

EXPLORING WORST-CASE MIXTURE OF COPULAS FOR OPTIMAL MEAN-CVAR PORTFOLIO ALLOCATION

João Ramos Jungblut

Orientador: Prof. Dr. Fernando Augusto Boeira Sabino da Silva

Universidade Federal do Rio Grande do Sul
Faculdade de Ciências Econômicas
Departamento de Economia e Relações Internacionais

Setembro 2023

Conteúdo

- 1 Introdução
 - Objetivos
 - Motivações
- 2 Teoria
 - CVaR
 - Cópuas
 - Problema de Otimização
- 3 Metodologia
 - Dados
 - Processo de Otimização
 - Estratégia Empírica
- 4 Resultados
- 5 Considerações Finais
- 6 Referências

Introdução

- O velho ditado "*não coloque todos os ovos na mesma cesta*" é um aviso sobre o perigo de colocar todos os seus recursos em uma única fonte. Isso é especialmente verdadeiro no mercado financeiro, onde decisões são tomadas sob um ambiente de incertezas;
- **Então, quantas cestas você precisaria para manter seus ovos seguros?**

Introdução

- A resposta a essa pergunta foi fornecida por Harry Markowitz em seu trabalho seminal que estabeleceu Moderna Teoria de Portfólios;
- Ele demonstrou que os investidores devem se preocupar com dois fatores principais ao construir uma carteira de investimentos: **retorno** e **risco**;
- Um investidor buscaria, portanto, alocar seu capital de forma a maximizar o retorno e minimizar o risco;
- Isso pode ser alcançado através da diversificação, que é a distribuição do capital em ativos financeiros com características diferentes.

Objetivos

- Construir uma carteira de investimentos que proteja contra condições negativas extremas de mercado e produza retornos positivo;
- Foi empregada uma otimização de portfólio Média-CVaR;
- A dependência estrutural entre os ativos foi modelada usando uma mistura de cópulas (*Clayton-t-Gumbel*) através da distribuição univariada ajustada por um modelo AR(1)-GARCH(1,1).

Motivações

- Na fundação da MPT, o **desvio-padrão** (σ) é usado como medida de risco;
- O seu uso considera tanto as variações positivas no retorno, *Upside Risk*, quanto as variações negativas, *Downside Risk*, (BACON, 2008);
- O **CVaR** foi priorizado por ser uma medida de risco coerente: (i) monotonicidade, (ii) invariância de translação, (iii) homogeneidade positiva e (iv) sub-aditividade.

Motivações

- Na fundação da MPT, a distribuição dos retornos é considerada **Normal** com média μ e variância σ^2 , $\tilde{R} \sim N(\mu, \sigma^2)$;
- Os retornos, contudo, apresentam fatos estilizados diferentes: (i) a distribuição não Normal, (ii) quase não há autocorrelação entre os retornos e (iii) existe dependência positiva entre retornos absolutos (ou quadráticos), (TAYLOR, 2011);
- Clusters de volatilidade podem fornecer uma explicação generalizada dos fatos estilizados. Principalmente, em períodos de crise quando diferentes ativos tendem a se mover conjuntamente

Conditional Value at Risk - CVaR

- Enquanto o VaR pergunta: *"O quão ruim as coisas podem ficar?"* O CVaR questiona: *"Se estamos na cauda, qual é a perda esperada?"*;
- O CVaR é a perda esperada durante um intervalo T condicionada a perda ser maior que β percentil da distribuição das perdas;
- É a média de perdas do VaR até $-\infty$;
- O **Conditional Value at Risk (CVaR)** é uma medida de risco coerente, demonstrado por (PFLUG, 2000), com propriedades atrativas, por exemplo, convexidade, segundo (ROCKAFELLAR; URYASEV, 2000);

Definição do CVaR

Seja w um vetor de pesos e r um vetor de eventos aleatórios. Seja $f(w, r)$ a função de perda da carteira. Assume-se que o vetor aleatório r possui uma função de densidade de probabilidade $p(r)$. Dessa forma, definimos $CVaR_\beta$ como em Wuertz et al. (2010), da seguinte maneira:

$$CVaR_\beta = \frac{1}{1-\beta} \int_{f(w,r) \leq VaR_\beta(w)} f(w, r) p(r) dr. \quad (1)$$

