$\mathbf{X} = \mathbf{G}(\mathbf{U}, \xi)$$

- Given observations & independent copies of $\mathbf{U}$,

$$\mathbf{x} = \mathbf{G}(\mathbf{U}^\star, \xi)$$

- A functional inverse of $\mathbf{G}$

$$\lim_{\varepsilon \downarrow 0} \left[\operatorname*{argmin}_{\xi} \|\mathbf{x} - \boldsymbol{G}(\mathbf{U}^\star, \xi)\| \mid \{\min_{\xi} \|\mathbf{x} - \boldsymbol{G}(\mathbf{U}^\star, \xi)\| \leq \varepsilon\} \right] \qquad (1)$$

- Generalized fiducial distribution (GFD) := distribution of (1)

Explicit limit

If $\mathbf{X} \in \mathbb{R}^n$ is continuous and $\xi \in \mathbb{R}^p$, the limit (1) has density [Hannig, Iyer, Lai & Lee (2015)]

$$r(\xi|\mathbf{x}) = \frac{f_{\mathbf{X}}(\mathbf{x}|\xi) J(\mathbf{x}, \xi)}{\int_{\Xi} f_{\mathbf{X}}(\mathbf{x}|\xi') J(\mathbf{x}, \xi') \, d\xi'},$$

where $J(\mathbf{x}, \xi) = D\left(\frac{d}{d\xi} \mathbf{G}(\mathbf{u}, \xi) \Big|_{\mathbf{u}=\mathbf{G}^{-1}(\mathbf{x},\xi)} \right)$

- $n = p$ gives $D(A) = |\mathrm{det} A|$
- $\|\cdot\|_2$ gives $D(A) = |\mathrm{det}(A^\top A)|^{1/2}$
  Compare to Fraser, Reid, Marras & Yi (2010)
- $\|\cdot\|_\infty$ gives $D(A) = \binom{n}{p}^{-1} \sum_{\mathbf{i}=(i_1,\ldots,i_p)} |\mathrm{det}(A)_{\mathbf{i}}|$ where $(A)_{\mathbf{i}}$ is the $p \times p$ matrix comprising of the $i_1, \ldots, i_p$th row of the $n \times p$ matrix A (Recommended)

Explicit limit

If $\mathbf{X} \in \mathbb{R}^n$ is continuous and $\xi \in \mathbb{R}^p$, the limit (1) has density [Hannig, Iyer, Lai & Lee (2015)]

$$r(\xi|\mathbf{x}) = \frac{f_{\mathbf{X}}(\mathbf{x}|\xi) J(\mathbf{x}, \xi)}{\int_{\Xi} f_{\mathbf{X}}(\mathbf{x}|\xi') J(\mathbf{x}, \xi')\, d\xi'},$$

where $J(\mathbf{x}, \xi) = D\left(\frac{d}{d\xi} \mathbf{G}(\mathbf{u}, \xi) \Big|_{\mathbf{u}=\mathbf{G}^{-1}(\mathbf{x},\xi)} \right)$

- $n = p$ gives $D(A) = |\det A|$
- $\|\cdot\|_2$ gives $D(A) = |\det(A^\top A)|^{1/2}$
  Compare to Fraser, Reid, Marras & Yi (2010)
- $\|\cdot\|_\infty$ gives $D(A) = \binom{n}{p}^{-1} \sum_{\mathbf{i}=(i_1,\ldots,i_p)} |\det(A)_{\mathbf{i}}|$ where $(A)_{\mathbf{i}}$ is the $p \times p$ matrix comprising of the $i_1, \ldots, i_p$th row of the $n \times p$ matrix A (Recommended)

