# **PerformanceAnalytics** Estimation of Higher Order Moments

Dries Cornilly and Brian G. Peterson

December 8, 2024

**Abstract**

This vignette gives a an overview of the estimation methods for coskewness and cokurtosis matrices contained in **PerformanceAnalytics**. We focus on explaining the different assumptions underlying each of the estimators and how to use them in **R**. The concepts are shown using the `edhec` dataset included in the package.

# 1 Introduction

The `PerformanceAnalytics` package was built to evaluate the performance and risk characteristics of financial assets or funds. Two important risk measurement functions are Value-at-Risk (VaR) and Expected Shortfall (ES). The modified versions, see Zangari (1996) and Boudt et al. (2008), incorporate departure from normality in the return series. In the univariate setting, this departure from normality is measured by skewness and kurtosis.

The portfolio third and fourth order central moments depend on the joint third and fourth order central moments of all underlying assets, these are gathered in the coskewness and cokurtosis matrices. Consider the random vector $\boldsymbol{R}$ with mean $\boldsymbol{\mu}$ and finite fourth order moments, then the coskewness matrix $\boldsymbol{\Phi}$ and cokurtosis matrix $\boldsymbol{\Psi}$ are defined by

$$\begin{aligned} \boldsymbol{\Phi} &= \mathbb{E}[(\boldsymbol{R}-\boldsymbol{\mu})(\boldsymbol{R}-\boldsymbol{\mu})' \otimes (\boldsymbol{R}-\boldsymbol{\mu})'], \\ \boldsymbol{\Psi} &= \mathbb{E}[(\boldsymbol{R}-\boldsymbol{\mu})(\boldsymbol{R}-\boldsymbol{\mu})' \otimes (\boldsymbol{R}-\boldsymbol{\mu})' \otimes (\boldsymbol{R}-\boldsymbol{\mu})'], \end{aligned} \tag{1}$$

where $\otimes$ denotes the Kronecker product.

For any portfolio with portfolio vector $\boldsymbol{w}$, the third and fourth order central moment are then given by

$$\begin{aligned} \mathbb{E}[(\boldsymbol{w}'\boldsymbol{R}-\boldsymbol{w}'\boldsymbol{\mu})^3] &= \boldsymbol{w}'\boldsymbol{\Phi}(\boldsymbol{w}\otimes\boldsymbol{w}), \\ \mathbb{E}[(\boldsymbol{w}'\boldsymbol{R}-\boldsymbol{w}'\boldsymbol{\mu})^4] &= \boldsymbol{w}'\boldsymbol{\Psi}(\boldsymbol{w}\otimes\boldsymbol{w}\otimes\boldsymbol{w}), \end{aligned} \tag{2}$$

which can then be used to compute the standardized skewness and excess kurtosis of the portfolio.

In addition, the contribution of each asset to the modified VaR (mVaR) or modified ES (mES) of the portfolio depends on the derivative with respect to the weights vector, which is also a function of the coskewness and cokurtosis matrices measuring the relationships between all assets.

Hence, the idea of improving the estimation of the higher order comoment matrices is that eventually we obtain more accurate estimates of mVaR and mES, or any other application involving the higher order comoment matrices.

In the next section we explain the underlying assumptions for all the estimation methods available in `PerformanceAnalytics` and demonstrate how to use them on the dataset `edhec` included in the package. In Section 3 we show how the different estimation techniques impacts the mVaR and mES values. It is not our aim however to draw any conclusions about which method is preferred.

Before starting, load the package and the dataset `edhec`. Below, the different asset names are shown, as well as the head and tail for select hedge fund style indices. As you can see, the data consists of monthly returns.

```
> data(edhec)
> colnames(edhec)
 [1] "Convertible Arbitrage"  "CTA Global"              "Distressed Secu
 [4] "Emerging Markets"       "Equity Market Neutral"   "Event Driven"
 [7] "Fixed Income Arbitrage" "Global Macro"            "Long/Short Equit
[10] "Merger Arbitrage"       "Relative Value"          "Short Selling"
[13] "Funds of Funds"
> head(edhec[, c(2,4,5,6,8,9,11,12)], n = 3)
           CTA Global Emerging Markets Equity Market Neutral Event Drive
1997-01-31     0.0393           0.0791                0.0189       0.021
1997-02-28     0.0298           0.0525                0.0101       0.008
1997-03-31    -0.0021          -0.0120                0.0016      -0.002
           Global Macro Long/Short Equity Relative Value Short Selling
1997-01-31       0.0573            0.0281         0.0180       -0.0166
1997-02-28       0.0175           -0.0006         0.0118        0.0426
1997-03-31      -0.0119           -0.0084         0.0010        0.0778
> tail(edhec[, c(2,4,5,6,8,9,11,12)], n = 3)
           CTA Global Emerging Markets Equity Market Neutral Event Drive
2021-03-31     0.0045          -0.0081                0.0102       0.010
2021-04-30     0.0250           0.0257                0.0147       0.027
2021-05-31     0.0164           0.0209                0.0035       0.012
           Global Macro Long/Short Equity Relative Value Short Selling
2021-03-31       0.0093            0.0075         0.0054         0.002
2021-04-30       0.0233            0.0249         0.0109         0.000
2021-05-31       0.0188            0.0085         0.0067         0.000
```

This data sometimes has missing values (NaN), as recently observed in the 'Short Selling' style indice. We choose to handle this by zero imputation, as we have very few. See the edhec documentation files. This will allow multi-target shrinkage estimators, among other computations to run smoothly in our examples.

