Fast Empirical Scenarios

FinEML Conference 2023

Michael Multerer, Paul Schneider, Rohan Sen

Università
della
Svizzera
italiana

Motivation

- Accurate sparse representation of large samples of data.
- Applications across disciplines faced with panel data with large cross-section and time-series dimensions.
- Scenarios reflecting sample moment information compatible with available sample data.
- Reconcile portfolio volatility from large samples parsimoniously.
- Capture risk landscape of asset returns efficiently.
- Moment matching within multivariate Gaussian mixture models framework.

Contribution

- Efficient solutions to scenario selection problem.
- Superior accuracy, computationally faster, fewer scenarios than extant algorithms.
- Provide a computational framework that resolves stability issues.
- Extend notion of choosing scenarios from quasi-Monte Carlo candidate points towards sampling data.
- Grant novel scenario-based representation for covariance matrices.
- Prove that ℓ_1-regularized least squares is not effective for scenario selection.

Preliminaries

- Samples

$$X := \left\{ \boldsymbol{x}_1, \dots, \boldsymbol{x}_N \right\} \subset \mathbb{R}^d.$$

- Empirical measure

$$\widehat{\mathbb{P}} := \frac{1}{N} \sum_{i=1}^{N} \delta_{\boldsymbol{x}_i}.$$

- Empirical moments

$$\widehat{y}_{\boldsymbol{\alpha}} := \int \boldsymbol{x}^{\boldsymbol{\alpha}} \, \mathrm{d}\widehat{\mathbb{P}}, \qquad \boldsymbol{\alpha} \in \mathbb{N}^d.$$

- Empirical moment sequence (up to degree $2q$)

$$\widehat{\boldsymbol{y}} = \left[\widehat{y}_{\boldsymbol{\alpha}} \right]_{|\boldsymbol{\alpha}| \leq 2q} \in \mathbb{R}^{m_{2q}}, \qquad m_{2q} = \binom{2q+d}{d}.$$

Preliminaries

- Scenarios

$$\Xi := \left\{\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_r\right\} \quad \text{with } r \ll N.$$

- Compressed measure

$$\mathbb{P}^\star := \sum_{j=1}^{r} \lambda_j \, \delta_{\boldsymbol{\xi}_j} \qquad \lambda_j \geq 0, \quad \sum_{j=1}^{r} \lambda_j = 1.$$

- Moment sequence (up to degree $2q$)

$$\boldsymbol{y}^\star := \left[y^\star_{\boldsymbol{\alpha}}\right]_{|\boldsymbol{\alpha}| \leq 2q} = \left[\int \boldsymbol{x}^{\boldsymbol{\alpha}} \, \mathrm{d}\mathbb{P}^\star\right]_{|\boldsymbol{\alpha}| \leq 2q} \in \mathbb{R}^{m_{2q}}.$$

Problem formulation

- Compressed measure

$$\mathbb{P}^\star := \sum_{j=1}^{r} \lambda_j \, \delta_{\boldsymbol{\xi}_j} \qquad \lambda_j \geq 0, \quad \sum_{j=1}^{r} \lambda_j = 1.$$

- Optimization problem

$$\operatorname*{argmin}_{r, \boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_r, \lambda_1, \ldots, \lambda_r} \left\| \, \boldsymbol{y}^\star - \widehat{\boldsymbol{y}} \, \right\|_2.$$

- Obtaining the sparsest solution is non-convex and NP-hard in general.

- ℓ_1-regularized least squares or LASSO does not solve the problem.

