Semiparametric Gaussian copula models: geometry and efficient rank-based estimation

Johan Segers[1], Ramon van den Akker[2], Bas J.M. Werker[2]

[1] Université catholique de Louvain

Institut de statistique, biostatistique et sciences actuarielles

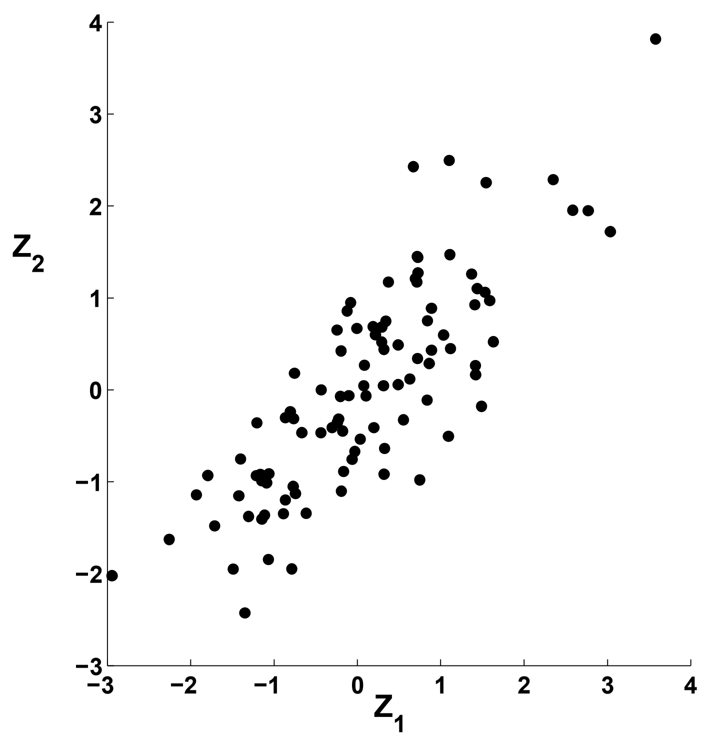
[2] Tilburg University

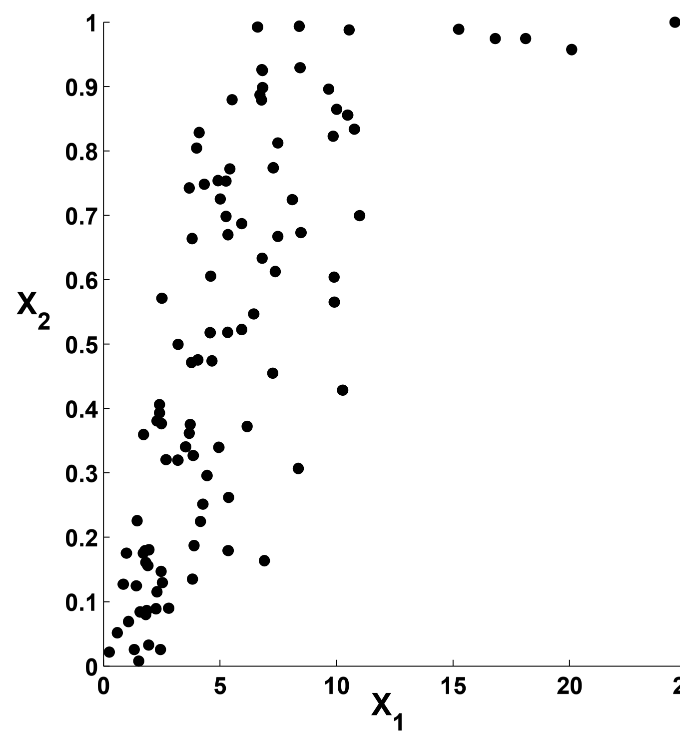
CentER

January 6, 2014

How to recover correlation matrix of latent Gaussian variables?

$$\mathbf{Z} = \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} \sim N_2\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \theta \\ & 1 \end{pmatrix} \right) \qquad \mathbf{X} = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} \eta_1(Z_1) \\ \eta_2(Z_2) \end{pmatrix}$$





Model (representation as transformation model)

Introduction
❖ model
❖ literature
❖ why rank-based?
❖ Outline
Main Results
Monte Carlo results
Remarks on proofs
Conclusion

Observe i.i.d. copies $\mathbf{X}_1, \ldots, \mathbf{X}_n$ of p-variate random vector $\mathbf{X}$, where

$$X_j = \eta_j(Z_j), \quad j = 1, \ldots, p$$
$$\mathbf{Z} = (Z_1, \ldots, Z_p)' \sim N_p(\mathbf{0}, R(\theta))$$

- $R(\theta)$ $p \times p$ correlation matrix indexed by $\theta \in \mathbb{R}^k$
- $\eta_j : \mathbb{R} \to \mathbb{R}$ unknown strictly increasing function

Main goal:
efficient estimation of θ in presence of infinite-dimensional nuisance parameters $\eta_1, \ldots \eta_p$

Introduction
❖ model
❖ literature
❖ why rank-based?
❖ Outline
Main Results
Monte Carlo results
Remarks on proofs
Conclusion

Model: structured correlation matrices

Some examples for $R(\theta)$:

- **unrestricted model**; for $p = 3$:

$$R(\theta) = \begin{pmatrix} 1 & \theta_1 & \theta_2 \\ & 1 & \theta_3 \\ & & 1 \end{pmatrix}, \qquad k = \frac{1}{2}p(p-1)$$

- **Toeplitz model**; for $p = 4$:

$$R(\theta) = \begin{pmatrix} 1 & \theta_1 & \theta_2 & \theta_3 \\ & 1 & \theta_1 & \theta_2 \\ & & 1 & \theta_1 \\ & & & 1 \end{pmatrix}, \qquad k = p-1$$

- exchangeable models, circular models, factor models, etc.

