Covariance Estimation via Fiducial Inference

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

1. University of North Carolina at Chapel Hill
2. University of California at Davis

June 17, 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)

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 $\mathbf{G}$

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

- 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 $\mathbf{G}$

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

- 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 $\mathbf{G}$

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

- 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 $D(A) = |\det A|$
- $\|\cdot\|_2$ 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 $J(\mathbf{x}, \xi) = D\left(\frac{d}{d\xi} \mathbf{G}(\mathbf{u}, \xi) \Big|_{\mathbf{u}=\mathbf{G}^{-1}(\mathbf{x},\xi)} \right)$

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

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

- $n = p$ gives $D(A) = |\det A|$
- $\|\bullet\|_2$ 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)

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

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

- 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

$$
\frac{\partial W_{\mathbf{i}}}{\partial A} =
\begin{array}{c}
\begin{array}{cccc} \quad B_1 \quad & \quad B_2 \quad & \quad \cdots \quad & \quad B_p \quad \end{array} \\
\begin{pmatrix}
(A^{-1}Y_{i_1})^T & & & \\
 & (A^{-1}Y_{i_1})^T & & \\
 & & \ddots & \\
 & & & (A^{-1}Y_{i_1})^T \\
\vdots & \vdots & \vdots & \vdots \\
(A^{-1}Y_{i_p})^T & & & \\
 & (A^{-1}Y_{i_p})^T & & \\
 & & \ddots & \\
 & & & (A^{-1}Y_{i_p})^T
\end{pmatrix}
\end{array}
\Rightarrow
\begin{array}{c}
\begin{array}{ccc} \quad B'_1 \quad & \quad \cdots \quad & \quad B'_p \quad \end{array} \\
\begin{pmatrix}
(A^{-1}Y_{i_1})^T & & \\
\vdots & & \\
(A^{-1}Y_{i_p})^T & & \\
 & \ddots & \\
 & & (A^{-1}Y_{i_1})^T \\
 & & \vdots \\
 & & (A^{-1}Y_{i_p})^T
\end{pmatrix}
\end{array}
$$

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

Special Case 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 longer IW
- 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 $A^{-1}Y_{i_k}$'s & sparse structure of A

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

General Case: possible fixed zeros in A

- No longer IW
- 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 $A^{-1}Y_{i_k}$'s & sparse structure of A

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

General Case: possible fixed zeros in A

- No longer IW
- 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 $A^{-1}Y_{i_k}$'s & sparse structure of A

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

General Case: possible fixed zeros in A

- No longer IW
- 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 $A^{-1}Y_{i_k}$'s & sparse structure of A

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

Clique model selection

The minimum description length (MDL) for a clique model A:

$$q(A) = \frac{1}{2}\left(\sum_{s=1}^{k} g_s^2\right)\log n + (p+1)\log k$$

Apply the MDL penalty & integrate out A, GFD for the model M

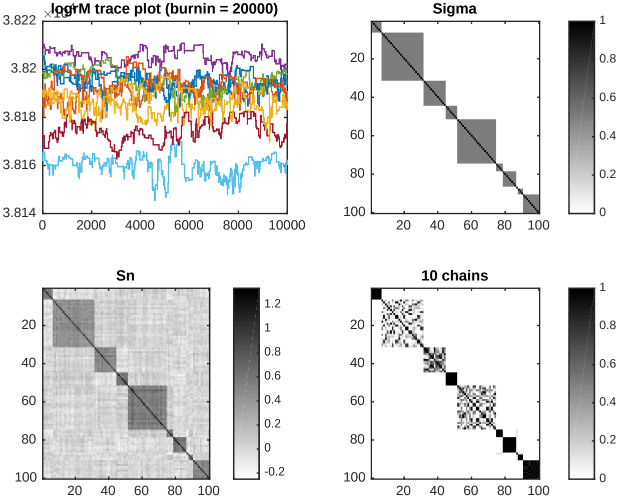
$$r(M|\mathbf{Y}) \propto \frac{\prod_{s=1}^{k}\left[\Gamma_{g_s}\left(\frac{n}{2}\right) C_s(\mathbf{Y})(2\pi)^{\frac{g_s(g_s-1)}{2}} \left|\det S_n^s\right|^{-\frac{n}{2}}\right]}{\exp\left\{\frac{1}{2}\left(\sum_{s=1}^{k} g_s^2\right)\log n + (p+1)\log k\right\}}$$

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 model selection

$k = 10, p = 100, n = 200.$



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 Markov chain Monte Carlo (MCMC) techniques

