Robust Portfolio Optimization with Multivariate Copulas: A Worst-Case CVaR Approach

Fernando Sabino da Silva[1], Flavio A. Ziegelmann[1,2]

[1]Department of Statistics - UFRGS, [2]Graduate Program in Economics - UFRGS

Introduction

- We address the uncertainty of the distribution of assets' returns in a CVaR minimization model by applying copula function and obtaining its robust counterpart
    - Specifically, we obtain a numerical estimate of the worst Copula CVaR, where the copulas belong to the class of Archimedean copula, and evaluate its out-of-sample performance with Gaussian copula CVaR model and a constant mixed portfolio.

Some Key Points

- Worst-case copula conditional Value-at-Risk optimization using linear programming.
- Risk management with higher dimension than those typically considered in the copula literature without pair-copula constructions.
- We select a diversified set of assets that can be useful during any market conditions.

Some Key Points

- Worst-case copula conditional Value-at-Risk optimization using linear programming.
- Risk management with higher dimension than those typically considered in the copula literature without pair-copula constructions.
- We select a diversified set of assets that can be useful during any market conditions.

Some Key Points

- Worst-case copula conditional Value-at-Risk optimization using linear programming.
- Risk management with higher dimension than those typically considered in the copula literature without pair-copula constructions.
- We select a diversified set of assets that can be useful during any market conditions.

Background

- A well-known problem of the Markowitz model is its sensitivity to the input parameters.
- This problem can be overcome by employing robust optimization and worst case techniques (Zhu and Fukushima, 2009) in which assumptions about the distributions of the random variables are relaxed.
  - Obtain the optimal portfolio solution by optimizing over a prescribed set and possible densities.
- Artzner *et al*. (1999) show that VaR has undesirable properties such as lack of sub-additivity, which implies that it is not a coherent measure.

Background

- A well-known problem of the Markowitz model is its sensitivity to the input parameters.
- This problem can be overcome by employing robust optimization and worst case techniques (Zhu and Fukushima, 2009) in which assumptions about the distributions of the random variables are relaxed.
  - Obtain the optimal portfolio solution by optimizing over a prescribed set and possible densities.
- Artzner *et al*. (1999) show that VaR has undesirable properties such as lack of sub-additivity, which implies that it is not a coherent measure.

Background

- Uryasev and Rockafellar (2001) show that an outright optimization with respect to CVaR is numerically difficult due to the dependence of the CVaR on the VaR.

- However, Rockafellar and Uryasev (2000) show that VaR and CVaR can be computed simultaneously by introducing auxiliary risk measures.
  - By combining this approach with scenario-based optimization problem, we can reduce this highly nonlinear problem to a LP problem.

- Kakouris and Rustem (2014) show how copula-based models can be introduced in the Worst Case CVaR framework.
  - This approach is motivated by an investor's desire to hedge against the worst possible scenario.

Background

- Uryasev and Rockafellar (2001) show that an outright optimization with respect to CVaR is numerically difficult due to the dependence of the CVaR on the VaR.

- However, Rockafellar and Uryasev (2000) show that VaR and CVaR can be computed simultaneously by introducing auxiliary risk measures.
  - By combining this approach with scenario-based optimization problem, we can reduce this highly nonlinear problem to a LP problem.

- Kakouris and Rustem (2014) show how copula-based models can be introduced in the Worst Case CVaR framework.
  - This approach is motivated by an investor's desire to hedge against the worst possible scenario.

Sklar's Theorem (1959)

Theorem 1

Let $X_1, ..., X_d$ be random variables with distribution functions $F_1, ..., F_d$, respectively. Then, there exists an d-copula C such that,

$$F(x_1, ..., x_d) = C(F_1(x_1), ..., F_d(x_d)), \tag{1}$$

for all $\mathbf{x} = (x_1, ..., x_d) \in \mathbb{R}^d$. If $F_1, ..., F_d$ are all continuous, then the function C is unique; otherwise C is determined only on $\operatorname{Im} F_1 \times ... \times \operatorname{Im} F_d$.

Why should we care about copulas?

