Testing Many Possibly Irregular Polynomial Constraints

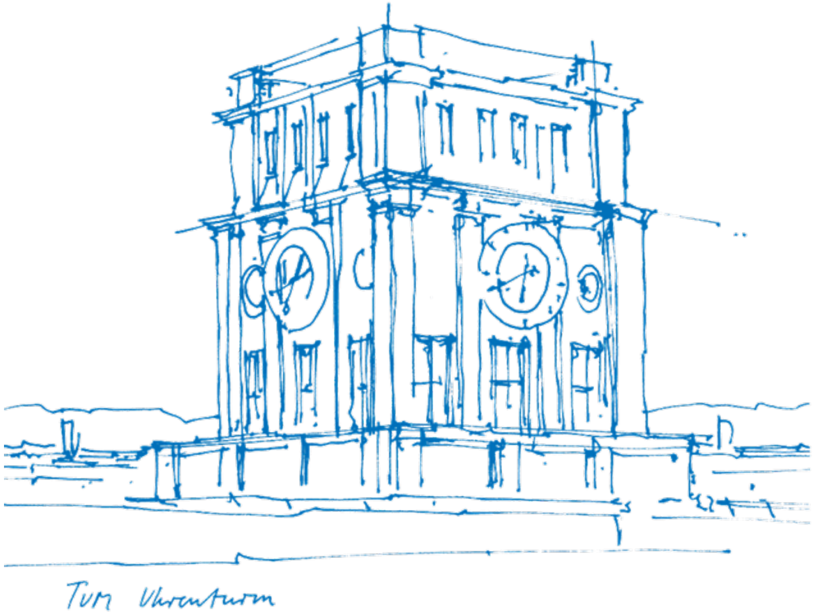
Nils Sturma

Research group Mathematical Statistics

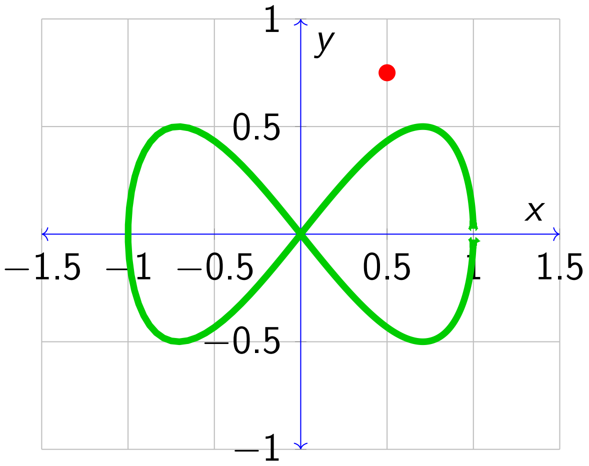
TUM School of Computation, Information and Technology

Technical University of Munich

(joint work with Mathias Drton and Dennis Leung)



Curve with a Singular Point: Lemniscate of Gerono



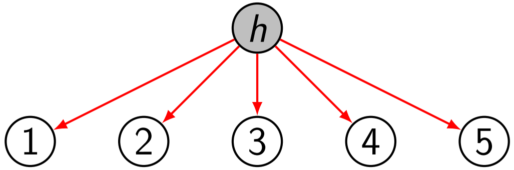
Parametrization

$$x = \frac{t^2 - 1}{t^2 + 1}, \qquad y = \frac{2t(t^2 - 1)}{(t^2 + 1)^2}$$

Characterization by Constraints

$$x^4 - x^2 + y^2 = 0$$

Statistical Example: One-Factor Analysis Model



$$X \sim N_k(0, \Sigma)$$

Parametrization

$$\Sigma = \Omega + \Gamma\Gamma^\top,$$

where $\Omega > 0$ diagonal and $\Gamma \in \mathbb{R}^{k \times 1}$.

Characterization by Constraints

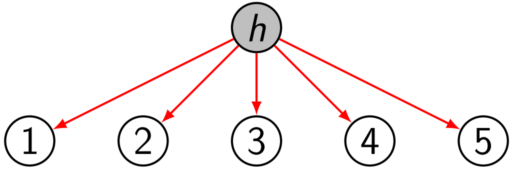
Equality constraints (tetrads):

$$\sigma_{uv}\sigma_{wz} - \sigma_{uw}\sigma_{vz} = 0.$$

Inequality constraints:

$$-\sigma_{uv}\sigma_{vw}\sigma_{uw} \leq 0, \qquad \sigma_{uv}^2\sigma_{vw}^2 - \sigma_{vv}^2\sigma_{uw}^2 \leq 0.$$

Statistical Example: One-Factor Analysis Model



$$X \sim N_k(0, \Sigma)$$

Parametrization

$$\Sigma = \Omega + \Gamma\Gamma^\top,$$

where $\Omega > 0$ diagonal and $\Gamma \in \mathbb{R}^{k\times 1}$.