Explicit limit

If $\mathbf{X} \in \mathbb{R}^n$ is continuous and $\xi \in \mathbb{R}^p$, the limit (1) has density [Hannig, Iyer, Lai & Lee (2015)]

$$r(\xi|\mathbf{x}) = \frac{f_{\mathbf{X}}(\mathbf{x}|\xi)J(\mathbf{x},\xi)}{\int_{\Xi} f_{\mathbf{X}}(\mathbf{x}|\xi')J(\mathbf{x},\xi')\,d\xi'},$$

where $J(\mathbf{x},\xi) = D\left(\left.\frac{d}{d\xi}\mathbf{G}(\mathbf{u},\xi)\right|_{\mathbf{u}=\mathbf{G}^{-1}(\mathbf{x},\xi)}\right)$

- $n = p$ gives $D(A) = |\det A|$
- $\|\bullet\|_2$ gives $D(A) = |\det(A^\top A)|^{1/2}$
  Compare to Fraser, Reid, Marras & Yi (2010)
- $\|\bullet\|_\infty$ gives $D(A) = \binom{n}{p}^{-1} \sum_{\mathbf{i}=(i_1,\ldots,i_p)} |\det(A)_{\mathbf{i}}|$ where $(A)_{\mathbf{i}}$ is the $p \times p$ matrix comprising of the $i_1,\ldots,i_p$th row of the $n \times p$ matrix A (Recommended)

Explicit limit

If $\mathbf{X} \in \mathbb{R}^n$ is continuous and $\xi \in \mathbb{R}^p$, the limit (1) has density [Hannig, Iyer, Lai & Lee (2015)]

$$r(\xi|\mathbf{x}) = \frac{f_{\mathbf{X}}(\mathbf{x}|\xi)J(\mathbf{x},\xi)}{\int_{\Xi} f_{\mathbf{X}}(\mathbf{x}|\xi')J(\mathbf{x},\xi')\,d\xi'},$$

where $J(\mathbf{x},\xi) = D\left(\left.\frac{d}{d\xi}\mathbf{G}(\mathbf{u},\xi)\right|_{\mathbf{u}=\mathbf{G}^{-1}(\mathbf{x},\xi)}\right)$

- $n = p$ gives $D(A) = |\det A|$
- $\|\cdot\|_2$ gives $D(A) = |\det(A^\top A)|^{1/2}$
  Compare to Fraser, Reid, Marras & Yi (2010)
- $\|\cdot\|_\infty$ gives $D(A) = \binom{n}{p}^{-1} \sum_{\mathbf{i}=(i_1,\ldots,i_p)} |\det(A)_{\mathbf{i}}|$ where $(A)_{\mathbf{i}}$ is the $p \times p$ matrix comprising of the $i_1,\ldots,i_p$th row of the $n \times p$ matrix A (Recommended)

Some important properties of GFD

- Aways proper
- Invariant to re-parametrization (same as Jeffreys)
- If $n > p$, NOT invariant to smooth transformation of the data
- Penalty needed for model selection [Hannig, Iyer, Lai & Lee (2015+)].

GFD for covariate A

DGE:

$$Y_i = AZ_i,\ Z_i \overset{iid}{\sim} N(0, I),\ (\Sigma = AA^T).$$

GFD of A: $r(A|\mathbf{Y}) \propto J(\mathbf{Y}, A) f(\mathbf{Y}, A)$, where

$$J(\mathbf{Y}, A) = \binom{n}{p}^{-1} \sum_{\substack{\mathbf{i}=(i_1,\cdots,i_p) \\ 1\le i_1<\cdots<i_p\le n}} \left|\det\left(\frac{\partial W_{\mathbf{i}}}{\partial A}\right)\right|,$$
$$f(\mathbf{Y}, A) = (2\pi)^{-\frac{np}{2}} |\det(A)|^{-n} \exp\left[-\frac{1}{2}\mathrm{tr}\{nS_n(AA^T)^{-1}\}\right],$$

$$W_{\mathbf{i}} = (Y_{i_1}^T, \cdots, Y_{i_p}^T)^T,\ Sn = \frac{1}{n}\sum_{i=1}^{n} Y_i^T Y_i.$$