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

- reversible jump MCMC (RJMCMC) methods
- Metroplois-Hastings with move types: update, birth, death
- acceptance rate based update proposal variance
- adaptive variance for birth proposal using *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 Markov chain Monte Carlo (MCMC) techniques

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

- reversible jump MCMC (RJMCMC) methods
- Metroplois-Hastings with move types: update, birth, death
- acceptance rate based update proposal variance
- adaptive variance for birth proposal using *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 Markov chain Monte Carlo (MCMC) techniques

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

- reversible jump MCMC (RJMCMC) methods
- Metroplois-Hastings with move types: update, birth, death
- acceptance rate based update proposal variance
- adaptive variance for birth proposal using *zeroth-order-method* [Brooks, Giudici & Roberts (2003)]

Access general case result

Definition 1

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

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

Six statistics: GFD, Dim, D2Sn, D2Sig, LogD, Eig1, Eig1/Eig2, Eigvec angle, Cond.

Multiple initial states

Access general case result

Definition 1

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

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

Six statistics: GFD, Dim, D2Sn, D2Sig, LogD, Eig1, Eig1/Eig2, Eigvec angle, Cond.

Multiple initial states

GC1: sparsity known

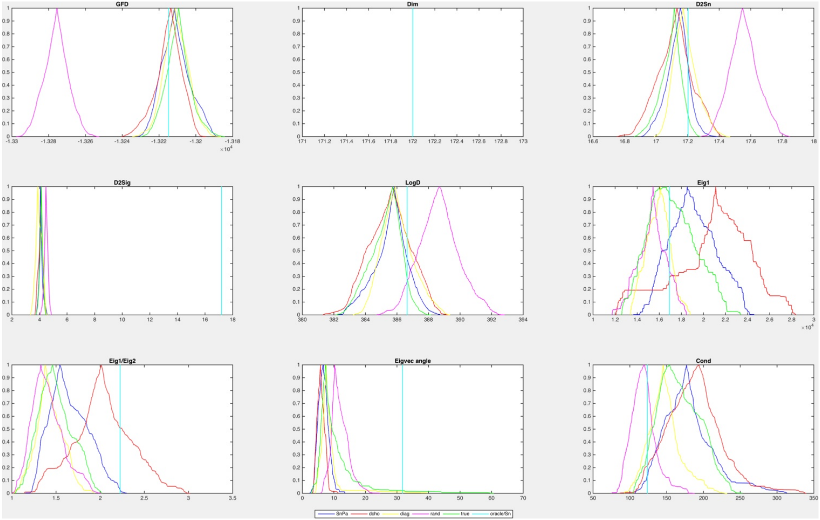


Figure: Confidence curves for $p = 50$, $n = 50$, $burnin = 100000$.

GC2: sparsity unknown

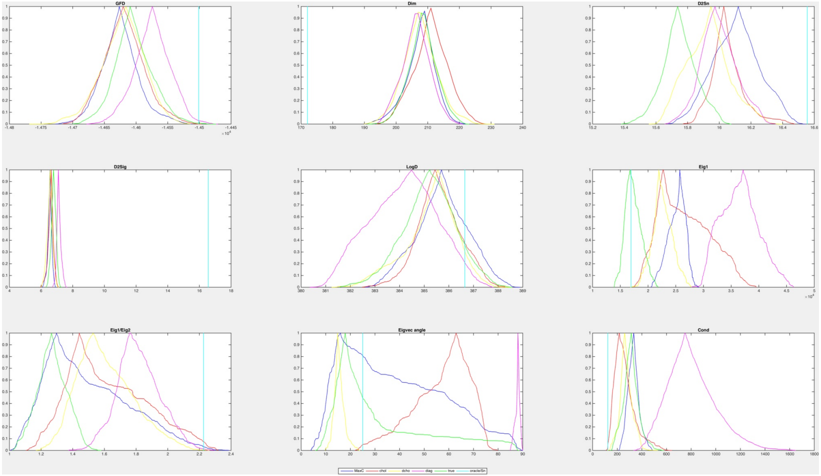


Figure: Confidence curves for $p = 50$, $n = 50$, $burnin = 100000$. Same data as before.

Recap

- GFI approach allows to obtain a measure on the parameter space without requiring priors.
- GFD is the distribution of the functional inverse of the data generating equation.
- GFD of the covariate matrix A is consistent if there is a one-to-one correspondence between A and $\Sigma = AA^T$.
- Sampling from GFD can be done via MCMC techniques. We propose an adaptive RJMCMC method for the general case with sparsity unknown.
- Confidence region can be defined based on the fiducial samples.
- A more sophisticated penalty term may be needed.

Thank You!