Characterization by Constraints

Equality constraints (tetrads):

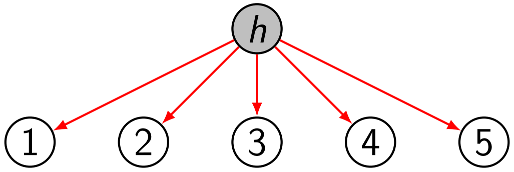
$$\sigma_{uv}\sigma_{wz} - \sigma_{uw}\sigma_{vz} = 0.$$

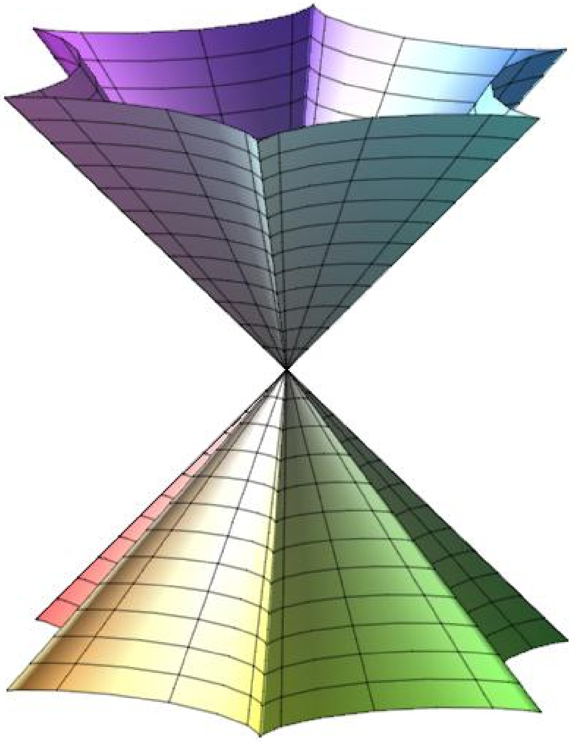
Inequality constraints:

$$-\sigma_{uv}\sigma_{vw}\sigma_{uw} \le 0, \qquad \sigma_{uv}^2\sigma_{vw}^2 - \sigma_{vv}^2\sigma_{uw}^2 \le 0.$$

Topic of the talk: Testing the goodness-of-fit based on samples $X_1, \ldots, X_n \sim N_k(0, \Sigma)$.

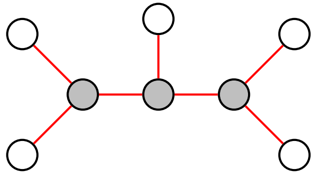
Statistical Example: One-Factor Analysis Model





Further Examples

- Gaussian Latent Tree Models

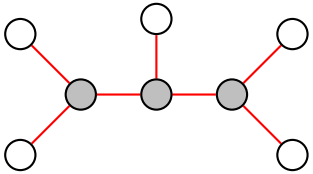


Characterized by vanishing of certain tetrads and inequality constraints on the covariance matrix. (Long paths $\longrightarrow$ small correlations)

Shiers, Zwiernik, Aston, Smith (2016).
The correlation space of Gaussian latent tree models and model selection without fitting. Biometrika, 103(3):531–545.

Further Examples

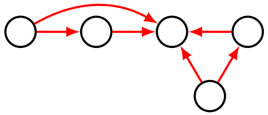
- Gaussian Latent Tree Models



Characterized by vanishing of certain tetrads and inequality constraints on the covariance matrix. (Long paths $\longrightarrow$ small correlations)

Shiers, Zwiernik, Aston, Smith (2016).
The correlation space of Gaussian latent tree models and model selection without fitting. Biometrika, 103(3):531–545.

- Linear Non-Gaussian Structural Equation Models



$$X = \Lambda^\top X + \varepsilon$$

Denote $S = (s_{ij})$ and $T = (t_{ijl})$ the second and third order moments of X.

$$\text{rk}\begin{pmatrix} s_{11} & s_{12} & \cdots & s_{1k} & s_{22} & s_{23} & \cdots & s_{kk} \\ t_{111} & t_{112} & \cdots & t_{11k} & t_{122} & t_{123} & \cdots & t_{1kk} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ t_{11k} & t_{12k} & \cdots & t_{1kk} & t_{22k} & t_{23k} & \cdots & t_{kkk} \end{pmatrix} = k$$

$$t_{111}t_{222}t_{333}t_{123} - (t_{222}t_{333}t_{112}t_{113} + t_{333}t_{111}t_{122}t_{223} + t_{111}t_{222}t_{333}t_{233}) - t_{123}(t_{111}t_{223}t_{233} + t_{222}t_{133}t_{113} + t_{333}t_{112}t_{122}) + \ldots = 0 \quad \text{(Aronhold invariant)}$$

Master Thesis Daniela Schkoda (2022).
Goodness-of-fit tests for non-Gaussian linear causal models.

- ...

General Setup: Testing Constraints on Statistical Models

Parametric family:

$$\mathcal{P} = \{P_\theta : \theta \in \Theta\}, \text{ where } \Theta \in \mathbb{R}^d.$$

Model:

$$\Theta_0 = \{\theta \in \Theta : \ f_j(\theta) \leq 0 \text{ for all } 1 \leq j \leq p\}.$$

Our main interest: Polynomial constraints f_j.