É evidente que o **CVaR contém o VaR (que não é convexo) em sua definição**. Rockafellar e Uryasev (2000) apresenta uma **função auxiliar**, da seguinte forma:

$$F_\beta(w, \alpha) = \alpha + \frac{1}{1-\beta} \int (f(w, r) - \alpha) p(r)^+ dr, \quad (2)$$

na qual o **CVaR não depende do VaR**. Eles também demonstram que $F_\beta(w, \alpha)$ é **convexa em relação a α** e que $\min F = \min CVaR$.

CVaR Discreto

Assumindo que a **representação analítica de $p(r)$ não está disponível**, a equação (2) pode ser **aproximada seguindo $k = 1, \dots, K$ cenários** que podem representar **retornos históricos ou simulados r** usando a equação abaixo:

$$F_{\beta}^d(w, \alpha) = \alpha + \frac{1}{(1 - \beta)J} \sum_{j=1}^J (f(w, r_j) - \alpha)^+. \quad (3)$$

Função de Cópula

- O uso de cópulas é uma **forma de lidar com** $p(r)$;
- A **mais simples** para modelar a dependência de ativos é **usar uma correlação linear** e matriz de covariância;
- No entanto, como Pfaff (2012) afirma, isso está correto apenas se os ativos estiverem distribuídos conjuntamente de forma elíptica. **Fatos estilizados**;
- Funções de cópula atendem às necessidades acima.

Função de Cópula: Definição

Cópulas são funções de distribuição multivariadas cujas margens unidimensionais são uniformemente distribuídas em $[0, 1]$. É possível usar cópulas para substituir funções de distribuição de probabilidade devido ao teorema de Sklar e seu corolário:

Theorem (Teorema de Sklar)

Seja F uma função de distribuição n -dimensional com margens F_1, \dots, F_n . Então existe uma n -cópula C tal que para todo $x \in R^n$,

$$F(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n)). \quad (4)$$

Além disso, se F_1, \dots, F_n forem contínuas, então C é única.

Função de Cópula: Definição

Corollary

Seja f a função de densidade de probabilidade multivariada da distribuição de probabilidade F e f_1, \dots, f_n as funções de densidade de probabilidade univariadas das margens F_1, \dots, F_n . A função de densidade de cópula de uma n -cópula C é a função $c: U[0, 1]^n \mapsto [0, \infty)$ tal que

$$c(u_1, \dots, u_n) = \frac{\partial^n C(u_1, \dots, u_n)}{\partial u_1 \dots \partial u_n} = \frac{f(x_1, \dots, x_n)}{\prod_{i=1}^n f_i(x_i)}. \quad (5)$$

A definição nos permite **separar a modelagem das margens $F_i(x_i)$ da estrutura de dependência multivariada** representada por C . A função de densidade de probabilidade da cópula é então a razão da função de probabilidade conjunta pelo que seria sob independência.

Famílias de Cópulas

Foi considerado uma **combinação linear** (média ponderada) de 3 funções de cópula:

- Cópula Clayton para avaliar a dependência na cauda inferior.
- Cópula t que possui uma dependência simétrica nas caudas inferior e superior.
- Cópula Gumbel com dependência na cauda superior.

Uma grande diversidade de possíveis estruturas de dependência entre ativos é considerada no modelo, com o objetivo de capturar melhor a dependência entre os ativos individuais. Com base em Pfaff (2012), Kakouris e Rustem (2014) e Hu (2006),

Cópula-CVaR

Kakouris e Rustem (2014) estabeleceram o quadro para usar funções de Cópula e CVaR. Seja $w \in W$ um vetor de decisão, $u \in U[0, 1]^n$ um vetor aleatório que segue uma distribuição contínua com função de densidade de cópula $c(\cdot)$ e $F(r) = (F_1(r_1), \dots, F_n(r_n))$ um conjunto de distribuições marginais onde $u = F(r)$. A equação correspondente de cópula para a equação (3) é

$$G_{\beta}^d(w, \alpha) = \alpha + \frac{1}{(1 - \beta)J^i} \sum_{j=1}^J (f(w, u_j^i) - \alpha)^+, \quad i = 1, 2, \dots, I. \quad (6)$$

A equação (6) pode então ser avaliada usando simulações de Monte Carlo. Fazemos isso amostrando realizações das cópulas $C_i(\cdot)$ usando as margens uniformes determinadas e o seu tamanho.