Introduction
❖ model
❖ literature
❖ why rank-based?
❖ Outline
Main Results
Monte Carlo results
Remarks on proofs
Conclusion

Intermezzo: copulas and Sklar's theorem

- **Copula (p-dimensional):**
  cdf of random vector $\mathbf{U} = (U_1, \ldots, U_p)'$ with uniform margins ($U_j \sim$ Un$[0, 1]$):

$$C(\mathbf{u}) = \mathbb{P}(U_1 \leq u_1, \ldots, U_p \leq u_p)$$

- **Sklar's theorem:**
  - if C is copula and $F_1, \ldots, F_p$ univariate cdfs, then

    $$F(\mathbf{x}) = C(F_1(x_1), \ldots, F_p(x_p)), \quad \mathbf{x} \in \mathbb{R}^p, \qquad (*)$$

    defines cdf F
  - every p-variate cdf F can be written as $(*)$ for copula C (unique if margins are continuous)
  - separation of dependence structure from margins

Introduction
❖ model
❖ literature
❖ why rank-based?
❖ Outline
Main Results
Monte Carlo results
Remarks on proofs
Conclusion

Intermezzo: Gaussian copula

- start with

$$\mathbf{Z} = (Z_1, \ldots, Z_p)' \sim N_p(\mathbf{0}, R(\theta))$$

- define $\mathbf{U} = (U_1, \ldots, U_p)'$ by

$$U_j = \Phi(Z_j), \quad j = 1, \ldots, p$$

- **Gaussian $R(\theta)$-copula** is cdf of $\mathbf{U}$:

$$\begin{aligned} C_\theta(\mathbf{u}) &= \mathbb{P}(U_1 \leq u_1, \ldots, U_p \leq u_p) \\ &= \Phi_\theta(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_p)), \end{aligned}$$

where Φ_θ denotes cdf of $N_p(\mathbf{0}, R(\theta))$

Introduction
❖ model
❖ literature
❖ why rank-based?
❖ Outline
Main Results
Monte Carlo results
Remarks on proofs
Conclusion

Model (representation as copula model)

- recall:

$$X_j = \eta_j(Z_j), \quad j = 1, \ldots, p$$
$$\mathbf{Z} = (Z_1, \ldots, Z_p)' \sim N_p(\mathbf{0}, R(\theta))$$

- margins:

$$F_j(x_j) = \mathbb{P}(X_j \le x_j) = \mathbb{P}(\eta_j(Z_j) \le x_j) = \Phi(\eta_j^{-1}(x_j))$$

- joint distribution:

$$\begin{aligned} F(\mathbf{x}) &= \mathbb{P}(X_1 \le x_1, \ldots, X_p \le x_p) \\ &= \mathbb{P}(\eta_1(Z_1) \le x_1, \ldots, \eta_p(Z_p) \le x_p) \\ &= \mathbb{P}(\Phi(Z_1) \le F_1(x_1), \ldots, \Phi(Z_p) \le F_p(x_p)) \\ &= C_\theta(F_1(x_1), \ldots, F_p(x_p)) \end{aligned}$$

Model

Introduction
❖ model
❖ literature
❖ why rank-based?
❖ Outline
Main Results
Monte Carlo results
Remarks on proofs
Conclusion

Consider i.i.d. copies $\mathbf{X}_1, \ldots, \mathbf{X}_1$ of $\mathbf{X}$, where

- copula of $\mathbf{X}$ is Gaussian $R(\theta)$-copula
- denote margins by $F_1, \ldots, F_p$

Assumptions:

- $\theta \in \Theta \subset \mathbb{R}^k$
  - mild, straightforward conditions on $\theta \mapsto R(\theta)$
- margins $F_1, \ldots, F_p$ are absolutely continuous

Main goal:
efficient (rank-based) estimation of θ

Introduction
❖ model
❖ literature
❖ why rank-based?
❖ Outline

Main Results

Monte Carlo results

Remarks on proofs

Conclusion

Relation to literature

If margins $F_1, \ldots, F_p$ **would have been** known

$$\hat{\theta}_n^{MLE} = \underset{\theta \in \Theta}{\operatorname{argmax}} \sum_{i=1}^{n} \Big\{ \log c_\theta(F_1(X_{i,1}), \ldots, F_p(X_{i,p})) + \sum_{j=1}^{p} \log f_j(X_{i,j}) \Big\}$$

- lot of semiparametric techniques to estimate θ; we are interested in efficient estimation

Relation to literature

Introduction
❖ model
❖ literature
❖ why rank-based?
❖ Outline
Main Results
Monte Carlo results
Remarks on proofs
Conclusion

Pseudo Likelihood Estimator (PLE):

$$\hat{\theta}_n^{PLE} = \operatorname*{argmax}_{\theta \in \Theta} \sum_{i=1}^{n} \log c_\theta(\hat{F}_{n1}(X_{i,1}), \ldots, \hat{F}_{np}(X_{i,p}))$$

where $\hat{F}_{nj} = n\hat{\mathbb{F}}_{nj}/(n+1)$

- introduced by OAKES (1994) and GENEST, GHOUDI & RIVEST (1995)
- efficient for $p = 2$ and $R_{12}(\theta) = \theta$
  - KLAASSEN & WELLNER (1997) and GENEST & WERKER (2002)
- not efficient in general
  - example ($p = 4$) in HOFF, NIU & WELLNER (2012)