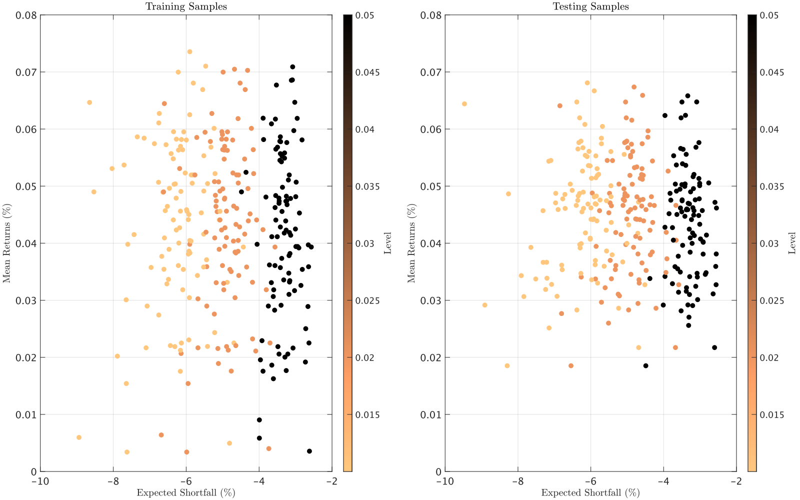
Relaxed formulation

- Relaxed convex formulation done in two stages.
- Scenario selection

$$\underset{\boldsymbol{\xi}_1,\ldots,\boldsymbol{\xi}_r}{\operatorname{argmin}} \left\| \boldsymbol{y}^\star - \widehat{\boldsymbol{y}} \right\|_2.$$

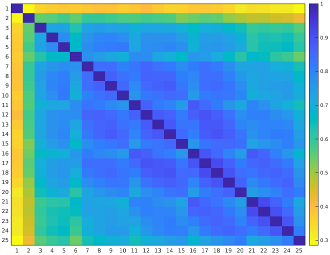
- Greedy choice of representative basis (evaluated at samples).
- Computational framework allows equivalent reformulation ensuring sparsity and stability.
- Retrieval of weights by enforcing probability constraints.

Portfolio optimization using scenarios

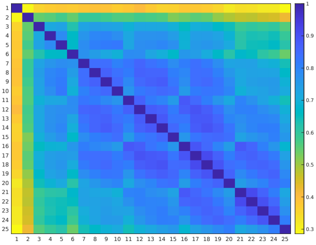


Portfolio optimization with expected shortfall constraint using scenarios

Reconciling portfolio volatility



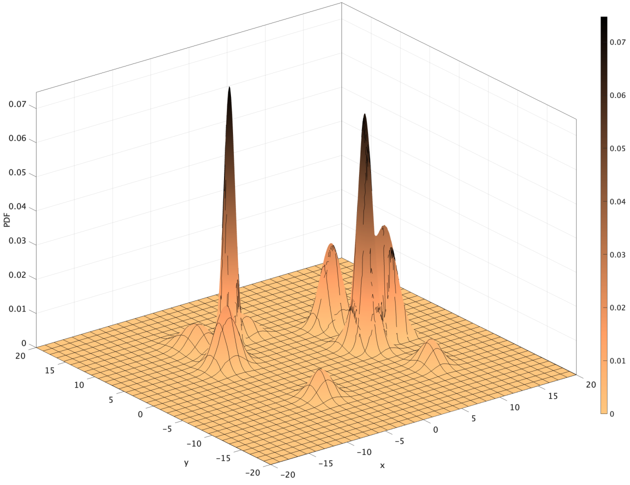
Sample correlation



Scenarios correlation

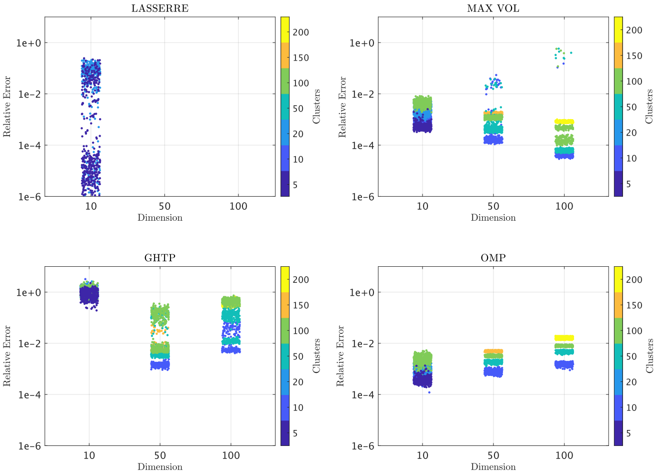
- Comparison of sample correlation with scenarios.
- Relative error of order 10^{-16} with 26 scenarios from 25000 observations.
- Covariance scenarios match exactly up to second moments.

