Copula Shrinkage and Portfolio Allocation in Ultra-High Dimensions

Stanislav Anatolyev[1] & Vladimir Pyrlik[1]

[1]CERGE-EI, Czechia

22^{nd} Conference of European Network for Business and Industrial Statistics
Trondheim, Norway
2022

Paper in a couple of words

- We consider Gaussian and t copulas characterized by correlation matrix parameters
- Apply covariance matrix shrinkage tools (Ledoit & Wolf, 00s–10s) to estimate correlation matrices, allowing use of (U)HD copulas
- Apply this to problem of portfolio allocation

Copulas

Copulas: convenient approach to generate flexible multivariate distributions

Briefly: copula is multivariate CDF with $U[0,1]$ marginals containing information about dependence

Theorem (Sklar 1959): CDF $F_X(x_1,\dots,x_p)$ with marginals $\{F_i(x_i)\}_{i=1}^p$ has copula, $C:\ [0,1]^p \to [0,1]$ such that

$$F_X(x) = C\big(F_1(x_1),\dots,F_p(x_p)\big)$$

and converse is true

Copula modeling: take copula, feed marginals in, obtain multivariate distribution

Copulas in (ultra) high dimensions

HD: number of variables p exceeds sample size n

UHD: datasets with up to thousands of variables that use up to 30 times lower sample sizes – well beyond what is studied in copula literature

Copulas & Curse of dimensionality:

- Archimedean copulas: easily extendable to HD, yet very rigid – $\dim\theta$ small and not tied to p
- Vine copulas: most flexible, but may lead to overfitting; arbitrary; estimation very hard to handle
- **Our choice:** elliptical copulas (Gaussian & t copulas) easily extendable to HD with use of shrinkage (as traditional estimators fail)

Copula estimation

Suppose $X \sim F_X(x) = C\big(F_1(x_1, \theta_1), \ldots, F_p(x_p, \theta_p), \Theta_C\big)$

Approaches to copula estimation:

- Pseudo-MLE (treating copula as distribution of *pseudo-observations*): close to full MLE estimates, but impractical in HD
  - estimate marginals $F_1(x_1), \ldots, F_p(x_p) \Rightarrow \hat{F}_1(x_1), \ldots, \hat{F}_p(x_p)$
  - use pseudo-observations $U_i = \hat{F}_i(x_i)$, $i = 1, \ldots, p$
- Method-of-moments estimators applied to pseudo-observations: very practical in HD, but ill-conditioned even in LD
- **Our approach**: apply MM together with shrinkage to estimate Θ_C when pure MM fails

Gaussian and t copulas

- Gaussian copula: $C_P^{\mathcal{N}}(u) = \Phi_P\big(\Phi^{-1}(u_1), \ldots \Phi^{-1}(u_p)\big)$

P – correlation matrix

$\Phi_P(x)$ – joint CDF of $\mathcal{N}(0^p, P)$

$\Phi^{-1}(u)$ – univariate $\mathcal{N}(0,1)$ quantile function

- Student's t copula: $C_{P,\nu}^{t}(u) = t_{P,\nu}\big(t_\nu^{-1}(u_1), \ldots t_\nu^{-1}(u_p)\big)$

P – correlation matrix

ν – degrees of freedom

$t_{P,\nu}(x)$ – joint CDF of multivariate Student's t

$t_\nu^{-1}(u)$ – univariate t_ν quantile function

NB: we have $p \times p$ correlation matrix P with $\frac{1}{2}p(p-1)$ distinct elements

Links between correlation matrix and observables

- Traditional: use Kendall's rank correlation of pseudo-observations

$$\tau_{ij} = \frac{2}{\pi} \arcsin(P_{ij})$$

$\Rightarrow$ estimate τ_{ij} for all pairs i, j and construct P

- Possibility: use approximate correlation of pseudo-observations

$$\mathrm{corr}(U) \approx P$$

with tiny approximation error (except when ν is small)
$\Rightarrow$ estimate $\mathrm{corr}(U)$ and equate to P

- Our proposal: use approximate correlation of pseudo-observations AND apply shrinkage

Quality of approximation $\mathrm{corr}(U) \approx P$ for $p = 2$

$$P = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \qquad \mathrm{cor}(u_1, u_2) \approx \rho$$