Introduction
- model
- literature
- why rank-based?
- Outline

Main Results

Monte Carlo results

Remarks on proofs

Conclusion

Relation to literature

CHEN, FAN & TSYRENNIKOV (2006): sieve-based approach

- rough idea is to replace unknown $\sqrt{f_j}(x_j)$ by $\sum_k a_k^j b_k(x_j)$, substitute this in parametric log-likelihood

$$\sum_{i=1}^{n} \left\{ \log c_\theta(F_1(X_{i,1}), \ldots, F_p(X_{i,p})) + \sum_{j=1}^{p} \log f_j(X_{i,j}) \right\}$$

  and maximize over θ and sieve coefficients
- efficient [SHEN (1997)]
  - under regularity conditions
  - also applicable for non-Gaussian models
- but: estimator is not rank-based

Invariance motivates rank-based inference

Introduction
- model
- literature
- why rank-based?
- Outline

Main Results

Monte Carlo results

Remarks on proofs

Conclusion

- applying strictly increasing transformations $T_1, \ldots, T_p$ yields

$$\begin{aligned}\mathbf{T}(\mathbf{X}) &= (T_1(X_1), \ldots, T_p(X_p)) \\ &= ((T_1 \circ \eta_1)(Z_1), \ldots, (T_p \circ \eta_p)(Z_p)), \quad \mathbf{Z} \sim N_p(\mathbf{0}, R(\theta))\end{aligned}$$

$\Longrightarrow$ parameter of interest θ remains same

- invariance principle leads to requirement:

$$\hat{\theta}_n(\mathbf{T}(\mathbf{X}_1), \ldots, \mathbf{T}(\mathbf{X}_n)) = \hat{\theta}_n(\mathbf{X}_1, \ldots, \mathbf{X}_n), \quad \text{for all } \mathbf{T}$$

  - $\Longrightarrow$ $\hat{\theta}_n$ depends on data only via vectors of component-wise ranks
- Hoff, Niu & Wellner (2012):
  - restricting to rank-based estimators should not yield loss of information

Outline

- main results
  - existence efficient rank-based estimator
  - characterization efficiency PLE
- Monte Carlo study
- some remarks on proofs
- conclusion

Main Result 1: Efficient rank-based estimator

Introduction

Main Results

❖ result 1

❖ result 2

Monte Carlo results

Remarks on proofs

Conclusion

We propose estimator $\hat{\theta}_n^{OSE}$ that is:

- rank-based
- efficient:

$$\sqrt{n}(\hat{\theta}_n^{OSE} - \theta) \xrightarrow{d} N_k(0, I^*(\theta)^{-1})$$

- if $\sqrt{n}(T_n - \theta) \xrightarrow{d} L$ and T_n is regular, we have

$$\operatorname{var} L \geq I^*(\theta)^{-1}$$

  - more precise formulation via convolution theorem
  - see Bickel, Klaassen, Ritov & Wellner (1993) and van der Vaart (2000) for semiparametric efficiency theory

Main Result 1: Efficient rank-based estimator

Introduction

Main Results

❖ result 1

❖ result 2

Monte Carlo results

Remarks on proofs

Conclusion

1. compute $\hat{F}_{nj}(X_{ij}) = R_{ij}/(n+1)$, $i = 1, \ldots, n$ and $j = 1, \ldots, p$
2. compute initial $\sqrt{n}$-consistent rank-based estimate $\tilde{\theta}_n$
   - e.g. PLE
   - for efficiency-proof: need to discretize to grid $n^{-1/2}\mathbb{Z}^k$
3. compute

$$\hat{\theta}_n^{OSE} = \tilde{\theta}_n + \frac{1}{n} \sum_{i=1}^{n} I^*(\tilde{\theta}_n)^{-1} \dot{\ell}_\theta^*(\hat{F}_{n,1}(X_{i1}), \ldots, \hat{F}_{n,p}(X_{ip}); \tilde{\theta}_n),$$

where

- $\dot{\ell}_\theta^*(\mathbf{u}; \theta)$ is efficient score for θ at uniform marginals
- $I^*(\theta)$ is efficient information matrix for estimation of θ

Introduction
Main Results
❖ result 1
❖ result 2
Monte Carlo results
Remarks on proofs
Conclusion

Main Result 1: Efficient rank-based estimator

$$\hat{\theta}_n^{OSE} = \tilde{\theta}_n + \frac{1}{n}\sum_{i=1}^{n} I^*(\tilde{\theta}_n)^{-1}\dot{\ell}_\theta^*(\hat{F}_{n,1}(X_{i1}),\ldots,\hat{F}_{n,p}(X_{ip});\tilde{\theta}_n)$$

where

$$\dot{\ell}_{\theta,m}^*(\mathbf{u};\theta) = \frac{1}{2}\mathbf{z}'A_m(\theta)\mathbf{z}, \quad \mathbf{z_j} = \Phi^{-1}(u_j)$$

with

$$A_m(\theta) = D_\theta(\mathbf{g}_m(\theta)) - \frac{\partial}{\partial\theta_m}R^{-1}(\theta)$$

$$D_\theta(\mathbf{b}) = R^{-1}(\theta)\,\mathrm{diag}(\mathbf{b}) + \mathrm{diag}(\mathbf{b})R^{-1}(\theta)$$

$$\mathbf{g}_m(\theta) = -(I_p + R(\theta)\circ R^{-1}(\theta))^{-1}\left[\frac{\partial}{\partial\theta_m}R(\theta)\circ R^{-1}(\theta)\right]\iota_p$$

and

$$I_{mm'}^*(\theta) = \frac{1}{2}\,\mathrm{tr}\left[R(\theta)A_m(\theta)R(\theta)A_{m'}(\theta)\right]$$

