

Post-selection Inference

Arun Kumar Kuchibhotla
University of Pennsylvania

Larry Brown



Andreas Buja



Junhui Cai



Linda Zhao



Ed George



Outline

Data Snooping: Effects and Examples

Formulation of the Problem

Three Solutions

Implications and Conclusions

Outline

Data Snooping: Effects and Examples

- Effects illustrated with stepwise selection.
- Data snooping in textbooks and practice.

Formulation of the Problem

- The Problem

Three Solutions

- Bonferroni, PoSI and UPoSI

Implications and Conclusions

Outline

Data Snooping: Effects and Examples

- Effects illustrated with stepwise selection.
- Data snooping in textbooks and practice.

Formulation of the Problem

- The Problem

Three Solutions

- Bonferroni, PoSI and UPoSI

Implications and Conclusions

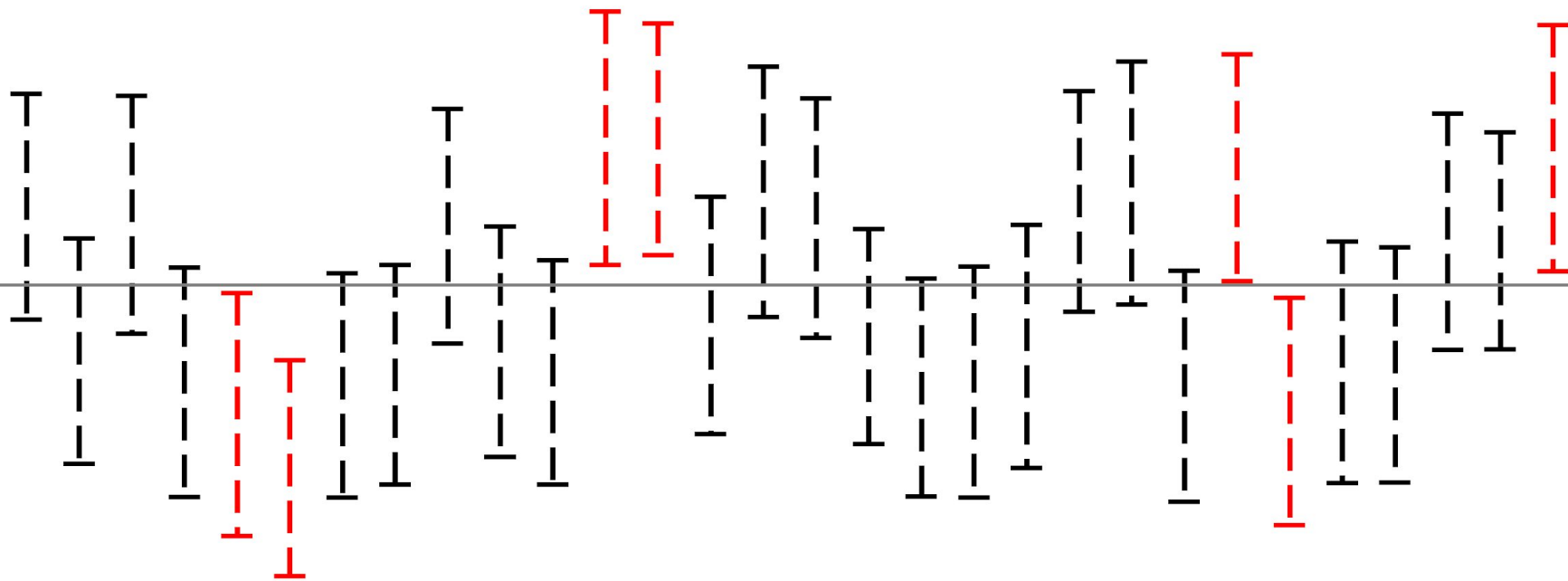
Effects of Data Snooping: Stepwise Selection

