# NEAREST COMOMENT ESTIMATION WITH UNOBSERVED FACTORS AND LINEAR SHRINKAGE

Kris Boudt, **Dries Cornilly**, Tim Verdonck

May 19, 2017

VUB VRIJE UNIVERSITEIT BRUSSEL

KU LEUVEN

# OUTLINE

# Higher Order Moments

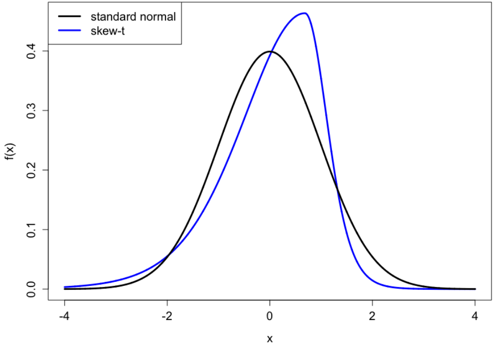
VRIJE UNIVERSITEIT BRUSSEL

KU LEUVEN

## UNIVARIATE MOMENTS

The moments describe the characteristics of the univariate distribution of portfolio returns $r_p$.

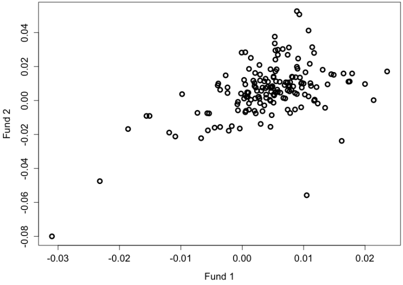
- mean, $\mu = \mathbb{E}[r_p]$
- variance, $\sigma^2 = \mathbb{E}[(r_p - \mu)^2]$
- skewness, $\phi = \mathbb{E}[(r_p - \mu)^3]$
- kurtosis, $\psi = \mathbb{E}[(r_p - \mu)^4]$



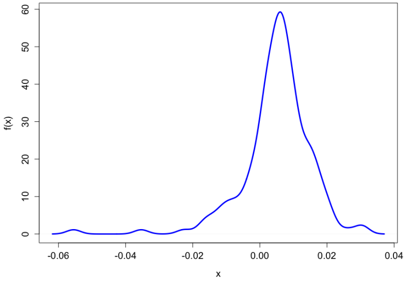
# UNIVARIATE VS. MULTIVARIATE

It holds that $r_p = \boldsymbol{w}'\boldsymbol{X}$,
$\boldsymbol{X}$ a $p$-dimensional random variable of asset returns.

(a) Sample of $\boldsymbol{X}$

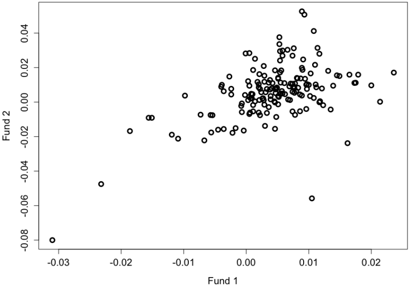



(b) $r_p = 0.5X_1 + 0.5X_2$



# MULTIVARIATE MOMENTS

## COVARIANCE MATRIX

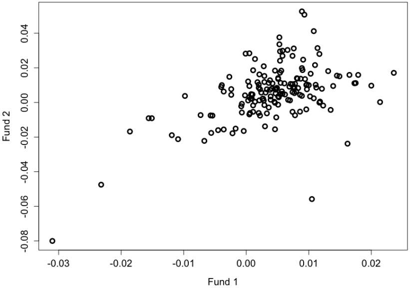



Covariance between 2 assets: $\sigma_{ij} = \mathbb{E}[(X_i - \mu_i)(X_j - \mu_j)]$

$$\mathbf{\Sigma} = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} = \begin{pmatrix} 5.344 & 5.937 \\ 5.937 & 25.181 \end{pmatrix} \times 10^{-5}, \qquad (\rho = 0.512)$$

# MULTIVARIATE MOMENTS

## COSKEWNESS MATRIX




Coskewness between 3 assets:

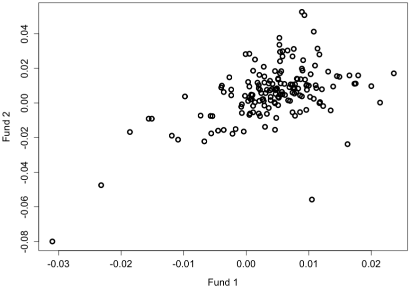
$$\phi_{ijk} = \mathbb{E}[(X_i - \mu_i)(X_j - \mu_j)(X_k - \mu_k)]$$

$$\mathbf{\Phi} = \begin{pmatrix} \phi_{111} & \boldsymbol{\phi_{112}} & \phi_{211} & \phi_{212} \\ \phi_{121} & \phi_{122} & \underline{\phi_{221}} & \phi_{222} \end{pmatrix}$$

$$= \begin{pmatrix} -0.458 & \mathbf{-1.107} & -1.107 & -1.950 \\ -1.107 & -1.950 & \underline{-1.950} & -4.522 \end{pmatrix} \times 10^{-6}$$