GFD for covariate A

DGE:

$$Y_i = AZ_i,\ Z_i \overset{iid}{\sim} N(0, I),\ (\Sigma = AA^T).$$

GFD of A: $r(A|\mathbf{Y}) \propto J(\mathbf{Y}, A) f(\mathbf{Y}, A)$, where

$$J(\mathbf{Y}, A) = \binom{n}{p}^{-1} \sum_{\substack{\mathbf{i}=(i_1,\cdots,i_p)\\ 1\le i_1<\cdots<i_p\le n}} \left|\det\left(\frac{\partial W_{\mathbf{i}}}{\partial A}\right)\right|,$$

$$f(\mathbf{Y}, A) = (2\pi)^{-\frac{np}{2}} |\det(A)|^{-n} \exp\left[-\frac{1}{2}\mathrm{tr}\{nS_n(AA^T)^{-1}\}\right],$$

$$W_{\mathbf{i}} = (Y_{i_1}^T, \cdots, Y_{i_p}^T)^T,\ Sn = \frac{1}{n}\sum_{i=1}^{n} Y_i^T Y_i.$$

Special Case I: no element of A is fixed at zero

$$\frac{\partial W_{\mathbf{i}}}{\partial A} = \begin{array}{c} \begin{array}{cccc} \quad B_{\mathbf{1}} \quad & \quad B_{\mathbf{2}} \quad & \quad \cdots \quad & \quad B_p \quad \end{array} \\ \begin{pmatrix} (A^{-1}Y_{i_1})^T & & & \\ & (A^{-1}Y_{i_1})^T & & \\ & & \ddots & \\ & & & (A^{-1}Y_{i_1})^T \\ \vdots & \vdots & \vdots & \vdots \\ (A^{-1}Y_{i_p})^T & & & \\ & (A^{-1}Y_{i_p})^T & & \\ & & \ddots & \\ & & & (A^{-1}Y_{i_p})^T \end{pmatrix} \end{array} \Rightarrow \begin{array}{c} \begin{array}{ccc} \quad B'_{\mathbf{1}} \quad & \quad \cdots \quad & \quad B'_p \quad \end{array} \\ \begin{pmatrix} (A^{-1}Y_{i_1})^T & & \\ \vdots & & \\ (A^{-1}Y_{i_p})^T & & \\ & \ddots & \\ & & (A^{-1}Y_{i_1})^T \\ & & \vdots \\ & & (A^{-1}Y_{i_p})^T \end{pmatrix} \end{array}$$

- Parameter space: $\mathbb{R}^{p \times p}$
- Easy calculation for $J(\mathbf{Y}, A)$
- $\Sigma \sim IW(n, nS_n)$

Special Case I: no element of A is fixed at zero

$$\frac{\partial W_{\mathbf{i}}}{\partial A} = \begin{array}{cccc} B_{\mathbf{1}} & B_{\mathbf{2}} & \cdots & B_p \end{array}$$
$$\begin{pmatrix} (A^{-\mathbf{1}}Y_{i_{\mathbf{1}}})^T & & & \\ & (A^{-\mathbf{1}}Y_{i_{\mathbf{1}}})^T & & \\ & & \ddots & \\ & & & (A^{-\mathbf{1}}Y_{i_{\mathbf{1}}})^T \\ \vdots & \vdots & \vdots & \vdots \\ (A^{-\mathbf{1}}Y_{i_p})^T & & & \\ & (A^{-\mathbf{1}}Y_{i_p})^T & & \\ & & \ddots & \\ & & & (A^{-\mathbf{1}}Y_{i_p})^T \end{pmatrix} \Rightarrow \begin{array}{ccc} B'_{\mathbf{1}} & \cdots & B'_p \end{array}$$
$$\begin{pmatrix} (A^{-\mathbf{1}}Y_{i_{\mathbf{1}}})^T & & \\ \vdots & & \\ (A^{-\mathbf{1}}Y_{i_p})^T & & \\ & \ddots & \\ & & (A^{-\mathbf{1}}Y_{i_{\mathbf{1}}})^T \\ & & \vdots \\ & & (A^{-\mathbf{1}}Y_{i_p})^T \end{pmatrix}$$