# 2 Estimation of the Coskewness and Cokurtosis matrices

In this section we cover the three main categories of estimation techniques implemented in PerformanceAnalytics, plug-in estimation, structured estimation and shrinkage estimation. This is not an exhaustive list and in the last subsection we refer the reader to other possibilities, such as EWMA estimation and MCA estimation (both available in PerformanceAnalytics); or the GO-GARCH model in rmgarch.

## 2.1 Plug-in estimation

Consider a sample $(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_n)$ of $p$-dimensional observations of the return variable $\boldsymbol{R}$. The most straightforward way of estimating the coskewness and cokurtosis matrices given in (1) consists of replacing each expectation by a sample average, i.e. the sample coskewness and cokurtosis matrices $\widehat{\boldsymbol{\Phi}}$ and $\widehat{\boldsymbol{\Psi}}$ have as elements

$$\begin{aligned} \widehat{\phi}_{ijk} &= \frac{1}{n}\sum_{t=1}^{n}(r_{ti}-\overline{r}_i)(r_{tj}-\overline{r}_j)(r_{tk}-\overline{r}_k), \\ \widehat{\psi}_{ijkl} &= \frac{1}{n}\sum_{t=1}^{n}(r_{ti}-\overline{r}_i)(r_{tj}-\overline{r}_j)(r_{tk}-\overline{r}_k)(r_{tl}-\overline{r}_l), \end{aligned} \tag{3}$$

where $\overline{\boldsymbol{r}}$ denotes the sample average. The way to estimate the plug-in sample coskewness and cokurtosis estimators is as follows

```
> m3 <- M3.MM(edhec)
> m4 <- M4.MM(edhec)
> dim(m3)
```

```
[1]  13 169
```

```
> dim(m4)
```

```
[1]   13 2197
```

The dimension of the matrices is correct, at $p \times p^2$ for the coskewness matrix and $p \times p^3$ for the cokurtosis matrix.

As shown in Equation (2), the portfolio third and fourth order central moment are easily obtained by left and right multiplication with the weight vector $w$. In the vignette we consider the equal weighted portfolio. In code, the moments are computed as

```
> p <- ncol(edhec)
> w <- rep(1 / p, p)
> m3port <- t(w) %*% m3 %*% (w %x% w)
> m4port <- t(w) %*% m4 %*% (w %x% w %x% w)
```

Computing the third and fourth order central moment of the portfolio can also be done by directly using the functions `portm3` and `portm4` in the `PerformanceAnalytics` package. This has the advantage that these functions also accept the reduced form vector with unique coskewness and cokurtosis elements, as obtained by setting the option `as.mat = F` in any of the estimators shown in this vignette, as shown below for the coskewness matrix. Computing the reduced form has the advantage of a faster computation time because it heavily reduces the memory burden and avoids unnecessary computations for the portfolio moments, especially in higher dimensions.

```
> m3port_2 <- PerformanceAnalytics:::portm3(w, M3.MM(edhec))
> m3port_3 <- PerformanceAnalytics:::portm3(w, M3.MM(edhec, as.mat = F))
> c(m3port, m3port_2, m3port_3)

[1] -1.559239e-06 -1.559239e-06 -1.559239e-06
```

One of the properties of the sample estimator is that computing the portfolio third and fourth order central moment from the plug-in sample coskewness and cokurtosis matrices is equivalent to directly computing the third and fourth order central moment on the portfolio returns, as shown below.

```
> portreturns <- edhec %*% w
> m3port_univ <- mean((portreturns - mean(portreturns))^3)
> m4port_univ <- mean((portreturns - mean(portreturns))^4)
> c(m3port, m3port_univ, m4port, m4port_univ)

[1] -1.559239e-06 -1.559239e-06  1.302823e-07  1.302823e-07
```

This will not be the case for the other estimators used in this vignette, explaining the use for better estimators of the multivariate coskewness and cokurtosis matrices in order to obtain a better univariate estimate of the portfolio skewness and kurtosis.