# MULTIVARIATE MOMENTS

## COKURTOSIS MATRIX




Cokurtosis between 4 assets:

$$\psi_{ijkl} = \mathbb{E}[(X_i - \mu_i)(X_j - \mu_j)(X_k - \mu_k)(X_l - \mu_l)]$$

$$\mathbf{\Psi} = \begin{pmatrix} \color{red}{\psi_{1111}} & \boldsymbol{\psi_{1112}} & \psi_{1211} & \underline{\boldsymbol{\psi_{1212}}} & \psi_{2111} & \psi_{2112} & \psi_{2211} & \psi_{2212} \\ \psi_{1121} & \psi_{1122} & \psi_{1221} & \psi_{1222} & \psi_{2121} & \psi_{2122} & \underline{\psi_{2221}} & \color{blue}{\psi_{2222}} \end{pmatrix}$$

$$= \begin{pmatrix} \color{red}{0.207} & \mathbf{0.351} & 0.351 & \underline{\mathbf{0.785}} & 0.351 & 0.785 & 0.785 & 1.678 \\ 0.351 & 0.785 & 0.785 & 1.678 & 0.785 & 1.678 & \underline{1.678} & \color{blue}{5.799} \end{pmatrix}$$

## WHY BOTHER?

1. Easy to compute portfolio moments:

$$\sigma_{r_p}^2 = \boldsymbol{w}' \boldsymbol{\Sigma}_{\boldsymbol{X}} \boldsymbol{w}$$

$$\phi_{r_p} = \boldsymbol{w}' \boldsymbol{\Phi}_{\boldsymbol{X}} (\boldsymbol{w} \otimes \boldsymbol{w})$$

$$\psi_{r_p} = \boldsymbol{w}' \boldsymbol{\Psi}_{\boldsymbol{X}} (\boldsymbol{w} \otimes \boldsymbol{w} \otimes \boldsymbol{w})$$

2. Risk budgets for portfolio optimization, e.g. mVaR, mES (Zangari (1996), Boudt, Peterson and Croux (2008))

$$\text{mVaR}(\alpha) = \mu_{r_p} + \sigma_{r_p}(z_\alpha + c_1 \boldsymbol{s}_{r_p} + c_2 \kappa_{r_p} - c_3 \boldsymbol{s}_{r_p}{}^2),$$

with $s_{r_p} = \phi_{r_p} / \sigma_{r_p}^3$ and $\kappa_{r_p} = \psi_{r_p} / \sigma_{r_p}^4 - 3$.

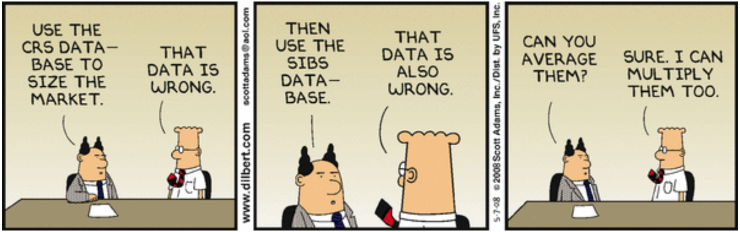
3. Discover variance, skewness and kurtosis hotspots in the portfolio, hedge when necessary. (Rubinstein (1973), Diacogiannis(1995), Litterman (1996))
4. Gradient-based optimization algorithms

# Estimation

# THEORY VS. PRACTICE

In practice: **estimation**

**Garbage in, garbage out**



# SAMPLE ESTIMATORS

Sample estimators:

$$\widehat{\Sigma}_s = \frac{1}{n}\sum_{i=1}^{n}(\boldsymbol{x}_i - \overline{\boldsymbol{x}})(\boldsymbol{x}_i - \overline{\boldsymbol{x}})',$$

$$\widehat{\Phi}_s = \frac{1}{n}\sum_{i=1}^{n}(\boldsymbol{x}_i - \overline{\boldsymbol{x}})(\boldsymbol{x}_i - \overline{\boldsymbol{x}})' \otimes (\boldsymbol{x}_i - \overline{\boldsymbol{x}})',$$

$$\widehat{\Psi}_s = \frac{1}{n}\sum_{i=1}^{n}(\boldsymbol{x}_i - \overline{\boldsymbol{x}})(\boldsymbol{x}_i - \overline{\boldsymbol{x}})' \otimes (\boldsymbol{x}_i - \overline{\boldsymbol{x}})' \otimes (\boldsymbol{x}_i - \overline{\boldsymbol{x}})'.$$