Multivariate Gaussian mixture models



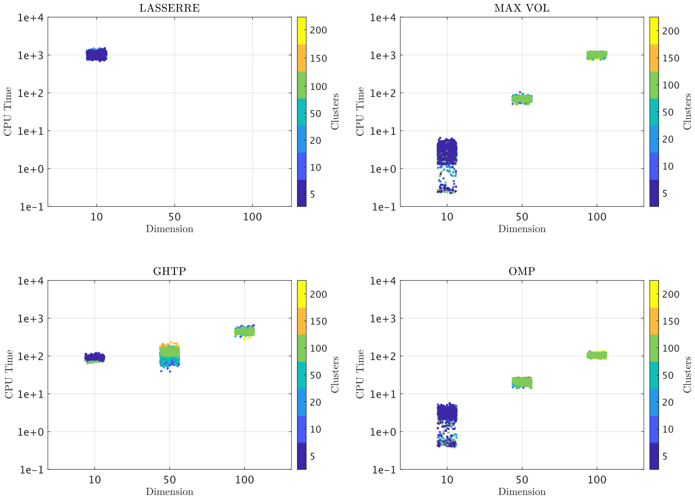
PDF of a Gaussian mixture distribution in $\mathbb{R}^2$ with 10 clusters

Relative error comparison



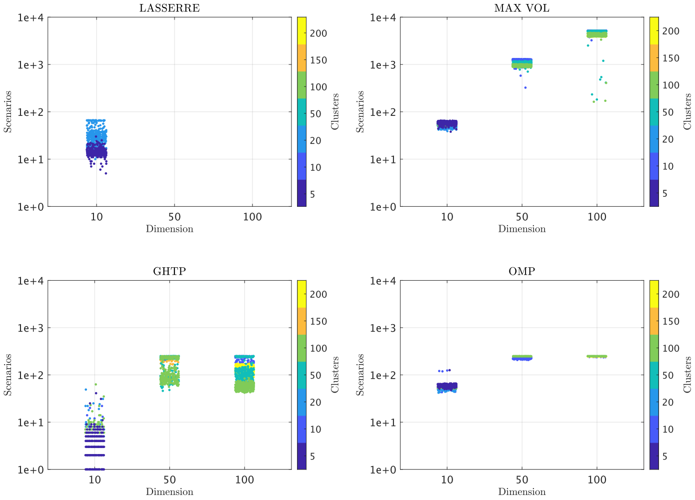
Comparison of relative errors of different algorithms

Computation time comparison



Comparison of computation times of different algorithms

Number of scenarios comparison



Comparison of number of scenarios extracted by the different algorithms

Acknowledgement

This project has been funded by the

Swiss National Science Foundation

Università della Svizzera italiana

References

Claudia Bittante, Stefano De Marchi, and Giacomo Elefante.
A new quasi-Monte Carlo technique based on nonnegative least squares and approximate Fekete points.
Numerical Mathematics: Theory, Methods and Applications, 9(4), 2016.

Jean-Luc Bouchot, Simon Foucart, and Pawel Hitczenko.
Hard thresholding pursuit algorithms: Number of iterations.
Applied and Computational Harmonic Analysis, 41(2), 2016.
Sparse Representations with Applications in Imaging Science, Data Analysis, and Beyond, Part II.