- Assuming that $F(\cdot)$ and $C(\cdot)$ are differentiable, by (1) we have

$$\frac{\partial^d F(x_1,...,x_d)}{\partial x_1 ... \partial x_d} \equiv f(x_1,...,x_d) = \frac{\partial^d C(F_1(x_1),...,F_d(x_d))}{\partial x_1 ... \partial x_d} \quad (2)$$

$$= c(u_1,...,u_d) \prod_{i=1}^{d} f_i(x_i), \quad (3)$$

where $u_i = F_i(x_i)$, $i = 1, ..., d$.

multivariate = "1-dim" (marginals) + "joint" (copula).

Background

- $f(x, y)$: loss function depending upon a decision vector x that belongs to any arbitrarily chosen subset $X \in \mathbb{R}^n$ and a random vector $y \in \mathbb{R}^m$.
  - Decision vector x: vector of portfolios' weights.
  - X: set of feasible portfolios, subject to linear constraints.
  - y: market variables that can affect the loss function.

- The optimization of the *CVaR* is difficult due to the presence of the *VaR* in its definition (it requires the use of the nonlinear function max).

$$CVaR_\beta(x) = \frac{1}{1-\beta} \int_{f(x,y) \geq VaR_\beta(x)} f(x, y) p(y) dy \quad (4)$$

- Rockafellar and Uryasev (2000) define a simpler auxiliary function

$$F_\beta(x, \alpha) = \alpha + \frac{1}{1-\beta} \int_{f(x,y) \geq \alpha} (f(x, y) - \alpha) p(y) dy \quad (5)$$

that can be used instead of the *CVaR* directly, without the need to compute the *VaR* first.

Background

- $f(x, y)$: loss function depending upon a decision vector x that belongs to any arbitrarily chosen subset $X \in \mathbb{R}^n$ and a random vector $y \in \mathbb{R}^m$.
  - Decision vector x: vector of portfolios' weights.
  - X: set of feasible portfolios, subject to linear constraints.
  - y: market variables that can affect the loss function.

- The optimization of the *CVaR* is difficult due to the presence of the *VaR* in its definition (it requires the use of the nonlinear function max).

$$CVaR_\beta(x) = \frac{1}{1-\beta} \int_{f(x,y) \geq VaR_\beta(x)} f(x,y)p(y)dy \qquad (4)$$

- Rockafellar and Uryasev (2000) define a simpler auxiliary function

$$F_\beta(x, \alpha) = \alpha + \frac{1}{1-\beta} \int_{f(x,y) \geq \alpha} (f(x,y) - \alpha)\, p(y)dy \qquad (5)$$

that can be used instead of the *CVaR* directly, without the need to compute the *VaR* first.

Background

- Assume that the analytical representation for the density $p(y)$ is not available, but we can approximate $F_\beta(x, \alpha)$ by using J scenarios, y_j, $j = 1, \ldots, J$, which are sampled from the density function $p(y)$.

- We approximate

$$F_\beta(x,\alpha) = \alpha + \frac{1}{1-\beta}\int_{f(x,y)\geq\alpha} (f(x,y)-\alpha)\, p(y)dy$$
$$= \alpha + \frac{1}{1-\beta}\int_{y\in\mathbb{R}^m} (f(x,y)-\alpha)^+\, p(y)dy \qquad (6)$$

by its discretized version

$$\widetilde{F}_\beta(x,\alpha) = \alpha + \frac{1}{(1-\beta)J}\sum_{j=1}^{J} (f(x,y_j)-\alpha)^+ .$$

Background

- If the loss function $f(x, y)$ is linear with respect to x, then the optimization problem reduces to the following linear programming (LP) problem:

$$\underset{x \in \mathbb{R}^n, z \in \mathbb{R}^J, \alpha \in \mathbb{R}}{\text{minimize}} \; \alpha + \frac{1}{(1-\beta)\, J} \sum_{j=1}^{J} z_j \tag{7}$$

$$\text{subject to} \quad x \in X, \tag{8}$$

$$z_j \geq f(x, y_j) - \alpha, \tag{9}$$

$$z_j \geq 0, \; j = 1, \ldots, J, \tag{10}$$

where the feasible set X is convex, and z_j are indicator variables.

- Thereby, the optimization problem can be solved by using algorithms that are capable of solving efficiently very large-scale problems as, for example, simplex or interior point methods.