Similar to the difference in $1/n$ and $1/(n-1)$ for covariance estimation, the plug-in sample coskewness estimator can be bias-corrected by using a different factor in front:

$$\widehat{\phi}_{ijk}^{(\text{unb})} = \frac{n}{(n-1)(n-2)} \sum_{t=1}^{n} (r_{ti} - \overline{r}_i)(r_{tj} - \overline{r}_j)(r_{tk} - \overline{r}_k). \tag{4}$$

The option is easily selected in `M3.MM` by adding `unbiased = T` to the arguments. Below we provide a small simulation study demonstrating unbiasedness by estimating the third order central moment of a standard Exponential random variable. We do $200\,000$ replications of sample size $50$. The true value equals $2$.

```
> m3 <- M3.MM(edhec, unbiased = TRUE)
> set.seed(201706)
> x <- matrix(rexp(10000000), nrow = 50)
> xc <- x - matrix(colMeans(x), nrow = 50, ncol = 200000, byrow = TRUE)
> skews_plugin <- colMeans(xc^3)
> mean(skews_plugin)

[1] 1.886832

> mean(skews_plugin) * 50^2 / (49 * 48)

[1] 2.005561
```

## 2.2 Structured estimation

The basis of structured estimation is to assume a certain data-generating process for the returns. Under a chosen process, the coskewness and cokurtosis matrices will be constrained structurally, for example by zero elements or other relations between the different elements. First, we take a look at the strong assumption of independence, followed by a more realistic multi-factor model. Additionally, the constant-correlation model proposed in Martellini and Ziemann (2010) is implemented and some additional coskewness structuresare shown at the end.

### 2.2.1 Independence

Though the assumption of independence might be unrealistic, it nicely demonstrates the effect that choosing a structure or data-generating process has on the coskewness and cokurtosis matrices. Define the marginal variances $\sigma_i^2$, then for $i < jk < l$, under the assumption of independence,

$$\begin{aligned} \phi_{iij} &= 0, \\ \phi_{ijk} &= 0, \\ \psi_{iiil} &= 0, \\ \psi_{iikk} &= \sigma_i^2 \sigma_k^2, \\ \psi_{ijkk} &= 0, \\ \psi_{ijkl} &= 0, \end{aligned} \tag{5}$$

and similar for permuations of the indices. Hence, only the elements $\phi_{iii}, \psi_{iiii}$ and $\psi_{iikk}$ are non-zero. Estimating the portfolio third and fourth order central moments by inserting zeros at the appropriate places and using the sample estimators for the non-zero elements clearly gives different results than before,

```
> m3 <- M3.struct(edhec, "Indep")
> m4 <- M4.struct(edhec, "Indep")
> m3port <- PerformanceAnalytics:::portm3(w, m3)
> m4port <- PerformanceAnalytics:::portm4(w, m4)
> c(m3port_univ, m3port, m4port_univ, m4port)

[1] -1.559239e-06 -9.149033e-09  1.302823e-07  4.740322e-09
```

This will be the case for all estimators to come. Hence, we will not show the different estimates each time but summarize them at the end of this section.

An even stronger assumption to make is assuming that the marginal distributions are independent and identical, then $\phi_{iii} = \phi$, $\psi_{iiii} = \psi$ and $\sigma_i^2 = \sigma^2$ for all $i$. The common moments are

estimated by taking the mean over the sample values, i.e.

$$\begin{aligned}\widehat{\phi}_{iiii}^{(\text{IndepId})} &= \frac{1}{p}\sum_{j=1}^{p}\widehat{\phi}_{jjjj},\\ \widehat{\psi}_{iikk}^{(\text{IndepId})} &= \frac{1}{p}\sum_{j=1}^{p}(\widehat{\sigma}_j^2)^2.\end{aligned} \tag{6}$$

This estimator is available by the option `struct = "IndepId"` for both the coskewness and cokurtosis matrices.

```
> m3 <- M3.struct(edhec, "IndepId")
> m4 <- M4.struct(edhec, "IndepId")
```

As demonstrated before, the plug-in sample coskewness matrix can be bias-corrected by using an alternative standardization constant. For the targets. The option to estimate the marginal third order central moment using the unbiased sample estimators is available by specifying `unbiasedMarg = T` as an option in the function `M3.struct`.

### 2.2.2 Factor model

In economics, linear factor models are very popular, e.g. the three-factor Fama-French model for explaining asset returns (Fama and French (1988)). The data-generating process of a multi-factor model is defined as follows

$$\boldsymbol{R} = \boldsymbol{B}\boldsymbol{F} + \boldsymbol{\varepsilon}, \tag{7}$$

with $\boldsymbol{F}$ the random vector with factors, $\boldsymbol{B}$ the matrix with factor loadings and $\boldsymbol{\varepsilon}$ the idiosyncratic term. Ususally, the components of $\boldsymbol{\varepsilon}$ are considered independent and independent from the factors. It is well known that under these assumptions, the covariance matrix of the observed returns can be decomposed as

$$\boldsymbol{\Sigma}_{\boldsymbol{R}} = \boldsymbol{B}\boldsymbol{\Sigma}_{\boldsymbol{F}}\boldsymbol{B}' + \boldsymbol{\Delta}, \tag{8}$$

with $\boldsymbol{\Sigma}_{\boldsymbol{F}}$ the covariance matrix of the factors and $\boldsymbol{\Delta}$ the covariance matrix of the idiosyncratic term.