$(X, Y) \sim N(0, I_{p+1}) \Rightarrow$ 500 observations

$X_{\hat{j}}$ most correlated with $Y \Rightarrow$ 95% CI for slope

$p = 5$

Unadjusted: **76.9%**



Effects of Data Snooping: Stepwise Selection

$$(X, Y) \sim N(0, I_{p+1})$$

\Rightarrow

500 observations

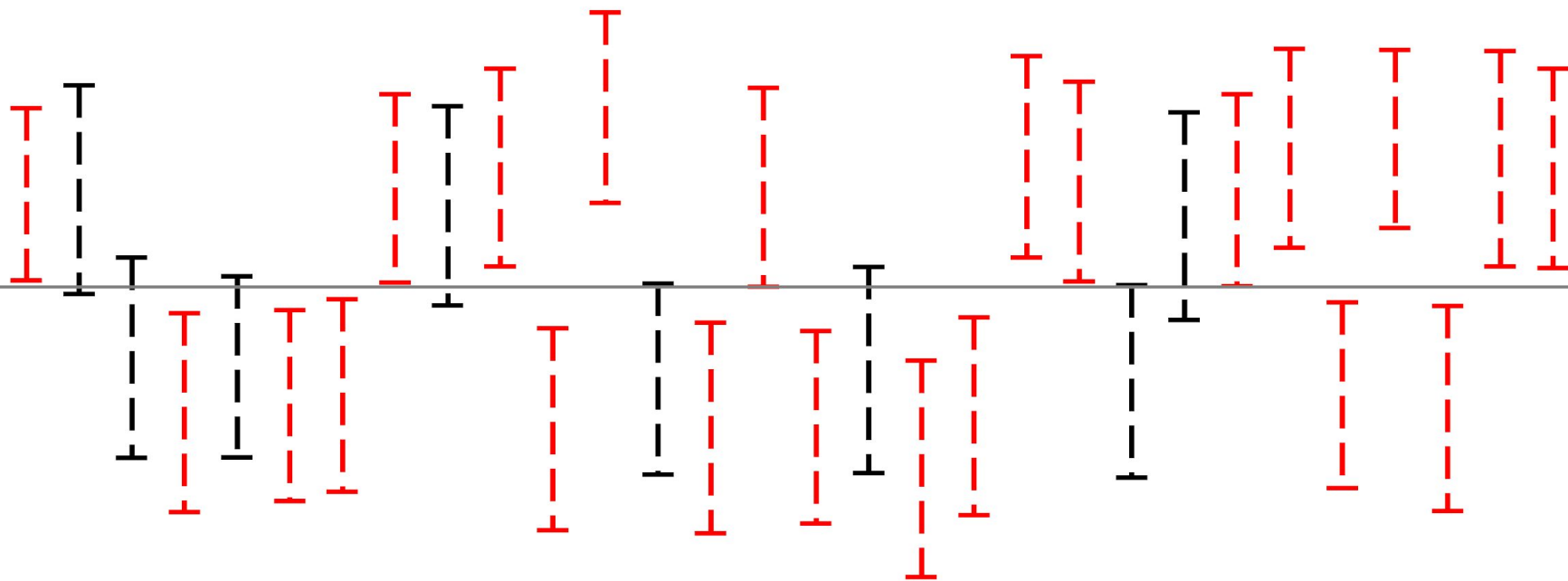
$X_{\hat{j}}$ most correlated with Y

\Rightarrow

95% CI for slope

$p = 20$

Unadjusted: **32.6%**

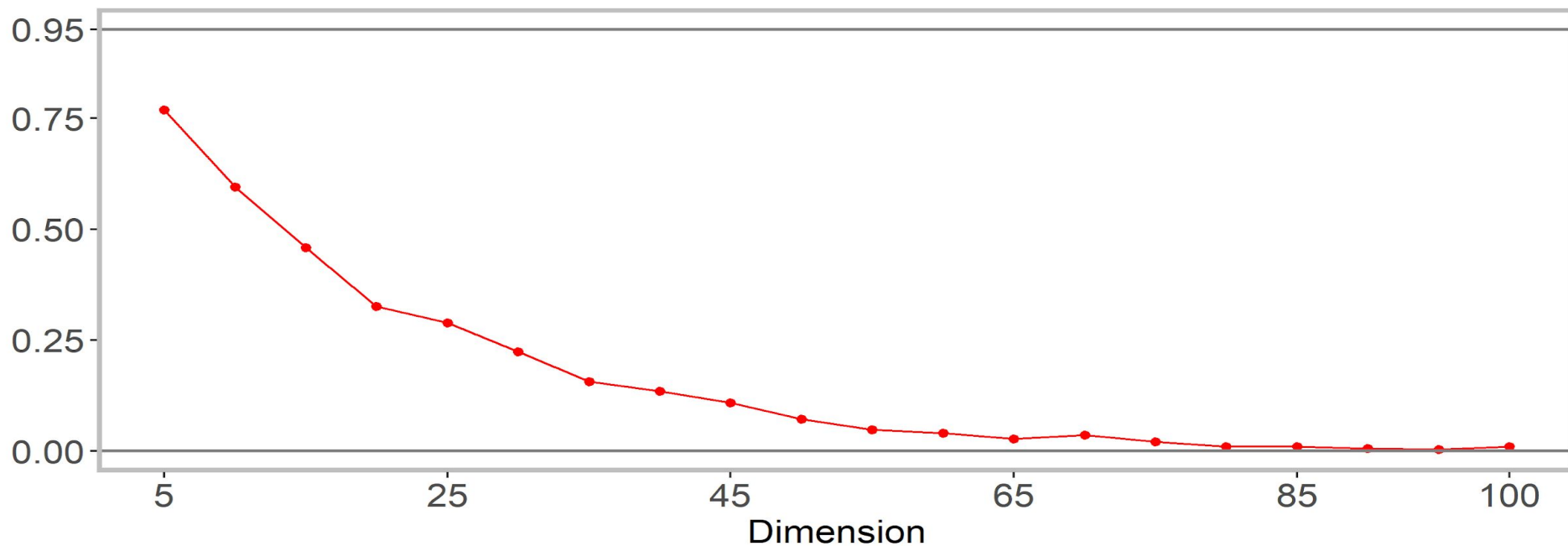


Effects of Data Snooping: Stepwise Selection

$(X, Y) \sim N(0, I_{p+1}) \Rightarrow$ 500 observations

$X_{\hat{j}}$ most correlated with $Y \Rightarrow$ 95% CI for slope

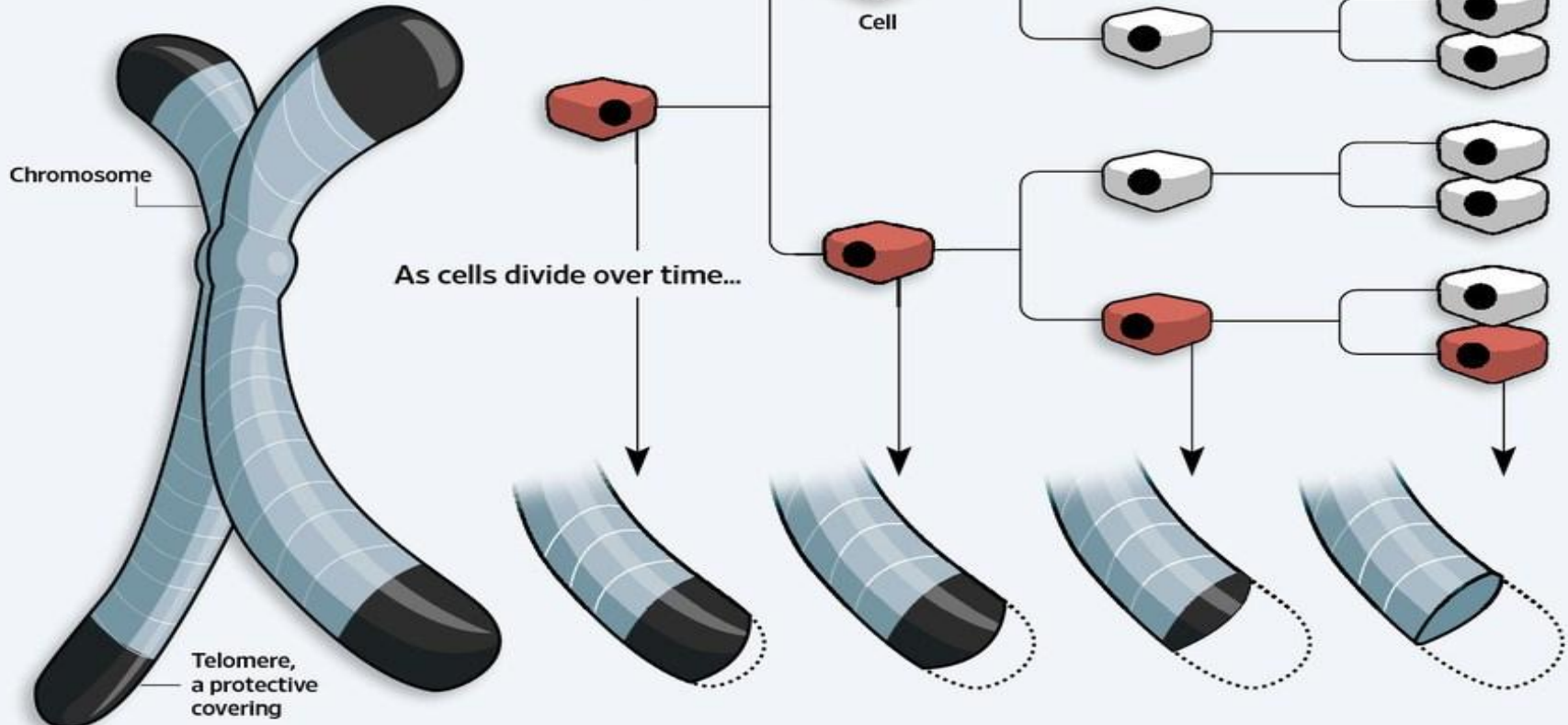
Empirical Coverage



Telomere Length Analysis

What We Lose With Age

As we grow older, telomeres at the end of our chromosomes shrink. New research suggests major depression also is linked to shorter telomeres, a sign of 'accelerated aging.'



Source: WSJ research

...telomeres shorten, and eventually cell division stops.

The Wall Street Journal

Case Study: Covariate Selection

nature.com

Scientific Reports, 2019

WGS-based telomere length analysis in Dutch family trios implicates stronger maternal inheritance and a role for *RRM1* gene

“The MLR models were tested by **sequential introduction of predictors and interaction terms.**

• • •

ultimately, from the **three best models** with similar adjusted R squared values the **simplest one was chosen.**”

Outline

Data Snooping: Effects and Examples

- Effects illustrated with stepwise selection.
- Data snooping in textbooks and practice.

Formulation of the Problem

- **The Problem**

Three Solutions

- Bonferroni, PoSI and UPoSI

Implications and Conclusions

Recap: Classical Inference

What does classical inference guarantee?

If 100 analysts report 100 confidence intervals from 100 studies, that is,

$$\mathbb{P}(\theta_j \in \widehat{\text{CI}}_j) \approx 0.95, \quad \forall \quad 1 \leq j \leq 100,$$

then about 95 of these confidence intervals actually contain their respective truths.

In the absence
of exploration:

$$\widehat{\text{CI}}_j := \hat{\theta}_j \pm 1.96 \frac{\hat{\sigma}_j}{\sqrt{n}}.$$

But the practice is ...

- Analysts start with the regression data, say.
- Explore the data. Find \widehat{M} and the estimator:

$$\widehat{\beta}_{\widehat{M}} := \operatorname{argmin}_{\theta} \frac{1}{n} \sum_{i=1}^n (Y_i - X_{i,\widehat{M}}^\top \theta)^2$$