Problema de Otimização CVaR com Cópula

Seguindo as suposições de convexidade e linearidade da função de perda $f(w, u)$ em relação a w , o problema de otimização pode ser modelado seguindo a abordagem de Rockafellar e Uryasev (2002) de programação linear do mean-CVaR:

$$\begin{aligned} \min_{w \in \mathbb{R}^n, z \in \mathbb{R}^J, \alpha \in \mathbb{R}} \quad & \alpha + \frac{1}{(1 - \beta)J} \sum_{i=1}^J z_i, \\ \text{s.a} \quad & z_i \geq f(w, u_j) - \alpha, i = 1, \dots, J \\ & z_i \geq 0, j = 1, \dots, d, \\ & w \in W, \\ & w^T \hat{\mu} \geq R, \\ & w^T \mathbf{1} = 1, \end{aligned} \tag{7}$$

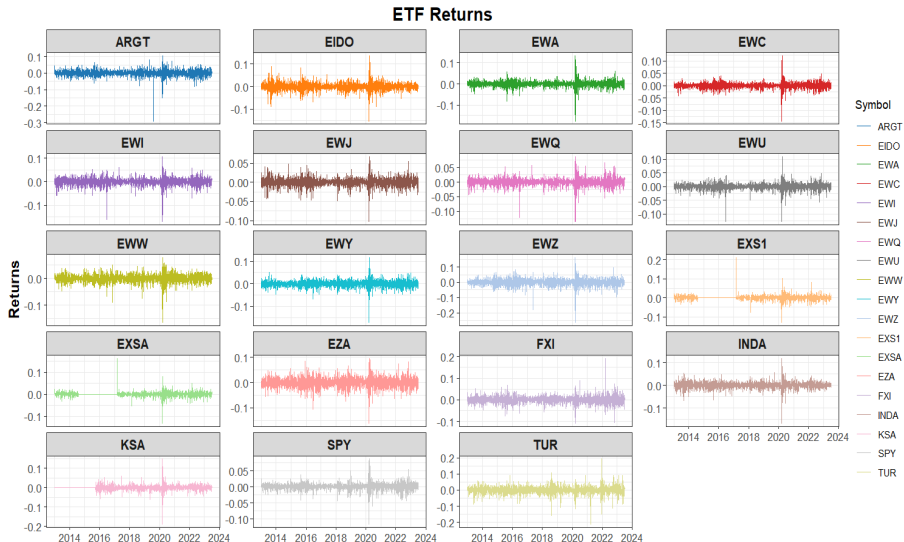
onde z_i são variáveis auxiliares fictícias e R é um **dado retorno diário esperado** para a carteira, **sem vendas a descoberto**.

Metodologia: Dados

Foram usados dados de 19 ETFs de índices de países, provenientes do *Historical Market Data - Stooq*, que abrangem o período de 02 de janeiro de 2013 até 30 de junho de 2023, totalizando 2692 dias de trading ($L = 2692$).

- ARGT;
- EWA;
- EWZ;
- EWC;
- FXI;
- EWQ;
- EXS1;
- INDA;
- EIDO;
- FWI;
- EWJ;
- EWW;
- KSA;
- EZA;
- EWY;
- TUR;
- EWU;
- SPY;
- EXSA.