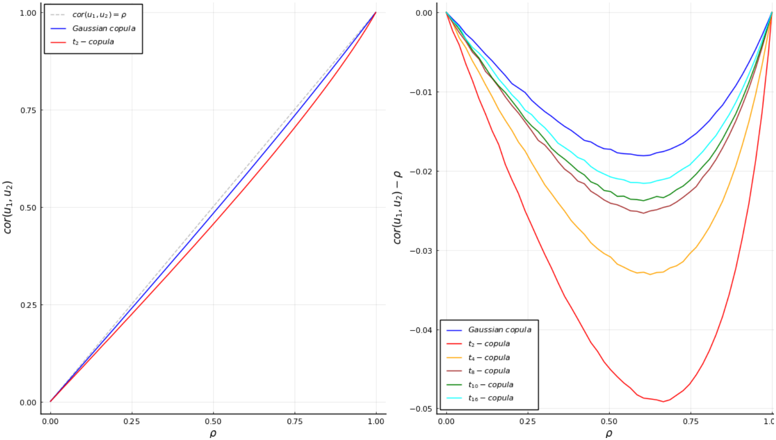


Figure: $\mathrm{cor}(u_1, u_2)$ versus ρ for Gaussian and t copulas with $\nu \in \{2, 4, 8, 10, 16\}$

Shrinkage of large covariance matrix

Olivier Ledoit & Michael Wolf (2004) Honey, I shrunk the sample covariance matrix. *Journal of Portfolio Management*

+ Ledoit & Wolf (2004 JMA, 2012 AnnStat, 2017 CSDA, 2018...2020)

Idea: sample covariance matrix $S_n = n^{-1}X_n'X_n$ for $\Sigma = \text{cov}(X)$ is ill-conditioned when $p \sim n$ and non-invertible when $p > n$

$\Rightarrow$ apply *linear shrinkage* of S_n towards identity:

$$S^* = \hat{\vartheta}\hat{\mu}I_p + (1 - \hat{\vartheta})S_n$$

where $\hat{\vartheta}$ is data-dependent optimal shrinkage intensity

$\Rightarrow$ S^* is positive definite and consistent under $p, n \to \infty$ and $p/n \to \bar{c}$

Shrinkage: eigenvalues story

If $\lambda_1, \ldots, \lambda_p$ are *overdispersed* eigenvalues of sample covariance matrix S_n then shrinkage shifts distribution of eigenvalues:

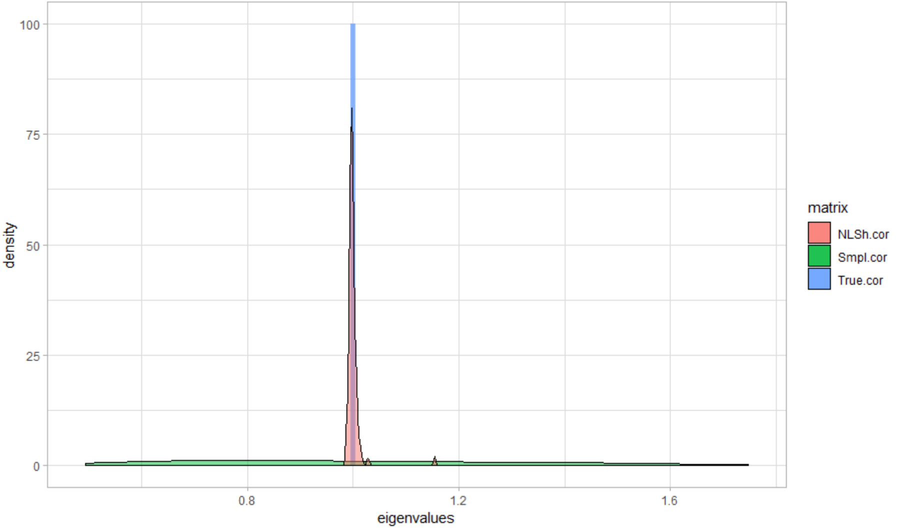
$$\lambda_i^* = \hat{\vartheta}\hat{\mu} + (1 - \hat{\vartheta})\lambda_i$$

where $\hat{\vartheta}$ is constant shrinkage intensity

Ledoit & Péché (2011) and Ledoit & Wolf (2012): optimal *nonlinear* shrinkage with variable shrinkage intensity

$$\lambda_i^* = \hat{\rho}(\lambda_i)\lambda_i$$

Example: true covariance matrix is identity



Note: simulated from multivariate normal with $p = 100$ and $n = 50$

Simulation design

- DGP
  - Dimensionality: $p \in \{10, 100, 1000\}$
  - p-to-n ratio: $\frac{1}{20} \leq \frac{p}{n} \leq 20$
  - Correlation matrix: identity OR random non-sparse positive definite
  - Other: skew-t marginals, $d.f. = 8$ for t copula
- Quality criteria
  - Positive definiteness of $\hat{P}$
  - Euclidean loss, $L_E(P, \hat{P}) = ||P - \hat{P}||^2$
  - Kullback–Leibler IC, $KLIC_{C_P|C_{\hat{P}}} = E_{C_P}[\log \frac{\partial C_P(u)}{\partial u} - \log \frac{\partial C_{\hat{P}}(u)}{\partial u}]$
- Estimators
  - Traditional: Kendall's rank correlation AND sample correlation
  - Proposed: linearly AND non-linearly shrunk sample correlation