Worst Case CVaR

- Zhu and Fukushima (2009) consider the case where the probability distribution π is only known to belong to a certain set, say $\mathcal{P}$. They defined the worst-case CVaR (WCVaR) as the CVaR when the worst-case probability distribution in the set $\mathcal{P}$ occurs.

Definition 2

Given a confidence level β, $\beta \in (0,1)$, the worst-case CVaR for fixed $w \in \mathcal{W}$ with respect to the uncertainty set $\mathcal{P}$ is defined as

$$\begin{aligned} WCVaR_{\beta}(\mathbf{w}) &\equiv \sup_{\pi \in \mathcal{P}} CVaR_{\beta}(\mathbf{w}) \\ &= \sup_{\pi \in \mathcal{P}} \min_{\alpha \in \mathbb{R}} F_{\beta}(\mathbf{w}, \alpha) \end{aligned} \tag{11}$$

Worst Case CVaR

- They showed that the WCVaR is a coherent risk measure and it clearly satisfies $WCVaR_{\beta}(\mathbf{w}) \geq CVaR_{\beta}(\mathbf{w}) \geq VaR_{\beta}(\mathbf{w})$. Thus, $WCVaR_{\beta}(\mathbf{w})$ can be effectively used as a risk measure.

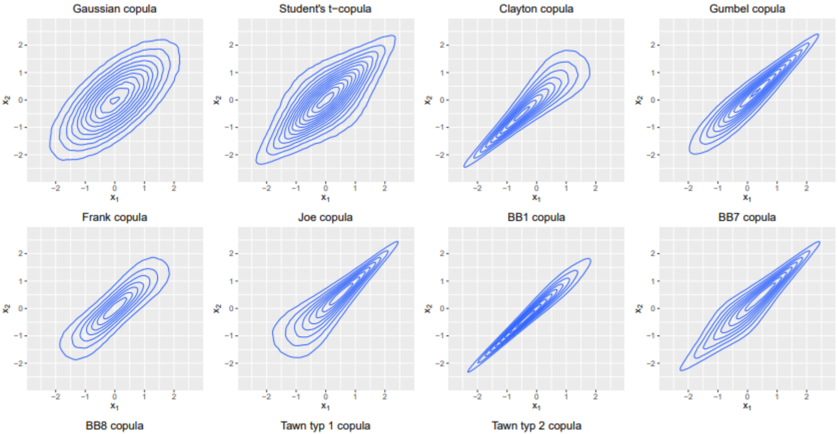
- Kakouris and Rustem (2014) extended their framework considering the case where the distribution of the returns belong to a set of copulas $C(\cdot) \in \mathcal{C}$ consisting of all mixtures of some possible copula distribution scenarios, i.e.,

$$c(\cdot) \in \mathcal{C}_M \equiv \left\{ \sum_{i=1}^{d} \pi_i c_i(\cdot) : \sum_{i=1}^{d} \pi_i = 1,\ \pi_i \geq 0,\ i = 1, \cdots, d \right\}. \qquad (12)$$

- Thus, the Worst Case Copula-CVaR is the mixture copula that produces the greatest CVaR, i.e., the worst performing copula combination in the set $\mathcal{C}_M$.

Archimedean mixture copula

Optimal linear combination of Clayton, Frank and Gumbel copulas $\Rightarrow$ to cover a large spectrum of possible dependence structures.



Worst Case Copula-CVaR

- By using the approach of Rockafellar and Uryasev (2000), the integral in

$$\begin{aligned} H_\beta(\mathbf{w},\alpha,\pi) &= \alpha + \frac{1}{1-\beta}\int_{\mathbf{u}\in[0,1]^n}\left(\widetilde{h}(\mathbf{w},\mathbf{u})-\alpha\right)^+\sum_{i=1}^{d}\pi_i c_i(\mathbf{u})\,d\mathbf{u} \\ &= \sum_{i=1}^{d}\pi_i H_\beta^i(\mathbf{w},\alpha), \end{aligned}$$

  is approximated by sample realizations from the copulas $C_i(\cdot)\in\mathcal{C}$, by using as inputs the filtered uniform margins.