$\Longrightarrow$ **OSE is easy to implement**

Introduction
Main Results
❖ result 1
❖ result 2
Monte Carlo results
Remarks on proofs
Conclusion

Main result 2: geometry

Exploit geometry (shapes of scores + tangent space) to

- provide easy conditions to check for efficiency given estimator
  - application to PLE:
    - efficient for unrestricted model [$k = p(p-1)/2$]
      - stated (without proof) in KLAASSEN & WELLNER (1997)
    - efficient for exchangeable models [$p \geq 3$]
      - for $p = 4$ in HOFF, NIU & WELLNER (2012)
    - PLE efficient for factor models
- easy conditions adaptivity, i.e. $I^*(\theta) = I(\theta)$

Introduction
Main Results
Monte Carlo results
❖ exchangeable
❖ circular
❖ high-dimensional
❖ Toeplitz
Remarks on proofs
Conclusion

Exchangeable model in $p = 3$ (PLE efficient)

- correlation structure:

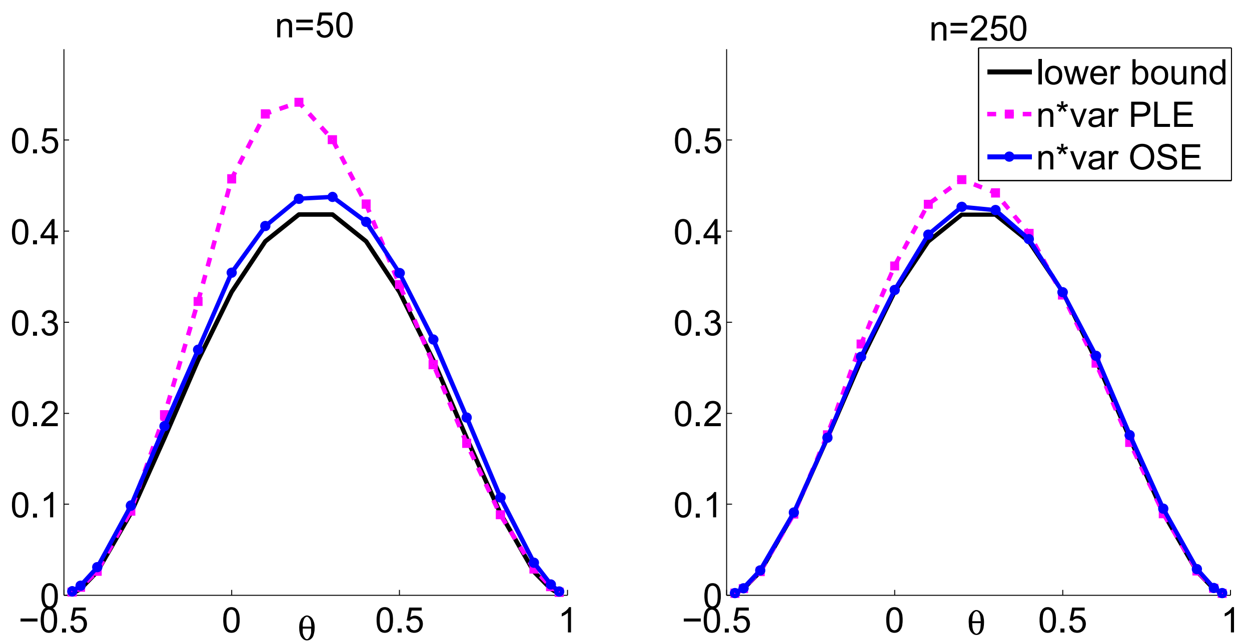
$$R(\theta) = \begin{pmatrix} 1 & \theta & \theta \\ & 1 & \theta \\ & & 1 \end{pmatrix}$$

- PLE efficient
- evaluate finite-sample performance PLE and OSE

Introduction
Main Results
Monte Carlo results
❖ exchangeable
❖ circular
❖ high-dimensional
❖ Toeplitz
Remarks on proofs
Conclusion

Exchangeable model in $p = 3$ (PLE efficient)

Simulated approximations* to $n\,\mathrm{var}(\hat\theta_n^{PLE})$ and $n\,\mathrm{var}(\hat\theta_n^{OSE})$:

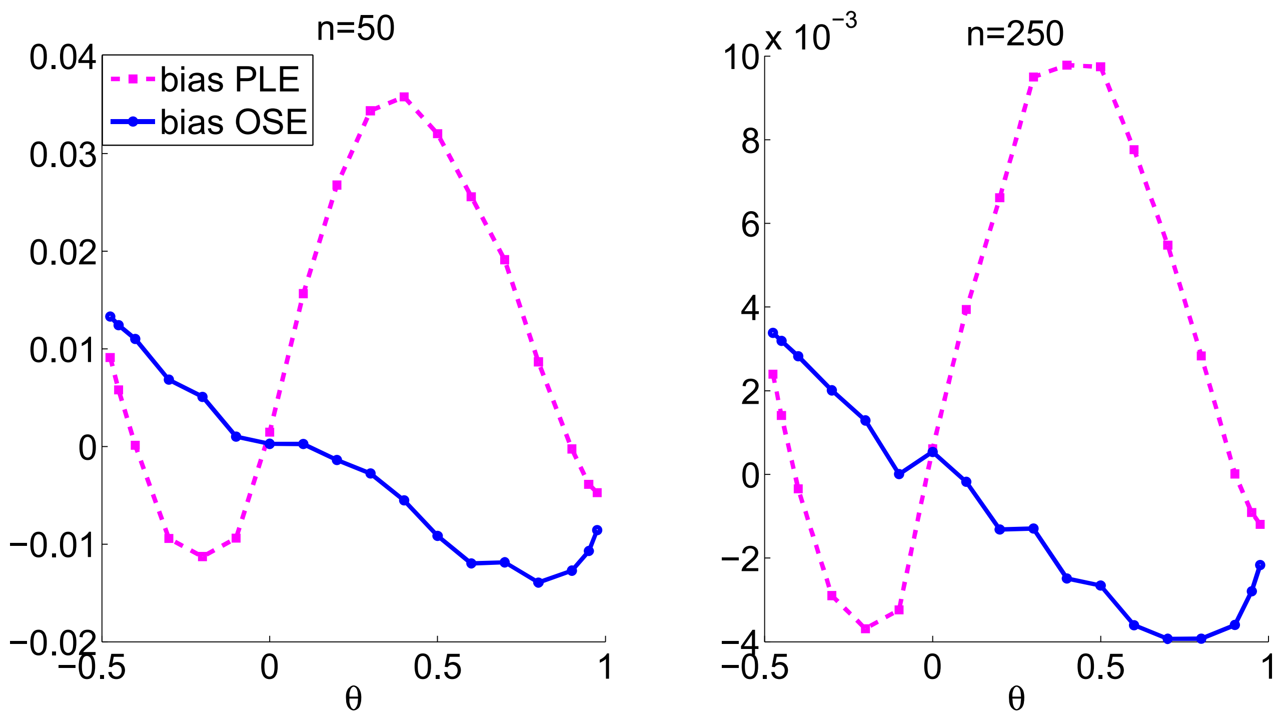


*: based on 15,000 replications

Introduction
Main Results
Monte Carlo results
❖ exchangeable
❖ circular
❖ high-dimensional
❖ Toeplitz
Remarks on proofs
Conclusion

Exchangeable model in $p = 3$ (PLE efficient)

Simulated approximations to $\mathbb{E}_\theta[\hat{\theta}_n^{PLE}] - \theta$ and $\mathbb{E}_\theta[\hat{\theta}_n^{OSE}] - \theta$:



$\Longrightarrow$ even if PLE is also efficient using OSE can be good idea

Circular model in $p = 4$ (PLE nearly efficient)

- correlation structure:

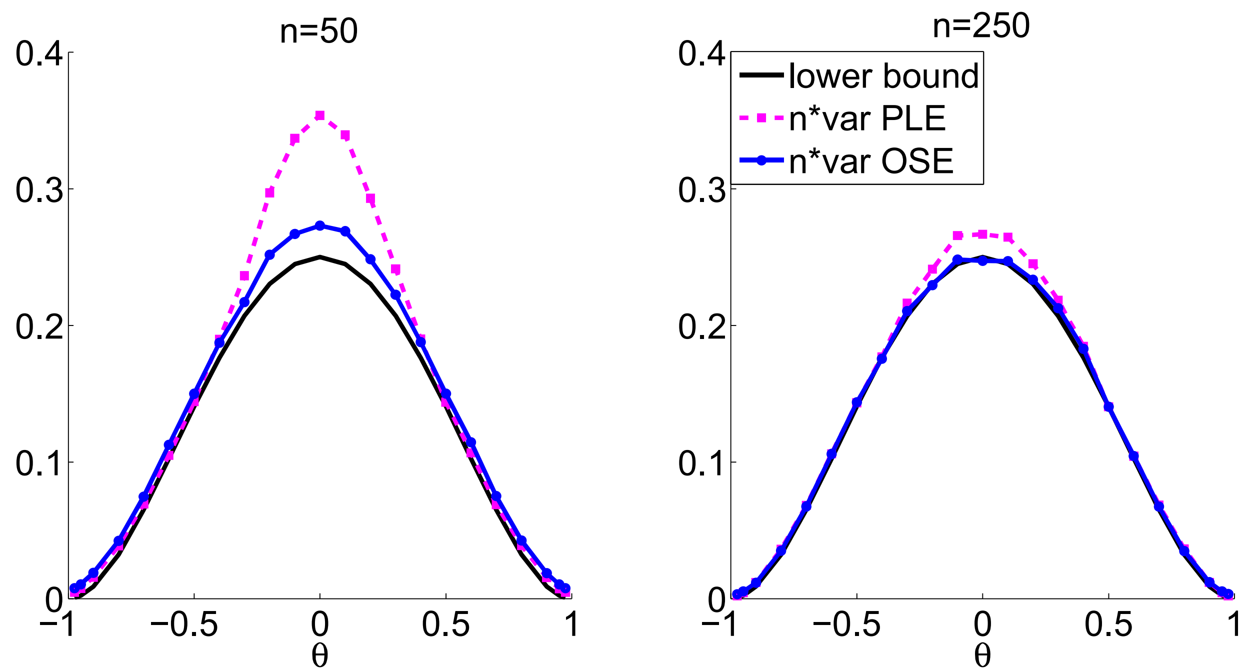
$$R(\theta) = \begin{pmatrix} 1 & \theta & \theta^2 & \theta \\ & 1 & \theta & \theta^2 \\ & & 1 & \theta \\ & & & 1 \end{pmatrix}$$

- PLE not efficient
- but ARE is close to 1 (for $|\theta| << 1$)
  - “nearly efficient”
- evaluate finite-sample performance PLE and OSE

Introduction
Main Results
Monte Carlo results
❖ exchangeable
❖ circular
❖ high-dimensional
❖ Toeplitz
Remarks on proofs
Conclusion

Circular model in $p = 4$ (PLE nearly efficient)