- Report the confidence interval for $\widehat{j} \in \widehat{M}$:

$$\widehat{CI}_{\widehat{j} \cdot \widehat{M}}^{\text{naive}} := \widehat{\beta}_{\widehat{j} \cdot \widehat{M}} \pm 1.96 \frac{\widehat{\sigma}_{\widehat{j} \cdot \widehat{M}}}{\sqrt{n}}.$$

The selection of \widehat{j} and \widehat{M} is not systematic, in practice.

Problem: Post-selection Inference

Following classical inference, we want to construct a CI based on $\hat{\beta}_{\widehat{M}}$, with correct coverage.

For each $M \subseteq \{1, 2, \dots, p\}$,

$$\beta_M := \operatorname{argmin}_{\theta \in \mathbb{R}^{|M|}} \frac{1}{n} \sum_{i=1}^n \mathbb{E} \left[(Y_i - X_i \theta)^2 \right]$$

Problem: Construct $\widehat{CI}_{\widehat{j} \cdot \widehat{M}}$, such that

$$\liminf_{n \rightarrow \infty} \mathbb{P} \left(\beta_{\widehat{j} \cdot \widehat{M}} \in \widehat{CI}_{\widehat{j} \cdot \widehat{M}} \right) \geq 1 - \alpha.$$

irrespective of how

\widehat{M} and $\widehat{j} \in \widehat{M}$ are chosen.

Literature Overview

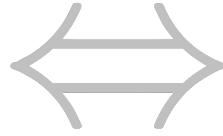
Buehler and Feddersen (1963); Olshen (1973); Sen (1979); Rencher and Pun (1980); Freedman (1983); Sen and Saleh (1987); Dijkstra and Veldkamp (1988); Hurvich and Tsai (1990); Potscher (1991); Pfeiffer, Redd and Carroll (2017).



Cox (1965); Kabaila (1998); Hjort and Claeskens (2003); Claeskens and Carroll (2007); **Berk et al. (2013)**; Lee, Sun, Sun and Taylor (2016); Tibshirani, Taylor, Lockhart and Tibshirani (2016); **Bachoc, Preinerstorfer and Steinberger (2019)**; Rinaldo, Wasserman, G'Sell and Lei (2019).

A Key Result

Simultaneous Inference
FWER Control



Post-selection
Inference

$$\mathbb{P} \left(\bigcap_{|M| \leq k, j \in M} \{ \beta_{j \cdot M} \in \widehat{\text{CI}}_{j \cdot M} \} \right) = \inf_{|\widehat{M}| \leq k, \widehat{j} \in \widehat{M}} \mathbb{P} \left(\beta_{\widehat{j} \cdot \widehat{M}} \in \widehat{\text{CI}}_{\widehat{j} \cdot \widehat{M}} \right)$$

Proof:

$$\begin{aligned} \prod_{|M| \leq k, j \in M} \mathbf{1}\{ \beta_{j \cdot M} \in \widehat{\text{CI}}_{j \cdot M} \} &= \min_{|M| \leq k, j \in M} \mathbf{1}\{ \beta_{j \cdot M} \in \widehat{\text{CI}}_{j \cdot M} \} \\ &= \mathbf{1}\{ \beta_{\widehat{j} \cdot \widehat{M}} \in \widehat{\text{CI}}_{\widehat{j} \cdot \widehat{M}} \}. \end{aligned}$$

Outline

Data Snooping: Effects and Examples

- Effects illustrated with stepwise selection.
- Data snooping in textbooks and practice.