Based on samples $X_1, \ldots, X_n \sim P_\theta$ test

$$H_0 : \theta \in \Theta_0 \text{ vs. } H_1 : \theta \in \Theta \setminus \Theta_0.$$

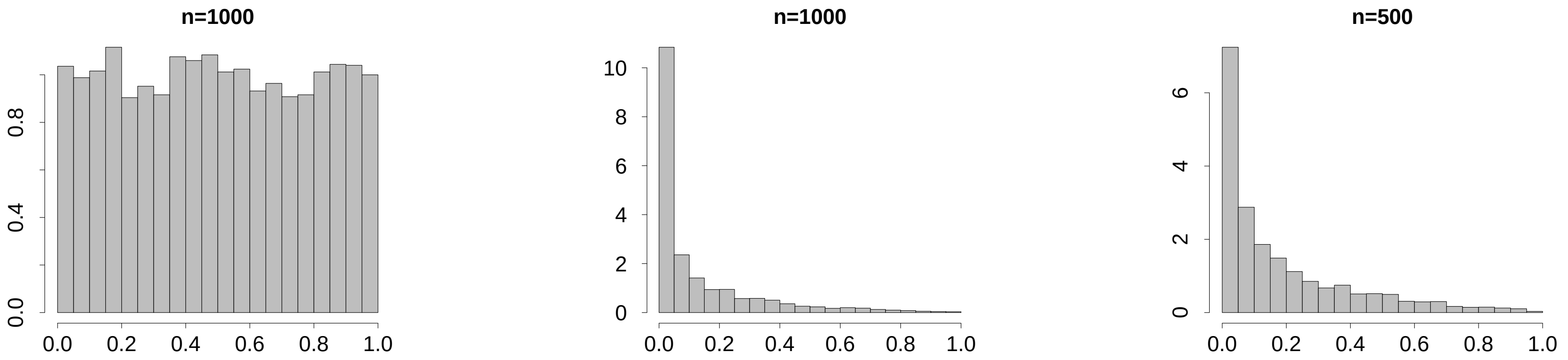
Likelihood-Ratio Test

$$\lambda_n = -2 \log \left(\frac{\sup_{\theta \in \Theta_0} \mathcal{L}_n(\theta)}{\sup_{\theta \in \Theta} \mathcal{L}_n(\theta)} \right).$$

Likelihood-Ratio Test

$$\lambda_n = -2 \log \left(\frac{\sup_{\theta \in \Theta_0} \mathcal{L}_n(\theta)}{\sup_{\theta \in \Theta} \mathcal{L}_n(\theta)} \right).$$

Simulated p-values (one-factor analysis model, Bartlett correction):



$k = 15$ observed variables, Σ **regular point**.

$k = 15$ observed variables, Σ close to a **singular point**.

$\mathbf{k = 200}$ observed variables, Σ regular point.

Wald Test

Tetrad: $f_1(\Sigma) = \sigma_{13}\sigma_{24} - \sigma_{23}\sigma_{14}$.

$$W_n = \frac{f_1(\hat{\Sigma})^2}{\widehat{\text{var}}[f_1(\hat{\Sigma})]} = \frac{n\, f_1(\hat{\Sigma})^2}{(\nabla f_1(\hat{\Sigma}))^\top V(\hat{\Sigma}) \nabla f_1(\hat{\Sigma})}, \qquad \text{where } \hat{\Sigma} = \frac{1}{n}\sum_{i=1}^{n} X_i X_i^\top.$$

Wald Test

Tetrad: $f_1(\Sigma) = \sigma_{13}\sigma_{24} - \sigma_{23}\sigma_{14}$.

$$W_n = \frac{f_1(\hat{\Sigma})^2}{\widehat{\text{var}}[f_1(\hat{\Sigma})]} = \frac{n\, f_1(\hat{\Sigma})^2}{(\nabla f_1(\hat{\Sigma}))^\top V(\hat{\Sigma}) \nabla f_1(\hat{\Sigma})}, \qquad \text{where } \hat{\Sigma} = \frac{1}{n}\sum_{i=1}^n X_i X_i^\top.$$

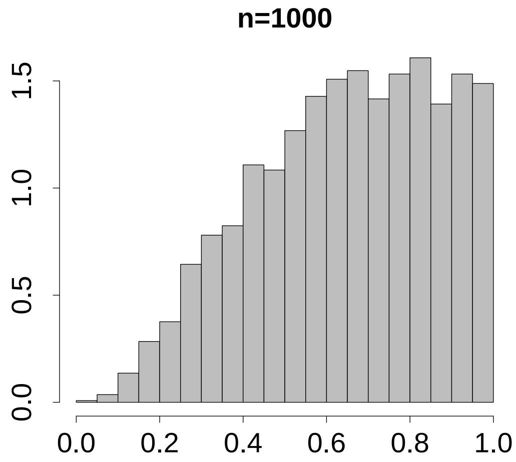
Limitations

- ✗ Invalid at singular points ($\nabla f_1(\Sigma) = 0$).

  $W_n \to_d F$ where $\frac{1}{4}\chi^2_1 \prec_{st} F \prec_{st} \chi^2_1$ (D. & Xiao, 2016)

- ✗ Only allows for low number of constraints ($p \leq d$).