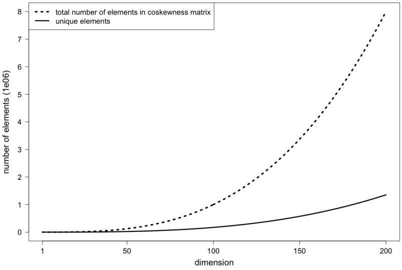
**Issues:**

- computational
- statistical

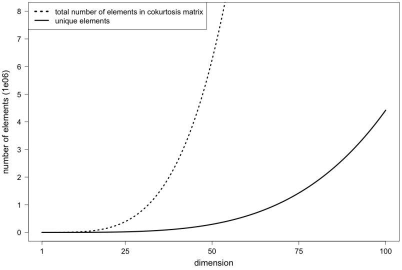
# SAMPLE ESTIMATORS

## NUMBER OF UNIQUE ELEMENTS

(a) Coskewness




(b) Cokurtosis



# SAMPLE ESTIMATORS

## NUMBER OF UNIQUE ELEMENTS

**Speed gain:**

current implementation in `PerformanceAnalytics` (Fortran) vs. estimation of only unique elements (`RcppArmadillo`)

- Coskewness matrix:

| $p$ | 10 | 20 | 30 | 50 | 100 |
|---|---|---|---|---|---|
| Absolute time (ms.) | 0.40 | 1.20 | 4.10 | 18.90 | 183.10 |
| Relative improvement (%) | 275.00 | 608.30 | 614.60 | 406.90 | 880.20 |

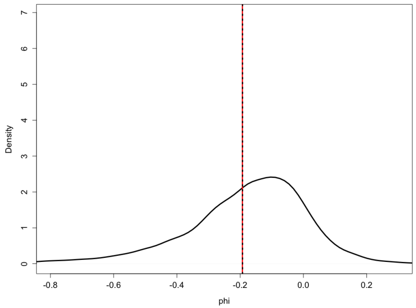
- Cokurtosis matrix:

| $p$ | 10 | 20 | 30 | 50 | 100 |
|---|---|---|---|---|---|
| Absolute time (ms.) | 0.90 | 11.10 | 65.30 | 346.50 | 4284.40 |
| Relative improvement (%) | 1200.00 | 993.70 | 1831.20 | 2908.20 | 3088.50 |

Comparison made with the package `rbenchmark`

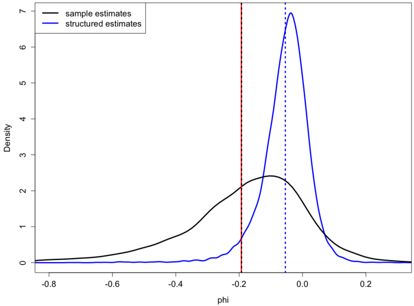
# SAMPLE ESTIMATORS

## STATISTICAL ISSUES




- consistent
- large mean squared error (MSE)
- many unique elements (error maximization)

# STRUCTURED ESTIMATION




Assume a certain data-generating process, e.g. independence, factor model, . . .

- imposes a certain structure on the covariance, coskewness and cokurtosis matrices
- reduces number of free parameters
- potentially reduces the MSE (when the model is not too misspecified)

# LINEAR SHRINKAGE

- Introduced by Ledoit and Wolf (2003) for covariance estimation.
- Extended to coskewness and cokurtosis estimation in Martellini and Ziemann (2010).
- Improved shrinkage estimation of the coskewness matrix in Boudt, Cornilly and Verdonck (2017a). Contributions:
  - Unbiased and consistent estimators leading to lower MSE,
  - optimized target matrices,
  - possibility for multi-target shrinkage (using multiple targets at once).

## LINEAR SHRINKAGE

$$\widehat{\boldsymbol{\Phi}}^{\mathsf{ST}}(\lambda) = (1-\lambda)\widehat{\boldsymbol{\Phi}} + \lambda\widehat{\boldsymbol{T}}, \qquad \lambda \in [0,1],$$

where

- $\widehat{\boldsymbol{\Phi}}$ is the unbiased sample estimator,
- $\widehat{\boldsymbol{T}}$ is a structured coskewness matrix