Formulation of the Problem

- The Problem

Three Solutions

- Bonferroni, PoSI and UPoSI

Implications and Conclusions

Solution 0: Bonferroni

The classical interval for $\beta_{j \cdot M}$ is 

$$\left\{ \theta : \left| \frac{\sqrt{n}(\hat{\beta}_{j \cdot M} - \theta)}{\hat{\sigma}_{j \cdot M}} \right| \leq z_{\alpha/2} \right\}.$$

For *simultaneity*, inflate the confidence regions:

$$\widehat{\text{CI}}_{j \cdot M}^{\text{Bonf}} := \left\{ \theta : \left| \frac{\sqrt{n}(\hat{\beta}_{j \cdot M} - \theta)}{\hat{\sigma}_{j \cdot M}} \right| \leq K_{\alpha}^{\text{Bonf}} \right\},$$

where $K_{\alpha}^{\text{Bonf}} := 1 - \frac{\alpha}{2} \left(\sum_{|M| \leq k, j \in M} 1 \right)^{-1}$ quantile of $Z \asymp \sqrt{k \log(ep/k)}$.

Solution 1: PoSI

Bonferroni doesn't account for dependence. The smallest inflation is given by

$$\widehat{\text{CI}}_{j \cdot M}^{\text{PoSI}} := \left\{ \theta : \left| \frac{\sqrt{n}(\widehat{\beta}_{j \cdot M} - \theta)}{\widehat{\sigma}_{j \cdot M}} \right| \leq K_{\alpha}^{\text{PoSI}} \right\},$$

where $K_{\alpha}^{\text{PoSI}} := (1 - \alpha)$ quantile of $\max_{|M| \leq k, j \in M} \left| \frac{n^{1/2}(\widehat{\beta}_{j \cdot M} - \beta_{j \cdot M})}{\widehat{\sigma}_{j \cdot M}} \right|$.

$$K_{\alpha}^{\text{Bonf}} \geq K_{\alpha}^{\text{PoSI}}$$

In the worst case,

$$K_{\alpha}^{\text{PoSI}} \asymp \sqrt{k \log(ep/k)}$$

Disadvantage

$$\widehat{\text{CI}}_{j \cdot \text{M}}^{\text{Bonf}} := \left\{ \theta : \left| \frac{\sqrt{n}(\hat{\beta}_{j \cdot \text{M}} - \theta)}{\hat{\sigma}_{j \cdot \text{M}}} \right| \leq K_{\alpha}^{\text{Bonf}} \right\},$$

$$\widehat{\text{CI}}_{j \cdot \text{M}}^{\text{PoSI}} := \left\{ \theta : \left| \frac{\sqrt{n}(\hat{\beta}_{j \cdot \text{M}} - \theta)}{\hat{\sigma}_{j \cdot \text{M}}} \right| \leq K_{\alpha}^{\text{PoSI}} \right\},$$

If we are considering selection of size up to 20,
then

Width for models of size 2

\approx

Width for models of size 20.

Solution 3: UPoSI

For every $M \subseteq \{1, 2, \dots, p\}$, define

$$\widehat{\text{CI}}_M^{\text{UPoSI}^*} := \left\{ \theta \in \mathbb{R}^{|M|} : \|\widehat{\Sigma}_M(\widehat{\beta}_M - \theta)\|_\infty \leq C_{xy} + C_{xx} \|\theta\|_1 \right\},$$

where $\widehat{\Sigma}_M := n^{-1} \sum_{i=1}^n X_{i,M} X_{i,M}^\top$ and

$$\mathbb{P} \left(\left\| \frac{1}{n} \sum_{i=1}^n X_i Y_i - \mathbb{E}[X_i Y_i] \right\|_\infty \leq C_{xy}, \left\| \frac{1}{n} \sum_{i=1}^n X_i X_i^\top - \mathbb{E}[X_i X_i^\top] \right\|_\infty \leq C_{xx} \right) \geq 1 - \alpha.$$

$$\mathbb{P} \left(\beta_{\widehat{M}} \in \widehat{\text{CI}}_{\widehat{M}}^{\text{UPoSI}^*} \right) \geq 1 - \alpha.$$

Key Steps in the Proof

Proof

For every $M \subseteq \{1, 2, \dots, p\}$,

$$\widehat{\Sigma}_M \widehat{\beta}_M = \frac{1}{n} \sum_{i=1}^n X_{i,M} Y_i$$

Proof

For every $M \subseteq \{1, 2, \dots, p\}$,

$$\hat{\Sigma}_M \hat{\beta}_M = \frac{1}{n} \sum_{i=1}^n X_{i,M} Y_i$$

$$\hat{\Sigma}_M (\hat{\beta}_M - \beta_M) = \frac{1}{n} \sum_{i=1}^n X_{i,M} Y_i - \frac{1}{n} \sum_{i=1}^n X_{i,M} X_{i,M}^\top \beta_M$$

Proof

For every $M \subseteq \{1, 2, \dots, p\}$,

$$\begin{aligned}\widehat{\Sigma}_M \widehat{\beta}_M &= \frac{1}{n} \sum_{i=1}^n X_{i,M} Y_i \\ \widehat{\Sigma}_M (\widehat{\beta}_M - \beta_M) &= \frac{1}{n} \sum_{i=1}^n X_{i,M} Y_i - \frac{1}{n} \sum_{i=1}^n X_{i,M} X_{i,M}^\top \beta_M \\ &\quad - \frac{1}{n} \sum_{i=1}^n \mathbb{E}[X_{i,M} Y_i] + \frac{1}{n} \sum_{i=1}^n \mathbb{E}[X_{i,M} X_{i,M}^\top] \beta_M\end{aligned}$$

Because

$$\beta_M = \operatorname{argmin}_{\theta \in \mathbb{R}^{|M|}} \frac{1}{n} \sum_{i=1}^n \mathbb{E}[(Y_i - X_{i,M}^\top \theta)^2] \quad \Rightarrow \quad \frac{1}{n} \sum_{i=1}^n \mathbb{E}[X_{i,M} (Y_i - X_{i,M}^\top \beta_M)] = 0.$$