- Parameter space: $\mathbb{R}^{p \times p}$
- Easy calculation for $J(\mathbf{Y}, A)$
- $\Sigma \sim IW(n, nS_n)$

Special Case II: Clique model

- Coordinates $l \& m$ of $\mathbf{Y}_i$ are correlated only if belong to the same clique
- Total k cliques with sizes $g_1, \cdots, g_k$, $\sum_{s=1}^{k} g_s = p$
- An extension of Special Case I

General Case: possible fixed zeros in A

No closed-form expression

$$r(A|\mathbf{Y}) \propto \frac{\exp\left[-\frac{1}{2}\mathrm{tr}\{nS_n(AA^T)^{-1}\}\right]}{|\det(A)|^n \binom{n}{p}} \sum_{\substack{\mathbf{i}=(i_1,\cdots,i_p)\\ 1\le i_1<\cdots<i_p\le n}} \prod_{i=1}^{p} \binom{p}{p_i} \overline{\left|\det\left(U_{\mathbf{i},i}\right)_{\mathbf{r}_i}\right|}$$

- $\overline{\left|\det\left(U_{\mathbf{i},i}\right)_{\mathbf{r}_i}\right|}$: defined by $A^{-1}Y_{i_\ell}$'s & sparse structure of A

Theorem 1 (Consistency)

If there is one-to-one correspondence between Σ and A, $r(A|\mathbf{Y})$ is asymptotically normal.

General Case: possible fixed zeros in A

No closed-form expression

$$r(A|\mathbf{Y}) \propto \frac{\exp\left[-\frac{1}{2}\mathrm{tr}\{nS_n(AA^T)^{-1}\}\right]}{|\det(A)|^n \binom{n}{p}} \sum_{\substack{\mathbf{i}=(i_1,\cdots,i_p) \\ 1\le i_1<\cdots<i_p\le n}} \prod_{i=1}^{p} \binom{p}{p_i} \overline{\left|\det\left(U_{\mathbf{i},i}\right)_{\mathbf{r}_i}\right|}$$

- $\overline{\left|\det\left(U_{\mathbf{i},i}\right)_{\mathbf{r}_i}\right|}$: defined by $A^{-1}Y_{i_\ell}$'s & sparse structure of A

Theorem 1 (Consistency)

If there is one-to-one correspondence between Σ and A, $r(A|\mathbf{Y})$ is asymptotically normal.

Fiducial Bernstein von-Mises Theorem [Sonderegger & Hannig (2013)]

Conditions:

(a) the Maximum Likelihood Estimator (MLE) is asymptotically normal

(b) the Bayesian posterior distribution becomes close to that of the MLE

(c) the fiducial distribution is close to the Bayesian posterior

⋆*Conditions (a) & (b) are easy to show.*

Condition (c) holds

Proposition 1

Let $L_n(\cdot)$ be the log-likelihood with $Y_1, \cdots, Y_n$. Then for any $\delta > 0$

$$\inf_{A \notin B(A_0,\delta)} \frac{\min_{\substack{\mathbf{i}=\{i_1,\cdots,i_p\} \\ 1\le i_1<\cdots<i_p\le n}} \log f(A, \mathbf{Y}_\mathbf{i})}{|L_n(A) - L_n(A_0)|} \xrightarrow{A_0} 0, \quad as\ n \to \infty.$$

Proposition 2

Let $\mathbf{Y}_0 = (Y_1, Y_2, \cdots, Y_p)$ and $\pi(A) = E_{A_0} J(\mathbf{Y}_0, A)$. Assume that there is a one-to-one correspondence between $\Sigma = AA^T$ and A. Then the Jacobian function $J(\mathbf{Y}, A) \xrightarrow{a.s.} \pi(A)$ uniformly on compacts in A.