Retorno Logarítmico



Estatísticas Descritivas

ETF	Min.	1st Q.	Median	Mean	3rd Q.	Max.	Skewness	Kurtosis
ARGT	-0.29	-0.0086	0.0006	0.0004	0.0101	0.10	-1.9764	27.6424
EIDO	-0.15	-0.0079	0.0000	-0.0000	0.0082	0.14	-0.5315	9.4203
EWA	-0.18	-0.0064	0.0003	0.0001	0.0071	0.13	-1.1350	22.8411
EWC	-0.14	-0.0052	0.0006	0.0002	0.0058	0.12	-1.1270	24.3630
EWI	-0.17	-0.0069	0.0007	0.0002	0.0083	0.11	-1.4798	15.4850
EWJ	-0.10	-0.0050	0.0004	0.0002	0.0060	0.07	-0.5206	6.3646
EWQ	-0.14	-0.0055	0.0006	0.0003	0.0068	0.09	-1.2070	15.0895
EWU	-0.13	-0.0051	0.0005	0.0001	0.0060	0.11	-1.3594	18.4232
EWV	-0.17	-0.0077	0.0000	0.0000	0.0086	0.08	-1.0387	9.7245
EWY	-0.17	-0.0078	0.0000	0.0001	0.0086	0.12	-0.7468	11.4362
...								

Estatísticas Descritivas

ETF	Min.	1st Q.	Median	Mean	3rd Q.	Max.	Skewness	Kurtosis
...								
EWZ	-0.26	-0.0116	0.0001	-0.0001	0.0123	0.16	-1.0798	13.1711
EXS1	-0.13	-0.0026	0.0000	0.0003	0.0039	0.21	1.8920	63.3013
EXSA	-0.13	-0.0019	0.0000	0.0003	0.0034	0.16	0.5871	56.3795
EZA	-0.16	-0.0109	0.0000	-0.0001	0.0112	0.10	-0.5740	5.4369
FXI	-0.11	-0.0088	0.0000	-0.0001	0.0089	0.19	0.4166	9.6157
INDA	-0.17	-0.0067	0.0003	0.0002	0.0080	0.12	-1.0409	16.2523
KSA	-0.19	-0.0027	0.0000	0.0002	0.0037	0.15	-0.8038	41.9807
SPY	-0.12	-0.0036	0.0005	0.0005	0.0055	0.09	-0.8170	14.4863
TUR	-0.21	-0.0108	0.0000	-0.0002	0.0120	0.19	-0.5261	8.8241

Processo de Otimização

Uma otimização em janelas rolantes de $T = 252, 504, 1260$, com base em Xi (2014), foi aplicada resultando em $L - T = 2440, 2188, 1432$ otimizações, respectivamente.

Aqui está um resumo do processo de otimização:

- Otimização 1: Os dados do dia 1 ao dia 252 foram usados para estimar o modelo de Copula de Pior Caso. Com base nesse modelo, os pesos da carteira Média-CVaR para o dia 253 foram determinados adequadamente;
- Otimização 2: Os dados do dia 2 ao dia 253 foram usados para estimar o modelo de Copula de Pior Caso, e os pesos da carteira Média-CVaR para o dia 254 foram determinados adequadamente;
- ...
- Otimização 2440 (para $T = 252$): Os dados do dia 2439 ao dia 2691 foram usados para estimar o modelo de Copula de Pior Caso, e os pesos da carteira Média-CVaR para o dia 2692 foram determinados adequadamente.

Estratégia Empírica

Os passos são apresentados abaixo e **repetidos para cada etapa de otimização** previamente demonstrada:

- 1 Ajustar um modelo $AR(1)$ - $GARCH(1,1)$ com distribuição t -assimétrica para cada ativo;
- 2 Usar o modelo paramétrico estimado para **construir um vetor de resíduos padronizados**;
- 3 Calcular a **distribuição pseudo-uniformes $[0,1]$** a partir dos **resíduos padronizados** usando a distribuição t -assimétrica do processo de erro $GARCH$;
- 4 **Estimar o modelo de mistura de cópulas Clayton- t -Gumbel multivariada** para os dados que foram transformados em distribuições marginais $[0,1]$ a partir da combinação linear de cópulas. **Nonlinear max. of log-likelihood**;
- 5 Utilizar a estrutura de dependência estimada pela mistura de cópulas para gerar $K = 1000$ **cenários de variáveis aleatórias para as variáveis pseudo-uniformemente distribuídas acumuladas**.