Proof

For every $M \subseteq \{1, 2, \dots, p\}$,

$$\begin{aligned}\widehat{\Sigma}_M \widehat{\beta}_M &= \frac{1}{n} \sum_{i=1}^n X_{i,M} Y_i \\ \widehat{\Sigma}_M (\widehat{\beta}_M - \beta_M) &= \frac{1}{n} \sum_{i=1}^n X_{i,M} Y_i - \frac{1}{n} \sum_{i=1}^n X_{i,M} X_{i,M}^\top \beta_M \\ &\quad - \frac{1}{n} \sum_{i=1}^n \mathbb{E}[X_{i,M} Y_i] + \frac{1}{n} \sum_{i=1}^n \mathbb{E}[X_{i,M} X_{i,M}^\top] \beta_M \\ &= \frac{1}{n} \sum_{i=1}^n X_{i,M} Y_i - \mathbb{E}[X_{i,M} Y_i] \\ &\quad - \frac{1}{n} \sum_{i=1}^n \{X_{i,M} X_{i,M}^\top - \mathbb{E}[X_{i,M} X_{i,M}^\top]\} \beta_M\end{aligned}$$

Proof

For every $M \subseteq \{1, 2, \dots, p\}$,

$$\widehat{\Sigma}_M \widehat{\beta}_M = \frac{1}{n} \sum_{i=1}^n X_{i,M} Y_i$$

$$\begin{aligned} \widehat{\Sigma}_M (\widehat{\beta}_M - \beta_M) &= \frac{1}{n} \sum_{i=1}^n X_{i,M} Y_i - \mathbb{E}[X_{i,M} Y_i] \\ &\quad - \frac{1}{n} \sum_{i=1}^n \{X_{i,M} X_{i,M}^\top - \mathbb{E}[X_{i,M} X_{i,M}^\top]\} \beta_M \end{aligned}$$

Therefore,

$$\|\widehat{\Sigma}_M (\widehat{\beta}_M - \beta_M)\|_\infty \leq \left\| \frac{1}{n} \sum_{i=1}^n X_{i,M} Y_i - \mathbb{E}[X_{i,M} Y_i] \right\|_\infty + \left\| \frac{1}{n} \sum_{i=1}^n X_{i,M} X_{i,M}^\top - \mathbb{E}[X_{i,M} X_{i,M}^\top] \right\|_\infty \|\beta_M\|_1.$$

Proof

For every $M \subseteq \{1, 2, \dots, p\}$,

$$\|\widehat{\Sigma}_M(\widehat{\beta}_M - \beta_M)\|_\infty \leq \left\| \frac{1}{n} \sum_{i=1}^n X_{i,M} Y_i - \mathbb{E}[X_{i,M} Y_i] \right\|_\infty + \left\| \frac{1}{n} \sum_{i=1}^n X_{i,M} X_{i,M}^\top - \mathbb{E}[X_{i,M} X_{i,M}^\top] \right\|_\infty \|\beta_M\|_1.$$

Recall

$$\mathbb{P} \left(\left\| \frac{1}{n} \sum_{i=1}^n X_i Y_i - \mathbb{E}[X_i Y_i] \right\|_\infty \leq C_{xy}, \left\| \frac{1}{n} \sum_{i=1}^n X_i X_i^\top - \mathbb{E}[X_i X_i^\top] \right\|_\infty \leq C_{xx} \right) \geq 1 - \alpha.$$

Proof

For every $M \subseteq \{1, 2, \dots, p\}$,

$$\|\widehat{\Sigma}_M(\widehat{\beta}_M - \beta_M)\|_\infty \leq \left\| \frac{1}{n} \sum_{i=1}^n X_{i,M} Y_i - \mathbb{E}[X_{i,M} Y_i] \right\|_\infty + \left\| \frac{1}{n} \sum_{i=1}^n X_{i,M} X_{i,M}^\top - \mathbb{E}[X_{i,M} X_{i,M}^\top] \right\|_\infty \|\beta_M\|_1.$$

Curious case of fixed covariates:

$$\begin{aligned} \|\widehat{\Sigma}_M(\widehat{\beta}_M - \beta_M)\|_\infty &\leq \left\| \frac{1}{n} \sum_{i=1}^n X_{i,M} Y_i - \mathbb{E}[X_{i,M} Y_i] \right\|_\infty. \\ &\leq \left\| \frac{1}{n} \sum_{i=1}^n X_i Y_i - \mathbb{E}[X_i Y_i] \right\|_\infty. \end{aligned}$$

Hence, $\mathbb{P} \left(\bigcap_M \{ \|\widehat{\Sigma}_M(\widehat{\beta}_M - \beta_M)\|_\infty \leq C_{xy} \} \right) \geq 1 - \alpha.$

Proof

For every $M \subseteq \{1, 2, \dots, p\}$,

$$\begin{aligned} \|\widehat{\Sigma}_M(\widehat{\beta}_M - \beta_M)\|_\infty &\leq \left\| \frac{1}{n} \sum_{i=1}^n X_{i,M} Y_i - \mathbb{E}[X_{i,M} Y_i] \right\|_\infty + \left\| \frac{1}{n} \sum_{i=1}^n X_{i,M} X_{i,M}^\top - \mathbb{E}[X_{i,M} X_{i,M}^\top] \right\|_\infty \|\beta_M\|_1. \\ &\leq \left\| \frac{1}{n} \sum_{i=1}^n X_i Y_i - \mathbb{E}[X_i Y_i] \right\|_\infty + \left\| \frac{1}{n} \sum_{i=1}^n X_i X_i^\top - \mathbb{E}[X_i X_i^\top] \right\|_\infty \|\beta_M\|_1. \end{aligned}$$

Hence

$$\mathbb{P} \left(\bigcap_M \left\{ \|\widehat{\Sigma}_M(\widehat{\beta}_M - \beta_M)\|_\infty \leq C_{xy} + C_{xx} \|\beta_M\|_1 \right\} \right) \geq 1 - \alpha.$$

$$\mathbb{P} \left(\bigcap_M \left\{ \beta_M \in \widehat{\text{CI}}_M^{\text{UPoSI}^*} \right\} \right) \geq 1 - \alpha.$$

Simpler Regions

For every $M \subseteq \{1, 2, \dots, p\}$,

$$\begin{aligned} \|\widehat{\Sigma}_M(\widehat{\beta}_M - \beta_M)\|_\infty &\leq \left\| \frac{1}{n} \sum_{i=1}^n X_{i,M} Y_i - \mathbb{E}[X_{i,M} Y_i] \right\|_\infty + \left\| \frac{1}{n} \sum_{i=1}^n X_{i,M} X_{i,M}^\top - \mathbb{E}[X_{i,M} X_{i,M}^\top] \right\|_\infty \|\beta_M\|_1. \\ &\leq \left\| \frac{1}{n} \sum_{i=1}^n X_i Y_i - \mathbb{E}[X_i Y_i] \right\|_\infty + \left\| \frac{1}{n} \sum_{i=1}^n X_i X_i^\top - \mathbb{E}[X_i X_i^\top] \right\|_\infty \|\beta_M\|_1. \\ &\lesssim \left\| \frac{1}{n} \sum_{i=1}^n X_i Y_i - \mathbb{E}[X_i Y_i] \right\|_\infty + \left\| \frac{1}{n} \sum_{i=1}^n X_i X_i^\top - \mathbb{E}[X_i X_i^\top] \right\|_\infty \|\widehat{\beta}_M\|_1. \end{aligned}$$