A similar decomposition can be made for the coskewness and cokurtosis matrices. Boudt et al. (2015) derives the following expressions

$$\begin{aligned}\boldsymbol{\Phi}_{\boldsymbol{R}} &= \boldsymbol{B}\boldsymbol{\Phi}_{\boldsymbol{F}}(\boldsymbol{B}' \otimes \boldsymbol{B}') + \boldsymbol{\Omega},\\ \boldsymbol{\Psi}_{\boldsymbol{R}} &= \boldsymbol{B}\boldsymbol{\Psi}_{\boldsymbol{F}}(\boldsymbol{B}' \otimes \boldsymbol{B}' \otimes \boldsymbol{B}' + \boldsymbol{\Gamma}),\end{aligned} \tag{9}$$

with $\boldsymbol{\Phi}_{\boldsymbol{F}}$ and $\boldsymbol{\Psi}_{\boldsymbol{F}}$ the coskewness and cokurtosis matrices of the factors and $\boldsymbol{\Omega}$ and $\boldsymbol{\Gamma}$ sparse residual matrices. For their exact definition we refer to Boudt et al. (2015).

Estimation of the factor coskewness and cokurtosis matrices is done by first estimating the factor loadings $\boldsymbol{B}$ and the residuals $\boldsymbol{\varepsilon}$ using Ordinary Least Squares (OLS). These estimates are then combined through Equation (9) with the factor coskewness and cokurtosis matrices estimated by the plug-in sample estimators. The `M3.struct` and `M4.struct` functions do this by specifying `struct = "observedfactor"` and providing the factor observations in `f`. Below we use the equal-weighted portfolio as the observed factor.

```
> f <- rowMeans(edhec, na.rm = TRUE)
> m3 <- M3.struct(edhec, "observedfactor", f)
> m4 <- M4.struct(edhec, "observedfactor", f)
```

### 2.2.3 Constant correlation

The constant correlation model for the covariance matrix, proposed in Elton and Gruber (1973), assumes that the covariance matrix has off-diagonal terms given by

$$\sigma_{ij} = \rho\sqrt{\sigma_{ii}\sigma_{jj}}, \tag{10}$$

and thus the correlation $\rho$ between all the assets is equal. Martellini and Ziemann (2010) propose extended correlation coefficients for third and fourth order dependence between assets. The constant correlation coskewness and cokurtosis matrices are then constructed by assuming these extended correlations are constant, similar to the constant correlation covariance matrix. These estimators are accessible as follows

```
> m3 <- M3.struct(edhec, "CC")
> m4 <- M4.struct(edhec, "CC")
```

### 2.2.4 Additional structured coskewness options

In Section 2.2.2, the coskewness matrix under a multi-factor model was shown. When considering the more simple model of a one-factor model with an idiosyncratic term following an elliptical distribution, Simaan (1993) derive a way to estimate the coskewness matrix without needing the factor observations. All the skewness and coskewness observed in the returns can be explained by a single underlying factor, since the idiosyncratic term does not contribute to the overall skewness and coskewness.

```
> m3 <- M3.struct(edhec, "latent1factor")
```

In general, a central-symmetric random variable (in particular elliptically distributed) has a coskewness matrix consisting of only zeros. This is the reason why the idiosyncratic term does not contribute to the coskewness matrix in the model of Simaan (1993). This extremely restricted coskewness matrix is obtained by

```
> m3 <- M3.struct(edhec, "CS")
```

## 2.3 Shrinkage estimation

We introduced the sample estimator in Section 2.1 and several structured estimators in 2.2. The sample estimators have a small bias (or unbiased), but a larger estimation variance and the structured estimators usually have a smaller estimation variance, but possibly a large bias. The bias of the structured models depends on how close the true data-generating process is to the one you assume it to be. Hence, combining both types of estimators could reduce the total mean squared error of estimation (MSE).

This is exactly the idea of shrinkage estimation. It was introduced in the literature of covariance estimation by Ledoit and Wolf (2003). Martellini and Ziemann (2010) generalized their approach to estimation of coskewness and cokurtosis matrices. In Boudt et al. (2017), the authors improve the estimators for the shrinkage intensity proposed in Martellini and Ziemann (2010) for the coskewness matrix and extend the methodology to a multi-target shrinkage setting, as explained later in Section 2.3.2.