- ✗ Difficult to handle inequality constraints.



Testing single tetrad,
Σ close to a **singular point**.

Connection to U-statistics

Tetrad: $f_1(\Sigma) = \sigma_{13}\sigma_{24} - \sigma_{23}\sigma_{14}$.

Observation:

$\widehat{f}_1 = \frac{n}{n-1} f_1(\hat{\Sigma}_n) = \frac{1}{\binom{n}{2}} \sum_{i<j} h_1(X_i, X_j)$ is a U-statistic with kernel

$$h_1(X_i, X_j) = \frac{1}{2}\{(X_{i1}X_{i3}X_{j2}X_{j4} - X_{i2}X_{i3}X_{j1}X_{j4}) + (X_{j1}X_{j3}X_{i2}X_{i4} - X_{j2}X_{j3}X_{i1}X_{i4})\}.$$

Connection to U-statistics

Tetrad: $f_1(\Sigma) = \sigma_{13}\sigma_{24} - \sigma_{23}\sigma_{14}$.

Observation:

$\widehat{f}_1 = \frac{n}{n-1} f_1(\hat{\Sigma}_n) = \frac{1}{\binom{n}{2}} \sum_{i<j} h_1(X_i, X_j)$ is a U-statistic with kernel

$$h_1(X_i, X_j) = \frac{1}{2}\{(X_{i1}X_{i3}X_{j2}X_{j4} - X_{i2}X_{i3}X_{j1}X_{j4}) + (X_{j1}X_{j3}X_{i2}X_{i4} - X_{j2}X_{j3}X_{i1}X_{i4})\}.$$

Asymptotics (one dimensional):

Gaussian approximation: $\sqrt{n}(\widehat{f}_1 - f_1(\Sigma)) \longrightarrow N(0, m^2\sigma_{g_1}^2)$

where m is the degree of the kernel h_1 and $\sigma_{g_1}^2$ is the variance of the Hájek projection

$$g_1(X_i) = \mathbb{E}[h_1(X_i, X_j)|X_i] = \frac{1}{2}\left\{(X_{i1}X_{i3}\sigma_{24} - X_{i2}X_{i3}\sigma_{14}) + (\sigma_{13}X_{i2}X_{i4} - \sigma_{23}X_{i1}X_{i4})\right\}.$$

Irregular points: $\sigma_{g_1}^2 = 0 \implies$ U-statistic is degenerate $\implies$ Gaussian approximations fails.

Estimable Constraints and U-statistics

Assumption: $f(\theta) = (f_1(\theta), \ldots, f_p(\theta))^\top$ is *estimable*.

That is, for some integer m there exists a measurable, symmetric function $h : \mathbb{R}^m \to \mathbb{R}^p$ such that

$$\mathbb{E}[h(X_1, \ldots, X_m)] = f(\theta) \quad \text{for all } \theta \in \Theta,$$

whenever $X_1, \ldots, X_m$ are i.i.d. with distribution P_θ.

Estimable Constraints and U-statistics

Assumption: $f(\theta) = (f_1(\theta), \ldots, f_p(\theta))^\top$ is *estimable*.

That is, for some integer m there exists a measurable, symmetric function $h : \mathbb{R}^m \to \mathbb{R}^p$ such that

$$\mathbb{E}[h(X_1, \ldots, X_m)] = f(\theta) \quad \text{for all } \theta \in \Theta,$$

whenever $X_1, \ldots, X_m$ are i.i.d. with distribution P_θ.

U-statistics: $U_n = \frac{1}{\binom{n}{m}} \sum_{(i_1,\ldots,i_m) \in I_{n,m}} h(X_{i_1}, \ldots, X_{i_m})$, where $I_{n,m} = \{(i_1, \ldots, i_m) : 1 \le i_1 < \ldots < i_m \le n\}$.

$\longrightarrow$ Reject for "large" values of $\max_{1 \le j \le p}(\sqrt{n}\, \hat{\sigma}_j^{-1}) U_{n,j}$.

Estimable Constraints and U-statistics

Assumption: $f(\theta) = (f_1(\theta), \ldots, f_p(\theta))^\top$ is *estimable*.

That is, for some integer m there exists a measurable, symmetric function $h : \mathbb{R}^m \to \mathbb{R}^p$ such that

$$\mathbb{E}[h(X_1, \ldots, X_m)] = f(\theta) \quad \text{for all } \theta \in \Theta,$$

whenever $X_1, \ldots, X_m$ are i.i.d. with distribution P_θ.

U-statistics: $U_n = \frac{1}{\binom{n}{m}} \sum_{(i_1,\ldots,i_m) \in I_{n,m}} h(X_{i_1}, \ldots, X_{i_m})$, where $I_{n,m} = \{(i_1, \ldots, i_m) : 1 \leq i_1 < \ldots < i_m \leq n\}$.

$\longrightarrow$ Reject for "large" values of $\max_{1 \leq j \leq p}(\sqrt{n}\, \hat{\sigma}_j^{-1}) U_{n,j}$.