Hence, if $k = o(n^{1/2} / \log(p))$

$$\liminf_{n \rightarrow \infty} \mathbb{P} \left(\bigcap_{|M| \leq k} \left\{ \|\widehat{\Sigma}_M(\widehat{\beta}_M - \beta_M)\|_\infty \leq C_{xy} + C_{xx} \|\widehat{\beta}_M\|_1 \right\} \right) \geq 1 - \alpha.$$

Solution 3: Recap

For every $M \subseteq \{1, 2, \dots, p\}$, define

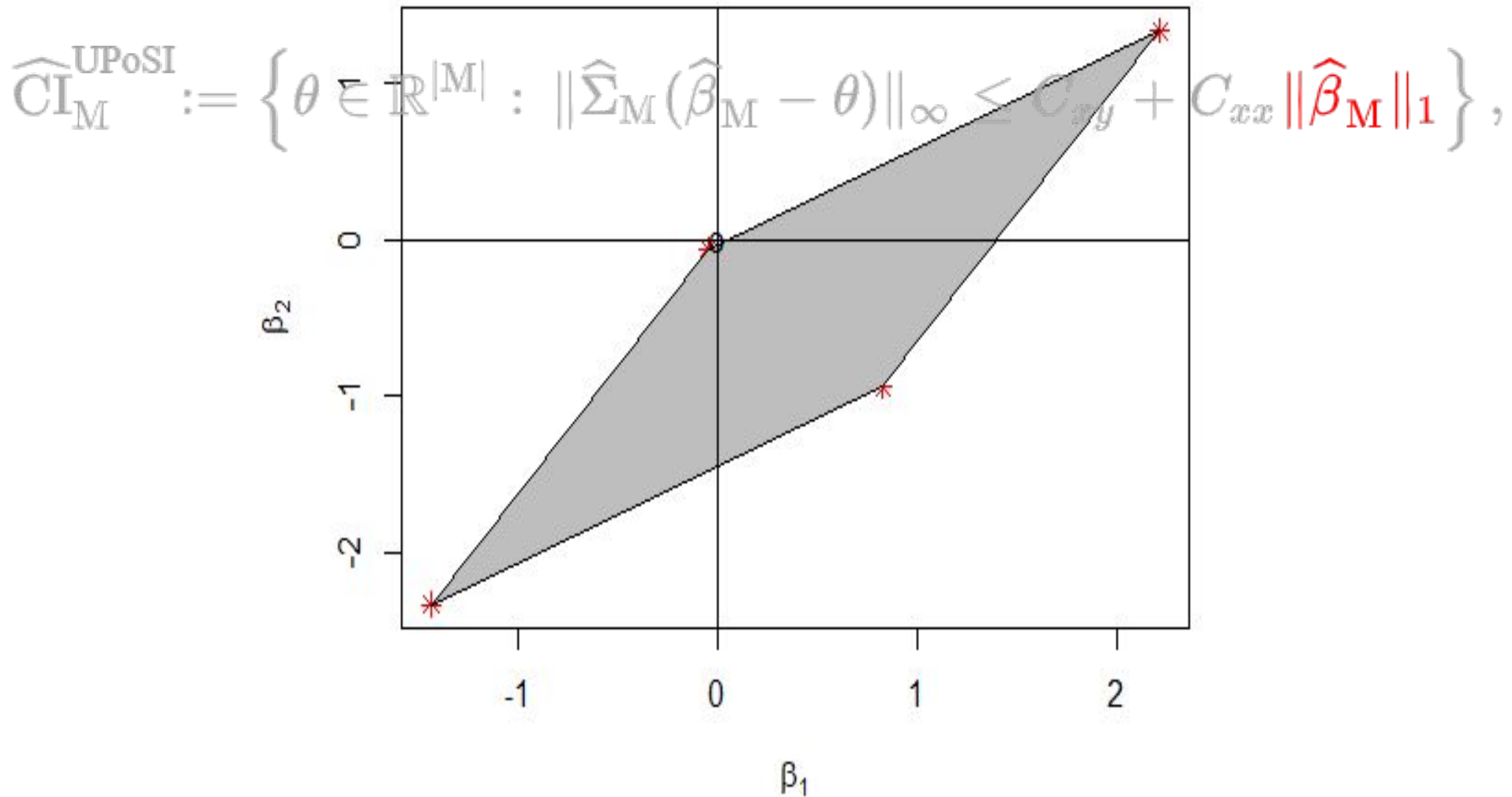
$$\widehat{\text{CI}}_M^{\text{UPoSI}^*} := \left\{ \theta \in \mathbb{R}^{|M|} : \|\widehat{\Sigma}_M(\widehat{\beta}_M - \theta)\|_\infty \leq C_{xy} + C_{xx} \|\theta\|_1 \right\},$$
$$\widehat{\text{CI}}_M^{\text{UPoSI}} := \left\{ \theta \in \mathbb{R}^{|M|} : \|\widehat{\Sigma}_M(\widehat{\beta}_M - \theta)\|_\infty \leq C_{xy} + C_{xx} \|\widehat{\beta}_M\|_1 \right\},$$