Clique model selection

Apply a sparse structure penalty $q(M)$ & marginalize out A,

$$r(M|\mathbf{Y}) \propto \prod_{s=1}^{k} \left[\Gamma_{g_s}\left(\frac{n}{2}\right) C_s(\mathbf{Y})(2\pi)^{\frac{g_s(g_s-1)}{2}} \left|\det S_n^s\right|^{-\frac{n}{2}} \right] \exp\left\{-q(M)\right\}$$

for clique s,

- S_n^s: $g_s \times g_s$ sample covariance matrix
- $C_s(\mathbf{Y})$: Jacobian constant
- $\Gamma_{g_s}(\frac{n}{2})$: multivariate gamma function

Clique result

Definition 1

A clique model M is **compatible** *with the true covariance matrix Σ_0, if $\Sigma_0^M = \Sigma_0$. Denote it as $M_0 \subset M$, where M_0 is the true model.*

Theorem 2

Assume the penalty $q(M)$ satisfies

i. *for any $M_0 \not\subset M$, $det\Sigma_0 < det\Sigma^M$, and $-an + q(M) - q(M_0) \to -\infty, \forall a$ as $n \to \infty$;*

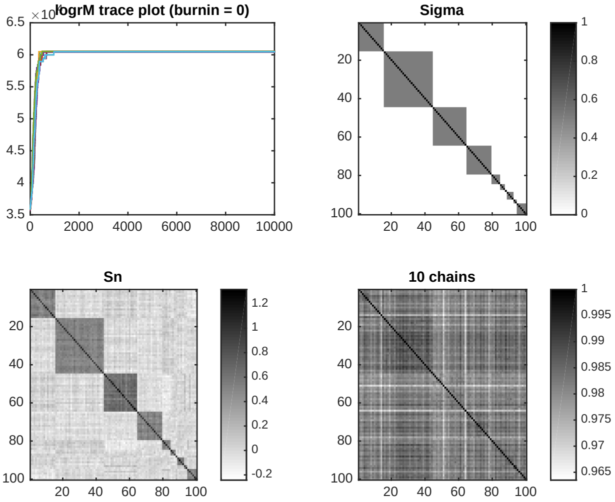
ii. *for any $M_0 \subset M$, $q(M) - q(M_0)$ is bounded.*

Then the penalized GFD

$$r(M_0|\mathbf{Y}) \xrightarrow{P} 1, \quad as\ n \to \infty$$

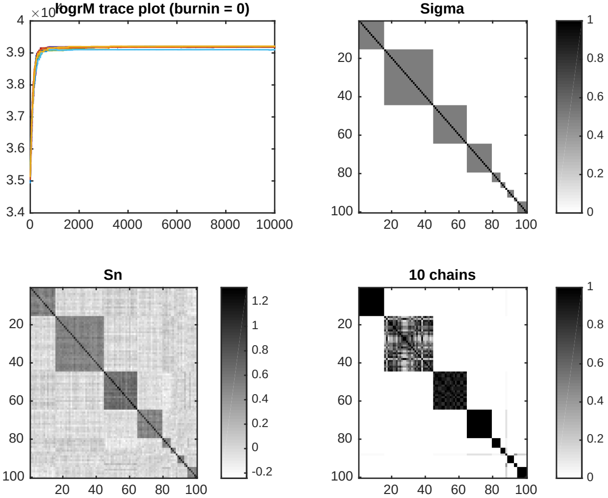
No penalty.

$q(M) = 0;\ k = 10, p = 100, n = 200.$



A minimum description length (MDL) penalty

$$q(M) = \frac{1}{3}\left(\sum_{s=1}^{k} g_s^2\right)\log n + (p+1)\log k;\ \ k = 10, p = 100, n = 200.$$



General case model selection

$$r(A|\mathbf{Y}) \propto \frac{\exp\left[-\frac{1}{2}\mathrm{tr}\{nS_n(AA^T)^{-1}\}\right]}{|\det(A)|^n \binom{n}{p}} \sum_{\substack{\mathbf{i}=(i_1,\cdots,i_p)\\ 1\le i_1<\cdots<i_p\le n}} \prod_{i=1}^{p} \binom{p}{p_i} \overline{\left|\det\left(U_{\mathbf{i},i}\right)_{\mathbf{r}_i}\right|}$$

GC 1. Known sparse structure of A.