Aim: choose $\lambda$ as optimal trade-off between

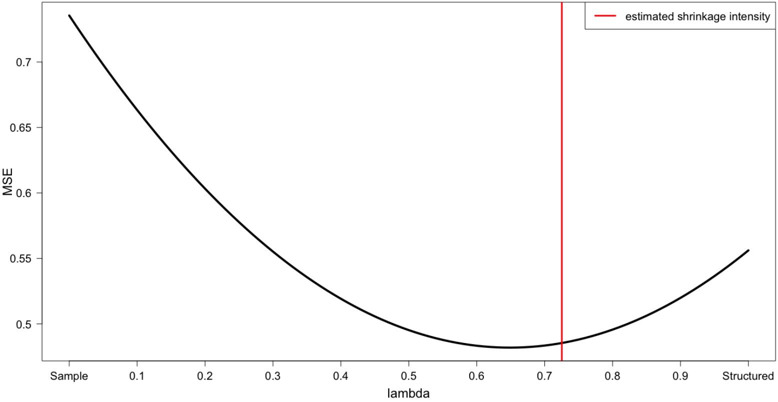
- unbiasedness and large estimation variance of $\widehat{\boldsymbol{\Phi}}$,
- biased but smaller estimation variance of $\widehat{\boldsymbol{T}}$.

Approach: minimize the MSE,

$$\lambda^{\star} = \arg\min_{\lambda\in[0,1]} \mathbb{E}\left[\|(1-\lambda)\widehat{\boldsymbol{\Phi}} + \lambda\widehat{\boldsymbol{T}} - \boldsymbol{\Phi}\|^2\right]$$

# LINEAR SHRINKAGE

Figure: Mean Squared Error



# NEAREST COMOMENT ESTIMATION

Suppose that the data follows a multi-factor model,

$$\boldsymbol{X} = \boldsymbol{B}\boldsymbol{F} + \varepsilon,$$

then (see Boudt, Lu and Peeters (2015)),

$$\begin{aligned}
\boldsymbol{\Sigma}_{\boldsymbol{X}} &= \boldsymbol{B}\boldsymbol{\Sigma}_{\boldsymbol{F}}\boldsymbol{B}' + \boldsymbol{\Delta} \\
\boldsymbol{\Phi}_{\boldsymbol{X}} &= \boldsymbol{B}\boldsymbol{\Phi}_{\boldsymbol{F}}(\boldsymbol{B}' \otimes \boldsymbol{B}') + \boldsymbol{\Omega} \\
\boldsymbol{\Psi}_{\boldsymbol{X}} &= \boldsymbol{B}\boldsymbol{\Psi}_{\boldsymbol{F}}(\boldsymbol{B}' \otimes \boldsymbol{B}' \otimes \boldsymbol{B}') + \boldsymbol{\Gamma}
\end{aligned}$$

- Factors are **observed**: structured estimation as in Boudt, Lu and Peeters (2015)
- Factors are **unobserved**: nearest comoment estimation as in Boudt, Cornilly and Verdonck (2017b) (working paper)

# Conclusion

# CONCLUSION

- Higher order moments matter
- Many practical applications
- Estimation matters more!
- Coming soon: update to `PerformanceAnalytics`
  - improve computational efficiency of the higher order moments
  - implementation of several state-of-the-art estimators

# REFERENCES


Boudt, K., Cornilly, D. Verdonck, T. “A Coskewness shrinkage approach for estimating the skewness of linear combinations of random variables.” (2017a). Available at SSRN: https://ssrn.com/abstract=2839781

Boudt, K., Cornilly, D. Verdonck, T. “Nearest comoment estimation with unobserved factors” (2017b). (working paper)

Boudt, K., Lu, W., Peeters, B. “Higher order comoments of multifactor models and asset allocation.” *Finance Research Letters* 13 (2015), 225–233

Boudt, K., Peterson, B. G., Croux, C. “Estimation and decomposition of downside risk for portfolios with non-normal returns.” *Journal of Risk* 11(2) (2008), 79–103

Diacogiannis, G. P. “Three-parameter asset pricing.” *Managerial and Decision Economics* 15(2) (1994), 149–158

# REFERENCES

Ledoit, O., Wolf, M. "Improved estimation of the covariance matrix of stock returns with an application to portfolio selection." *Journal of empirical finance* 10(5) (2003), 603–621

Litterman, R. "Hot Spots™ and hedges." *The Journal of Portfolio Management* 23(5) (1996), 52–75

Martellini, L., Ziemann, V. "Improved estimates of higher-order comoments and implications for portfolio selection." *Review of Financial Studies* 23(4) (2010), 1467–1502

Rubinstein, M. E. "The fundamental theorem of parameter-preference security valuation." *Journal of Financial and Quantitative Analysis* 8(1) (1973), 61–69

Zangari, P. "A VaR methodology for portfolios that include options." *RiskMetrics Monitor* (1st quarter 1996), 4–12