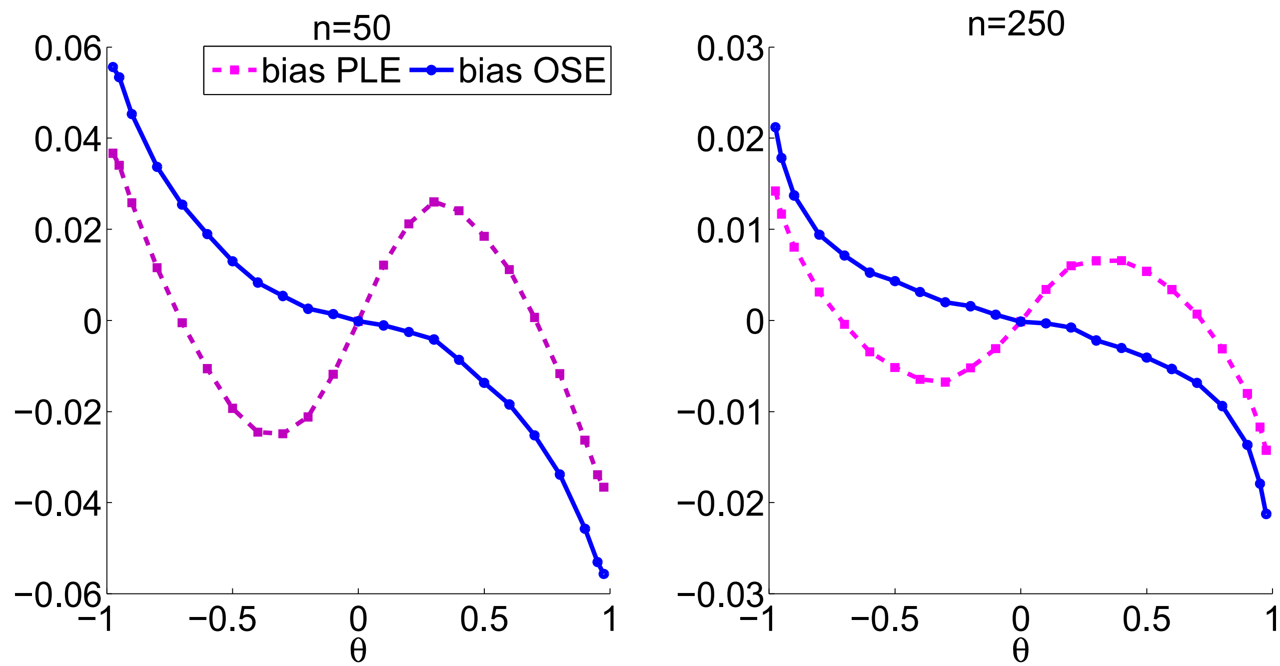
Simulated approximations to $n\,\mathrm{var}(\hat{\theta}_n^{PLE})$ and $n\,\mathrm{var}(\hat{\theta}_n^{OSE})$:



Introduction
Main Results
Monte Carlo results
❖ exchangeable
❖ circular
❖ high-dimensional
❖ Toeplitz
Remarks on proofs
Conclusion

Examples for which PLE is (nearly) efficient

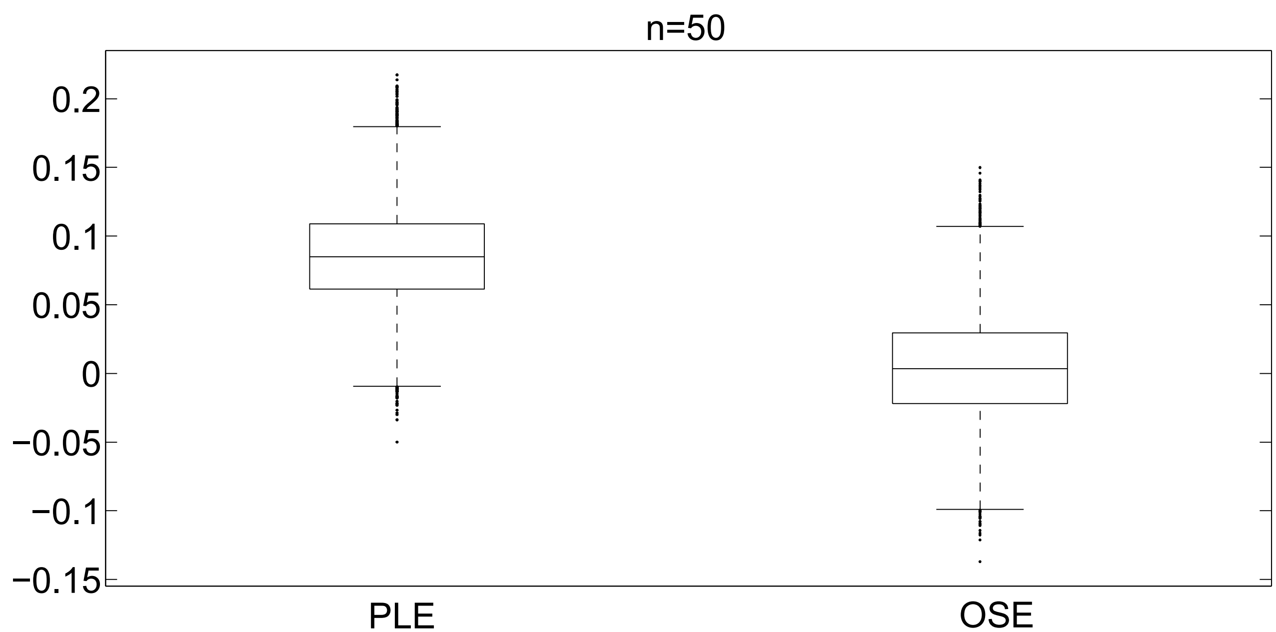
Simulated approximations to $\mathbb{E}_\theta[\hat{\theta}_n^{PLE}] - \theta$ and $\mathbb{E}_\theta[\hat{\theta}_n^{OSE}] - \theta$:



High-dimensional example: exchangeable in $p = 100$ (PLE efficient)

Introduction
Main Results
Monte Carlo results
❖ exchangeable
❖ circular
❖ high-dimensional
❖ Toeplitz
Remarks on proofs
Conclusion

- exchangeable model for $p = 100$
- PLE is efficient
- boxplot of $\hat{\theta}_n^{PLE} - \theta$ and $\hat{\theta}_n^{OSE} - \theta$ for $n = 50$ and at $\theta = .25$: (15,000 replications)



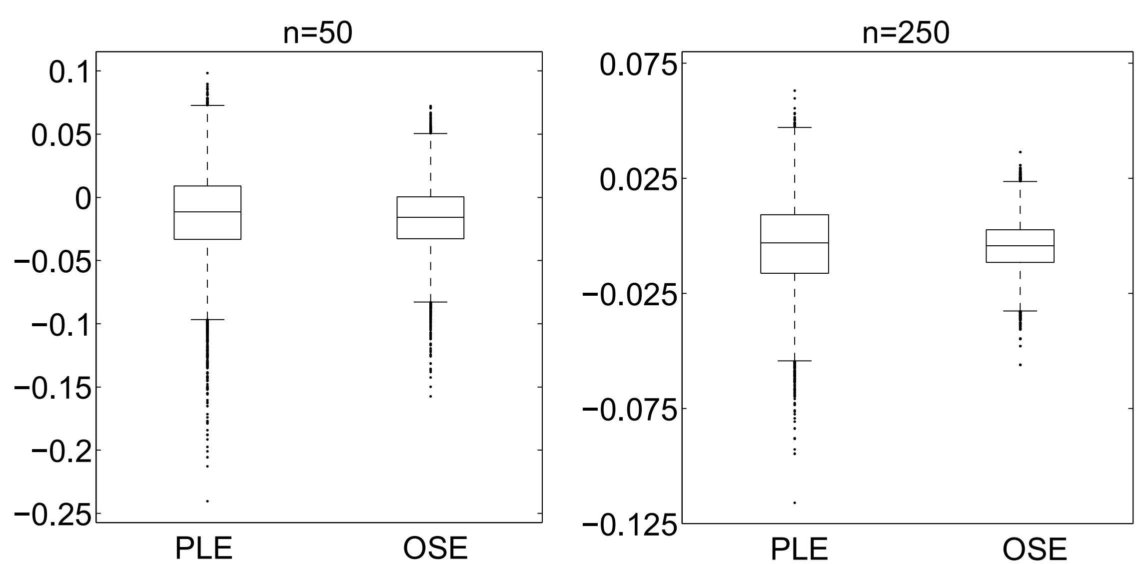
$\Longrightarrow$ variances close, but OSE less biased

Toeplitz example: PLE quite inefficient

- Toeplitz model in $p = 4$, i.e.

$$R(\theta) = \begin{pmatrix} 1 & \theta_1 & \theta_2 & \theta_3 \\ & 1 & \theta_1 & \theta_2 \\ & & 1 & \theta_1 \\ & & & 1 \end{pmatrix} \quad (k = 3)$$

- at $\theta = (0.4945, -0.4593, -0.8462)'$ ARE of PLE is 18%
- boxplots of $\hat{\theta}_{n,1}^{PLE} - \theta$ and $\hat{\theta}_{n,1}^{OSE} - \theta$ for $n = 50, 250$:



Introduction
Main Results
Monte Carlo results
Remarks on proofs
❖ efficient score
❖ efficiency OSE
❖ geometry
Conclusion

Efficient score

$$\dot{\ell}^*_{\theta,m}(\mathbf{u};\theta) = \frac{1}{2}\mathbf{z}' A_m(\theta)\mathbf{z}, \quad \mathbf{z_j} = \Phi^{-1}(u_j)$$

with

$$A_m(\theta) = D_\theta(\mathbf{g}_m(\theta)) - \frac{\partial}{\partial\theta_m} R^{-1}(\theta)$$

$$D_\theta(\mathbf{b}) = R^{-1}(\theta)\,\mathrm{diag}(\mathbf{b}) + \mathrm{diag}(\mathbf{b}) R^{-1}(\theta)$$

$$\mathbf{g}_m(\theta) = -(I_p + R(\theta)\circ R^{-1}(\theta))^{-1}\left[\frac{\partial}{\partial\theta_m} R(\theta)\circ R^{-1}(\theta)\right]\iota_p$$

Efficient score

Derivation efficient score:

- calculate (parametric) score for θ and nonparametric tangent-space for copula model
- important:

$$\text{score}(\mathbf{X}; \mathbb{P}_{\theta,F_1,\ldots,F_p}) = \text{score}(F_1(U_1),\ldots,F_p(U_p);\mathbb{P}_\theta)$$

  where $\mathbb{P}_\theta = \mathbb{P}_{\theta,\text{Un}[0,1],\ldots,\text{Un}[0,1]}$
  - $I^*(\theta)$ does not depend on $F_1,\ldots,F_p$
- efficient score is determined by projection of score on orthocomplement of nonparametric tangent-space
  - orthogonality conditions imply system of coupled Sturm-Liouville equations
  - solving system yields efficient score

Introduction

Main Results

Monte Carlo results

Remarks on proofs

❖ efficient score

❖ efficiency OSE

❖ geometry

Conclusion

Efficiency OSE

Efficient (at $\mathbb{P}_\theta$) if:

$$\hat{\theta}_n = \theta + \frac{1}{n}\sum_{i=1}^{n} I^{*-1}(\theta)\,\dot{\ell}^*_\theta(F_1(X_{i1}),\ldots,F_p(X_{ip});\theta) + o_P(1/\sqrt{n})$$