where

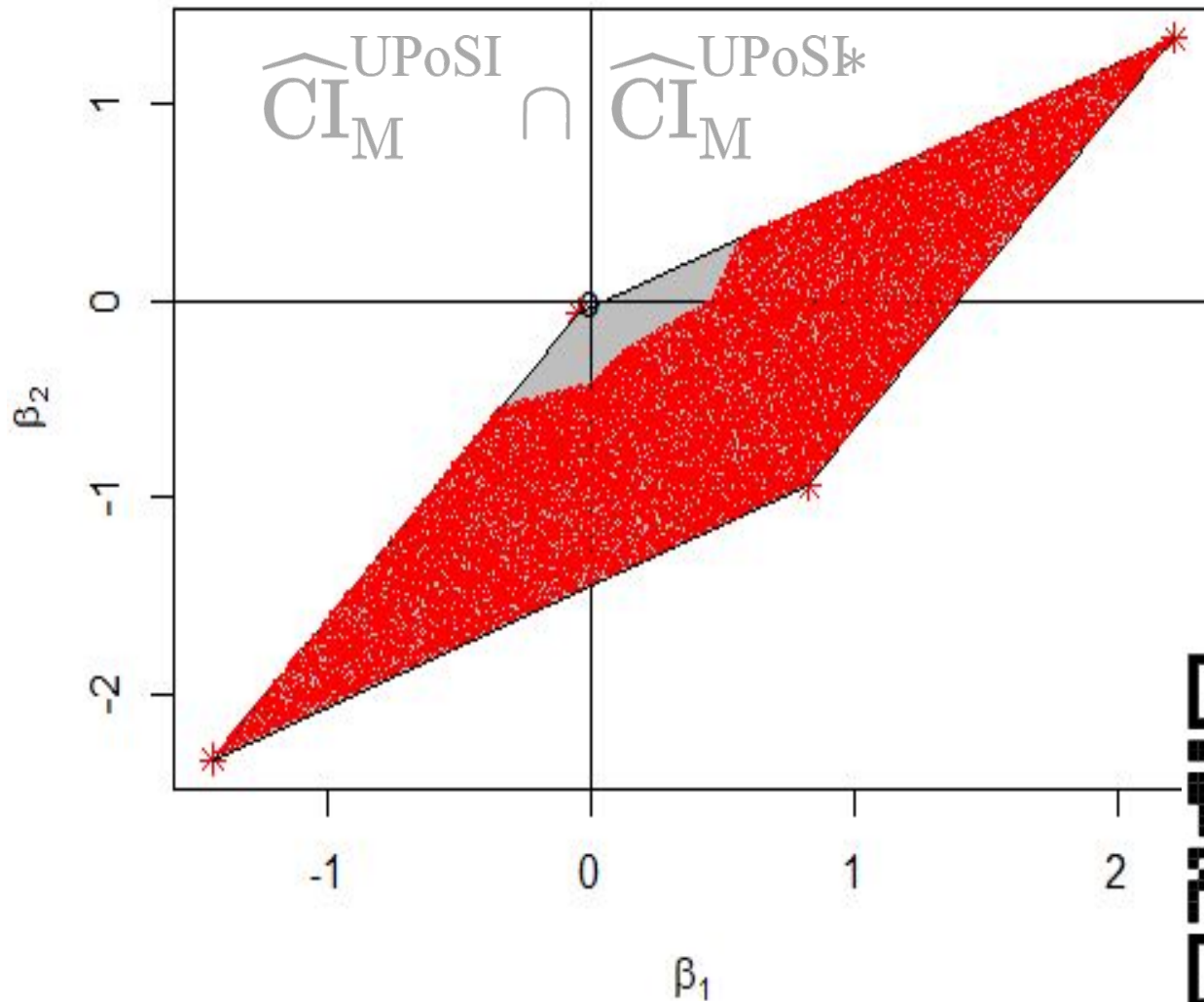
$$\mathbb{P} \left(\left\| \frac{1}{n} \sum_{i=1}^n X_i Y_i - \mathbb{E}[X_i Y_i] \right\|_\infty \leq C_{xy}, \left\| \frac{1}{n} \sum_{i=1}^n X_i X_i^\top - \mathbb{E}[X_i X_i^\top] \right\|_\infty \leq C_{xx} \right) \geq 1 - \alpha.$$

$$\liminf_{n \rightarrow \infty} \mathbb{P} \left(\beta_{\widehat{M}} \in \widehat{\text{CI}}_{\widehat{M}}^{\text{UPoSI}^*} \cap \widehat{\text{CI}}_{\widehat{M}}^{\text{UPoSI}} \right) \geq 1 - \alpha.$$

Solution 3: Illustration



Solution 3: Illustration



Outline

Data Snooping: Effects and Examples

- Effects illustrated with stepwise selection.
- Data snooping in textbooks and practice.

Formulation of the Problem

- The Problem

Three Solutions

- Bonferroni, PoSI and UPoSI

Implications and Conclusions

Recap

$$\widehat{\text{CI}}_{j \cdot \mathbf{M}}^{\text{Bonf}} := \left\{ \theta : \left| \frac{\sqrt{n}(\hat{\beta}_{j \cdot \mathbf{M}} - \theta)}{\hat{\sigma}_{j \cdot \mathbf{M}}} \right| \leq K_{\alpha}^{\text{Bonf}} \right\},$$

$$\widehat{\text{CI}}_{j \cdot \mathbf{M}}^{\text{PoSI}} := \left\{ \theta : \left| \frac{\sqrt{n}(\hat{\beta}_{j \cdot \mathbf{M}} - \theta)}{\hat{\sigma}_{j \cdot \mathbf{M}}} \right| \leq K_{\alpha}^{\text{PoSI}} \right\},$$

$$\widehat{\text{CI}}_{\mathbf{M}}^{\text{UPoSI}} := \left\{ \theta \in \mathbb{R}^{|\mathbf{M}|} : \|\hat{\Sigma}_{\mathbf{M}}(\hat{\beta}_{\mathbf{M}} - \theta)\|_{\infty} \leq C_{xy} + C_{xx} \|\hat{\beta}_{\mathbf{M}}\|_1 \right\},$$

In the worst case,

$$\text{Leb} \left(\widehat{\text{CI}}_{\mathbf{M}}^{\text{Bonf}} \right) \asymp \text{Leb} \left(\widehat{\text{CI}}_{\mathbf{M}}^{\text{PoSI}} \right) \asymp \left(\sqrt{k \log(ep/k)/n} \right)^{|\mathbf{M}|}$$

$$\text{Leb} \left(\widehat{\text{CI}}_{\mathbf{M}}^{\text{UPoSI}} \right) \asymp \left(\sqrt{|\mathbf{M}| \log(p)/n} \right)^{|\mathbf{M}|}.$$