Asymptotics: $\sqrt{n}\,(U_n - f(\theta)) \longrightarrow N_p(0, \Gamma_g)$, where $\Gamma_g = \mathrm{Cov}[g(X_1)]$ and g Hájek projection.

U-statistic is degenerate at irregular points $\Longrightarrow$ Gaussian approximation fails.

Independent Sums

Observation: $h(X_{(i-1)m+1}, \ldots, X_{im})$ are independent.

$$H_n = \frac{m}{n} \sum_{i=1}^{m} h(X_{(i-1)m+1}, \ldots, X_{im}).$$

Test statistic:

$$\max_{1 \le j \le p} (\sqrt{n}\ \hat{\sigma}_j^{-1}) H_{n,j}.$$

Independent Sums

Observation: $h(X_{(i-1)m+1}, \ldots, X_{im})$ are independent.

$$H_n = \frac{m}{n} \sum_{i=1}^{m} h(X_{(i-1)m+1}, \ldots, X_{im}).$$

Test statistic:

$$\max_{1 \le j \le p} (\sqrt{n}\, \hat{\sigma}_j^{-1}) H_{n,j}.$$

Asymptotics: $\sqrt{n/m}\,(H_n - f(\theta)) \longrightarrow N(0, \Gamma_h)$, where $\Gamma_h = \mathrm{Cov}[h(X_1, \ldots, X_m)]$.

- ✔ High-dimensional approximation of test statistic ($p \gg n$). (Chernozhukov et al., 2013)
- ✔ Non-degenerate limit at every parameter.
- ✗ inefficient . . . sum is only over $\frac{n}{m}$ elements.

Independent sums guard against degeneracy, but can we do better/use more kernel evaluations?

Proposal: Randomized Incomplete U-statistics

$$U'_{n,N} = \frac{1}{\hat{N}} \sum_{\iota=(i_1,\ldots,i_m)\in I_{n,m}} Z_\iota h(X_{i_1},\ldots,X_{i_m})$$

- $I_{n,m} = \{(i_1,\ldots,i_m) : 1 \le i_1 < \ldots < i_m \le n\}$.
- Computational budget parameter $N \le \binom{n}{m}$.
- $\{Z_\iota : \iota \in I_{n,m}\}$ are i.i.d. $\mathrm{Ber}(p_n)$ with $p_n = N/\binom{n}{m}$.
- $\hat{N} = \sum_{\iota\in I_{n,m}} Z_\iota$ is the number of successes.

Proposal: Randomized Incomplete U-statistics

$$U'_{n,N} = \frac{1}{\hat{N}} \sum_{\iota=(i_1,\dots,i_m)\in I_{n,m}} Z_\iota h(X_{i_1},\dots,X_{i_m})$$

- $I_{n,m} = \{(i_1,\dots,i_m) : 1 \le i_1 < \dots < i_m \le n\}$.
- Computational budget parameter $N \le \binom{n}{m}$.
- $\{Z_\iota : \iota \in I_{n,m}\}$ are i.i.d. $\mathrm{Ber}(p_n)$ with $p_n = N/\binom{n}{m}$.
- $\hat{N} = \sum_{\iota\in I_{n,m}} Z_\iota$ is the number of successes.

Asymptotics: $\sqrt{n}(U'_{n,N} - f(\theta)) \approx N(0, m^2\Gamma_g + \frac{n}{N}\Gamma_h)$.

Choose $N = \mathcal{O}(n)$ to guard against degeneracy!

Proposed Test

Test statistic

$$\mathcal{T} = \max_{1\leq j\leq p}(\sqrt{n}\,\widehat{\sigma}_j^{-1})U'_{n,N,j}.$$

Critical value

1. Approximate distribution of $\mathcal{T}$ by maximum of Gaussian random vector $Y \sim N_p(0,\Gamma)$, where $\Gamma = m^2\Gamma_g + \frac{n}{N}\Gamma_h$.
2. Construct an estimate $\widehat{\Gamma}$ of the true asymptotic covariance matrix Γ in a Gaussian multiplier bootstrap method. Then $W \sim N_p(0,\widehat{\Gamma})$ is "close" to $Y \sim N_p(0,\Gamma)$.
3. Critical value: Quantile $c_{W_0}(1-\alpha)$ of $W_0 = \max_{1\leq j\leq p}\widehat{\sigma}_j^{-1}W_j$.

Proposed Test

Test statistic

$$\mathcal{T} = \max_{1 \le j \le p} (\sqrt{n}\, \hat{\sigma}_j^{-1}) U'_{n,N,j}.$$

Critical value

1. Approximate distribution of $\mathcal{T}$ by maximum of Gaussian random vector $Y \sim N_p(0, \Gamma)$, where $\Gamma = m^2 \Gamma_g + \frac{n}{N} \Gamma_h$.
2. Construct an estimate $\hat{\Gamma}$ of the true asymptotic covariance matrix Γ in a Gaussian multiplier bootstrap method. Then $W \sim N_p(0, \hat{\Gamma})$ is "close" to $Y \sim N_p(0, \Gamma)$.
3. Critical value: Quantile $c_{W_0}(1-\alpha)$ of $W_0 = \max_{1 \le j \le p} \hat{\sigma}_j^{-1} W_j$.