- standard MCMC

GC 2. Sparse structure of A unknown (model selection).

- apply an MDL based penalty
- Gibbs w/ moves: update, birth, death (RJMCMC)
- adaptive birth proposal variance
  - *zeroth-order-method* [Brooks, Giudici & Roberts (2003)]

General case model selection

$$r(A|\mathbf{Y}) \propto \frac{\exp\left[-\frac{1}{2}\mathrm{tr}\{nS_n(AA^T)^{-1}\}\right]}{|\det(A)|^n \binom{n}{p}} \sum_{\substack{\mathbf{i}=(i_1,\cdots,i_p) \\ 1\leq i_1<\cdots<i_p\leq n}} \prod_{i=1}^{p} \binom{p}{p_i} \overline{\left|\det\left(U_{\mathbf{i},i}\right)_{\mathbf{r}_i}\right|}$$

GC 1. Known sparse structure of A.

- standard MCMC

GC 2. Sparse structure of A unknown (model selection).

- apply an MDL based penalty
- Gibbs w/ moves: update, birth, death (RJMCMC)
- adaptive birth proposal variance
  - *zeroth-order-method* [Brooks, Giudici & Roberts (2003)]

General case model selection

$$r(A|\mathbf{Y}) \propto \frac{\exp\left[-\frac{1}{2}\mathrm{tr}\{nS_n(AA^T)^{-1}\}\right]}{|\det(A)|^n \binom{n}{p}} \sum_{\substack{\mathbf{i}=(i_1,\cdots,i_p)\\ 1\le i_1<\cdots<i_p\le n}} \prod_{i=1}^{p} \binom{p}{p_i} \overline{\left|\det\left(U_{\mathbf{i},i}\right)_{\mathbf{r}_i}\right|}$$

GC 1. Known sparse structure of A.

- standard MCMC

GC 2. Sparse structure of A unknown (model selection).

- apply an MDL based penalty
- Gibbs w/ moves: update, birth, death (RJMCMC)
- adaptive birth proposal variance
  - *zeroth-order-method* [Brooks, Giudici & Roberts (2003)]

General case model selection

$$r(A|\mathbf{Y}) \propto \frac{\exp\left[-\frac{1}{2}\mathrm{tr}\{nS_n(AA^T)^{-1}\}\right]}{|\det(A)|^n \binom{n}{p}} \sum_{\substack{\mathbf{i}=(i_1,\cdots,i_p) \\ 1\le i_1<\cdots<i_p\le n}} \prod_{i=1}^{p} \binom{p}{p_i} \overline{\left|\det\left(U_{\mathbf{i},i}\right)_{\mathbf{r}_i}\right|}$$

GC 1. Known sparse structure of A.

- standard MCMC

GC 2. Sparse structure of A unknown (model selection).

- apply an MDL based penalty
- Gibbs w/ moves: update, birth, death (RJMCMC)
- adaptive birth proposal variance
  - *zeroth-order-method* [Brooks, Giudici & Roberts (2003)]

General case model selection

$$r(A|\mathbf{Y}) \propto \frac{\exp\left[-\frac{1}{2}\mathrm{tr}\{nS_n(AA^T)^{-1}\}\right]}{|\det(A)|^n \binom{n}{p}} \sum_{\substack{\mathbf{i}=(i_1,\cdots,i_p) \\ 1\le i_1<\cdots<i_p\le n}} \prod_{i=1}^{p} \binom{p}{p_i} \overline{\left|\det\left(U_{\mathbf{i},i}\right)_{\mathbf{r}_i}\right|}$$

GC 1. Known sparse structure of A.

- standard MCMC

GC 2. Sparse structure of A unknown (model selection).