Estratégia Empírica

- 6 Calcular quantis t -assimétricos para essas simulações de Monte Carlo, $z_{j,t}$, para $j = 1, \dots, 8$ e $t = 1, \dots, (L - T)$, onde $Q(F(z)) = z$.
- 7 Para cada ativo j , determinar os cenários K de retornos log-diários simulados para fora da amostra para o dia seguinte previsto,

$$r_{j,t} = X_{j,t} + \epsilon_{j,t}, \quad (8)$$

onde $X_{j,t}$ é fornecido pelo modelo AR(1),

$$X_{j,t} = \epsilon_{j,t} + \phi_{j,i} X_{j,t-i} \quad (9)$$

e $\epsilon_{j,i}$ é o termo de erro seguindo um processo GARCH(1,1) dado por

$$\begin{aligned} \epsilon_{j,t} &= \sigma_{j,t} z_{j,t} \\ \sigma_{j,t}^2 &= \alpha_{j,0} + \alpha_{j,1} \epsilon_{j,t-1}^2 + \beta_{j,1} \sigma_{j,t-1}^2 \end{aligned} \quad (10)$$

Estratégia Empírica

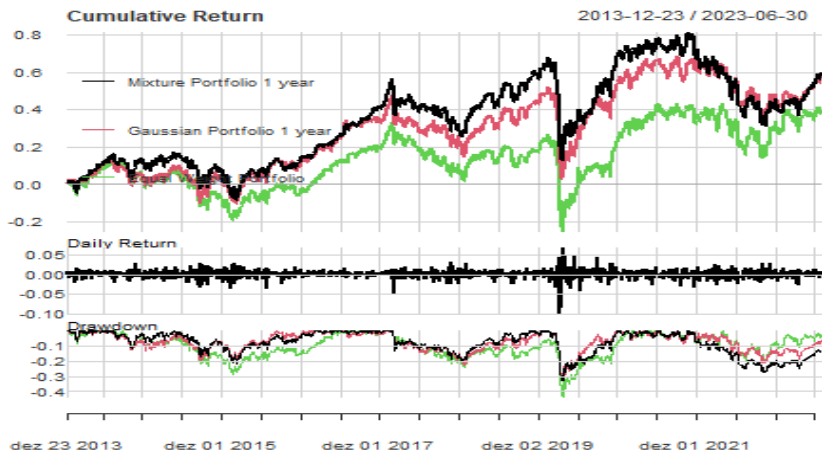
- 8 **Utilizar os dados simulados como entradas** ao otimizar a carteira. Isso é feito usando os trabalhos de Wuertz et al. (2010), nos quais o método de Rockafellar e Uryasev (2002) é aplicado para otimizar o CVaR com programação linear.
- Para avaliar o desempenho da otimização, rodamos a otimização para **meta de retorno diário** de 0.0003 ou 0.03%;
- Passos semelhantes são aplicados para otimizar uma **Carteira de cópula Gaussiana** para cada período. Também consideramos para fins de comparação uma **carteira igualmente ponderada (EWP)**;
- Após as otimizações, **calculamos os retornos simples da carteira fora da amostra**, que são usados para calcular **medidas de performance** como em Bacon (2008).

Resultados

Metrics	MCP 1Y	MCP 2Y	MCP 5Y	GCP 1Y	GCP 2Y	GCP 5Y	EWP
Annualized Return	0.0494	0.0790	0.1206	0.0490	0.0692	0.0648	0.0350
Annualized Std. Dev.	0.1498	0.1578	0.1730	0.1516	0.1630	0.1710	0.1798
Sharpe Ratio	0.3300	0.5005	0.6974	0.3232	0.4245	0.3791	0.1947
Sortino Ratio	0.0337	0.0484	0.0653	0.0329	0.0422	0.0383	0.0239
Omega Sharpe Ratio	0.0774	0.1112	0.1506	0.0755	0.0993	0.0896	0.0539
VaR (97.5%)	-0.0195	-0.0198	-0.0214	-0.0202	-0.0209	-0.0207	-0.0223
CVaR (97.5%)	-0.0294	-0.0307	-0.0338	-0.0308	-0.0331	-0.0349	-0.0348
Semi-Deviation	0.0071	0.0074	0.0080	0.0073	0.0077	0.0082	0.0085
Worst Drawdown	0.3268	0.3215	0.3057	0.3253	0.3372	0.3146	0.4380