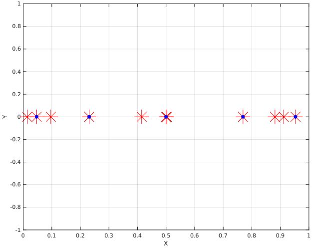
Stephen Boyd, Neal Parikh, Eric Chu, Borja Peleato, and Jonathan Eckstein.
Distributed optimization and statistical learning via the alternating direction method of multipliers.
Foundations and Trends in Machine Learning, 3, 01 2011.

H. Harbrecht, M. Peters, and R. Schneider.
On the low-rank approximation by the pivoted Cholesky decomposition.
Applied Numerical Mathematics, 62:28–440, 2012.

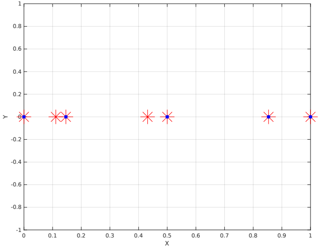
Jean Bernard Lasserre.
Moments, Positive Polynomials and Their Applications.
Imperial College Press, London, 2010.

Ernest K. Ryu and Stephen P. Boyd.
Extensions of Gauss Quadrature Via Linear Programming.
Foundations of Computational Mathematics, 15(4), Aug 2015.

Benchmark quadrature rules



Gauss-Legendre nodes



Clenshaw-Curtis nodes

- Comparison of scenarios with Gaussian and Clenshaw-Curtis nodes in $[0, 1]$.
- Exactly integrates up to $2q$ and q moments respectively.
- Successful recoveries with sparse scenarios from $10,000$ samples.

Orthogonal matching pursuit

input: kernel matrix $\boldsymbol{K} \in \mathbb{R}^{N \times N}$, vector $\boldsymbol{h} \in \mathbb{R}^{N}$, tolerance $\varepsilon > 0$

output: index set ind, low-rank approximation $\boldsymbol{K} \approx \boldsymbol{L}\boldsymbol{L}^\top$

and bi-orthogonal basis $\boldsymbol{B}$ such that $\boldsymbol{B}^\top \boldsymbol{L} = \boldsymbol{I}_m$

initialization: set $\boldsymbol{L}_0 = \boldsymbol{B}_0 = \text{ind} = [\,]$, $\boldsymbol{d}_0 = \text{diag}(\boldsymbol{K}), \boldsymbol{h}_0 = \boldsymbol{h}$, $\text{err} = 1$, $m = 1$
while err $> \varepsilon$
$\quad p_m := \arg\max_{1 \le i \le m_{2q}} d_{m-1,i}, \quad \text{ind} := [\text{ind},\ p_m]$
$\quad \boldsymbol{\ell}_m := \frac{1}{\sqrt{d_{m-1,p_m}}}\Big(\boldsymbol{K} - \boldsymbol{L}\boldsymbol{L}^\top\Big)\boldsymbol{e}_{p_m}, \quad \boldsymbol{b}_m := \frac{1}{\sqrt{d_{m-1,p_m}}}\Big(\boldsymbol{I} - \boldsymbol{B}\boldsymbol{L}^\top\Big)\boldsymbol{e}_{p_m}$
$\quad$ set $\boldsymbol{L} := [\boldsymbol{L},\ \boldsymbol{\ell}_m], \quad \boldsymbol{B} := [\boldsymbol{B},\ \boldsymbol{b}_m]$
$\quad$ set $\boldsymbol{d}_m := \boldsymbol{d}_{m-1} - \boldsymbol{\ell}_m \odot \boldsymbol{\ell}_m$
$\quad$ set $\boldsymbol{h}_m := \boldsymbol{h}_{m-1} - (\boldsymbol{h}^\top \boldsymbol{b}_m)\boldsymbol{\ell}_m$
$\quad$ set $\text{err} := \|\boldsymbol{h}_m\|_2 \,/\, \|\boldsymbol{h}\|_2$
$\quad$ set $m := m + 1$
end