- apply an MDL based penalty
- Gibbs w/ moves: update, birth, death (RJMCMC)
- adaptive birth proposal variance
  - *zeroth-order-method* [Brooks, Giudici & Roberts (2003)]

Access general case result

Definition 2

Suppose the empirical cumulative density function is f_{ecdf}, then the **confidence curve function** f_{cc} *is defined as the following:*

$$f_{cc}(x) = \begin{cases} 2(1 - f_{ecdf}(x)), & \textit{if } f_{ecdf}(x) \geq 0.5, \\ 2f_{ecdf}(x), & \textit{otherwise.} \end{cases}$$

Various statistics: GFD, D2Sig, LogD, Eigvec angle, $\cdots$.

Access general case result

Definition 2

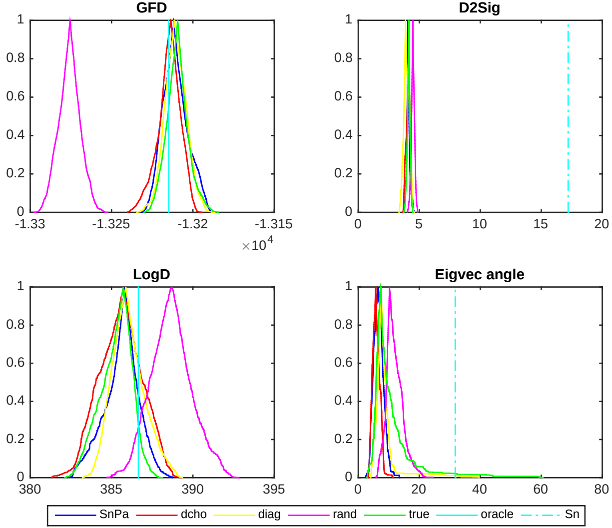
Suppose the empirical cumulative density function is f_{ecdf}, then the **confidence curve function** f_{cc} *is defined as the following:*

$$f_{cc}(x) = \begin{cases} 2(1 - f_{ecdf}(x)), & \text{if } f_{ecdf}(x) \geq 0.5, \\ 2f_{ecdf}(x), & \text{otherwise.} \end{cases}$$

Various statistics: GFD, D2Sig, LogD, Eigvec angle, $\cdots$.

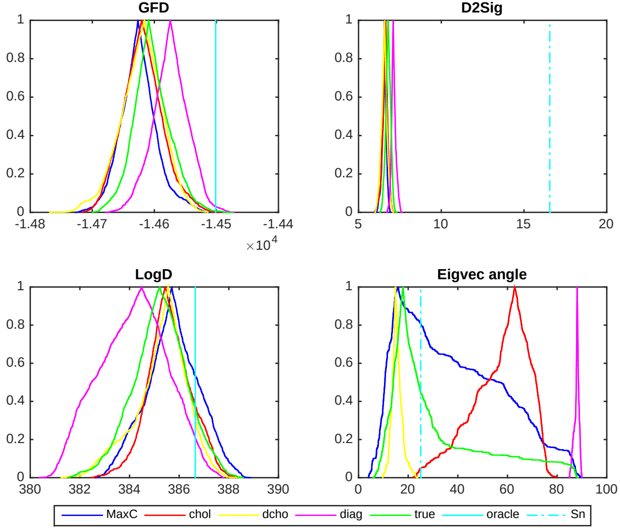
GC1: sparsity known

Confidence curves for $p = 50$, $n = 50$, $burnin = 100000$.



GC2: sparsity unknown

$p = 50$, $n = 50$, $burnin = 100000$. Same data as before.



- Recap
  - GFI does NOT require priors.
  - Functional inverse of DGE $\Rightarrow$ GFD.
  - $r(A|\mathbf{Y})$ is consistent if $A \leftrightarrow \Sigma$ is 1-to-1.
  - The clique result can be extended to $p \to \infty$, with $p/\sqrt{n} \to 0$.
  - Sample from GFD via MCMC/RJMCMC.
  - Fiducial samples $\Rightarrow$ confidence regions.
- Ongoing work
  - Improve sampling efficiency, e.g. BDMCMC instead of RJMCMC [Stephens (2000)].
  - Penalty tuning.

This project is supported by NSF–1007543, 106441.

Thank You!