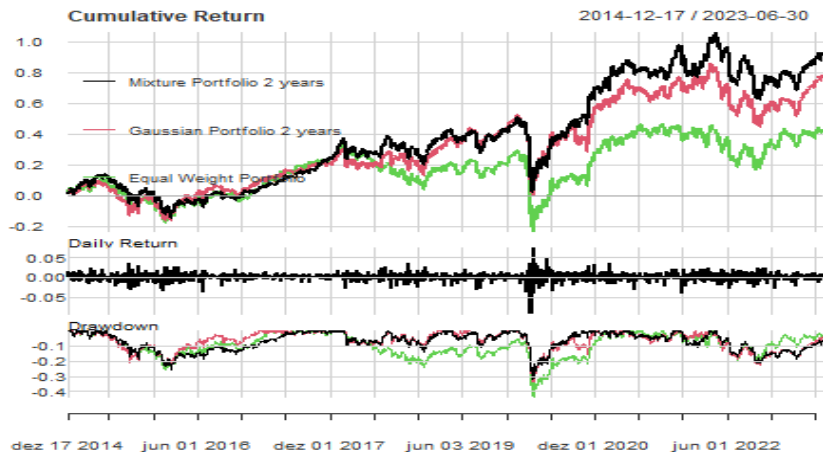
Performance

Performance chart



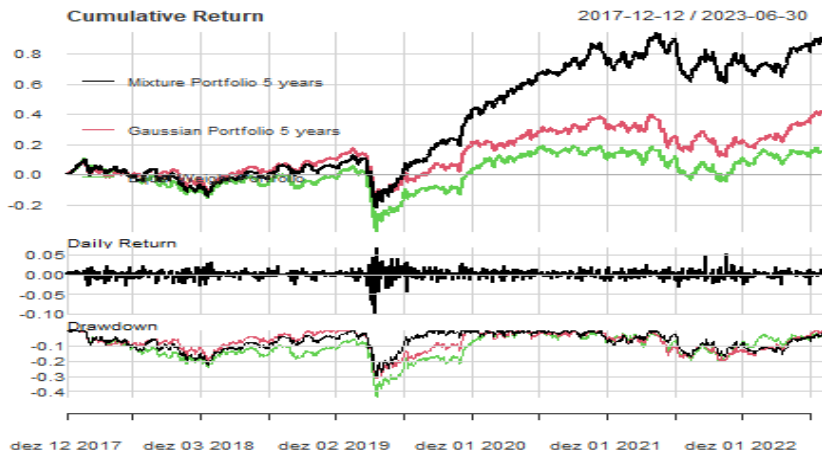
Performance

Performance chart



Performance

Performance chart



Considerações Finais

- A carteira gerada a partir dos retornos simulados com mistura de cópulas apresentou os melhores resultados quando avaliados por medidas de risco-retorno (índice Sharpe);
- Foram os portfólios que apresentaram os menores riscos e perdas financeiras no período em questão, ao mesmo tempo que obtiveram retorno iguais e até mesmo superiores;
- Em contrapartida, constatou-se que janelas de tempo maiores possibilitam estimativas mais precisas dos retornos e, consequentemente, melhores resultados;
- Isso ocorre porque janelas de tempo maiores possuem mais dados, que podem ser usados para identificar mais claramente padrões e tendências, permitindo que melhores decisões de investimento sejam tomadas;
- **Custos de transação** foram considerados, e foram um fator significativo a afetar a performance dos portfólios.