Results: summary

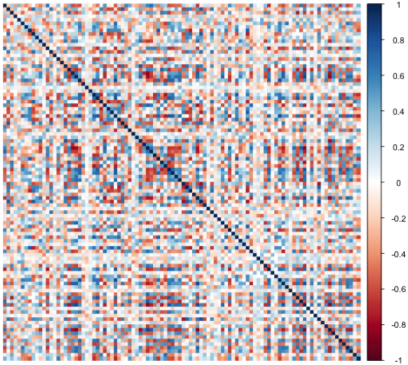
Conditioning of correlation matrices:

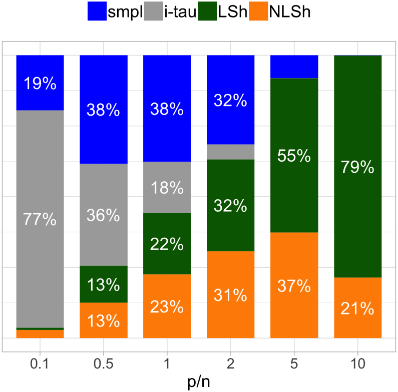
- Traditional estimators are positive-definite only under LD (and even then non-guaranteed)
- Shrinkage-based estimators guarantee positive-definiteness

Distance measures

- Shrinkage-based estimators generally outperform traditional ones, very significantly in HD
- Non-linear shrinkage tends to outperform linear one

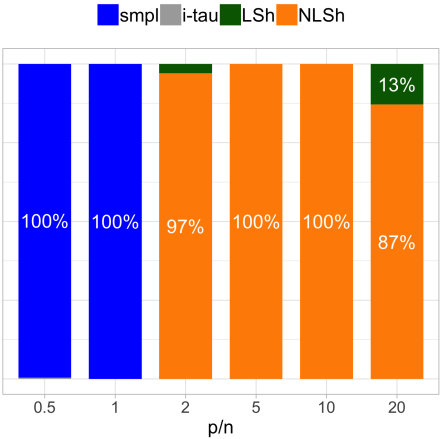
One particular case





KLIC: Random correlation matrix, Gaussian copula, $p = 100$

Another particular case



KLIC: Identity correlation matrix, t copula, $p = 1000$

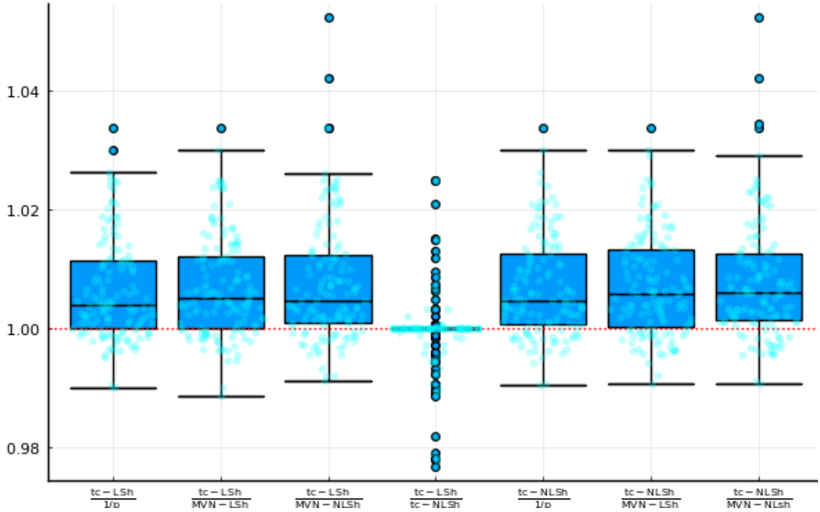
Portfolio allocation: setup

Portfolio allocation – most popular application for models of joint distribution of assets prices

This is UHD application!

- We use listing of Wilshire 5000 index and obtain data on 4980 assets from CRSP, 9 months of 2017
- 135 randomly chosen portfolios of $p = 3600$ assets
- 6 months of 2017 are used for estimation $\Rightarrow$ $n = 120$
- Dimensionality is $p/n = 30$!
- Simple marginal models of returns dynamics: EDF of ARMA-EGARCH-filtered residuals
- t copula for dependence

Portfolio allocation: relative cumulative return



Portfolios performance: t copula, $p = 3600$, $n = 120$

Summary

Conclusions:

- Shrinkage is powerful tool of elliptical copula estimation in UHD
- UHD copula modeling is beneficial in portfolio allocation

Possible extensions:

- Possibly, shrinkage towards other goals (to equicorrelation, to PCA)
- Skewed copula versions: no straight relation of P to observables; additional HD skewness parameters