- If the sampling generates a collection of values $\left(\mathbf{u}_i^{[1]},\mathbf{u}_i^{[2]},\ldots,\mathbf{u}_i^{[J]}\right)$, where $\mathbf{u}_i^{[j]}$ and S^i are the j-th sample drawn from copula $C_i(\cdot)$, we can approximate $H_\beta^i(\mathbf{w},\alpha)$ by

$$\widetilde{H}_\beta^i(\mathbf{w},\alpha) = \alpha + \frac{1}{(1-\beta)S^i}\sum_{j=1}^{S^i}\left(\widetilde{h}(\mathbf{w},\mathbf{u}_i^{[j]})-\alpha\right)^+, i=1,\ldots,d. \qquad (13)$$

Optimization Problem

- Assuming that the allowable set $\mathcal{W}$ is convex and the loss function $\widetilde{h}(\mathbf{w}, \mathbf{u})$ is linear with respect to $\mathbf{w}$, then optimization problem reduces to the following LP problem:

$$\underset{\mathbf{w}\in\mathbb{R}^n, v\in\mathbb{R}^m, \alpha\in\mathbb{R}}{\text{minimize}} \alpha + \frac{1}{(1-\beta)\, S^i} \sum_{i=1}^{S^i} v_i \quad (14)$$

$$\text{subject to} \quad \mathbf{w} \in \mathcal{W}, \quad (15)$$

$$v_i^{[j]} \geq \widetilde{h}(\mathbf{w}, \mathbf{u}_i^{[j]}) - \alpha, \quad (16)$$

$$v_i^{[j]} \geq 0,\ j = 1, \dots, J;\ i = 1, \dots, d, \quad (17)$$

where v_i, $i = 1, \dots, d$, are auxiliary indicator (dummy) variables and $m = \sum_{i=1}^{d} S^i$.

- By solving the LP problem, we find the optimal decision vector, $\mathbf{w}^*$, and at "one shot" the optimal VaR, α^*, and the optimal $CVaR$, $\widetilde{H}_\beta\,(\mathbf{w} = \mathbf{w}^*, \alpha = \alpha^*)$.

Data

- **Sources**: Adjusted closing prices for each stock.
- **Universe**: All shares that belongs to the S&P 500 market index.
- **Period**: July 2nd, 1990 to December 31st, 2015.
- **Total**: 1100 stocks over a 6426-day sample period.

Methodology

- Our optimization strategy adopts a sliding window of calibration of four years of daily data.
- We rebalance our portfolio every six months.
- We select, among all listed stocks in each formation period, a set of 50 stocks based on the ranked sum of absolute spreads (the 25 largest and the 25 smallest) between the normalized daily closing prices deviations of the S&P 500 index and all shares.
- We use day 1 to T, where T is the sliding window, to estimate the parameters of all models and determine the portfolio weights for day $T+1$ and then repeat the process including the latest observation and removing the oldest until we reach the end of the time series.

Methodology

- Our optimization strategy adopts a sliding window of calibration of four years of daily data.
- We rebalance our portfolio every six months.
- We select, among all listed stocks in each formation period, a set of 50 stocks based on the ranked sum of absolute spreads (the 25 largest and the 25 smallest) between the normalized daily closing prices deviations of the S&P 500 index and all shares.
- We use day 1 to T, where T is the sliding window, to estimate the parameters of all models and determine the portfolio weights for day $T+1$ and then repeat the process including the latest observation and removing the oldest until we reach the end of the time series.

Worst Case Copula-CVaR Portfolio Optimization

1. We fit an AR(1)-GARCH(1,1) model with skew t-distributed innovations to each univariate time series.

2. Using the estimated parametric model, we construct the standardized residuals vectors given, for each $i = 1, \ldots, 50$ and $t = 1, \ldots, L - T - 1$, where L is the data set sample period, by

$$\frac{\widehat{\varepsilon}_{i,t}}{\widehat{\sigma}_{i,t}}.$$

   The standardized residuals vectors are then converted to pseudo-uniform observations $z_{i,t} = \frac{n}{n+1} F_i \left(\widehat{\varepsilon}_{i,t}\right)$, where F_i is their empirical distribution function.