Our analysis. . .

If $N = \mathcal{O}(n)$ then the proposed test based on an incomplete U-statistic is asymptotically valid (controls type I error) in high dimensions $p \gg n$ and under *mixed degeneracy*:

$$P(\mathcal{T} > c_{W_0}(1-\alpha)) \le \alpha.$$

Mixed Degeneracy

Background on high-dimensional Gaussian approximation

Chernozhukov, Chetverikov, Kato (2013). *Gaussian approximations and multiplier bootstrap for maxima of sums of high-dimensional random vectors.* Ann. Statist., 41(6):2786–2819.

Mixed Degeneracy

Background on high-dimensional Gaussian approximation

Chernozhukov, Chetverikov, Kato (2013). *Gaussian approximations and multiplier bootstrap for maxima of sums of high-dimensional random vectors.* Ann. Statist., 41(6):2786–2819.

Chen (2018). *Gaussian and bootstrap approximations for high-dimensional U-statistics and their applications.* Ann. Statist., 46(2):642–678.

Assumption: Non-degenerate: There exists $c > 0$ such that $\sigma_{g_j}^2 \geq c$ for all $j = 1, \ldots, p$.

Mixed Degeneracy

Background on high-dimensional Gaussian approximation

Chernozhukov, Chetverikov, Kato (2013). *Gaussian approximations and multiplier bootstrap for maxima of sums of high-dimensional random vectors.* Ann. Statist., 41(6):2786–2819.

Chen (2018). *Gaussian and bootstrap approximations for high-dimensional U-statistics and their applications.* Ann. Statist., 46(2):642–678.

Assumption: Non-degenerate: There exists $c > 0$ such that $\sigma_{g_j}^2 \geq c$ for all $j = 1, \ldots, p$.

Chen, Kato (2019). *Randomized incomplete U-statistics in high dimensions.* Ann. Statist., 47(6):3127–3156.

Assumption: **Either** non-degenerate: There exists $c > 0$ such that $\sigma_{g_j}^2 \geq c$ for all $j = 1, \ldots, p$.

Or degenerate: $\sigma_{g_j}^2 = 0$ for all $j = 1, \ldots, p$.

Mixed Degeneracy

Background on high-dimensional Gaussian approximation

Chernozhukov, Chetverikov, Kato (2013). *Gaussian approximations and multiplier bootstrap for maxima of sums of high-dimensional random vectors.* Ann. Statist., 41(6):2786–2819.

Chen (2018). *Gaussian and bootstrap approximations for high-dimensional U-statistics and their applications.* Ann. Statist., 46(2):642–678.

Assumption: Non-degenerate: There exists $c > 0$ such that $\sigma_{g_j}^2 \geq c$ for all $j = 1, \ldots, p$.

Chen, Kato (2019). *Randomized incomplete U-statistics in high dimensions.* Ann. Statist., 47(6):3127–3156.

Assumption: **Either** non-degenerate: There exists $c > 0$ such that $\sigma_{g_j}^2 \geq c$ for all $j = 1, \ldots, p$.

Or degenerate: $\sigma_{g_j}^2 = 0$ for all $j = 1, \ldots, p$.

Mixed degeneracy assumption

Let $p_1, p_2 \in \mathbb{N}$ such that $p_1 + p_2 = p$ and assume:

(A) There exists $c > 0$ such that $\sigma_{g_j}^2 \geq c$ for all $j = 1, \ldots, p_1$.

(B) There exists $k > 0$ and $\beta > 0$ such that $\|g_j(X_1) - f_j(\theta)\|_{\psi_\beta} \leq Cn^{-k}$ for all $j = p_1 + 1, \ldots, p$.

$\implies \sigma_{g_j}^2 \leq \tilde{C} n^{-2k}$

High-dimensional Gaussian Approximation

Theorem

Under mixed degeneracy (and additional moment conditions on h), we have the ***Gaussian approximation*** *on the hyperrectangles*

$$\sup_{R \in \mathbb{R}^p_{\mathrm{re}}} |P(\sqrt{n}(U'_{n,N} - f(\theta)) \in R) - P(Y \in R)| \leq C\{\omega_{n,1} + \omega_{n,2} + \omega_{n,3}\},$$

where $Y \sim N_p(0, m^2\Gamma_g + \frac{n}{N}\Gamma_h)$ *and*

$$\omega_{n,1} = \left(\frac{m^{2/\beta} \log(pn)^{1+6/\beta}}{n \wedge N}\right)^{1/6}, \qquad \omega_{n,2} = \frac{N^{1/2} m^2 \log(pn)^{1/2+2/\beta}}{n^{\min\{1/2+k,\, 5/6\}}}, \qquad \omega_{n,3} = \left(\frac{Nm^2 \log(p)^2}{n^{\min\{1+k, m\}}}\right)^{1/3}.$$

Note:

If $N = \mathcal{O}(n)$ and $k \geq 1/3$ is fixed, then the bound vanishes asymptotically if $\log(pn)^{3/2+6/\beta} = o(n)$.