Recall:

$$\hat{\theta}_n^{OSE} = \tilde{\theta}_n + \frac{1}{n}\sum_{i=1}^{n} I^*(\tilde{\theta}_n)^{-1}\dot{\ell}^*_\theta(\hat{F}_{n,1}(X_{i1}),\ldots,\hat{F}_{n,p}(X_{ip});\tilde{\theta}_n)$$

So need to prove:

$$\sqrt{n}(\tilde{\theta}_n - \theta) + \frac{1}{\sqrt{n}}\sum_{i=1}^{n} I^*(\tilde{\theta}_n)^{-1}\dot{\ell}^*_\theta(\hat{F}_{n,1}(X_{i1}),\ldots,\hat{F}_{n,p}(X_{ip});\tilde{\theta}_n)$$
$$-\frac{1}{\sqrt{n}}\sum_{i=1}^{n} I^{*-1}(\theta)\,\dot{\ell}^*_\theta(F_1(X_{i1}),\ldots,F_p(X_{ip});\theta) = o_P(1)$$

Introduction
Main Results
Monte Carlo results
Remarks on proofs
❖ efficient score
❖ efficiency OSE
❖ geometry
Conclusion

Efficiency OSE

Following steps are crucial:

1. for any sequence $\theta_n = \theta + h_n/\sqrt{n}$, with h_n bounded,

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \dot{\ell}_\theta^*(\mathbf{U}_i; \theta_n) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \dot{\ell}_\theta^*(\mathbf{U}_i; \theta) - I^*(\theta) h_n + o(1; \mathbb{P}_\theta);$$

- relatively easy using result from VAN DER VAART (1988)

2. for any sequence $\theta_n = \theta + h_n/\sqrt{n}$, with h_n bounded,

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \dot{\ell}_\theta^*\big(\hat{F}_{n,1}(U_{i1}), \ldots, \hat{F}_{n,p}(U_{ip}); \theta_n\big)$$
$$= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \dot{\ell}_\theta^*(\mathbf{U}_i; \theta_n) + o(1; \mathbb{P}_\theta).$$

Introduction
Main Results
Monte Carlo results
Remarks on proofs
❖ efficient score
❖ efficiency OSE
❖ geometry
Conclusion

Efficiency OSE

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\dot{\ell}_\theta^*\big(\hat{F}_{n,1}(U_{i1}),\ldots,\hat{F}_{n,p}(U_{ip});\theta_n\big)-\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\dot{\ell}_\theta^*(\mathbf{U}_i;\theta_n)=o(1;\mathbb{P}_\theta).$$

Crux:

$$\mathbb{E}_\theta\left[\frac{\partial}{\partial u_j}\dot{\ell}_\theta^*(\mathbf{U}_i;\theta)\mid U_j\right]=0$$

Implies:

$$\mathbb{E}_{\theta_n}\Big[\frac{\partial}{\partial u_j}\dot{\ell}_\theta^*(\mathbf{U}_i;\theta_n)(\hat{F}_{nj}(U_{ij})-U_{ij})\mid U_{1,j},\ldots,U_{n,j}\Big]=0$$

$$\mathrm{var}_{\theta_n}\Big[\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\partial}{\partial u_j}\dot{\ell}_\theta^*(\mathbf{U}_i;\theta_n)(\hat{F}_{nj}(U_{ij})-U_{ij})\mid U_{1,j},\ldots,U_{n,j}\Big]$$

$$=\frac{1}{n}\sum_{i=1}^{n}A_{ni}(\hat{F}_{nj}(U_{ij})-U_{ij})^2$$

Introduction
Main Results
Monte Carlo results
Remarks on proofs
❖ efficient score
❖ efficiency OSE
❖ geometry
Conclusion

Geometry

- all relevant score and influence functions turn out to be centered quadratic forms in the Gaussianized observations $\mathbf{Z}$

$$q_A(z) = \frac{1}{2}\mathbf{z}'A\mathbf{z} - \frac{1}{2}\mathbb{E}_\theta \mathbf{Z}'A\mathbf{Z}$$

- identifying q_A with A (symmetric $p \times p$) leads to an inner product for matrices that also appeared in the efficient information matrix:

$$\langle A, B\rangle_\theta = \mathrm{cov}_\theta(q_A(\mathbf{Z}), q_B(\mathbf{Z})) = \frac{1}{2}\,\mathrm{tr}[R(\theta)AR(\theta)B]$$

- statistical interpretation of reduction to quadratic forms: least favourable submodel related to Gaussian with unknown variances [HOFF, NIU & WELLNER (2013)]
- identification with matrices yields convenient criteria for:
  - (in)efficiency of the pseudo-likelihood estimator
  - adaptivity

Introduction
Main Results
Monte Carlo results
Remarks on proofs
Conclusion

Conclusion

Contributions:

- One-Step Estimator:
  - rank-based and (semiparametrically) efficient
  - outperforms popular PLE both asymptotically and in finite samples
  - easy to implement [code available upon request]
- Geometry:
  - straightforward conditions to verify efficiency of given estimator
    - new: PLE is efficient in exchangeable models and factor models
  - adaptivity

Conclusion

Future work:

- general semiparametric copula models
  - OSE has same form, but efficient score cannot be calculated explicitly
    - proofs using implicit representations of efficient score
    - numerical approximations to implement efficient score function
- (univariate) Markovian models
  - serial dependence described via (Gaussian) copula and stationary distribution as nuisance parameter
  - calculations are related, but different due to different shape scores (tangent-space)