Worst Case Copula-CVaR Portfolio Optimization

3. Estimate the copula model, i.e., fits the multivariate Clayton-Frank-Gumbel (CFG) Mixture Copula to data that has been transformed to $[0,1]$ margins by

$$C^{CFG}(\Theta, \mathbf{u}) = \pi_1 C^C(\theta_1, \mathbf{u}) + \pi_2 C^F(\theta_2, \mathbf{u}) + (1 - \pi_1 - \pi_2) C^G(\theta_3, \mathbf{u}) \quad (18)$$

   where $\Theta = (\alpha, \beta, \delta)^\top$ are the Clayton, Frank and Gumbel copula parameters, respectively, and π_1, $\pi_2 \in [0,1]$.

   - Probability density function for multivariate Archimedean copula is computed as described in McNeil and Neshelova (2009).

4. Use the dependence structure determined by the estimated copula for generating J scenarios. To simulate data from the three Archimedean copulas we use the sampling algorithms provided in Melchiori (2006).

5. Compute t-quantiles for these Monte Carlo draws.

Worst Case Copula-CVaR Portfolio Optimization

6. Compute the standard deviation $\widehat{\sigma}_{i,t}$ using the estimated GARCH model.

7. Determine the simulated daily asset log-returns, i.e., determine the simulated daily log-returns as $r_{i,t}^{sim} = \widehat{\mu}_t + \widehat{\sigma}_{i,t} z_{i,t}$.

8. Finally, use the simulated data as inputs when optimizing portfolio weights by minimizing CVaR for a given confidence level and a given minimum expected return.

Worst Case Copula-CVaR Portfolio Optimization

- For each of the three copulas, we run 1000 return scenarios from the estimated multivariate CFG Mixture Copula model.
- The weights are recalibrated at a daily, weekly and monthly basis. We assume that the feasible set $\mathcal{W}$ attends the following linear constraints:
  - The sum of the weights should be 1, i.e., $\mathbf{w}^\top \mathbf{1} = 1$;
  - No short-selling: $\mathbf{w} \geq 0$;
  - The expected return should be bound below by an arbitrary value $\bar{r}$, i.e., $\mathbf{w}^\top \mathbb{E}\left(\mathbf{r}_{p,t+1}\right) \geq \bar{r}$, where $\bar{r}$ represents the target mean return;
- The confidence level β is set at $\beta = 0.95$.

Benchmarks

- Gaussian Copula Portfolio.
- Equally-weighted portfolio $w_i = 1/N$, i.e. $\mathbf{w} = \left(\frac{1}{N}, \ldots, \frac{1}{N}\right)^\top$ for any rebalancing date t.
- S&P 500 index as a proxy for market return.
- More benchmarks (to come): extended (robust) versions of the classic Markowitz model, shrinkage portfolios, VaR and CVaR optimal portfolios under multivariate normality, etc.

Performance Measures

- Portfolio mean excess return and standard deviation;
- Sharpe and Sortino Ratios;
- Maximum drawdown;
- Turnover.
  - The higher the turnover, the higher the transaction cost that the portfolio incurs at each rebalancing day.
- Break-even transaction cost: level of transaction costs that makes the investor indifferent between the dynamic and the static strategies.

Performance Measures

- Capital requirement loss function (CR)
  - Regulatory loss function to evaluate VaR forecasts

$$CR_t = \max\left[\left(\left(\frac{3+\delta}{60}\right)\sum_{i=0}^{59} VaR_{\beta,t-i}\right), VaR_{\beta,t}\right], \tag{19}$$

where δ is a multiplicative factor that depends on the number of violations of the VaR in the previous 250 trading days.

  - To mitigate data-snooping problems we apply a test for superior predictive ability (SPA) proposed by Hansen (2005) to determine which model significantly minimizes the expected loss function.

Empirical Results

Table 1: Excess returns of Worst Case Copula-CVaR (WCCVaR), Gaussian Copula-CVaR (GCCVaR) and Equal Weights portfolios without a minimum expected return constraint

