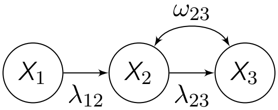
Leopold Mareis (PCIC 24)

Optimizing Experimental Design for Causal Effect Estimation with Partial Measurements

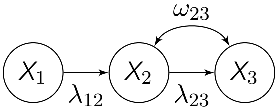
Fraunhofer IKS
Fraunhofer-Institut für Kognitive Systeme IKS

Gaussian Instrumental Variable Model - I



- X_1, X_2, X_3: Instrument, Treatment, Outcome

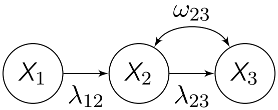
Gaussian Instrumental Variable Model - I



- X_1, X_2, X_3: Instrument, Treatment, Outcome
- Structural Equation Model (SEM)

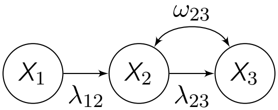
$$\begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ \lambda_{12} \neq 0 & 0 & 0 \\ 0 & \lambda_{23} & 0 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} + \varepsilon, \; \varepsilon \sim N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \omega_{11} & 0 & 0 \\ 0 & \omega_{22} & \omega_{23} \\ 0 & \omega_{23} & \omega_{33} \end{pmatrix} \right)$$

Gaussian Instrumental Variable Model - II



- Full Dataset: $D = (X_1^{(i)}, X_2^{(i)}, X_3^{(i)})_{i \in [n]}$. $\Sigma = Var(X)$. Estimator: $\hat{\lambda}_{23} = \hat{\sigma}_{13} / \hat{\sigma}_{12}$

Gaussian Instrumental Variable Model - II



- Full Dataset: $D = (X_1^{(i)}, X_2^{(i)}, X_3^{(i)})_{i \in [n]}$. $\Sigma = Var(X)$. Estimator: $\hat{\lambda}_{23} = \hat{\sigma}_{13}/\hat{\sigma}_{12}$
- Full and Partial Data: $(X_1^{(j)}, X_2^{(j)}, X_3^{(j)})_{j \in [m_1]}$ for $m_1 c_1$ and $(X_1^{(k)}, X_2^{(k)})_{k \in [m_2]}$ for $m_2 c_2$.
  Estimator:

$$\hat{\lambda}_{23} = \hat{\sigma}_{13}/(\frac{1}{1+\gamma}\hat{\sigma}_{12} + \frac{\gamma}{1+\gamma}\hat{\sigma}_{12}), \quad \gamma = m_2/m_1$$

Gaussian Instrumental Variable Model - II

$$\begin{tikzcd} X_1 \arrow[r, "\lambda_{12}"'] & X_2 \arrow[r, "\lambda_{23}"'] \arrow[r, leftrightarrow, bend left, "\omega_{23}"] & X_3 \end{tikzcd}$$

- Full Dataset: $D = (X_1^{(i)}, X_2^{(i)}, X_3^{(i)})_{i \in [n]}$. $\Sigma = Var(X)$. Estimator: $\hat{\lambda}_{23} = \hat{\sigma}_{13}/\hat{\sigma}_{12}$
- Full and Partial Data: $(X_1^{(j)}, X_2^{(j)}, X_3^{(j)})_{j \in [m_1]}$ for $m_1 c_1$ and $(X_1^{(k)}, X_2^{(k)})_{k \in [m_2]}$ for $m_2 c_2$. Estimator:

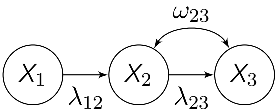
$$\hat{\lambda}_{23} = \hat{\sigma}_{13}/(\frac{1}{1+\gamma}\hat{\sigma}_{12} + \frac{\gamma}{1+\gamma}\hat{\sigma}_{12}), \quad \gamma = m_2/m_1$$

Optimization problem: $\boxed{\min_{m_1, m_2 > 0} Var(\hat{\lambda}_{23}) \quad \text{s.t.} \quad m_1 c_1 + m_2 c_2 \leq b.}$

Asymptotic Variance - Without Budget

Corollary: The asymptotic variance function $v(\gamma)$ *of* $\sqrt{m_1}(\hat{\lambda}_{23} - \lambda_{23})$ *is decreasing for* $\gamma > 0$ *if and only if*

$$\lambda_{12}^2 < -\frac{\omega_{22} + 2\omega_{23}/\lambda_{23}}{\omega_{11}}.$$

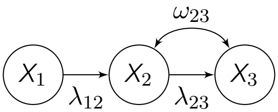


Asymptotic Variance - Without Budget

Corollary: The asymptotic variance function $v(\gamma)$ of $\sqrt{m_1}(\hat{\lambda}_{23} - \lambda_{23})$ is decreasing for $\gamma > 0$ if and only if

$$\lambda_{12}^2 < -\frac{\omega_{22} + 2\omega_{23}/\lambda_{23}}{\omega_{11}}.$$

- This can only be the case if $\omega_{23}\lambda_{23} < 0$, so counteracting effect paths.
-

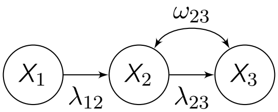


Asymptotic Variance - Without Budget

Corollary: The asymptotic variance function $v(\gamma)$ of $\sqrt{m_1}(\hat{\lambda}_{23} - \lambda_{23})$ is decreasing for $\gamma > 0$ if and only if

$$\lambda_{12}^2 < -\frac{\omega_{22} + 2\omega_{23}/\lambda_{23}}{\omega_{11}}.$$

- This can only be the case if $\omega_{23}\lambda_{23} < 0$, so counteracting effect paths.
- A weak instrument scenario ($\lambda_{12}^2\omega_{11}$ small) is necessary when requesting more X_{12} samples.