Let $\widehat{\mathbf{\Phi}}$ be the plug-in sample estimator and $\widehat{\boldsymbol{T}}$ some structured coskewness estimator. The true coskewness matrix is denoted by $\mathbf{\Phi}$. The shrinkage estimator combining the two is a convex combination of both estimators, i.e.

$$\widehat{\mathbf{\Phi}}^{(\mathrm{ST})}(\lambda) = (1-\lambda)\widehat{\mathbf{\Phi}} + \lambda\widehat{\boldsymbol{T}}, \tag{11}$$

with $\lambda \in [0,1]$. The value of $\lambda$ is optimized using some criterion. In the literature, minimizing the MSE is popular, this results in

$$\lambda^{\star} = \arg\min_{\lambda\in[0,1]} \mathbb{E}\left[\|\widehat{\mathbf{\Phi}}^{(\mathrm{ST})}(\lambda) - \mathbf{\Phi}\|^2\right]. \tag{12}$$

The MSE loss function can be estimated consistently, and even in some cases unbiasedly, as shown in Boudt et al. (2017).

### 2.3.1 Single-target shrinkage estimation

In shrinkage estimation, the main choice to make is which structured estimator to use. The functions `M3.shrink` and `M4.shrink` incorporate all the structured estiamtors given in Section 2.2. Besides the estimated coskewness and cokurtois matrices, the functions return the estimated shrinkage intensity and the estimated coefficients of the MSE loss function determining the shrinkage intensity. For more information about these coefficients we refer to Ledoit and Wolf (2003) and Boudt et al. (2017).

The different coskewness and cokurtosis single-target shrinkage estimators can be accessed by

```
> # target "Indep"
> m3 <- M3.shrink(edhec, 1)$M3sh
> m4 <- M4.shrink(edhec, 1)$M4sh
```

```
> # target "IndepId"
> m3 <- M3.shrink(edhec, 2)$M3sh
> m4 <- M4.shrink(edhec, 2)$M4sh
> # target "1-factor"
> m3 <- M3.shrink(edhec, 3, f)$M3sh
> m4 <- M4.shrink(edhec, 3, f)$M4sh
> # target "constant correlation"
> m3 <- M3.shrink(edhec, 4)$M3sh
> m4 <- M4.shrink(edhec, 4)$M4sh
> # target "latent 1-factor model of Simaan (1993)"
> m3 <- M3.shrink(edhec, 5)$M3sh
> # target "central-symmetry"
> m3 <- M3.shrink(edhec, 6)$M3sh
```

We remark that only shrinking towards an observed one-factor model is possible. If $\boldsymbol{T}_3$ is selected and multiple factor observations are provided, then a single-factor model is estimated for each of the factors and all the single-factor coskewness matrices are used as targets in a multi-target shrinkage estimator, as detailled in Section 2.3.2. The reason for accessing the targets estimators by a vector of `TRUE/FALSE` will become clear in Section 2.3.2.

### 2.3.2 Multi-target shrinkage estimation

Single-target shrinkage has the nice property that it reduces the MSE of estimation, but requires the user to choose a single structured estimator. In Boudt et al. (2017) the possibility of using multiple structured coskewness and cokurtosis estimators at once is explored. The multi-target shrinkage estimator of the coskewness matrix is then defined as

$$\widehat{\boldsymbol{\Phi}}^{(\mathrm{MT})}(\boldsymbol{\lambda}) = \left(1 - \sum_{k=1}^{K} \lambda_k\right) \widehat{\boldsymbol{\Phi}} + \sum_{k=1}^{K} \lambda_k \widehat{\boldsymbol{T}}_k, \tag{13}$$

with the vector of shrinkage intensities $\boldsymbol{\lambda}$ again minimizing the MSE between the estimator and the true coskewness matrix

$$\boldsymbol{\lambda}^{\star} = \arg\min_{\boldsymbol{\lambda}} \mathbb{E}\left[\|\widehat{\boldsymbol{\Phi}}^{(\mathrm{MT})}(\boldsymbol{\lambda}) - \boldsymbol{\Phi}\|^2\right]. \tag{14}$$

The optimization is done over all vectors $\boldsymbol{\lambda}$ satisfying $\lambda_k \geq 0$ and $\sum_{k=1}^{K} \lambda_k \leq 1$.

To use the multi-target estimators, simply select the structured estimators you want to include. For example, suppose we want to mix the structured estimators `"Indep"`, `"observedfactor"` and `"CC"`. This is done by

```
> m3 <- M3.shrink(edhec, c(1, 3, 4), f)$M3sh
> m4 <- M4.shrink(edhec, c(1, 3, 4), f)$M4sh
```

### 2.3.3 Unbiased estimatino of the MSE loss function

All the shrinkage estimators shown above use consistent, but biased estimators to estimate the MSE loss function determining the shrinkage intensities to use. Boudt et al. (2017) derive bias-corrected estimators for shrinkage estimation of the coskewness matrix with structured matrices `"Indep"`, `"IndepId"` and `"CS"`. The bias-corrected estimators improve the MSE compared to the biased estimators when the sample size is rather small. The option of using the unbiased estimators can be selected by the argument `unbiasedMarg = T` whenever estimating the coskewness matrix with `M3.shrink` and any combination of the three targets given above. For example, multi-target shrinkage estimation of the coskewness matrix using bias-corrected estimators for the MSE loss function with all the available structured coskwenss matrices is done by

```
> m3 <- M3.shrink(edhec, c(1, 2, 6), unbiasedMSE = T)$M3sh
```

## 2.4 Other estimation techniques

### 2.4.1 EWMA estimators

All the comoment estimators in Sections 2.1, 2.2 and 2.3 implicitely assume independent and identically distributed observations. However, especially when using portfolio returns, this is a strong assumption to make. To address the issue of time-variability, using a rolling window when estimating the moments is a frequently used approach.