Recap

$$\widehat{\text{CI}}_{j \cdot \mathbf{M}}^{\text{Bonf}} := \left\{ \theta : \left| \frac{\sqrt{n}(\widehat{\beta}_{j \cdot \mathbf{M}} - \theta)}{\widehat{\sigma}_{j \cdot \mathbf{M}}} \right| \leq K_{\alpha}^{\text{Bonf}} \right\},$$

$$\widehat{\text{CI}}_{j \cdot \mathbf{M}}^{\text{PoSI}} := \left\{ \theta : \left| \frac{\sqrt{n}(\widehat{\beta}_{j \cdot \mathbf{M}} - \theta)}{\widehat{\sigma}_{j \cdot \mathbf{M}}} \right| \leq K_{\alpha}^{\text{PoSI}} \right\},$$

$$\widehat{\text{CI}}_{\mathbf{M}}^{\text{UPoSI}} := \left\{ \theta \in \mathbb{R}^{|\mathbf{M}|} : \|\widehat{\Sigma}_{\mathbf{M}}(\widehat{\beta}_{\mathbf{M}} - \theta)\|_{\infty} \leq C_{xy} + C_{xx} \|\widehat{\beta}_{\mathbf{M}}\|_1 \right\},$$

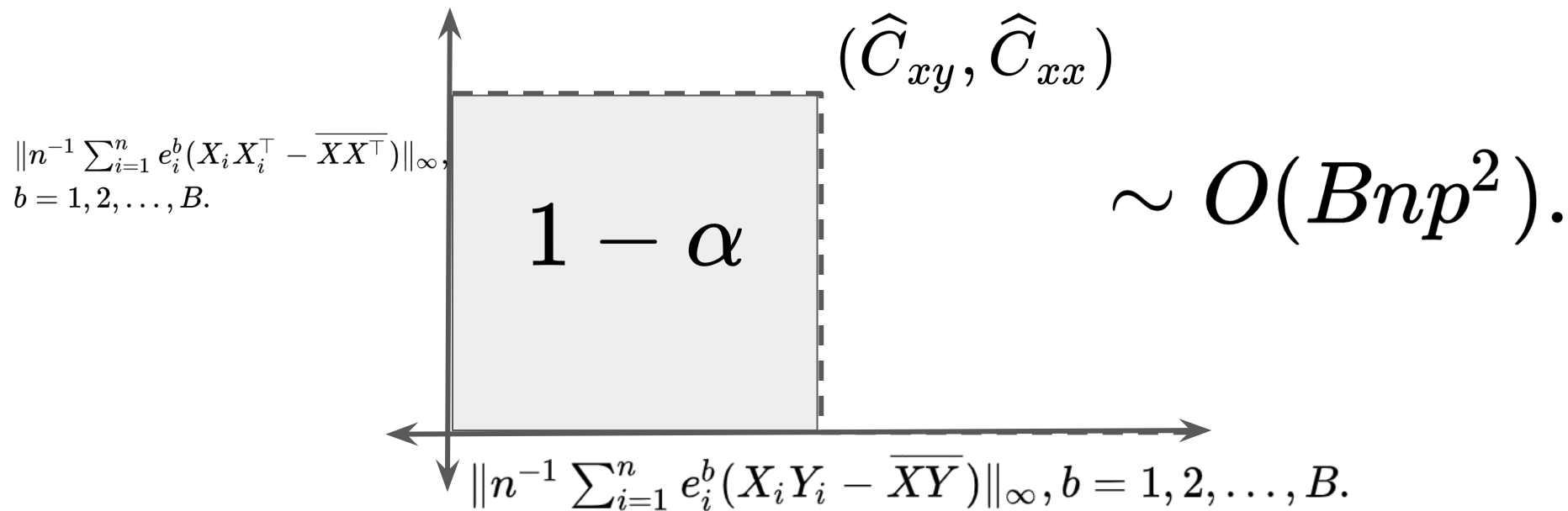
Interestingly, for fixed covariates ($C_{xx} = 0$) and

$$\text{Leb} \left(\widehat{\text{CI}}_{\mathbf{M}}^{\text{UPoSI}} \right) \asymp \left(\sqrt{\log(p)/n} \right)^{|\mathbf{M}|}.$$

Computation

$$\widehat{\text{CI}}_M^{\text{UPoSI}} := \left\{ \theta \in \mathbb{R}^{|\mathbf{M}|} : \|\widehat{\Sigma}_M(\widehat{\beta}_M - \theta)\|_\infty \leq C_{xy} + C_{xx} \|\widehat{\beta}_M\|_1 \right\},$$

$$\mathbb{P} \left(\left\| \frac{1}{n} \sum_{i=1}^n X_i Y_i - \mathbb{E}[X_i Y_i] \right\|_\infty \leq C_{xy}, \left\| \frac{1}{n} \sum_{i=1}^n X_i X_i^\top - \mathbb{E}[X_i X_i^\top] \right\|_\infty \leq C_{xx} \right) \geq 1 - \alpha.$$



Conclusions

$$\widehat{\text{CI}}_M^{\text{UPoSI}} := \left\{ \theta \in \mathbb{R}^{|\mathbf{M}|} : \|\widehat{\Sigma}_M(\widehat{\beta}_M - \theta)\|_\infty \leq C_{xy} + C_{xx} \|\widehat{\beta}_M\|_1 \right\},$$

- The regions are computationally efficient.
- They regions are also statistically “efficient”.
- The construction is only for OLS estimators.
- Applies for random/fixed design and dependence.
- Is there a similar construction for GLMs?

Conclusions (Contd.)

$$\widehat{\text{CI}}_M^{\text{UPoSI}} := \left\{ \theta \in \mathbb{R}^{|\mathbf{M}|} : \|\widehat{\Sigma}_M(\widehat{\beta}_M - \theta)\|_\infty \leq C_{xy} + C_{xx} \|\widehat{\beta}_M\|_1 \right\},$$

- There do exist statistically efficient PoSI for GLMs.
- These are based on pre-pivoting (Beran, 1988).
- They are not computationally efficient, readily.
- Randomized algorithms provide comp. efficiency.

References

Kuchibhotla A., Brown L., Buja A., Cai J., George E., Zhao L. (2019)
Valid Post-selection Inference in Model-free Linear Regression.
Annals of Statistics (Forthcoming).

Berk R., Brown L., Buja A., Zhang K., and Zhao L. (2013)
Valid Post-selection Inference. *Annals of Statistics*.

Bachoc F., Preinerstorfer D., and Steinberger L. (2016)
Uniformly valid confidence intervals post-model-selection.
Annals of Statistics (Forthcoming).

Chernozhukov V., Chetverikov D., Kato K., and Koike Y. (2019)
Improved Central Limit Theorem and bootstrap approximations in high
dimensions. *arXiv:1912.10529*.

Thank You for Your Attention.