	WCCVaR	**GCCVaR**	**1/N**	**S&P 500**
Daily Rebalancing				
Mean Return (%)	11.29	10.59	10.52	6.61
Standard Deviation (%)	15.35	15.05	23.86	19.01
Sharpe Ratio	0.70	0.67	0.42	0.34
Sortino Ratio	1.14	1.09	0.68	0.54
Turnover	0.6868	0.6425	0.0001	
Break-even (%)	0.0618	0.0622	275.09	
MDD1 (%)	-13.20	-12.89	-18.00	-13.20
MDD2 (%)	-38.94	-41.66	-71.84	-52.61
$VaR_{0.95}$	0.0091	0.0109	0.0209	0.0170
$CVaR_{0.95}$	0.0109	0.0135	0.0323	0.0283
$CR_{0.05}$(%)	13.00^{***}	15.63	29.71	23.91
Weekly Rebalancing				
Mean Return (%)	10.57	10.04	10.52	6.61
Standard Deviation (%)	15.48	15.21	23.86	19.01
Sharpe Ratio	0.65	0.63	0.42	0.34
Sortino Ratio	1.06	1.02	0.68	0.54
Turnover	0.1535	0.1564	0.0001	
Break-even (%)	0.2596	0.2427	275.09	
MDD1 (%)	-12.59	-10.39	-18.00	-13.20
MDD2 (%)	-39.78	-41.32	-71.84	-52.61
$VaR_{0.95}$	0.0091	0.0109	0.0209	0.0170
$CVaR_{0.95}$	0.0109	0.0135	0.0323	0.0283
$CR_{0.05}$ (%)	12.86^{***}	15.57	29.71	23.91

Empirical Results

Table 2: Excess returns of Worst Case Copula-CVaR (WCCVaR), Gaussian Copula-CVaR (GCCVaR) and Equal Weights portfolios without a daily minimum expected return constraint

	WCCVaR	GCCVaR	1/N	S&P 500
Monthly Rebalancing				
Mean Return (%)	9.22	9.42	10.52	6.61
Standard Deviation (%)	15.69	15.55	23.86	19.01
Sharpe Ratio	0.56	0.58	0.42	0.34
Sortino Ratio	0.91	0.94	0.68	0.54
Turnover	0.0447	0.0499	0.0001	
Break-even (%)	0.7829	0.7155	275.09	
MDD1 (%)	-13.21	-13.22	-18.00	-13.20
MDD2 (%)	-38.58	-46.13	-71.84	-52.61
$VaR_{0.95}$	0.0091	0.0109	0.0209	0.0170
$CVaR_{0.95}$	0.0109	0.0134	0.0323	0.0283
$CR_{0.05}$ (%)	12.22^{***}	14.67	29.71	23.91

Note: Out-of-sample performance statistics between July 1994 and December 2015 (5414 observations). The rows labeled MDD1 and MDD2 compute the largest drawdown in terms of maximum percentage drop between two consecutive days and between two days within a period of maximum six months, respectively. Returns, standard deviation, Sharpe ratio and Sortino ratio are annualized.

***, **, * significant at 1%, 5% and 10% levels, respectively.

Cumulative Excess Returns

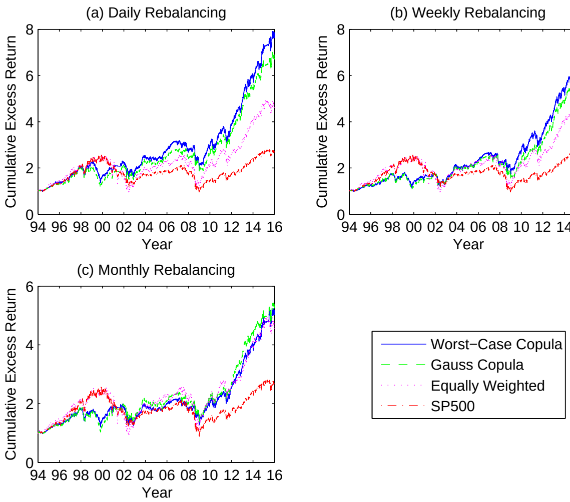


Figure 1: Cumulative excess returns of the portfolio strategies without daily target mean return

Conclusions

- By selecting a diversified set of assets over a long-term period we found that copula-based approaches offer better hedges against losses than the 1/N portfolio.
- The WCCVaR approach generates portfolios with better downside risk statistics for any rebalancing period and it is more profitable than the Gaussian Copula-CVaR for daily and weekly rebalancing.

Extensions

- Machine Learning and AI-based solutions.
  - Optimizer Search with Reinforcement Learning.
    - Stability issues.
    - Finding the portfolio with the highest downside risk-adjusted performance (Sortino ratio, Omega, Kappa).

Thanks! Any questions?