An alternative approach popularized by JPMorgan (1996) for covariance estimation is using exponentially decaying weights. Assume $\boldsymbol{\varepsilon}$ to denote the excess portfolio returns, thus having a mean of zero. Then the covariance estimate at time $t$ is updated by the observed excess returns at time $t$ as

$$\widehat{\boldsymbol{\Sigma}}_t^{(\text{EWMA})} = (1-\lambda)\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t' + \lambda\widehat{\boldsymbol{\Sigma}}_{t-1}^{(\text{EWMA})}, \tag{15}$$

with $\lambda \in (0,1)$ the parameter governing how fast the covariance estimate adapts to new observations.

Straightforward extensions to coskewness and cokurtosis can be defined as

$$\begin{aligned}\widehat{\boldsymbol{\Phi}}_t^{(\text{EWMA})} &= (1-\lambda)\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t' \otimes \boldsymbol{\varepsilon}_t' + \lambda\widehat{\boldsymbol{\Phi}}_{t-1}^{(\text{EWMA})},\\ \widehat{\boldsymbol{\Psi}}_t^{(\text{EWMA})} &= (1-\lambda)\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t' \otimes \boldsymbol{\varepsilon}_t' \otimes \boldsymbol{\varepsilon}_t' + \lambda\widehat{\boldsymbol{\Psi}}_{t-1}^{(\text{EWMA})}\end{aligned} \tag{16}$$

The functions `M3.ewma` and `M4.ewma` give the option to estimate the EWMA-comoments when providing a dataset with excess returns $R$, or updating the last known comoment matrix with the new observations in $R$ using the option `last.M3 = T` in case of the coskewness matrix or `last.M4 = T` for the cokurtosis matrix. The functions take both the full comoment matrices or the vectors with unique elements as input and return the full matrices by default or the vectors with only the unique elements by adding the option `as.mat = F` as for all the previously discussed estimators.

```
> m3 <- M3.ewma(edhec, lambda = 0.97)
> m4 <- M4.ewma(edhec, lambda = 0.97)
```

In the above code we used the default value of $\lambda = 0.97$ which is recommended for monthly returns by JPMorgan (1996).

### 2.4.2 Moment Component Analysis

A different approach to estimating the comoment matrices can be found in Jondeau and Jurczenko (2017). Here the aim is to find a subspace of $\mathbb{R}^p$ that explains most of observed coskewness or cokurtosis, much like PCA searches for the subspace explaining most of the variance. To make this formal, the methods seeks the $k$ dimensional subspace spanned by the columns of $\boldsymbol{U}$ obtained by

$$\begin{aligned} \boldsymbol{U}_3 &= \arg\min_{\boldsymbol{U}} \|\widehat{\boldsymbol{\Phi}} - \boldsymbol{U}\boldsymbol{\Phi}_{sub}(\boldsymbol{U} \otimes \boldsymbol{U})'\|^2 \\ \boldsymbol{U}_4 &= \arg\min_{\boldsymbol{U}} \|\widehat{\boldsymbol{\Psi}} - \boldsymbol{U}\boldsymbol{\Psi}_{sub}(\boldsymbol{U} \otimes \boldsymbol{U} \otimes \boldsymbol{U})'\|^2, \end{aligned} \tag{17}$$

where the optimization is done over matrices $\boldsymbol{U} \in \mathbb{R}^{p\times k}$ such that $\boldsymbol{U}'\boldsymbol{U} = \boldsymbol{I}$. The matrices $\boldsymbol{\Phi}_{sub}$ and $\boldsymbol{\Psi}_{sub}$ are any valid coskewness and cokurtosis matrices and are part of the optimization procedure. The optimisation in (17) is equivalent to finding $\boldsymbol{U} \in \mathbb{R}^{p\times k}$ such that $\boldsymbol{U}'\boldsymbol{U} = \boldsymbol{I}$, maximizing

$$\begin{aligned} \boldsymbol{U}_3 &= \arg\max_{\boldsymbol{U}} \|\boldsymbol{U}'\widehat{\boldsymbol{\Phi}}(\boldsymbol{U} \otimes \boldsymbol{U})\|^2 \\ \boldsymbol{U}_4 &= \arg\max_{\boldsymbol{U}} \|\boldsymbol{U}'\widehat{\boldsymbol{\Psi}}(\boldsymbol{U} \otimes \boldsymbol{U} \otimes \boldsymbol{U})\|^2, \end{aligned} \tag{18}$$