High-dimensional Bootstrap Approximation

Recall: $Y = m\, Y_g + \sqrt{n/N}\, Y_h,$ where $Y_g \sim N_p(0, \Gamma_g)$ and $Y_h \sim N_p(0, \Gamma_h)$ are independent.

High-dimensional Bootstrap Approximation

Recall: $Y = m\,Y_g + \sqrt{n/N}\,Y_h, \quad$ where $Y_g \sim N_p(0, \Gamma_g)$ and $Y_h \sim N_p(0, \Gamma_h)$ are independent.

Approach: Construct W_g, W_h such that, given the data, both are independent and approximate Y_g, Y_h.

High-dimensional Bootstrap Approximation

Recall: $Y = m\,Y_g + \sqrt{n/N}\,Y_h$, where $Y_g \sim N_p(0, \Gamma_g)$ and $Y_h \sim N_p(0, \Gamma_h)$ are independent.

Approach: Construct W_g, W_h such that, given the data, both are independent and approximate Y_g, Y_h.

Gaussian Multiplier Bootstrap:

$$W_h = \frac{1}{\sqrt{\hat{N}}} \sum_{\iota=(i_1,\ldots,i_m)\in I_{n,m}} \xi_\iota\, Z_\iota\, (h(X_{i_1},\ldots,X_{i_m}) - U'_{n,N}),$$

where $\{\xi_\iota : \iota \in I_{n,m}\}$ are a collection of independent $N(0,1)$ r.v.'s.

$\Longrightarrow$ Given the data, we have $W_h \approx Y_h$.

High-dimensional Bootstrap Approximation

Recall: $Y = m\, Y_g + \sqrt{n/N}\, Y_h, \quad$ where $Y_g \sim N_p(0, \Gamma_g)$ and $Y_h \sim N_p(0, \Gamma_h)$ are independent.

Approach: Construct W_g, W_h such that, given the data, both are independent and approximate Y_g, Y_h.

Gaussian Multiplier Bootstrap:

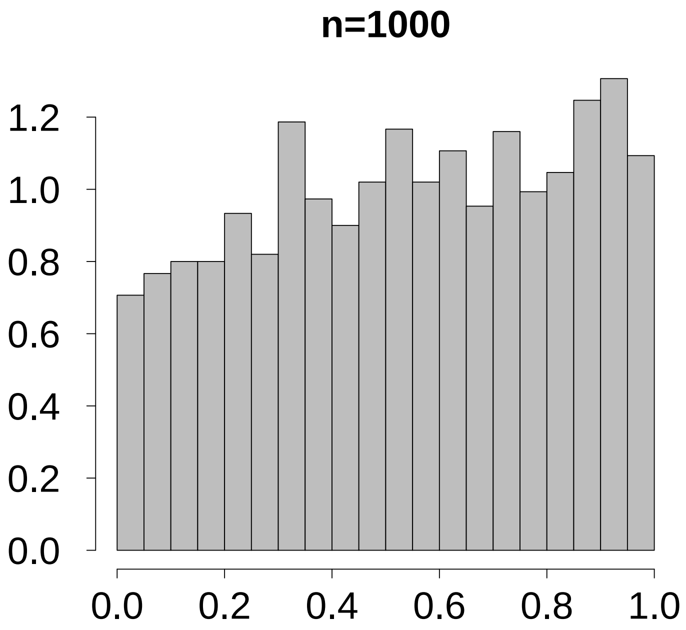
$$W_h = \frac{1}{\sqrt{\hat{N}}} \sum_{\iota=(i_1,\ldots,i_m)\in I_{n,m}} \xi_\iota\, Z_\iota \left(h(X_{i_1}, \ldots, X_{i_m}) - U'_{n,N}\right),$$

where $\{\xi_\iota : \iota \in I_{n,m}\}$ are a collection of independent $N(0,1)$ r.v.'s.

$\Longrightarrow$ Given the data, we have $W_h \approx Y_h$.

1. Similarly, we construct W_g, such that, given the data, $W_g \approx Y_g$.
2. Finite sample Berry Esseen type bound for the approximation $Y \approx W = m\, W_g + \sqrt{n/N}\, W_h$.
3. Control studentization.
4. Establish asymptotic validity (control of type I error).

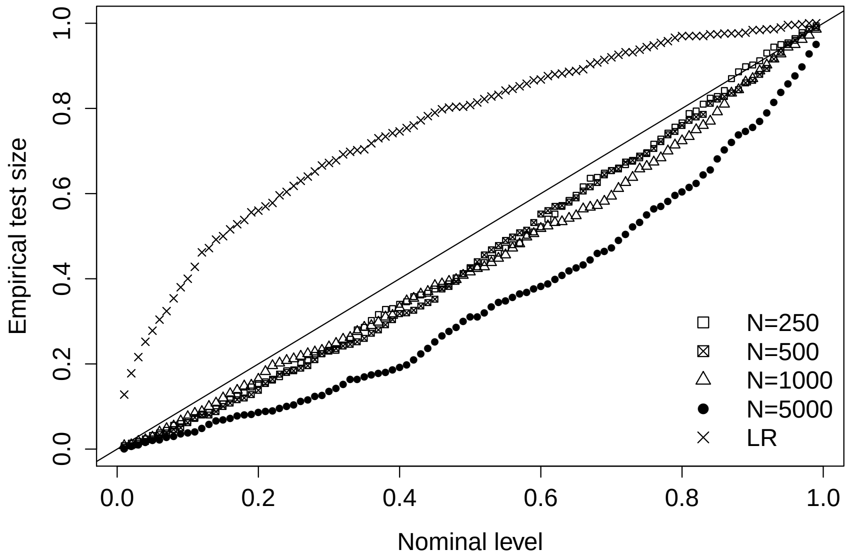
Our Test at Irregular Points



Simulated p-values for testing tetrads with $k = 15$ observed variables close to a singular point.
Computational budget parameter $N = 2n$.