Trabalhos Futuros

- Limitar a frequência de rebalanceamentos;
- Adicionar restrição de cardinalidade, como em (RAMOS et al., 2023), que pode limitar o número de ativos no portfólio reduzindo diversificação excessiva;
- Diferentes alvos de retorno;
- Cópulas dinâmicas (AUSIN; LOPES, 2010) e Cópulas Vine (XI, 2014).

Repositório GitHub

Por transparência e propósito de reprodutibilidade, o código usado para os cálculos presente nesse trabalho podem ser encontrados no repositório do GitHub:

<https://github.com/JoaoJungblut/PortfolioRiskOpt>

O repositório contém os scripts e arquivos necessários para gerar os cálculos de otimização de portfólio discutidos nessa pesquisa. Também estão incluídos códigos adicionais para pré-processamento de dados, análises e visualização. Ao acessar o repositório no GitHub, leitores podem revisar, validar e executar o código independentemente para verificar os resultados ou adaptá-lo para suas necessidades específicas.

Por favor, note que o repositório pode estar sujeito a atualizações ou melhorias no decorrer do tempo. Se você encontrar qualquer problema, ou tiver alguma pergunta relacionada, sinta-se a vontade para contatar o dono.

References

AUSIN, M. Concepcion; LOPES, Hedibert F. Time-varying joint distribution through copulas. **Computational Statistics & Data Analysis**, v. 54, n. 11, p. 2383–2399, nov. 2010. Disponível em: <<https://ideas.repec.org/a/eee/csdana/v54y2010i11p2383-2399.html>>.

BACON, C.R. **Practical Portfolio Performance Measurement and Attribution**. [S.l.]: Wiley, 2008. (The Wiley Finance Series). ISBN 9780470778050. Disponível em: <<https://books.google.com.br/books?id=exITHIfwuyUC>>.

HU, Ling. Dependence patterns across financial markets: a mixed copula approach. **Applied Financial Economics**, v. 16, n. 10, p. 717–729, 2006. Disponível em: <<https://EconPapers.repec.org/RePEc:taf:apfie:v:16:y:2006:i:10:p:717-729>>.

KAKOURIS, Iakovos; RUSTEM, Berc. Robust Portfolio Optimization with Copulas. **European Journal of Operational Research**, v. 235, n. 1, p. 28–37, 2014. Disponível em: <<https://doi.org/10.1016/j.ejor.2013.12.022>>.

PFAFF, B. **Financial Risk Modelling and Portfolio Optimization with R**. [S.l.: s.n.], out. 2012. P. 1–356. DOI: 10.1002/9781118477144.

PFLUG, Georg Ch. Some Remarks on the Value-at-Risk and the Conditional Value-at-Risk. In: **Probabilistic Constrained Optimization: Methodology and Applications**. Edição: Stanislav P. Uryasev. Boston, MA: Springer US, 2000. P. 272–281. ISBN 978-1-4757-3150-7. DOI: 10.1007/978-1-4757-3150-7_15. Disponível em: <https://doi.org/10.1007/978-1-4757-3150-7_15>.

References

- RAMOS, Henrique Pinto et al. A comparison of risk measures for portfolio optimization with cardinality constraints. **Expert Systems with Applications**, Elsevier, v. 228, p. 120412, 2023.
- ROCKAFELLAR, R. Tyrrell; URYASEV, Stanislav. Conditional value-at-risk for general loss distributions. **Journal of Banking and Finance**, p. 1443–1471, 2002.
- _____. Optimization of Conditional Value-at-Risk. **Journal of Risk**, v. 2, p. 21–41, 2000.
- TAYLOR, Stephen J. **Asset price dynamics, volatility, and prediction**. [S.l.]: Princeton university press, 2011.
- WUERTZ, Diethelm et al. **Portfolio Optimization with R/Rmetrics**. Edição: Wuertz. [S.l.]: Rmetrics Association Finance Online, www.rmetrics.org, abr. 2010. R package version 2130.80.
- XI, Linda Mon. **Portfolio Optimization with PCC-GARCH-CVaR model**. 2014. Disponível em: <<https://bora.uib.no/handle/1956/8555>>.