By projecting the data onto the subspace spanned by the columns of $\boldsymbol{U}$, the data is represented by means of the $k$ principal coskewness / cokurtosis components. As with PCA, these components can be used to construct estimates of the comoment matrices, ignoring the information not embedded in the first $k$ components. This is done as follows

$$\begin{aligned} \widehat{\boldsymbol{\Phi}}_{mca} &= \boldsymbol{U}_3 \left(\boldsymbol{U}_3'\widehat{\boldsymbol{\Phi}}(\boldsymbol{U}_3 \otimes \boldsymbol{U}_3)\right)(\boldsymbol{U}_3 \otimes \boldsymbol{U}_3)', \\ \widehat{\boldsymbol{\Psi}}_{mca} &= \boldsymbol{U}_4 \left(\boldsymbol{U}_4'\widehat{\boldsymbol{\Psi}}(\boldsymbol{U}_4 \otimes \boldsymbol{U}_4 \otimes \boldsymbol{U}_4)\right)(\boldsymbol{U}_4 \otimes \boldsymbol{U}_4 \otimes \boldsymbol{U}_4)'. \end{aligned} \tag{19}$$

We remark that the MCA procedure works on any estimates $\widehat{\boldsymbol{\Phi}}$ and $\widehat{\boldsymbol{\Psi}}$ of the coskewness and cokurtosis matrices, not only using the sample estimator.

To obtain these estimates for $k = 3$ components, do

```
> m3 <- M3.MCA(edhec, k = 3)$M3mca
> m4 <- M4.MCA(edhec, k = 3)$M4mca
```

The MCA functions output the coskewness or cokurtosis matrix, a boolean indicating convergence, the number of iterations the optimization algorithm has gone through and the optimal $\boldsymbol{U}$.

### 2.4.3 Other approaches

The other models available in R we are aware of were implemented by Alexios Ghalanos. The GO-GARCH model of Van der Weide (2002) is available in the rmgarch package, for details see see Ghalanos (2015). The model allows for explicit extraction of fitted and forecasted coskewness and cokurtosis matrices. However, the model does not explicitely model the higher order moments.

Contrary to this, the IFACD model of Ghalanos et al. (2015) does model the coskewness and cokurtosis matrices explicitely. Though the code is only available for the univariate model through the package racd, as far as we are aware.

## 2.5 Summary of estimated portfolio moments

All the methods implemented in PerformanceAnalytics and presented above provide different estimates of the third and fourth order central moments of the portfolio. A summary is given below:

```
> summary_momentestimates
```

| | m3 (1e-07) | m4 (1e-07) |
|---|---|---|
| sample | -15.59239183 | 1.30282262 |
| struct - Indep | -0.09149033 | 0.04740322 |
| struct - IndepInd | -0.09149033 | 0.09623759 |
| struct - 1f | -15.47136337 | 1.46863267 |
| struct - CC | -27.11372899 | 18.44513281 |
| struct - Simaan | -8.13421227 | NA |
| shrink - Indep | -9.00709949 | 0.87548773 |
| shrink - IndepInd | -8.99438247 | 0.90207989 |
| shrink - 1f | -15.51875113 | 1.39375468 |
| shrink - CC | -19.54857944 | 1.54315415 |
| shrink - Simaan | -15.55460195 | NA |
| shrink - CS | -8.93641335 | NA |
| shrink - Indep/1f/CC | -12.74539554 | 1.30888169 |
| shrink - Indep/IndepId/CS | -9.26490284 | NA |
| EWMA | -8.85057569 | 3.33865253 |
| MCA - 1 factor | -8.73356469 | 0.32814264 |
| MCA - 3 factors | -15.91419781 | 1.30490213 |

As you can see from the results, the shrinkage estimators are usually a compromise between the values obtained by the sample estimator and the structured estimators.

# 3 Impact on risk measures

When using the risk measures `VaR` or `ES` in `PerformanceAnalytics` there is the option of providing the coskewness and cokurtosis matrices explicitely. This is especially useful when we are interested in the risk measure at portfolio level and the component contribution of each asset.

In the previous section we have demonstrated all available estimation techniques in `PerformanceAnal` for the coskewness and cokurtosis matrices. The estimated portfolio third and fourth order central moments differed largely depending on which approach was used. It is thus logical to look at the impact this possibly has on the portfolio VaR and portfolio ES.

Though the functions `M2.struct`, `M2.shrink` and `M2.ewma` provide the structured, shrinkage and EWMA estimators for the covariance matrix, we will use the sample covariance matrix `cov` in the comparison. The reason for this is to isolate the influence of the coskewness and cokurtosis matrix in this example.