Size vs. Power

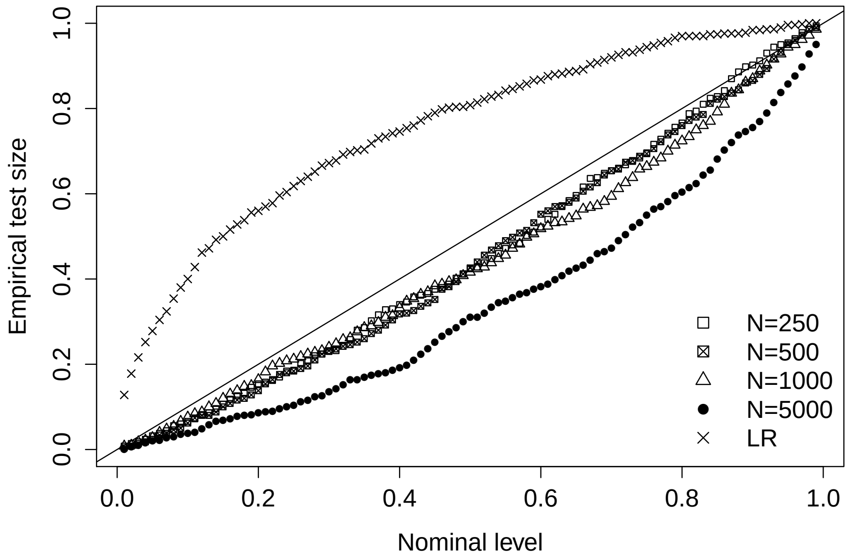
$n = 500$



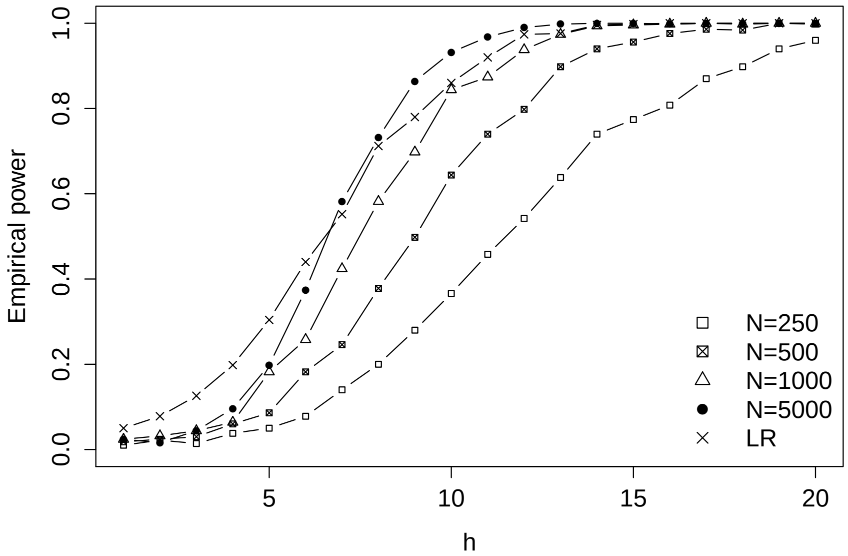
Empirical sizes vs. nominal levels for testing tetrads with $k = 15$ observed variables. True parameter is close to a **singular point**.

Size vs. Power

$n = 500$



Empirical sizes vs. nominal levels for testing tetrads with $k = 15$ observed variables. True parameter is close to a **singular point**.



Empirical power for different local alternatives for testing tetrads with $k = 15$ observed variables ($\alpha = 0.05$). True parameter is a **regular point**.

Trade-off between efficiency and guarding against singularities.

Conclusion

- ✔ General strategy for simultaneous testing of many constraints ($p \gg n$).
- ✔ Equality and inequality constraints.
- ✔ Optimization free.
  Although computationally demanding for large p and large computational budget N.
- ✔ Accommodate irregular settings where the incomplete U-statistics is mixed degenerate via $N = \mathcal{O}(n)$.

Our paper and background reading:

Sturma, Drton, Leung (2022).
Testing Many and Possibly Singular Polynomial Constraints. arXiv:2208.11756.

Leung, Drton (2018).
Algebraic tests of general Gaussian latent tree models. NeurIPS 2018.

Drton (2009).
Likelihood ratio tests and singularities. Ann. Statist., 37(2):979–1012

European Research Council
Established by the European Commission