Asymptotic Variance - With Budget

Theorem: The solution to the optimization problem is given by $(b/c_1, 0)$ *if*

$$\chi_1 := \frac{c_1 - c_2}{c_2}(\sigma_{11}\sigma_{13}^2\sigma_{22} - 2\,\sigma_{11}\sigma_{12}\sigma_{13}\sigma_{23}) \leq 2\,\sigma_{12}^2\sigma_{13}^2 + \sigma_{11}\sigma_{12}^2\sigma_{33} =: \chi_2.$$

Otherwise, a positive amount of X_{12} *samples are requested and the solution reads*

$$\left(\frac{b - \xi c_2 n}{c_1 + \xi c_2}, \frac{\xi(b + c_1 n)}{c_1 + \xi c_2}\right), \text{ where } \xi = \min\left(-1 + \frac{\sqrt{\chi_1\chi_2}}{\chi_2}, \frac{b}{c_2 n}\right).$$

Asymptotic Variance - With Budget

Theorem: The solution to the optimization problem is given by $(b/c_1, 0)$ *if*

$$\chi_1 := \frac{c_1 - c_2}{c_2}(\sigma_{11}\sigma_{13}^2\sigma_{22} - 2\,\sigma_{11}\sigma_{12}\sigma_{13}\sigma_{23}) \leq 2\,\sigma_{12}^2\sigma_{13}^2 + \sigma_{11}\sigma_{12}^2\sigma_{33} =: \chi_2.$$

Otherwise, a positive amount of X_{12} *samples are requested and the solution reads*

$$\left(\frac{b - \xi c_2 n}{c_1 + \xi c_2}, \frac{\xi(b + c_1 n)}{c_1 + \xi c_2}\right), \textit{ where } \xi = \min\left(-1 + \frac{\sqrt{\chi_1\chi_2}}{\chi_2}, \frac{b}{c_2 n}\right).$$

- Only model parameters and the ratio c_1/c_2 determine whether additional partial samples are requested.

Asymptotic Variance - With Budget

Theorem: The solution to the optimization problem is given by $(b/c_1, 0)$ *if*

$$\chi_1 := \frac{c_1 - c_2}{c_2}(\sigma_{11}\sigma_{13}^2\sigma_{22} - 2\,\sigma_{11}\sigma_{12}\sigma_{13}\sigma_{23}) \leq 2\,\sigma_{12}^2\sigma_{13}^2 + \sigma_{11}\sigma_{12}^2\sigma_{33} =: \chi_2.$$

Otherwise, a positive amount of X_{12} *samples are requested and the solution reads*

$$\left(\frac{b - \xi c_2 n}{c_1 + \xi c_2}, \frac{\xi(b + c_1 n)}{c_1 + \xi c_2}\right), \text{ where } \xi = \min\left(-1 + \frac{\sqrt{\chi_1\chi_2}}{\chi_2}, \frac{b}{c_2 n}\right).$$

- Only model parameters and the ratio c_1/c_2 determine whether additional partial samples are requested.
- Weakening the instrument λ_{12}^2 weakens the condition $\chi_1 < \chi_2$.

Applications

- Provide p-value and power calculations for the test $H_0 : \lambda_{23} \leq 0, H_A : \lambda_{23} > 0$.

Applications

- Provide p-value and power calculations for the test $H_0 : \lambda_{23} \leq 0, H_A : \lambda_{23} > 0$.
- Need primary information on the model (Λ, Ω). Either through expert knowledge or an initial dataset.

Applications

- Provide p-value and power calculations for the test $H_0 : \lambda_{23} \leq 0, H_A : \lambda_{23} > 0$.
- Need primary information on the model (Λ, Ω). Either through expert knowledge or an initial dataset.

Real Data:

Applications

- Provide p-value and power calculations for the test $H_0 : \lambda_{23} \leq 0, H_A : \lambda_{23} > 0$.
- Need primary information on the model (Λ, Ω). Either through expert knowledge or an initial dataset.

Real Data:

- Aggressive Driving: (Lighting, Driving style, Speed). 21% variance reduction. $m^* = (211, 293)$.
- ICU Data: (Caregiver rate, Lasix, Urine). 68% variance reduction. $m^* = (88, 77)$.
- Assembly Line [**?**]: No partial data helpful.

References

Drton, M. et al. (2018).
Algebraic problems in structural equation modeling.
In *The 50th anniversary of Gröbner bases*, volume 77, pages 35–87. Mathematical Society of Japan Tokyo.

Göbler, K., Windisch, T., Drton, M., Pychynski, T., Roth, M., and Sonntag, S. (2024).
causalassembly: Generating realistic production data for benchmarking causal discovery.
In *Causal Learning and Reasoning*, pages 609–642. PMLR.

Henckel, L., Perković, E., and Maathuis, M. H. (2022).
Graphical criteria for efficient total effect estimation via adjustment in causal linear models.
Journal of the Royal Statistical Society Series B: Statistical Methodology, 84(2):579–599.

Murphy, S. A. (2005).
An experimental design for the development of adaptive treatment strategies.
Statistics in medicine, 24(10):1455–1481.

Thank you for your attention

Fraunhofer-Institut für Kognitive Systeme IKS

Outlook

- Procedure also applicable for general linear SEMs with arbitrary structure under identifiability.
- Gaussianity can be relaxed towards asymptotic normality (via, e.g., the maximum likelihood estimator).