The $95\%$ VaR estimate of the equal-weighted portfolio when using sample estimators is computed by

```
> w <- rep(1 / ncol(edhec), ncol(edhec))
> p <- 0.95
> m <- colMeans(edhec)
> sigma <- cov(edhec)
> m3 <- M3.MM(edhec)
> m4 <- M4.MM(edhec)
> VaR95 <- VaR(p = p, method = "modified", portfolio_method = "componen
+               weights = w, mu = m, sigma = sigma, m3 = m3, m4 = m4)
> VaR95

$MVaR
[1] 0.01492109

$contribution
 [1]  0.0017255278  0.0003006205  0.0018530585  0.0032292640  0.0005185
 [6]  0.0020168094  0.0010602063  0.0008747133  0.0018608595  0.0008423
[11]  0.0012362120 -0.0022797885  0.0016826883

$pct_contrib_MVaR
 [1]  0.11564358  0.02014736  0.12419059  0.21642285  0.03475055  0.135
 [7]  0.07105423  0.05862263  0.12471341  0.05645694  0.08285000 -0.152
[13]  0.11277251
```

and similarly for the $95\%$ ES

```
> ES95 <- ES(p = p, method = "modified", portfolio_method = "component",
+            weights = w, mu = m, sigma = sigma, m3 = m3, m4 = m4)
> ES95

$MES
[1] 0.03625114

$contribution
 [1]  0.0059762619 -0.0028081561  0.0052816918  0.0062522301  0.0009068
 [6]  0.0058974265  0.0034627291  0.0001339207  0.0036259163  0.0034602
[11]  0.0037776798 -0.0031537633  0.0034381082

$pct_contrib_MES
 [1]  0.164857231 -0.077463949  0.145697277  0.172469910  0.025016152
 [6]  0.162682530  0.095520570  0.003694249  0.100022144  0.095451544
[11]  0.104208592 -0.086997642  0.094841391
```

We gather the predicted $95\%$ VaR and ES as in Section 2.5 and summarize the results below,

```
> rm95
```

| | VaR | ES |
|---|---|---|
| sample | 0.01492109 | 0.03625114 |
| struct - Indep | 0.01346559 | 0.01346559 |
| struct - IndepId | 0.01338954 | 0.01338954 |
| struct - 1f | 0.01463852 | 0.03862755 |
| struct - CC | -0.00961889 | -0.00961889 |
| shrink - Indep | 0.01420925 | 0.02771503 |
| shrink - IndepId | 0.01416508 | 0.02810111 |
| shrink - 1f | 0.01476466 | 0.03761155 |
| shrink - CC | 0.01532341 | 0.04043059 |
| shrink - Indep/1f/CC | 0.01432918 | 0.03536423 |
| EWMA | 0.01033940 | 0.01033940 |
| MCA - 1 factor | 0.01500212 | 0.01728611 |
| MCA - 3 factors | 0.01498244 | 0.03633400 |

Clearly, the constant correlation estimators for the coskewness and cokurtosis matrices are no good in combination with the sample covariance estimator. When using `M2.struct(edhec, "CC")`, then the $95\%$ VaR and ES estimates under the constant correlation estimators becomes `0.00739374` and `0.00739374` respectively, still unable to account for the difference between VaR and ES.

The differences in estimated risk measures are noticable. We let the reader decide whether or not the differences are large or small.

# 4 Conclusion

This vignette demonstrated how to use all the coskewness and cokurtosis estimators available in `PerformanceAnalytics`. We also showed the influence a certain choice of estimator actually has on the resulting portfolio moments and as a consequence on Value-at-Risk and Expected Shortfall predictions.

Since the purpose of this vignette is to illustrate the implemented estimators, we do not make any claim about which estimator is 'better' or which values to trust more. This depends on the application, the type of data and other factors.

# References


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

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

K. Boudt, D. Cornilly, and T. Verdonck. A coskewness shrinkage approach for estimating the skewness of linear combinations of random variables. 2017.

E. J. Elton and M. J. Gruber. Estimating the dependence structure of share prices - implications for portfolio selection. *Journal of Finance*, 28(5):1203–1232, 1973.

E. F. Fama and K. R. French. Permanent and temporary components of stock prices. *Journal of political Economy*, 96(2):246–273, 1988.

A. Ghalanos. The rmgarch models: Background and properties. (version 1.3-0), 2015.

A. Ghalanos, E. Rossi, and G. Urga. Independent factor autoregressive conditional density model. *Econometric Reviews*, 34(5):594–616, 2015.

E. Jondeau and E. Jurczenko. Moment component analysis: an illustration with international stock markets. *Journal of Business and Economic Statistics*, 2017.

JPMorgan. Riskmetrics technical document. Technical report, JPMorgan, 1996.

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

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

Y. Simaan. Portfolio selection and asset pricing-three parameter framework. *Management Science*, 39(5):568–577, 1993.

R. Van der Weide. Go-garch: a multivariate generalized orthogonal garch model. *Journal of Applied Econometrics*, 17(5):549–564, 2002.

P. Zangari. A var methodology for portfolios that include options. Technical report, RiskMetrics Monitor, 1996.