

Universidade Federal do Rio Grande do Sul
Instituto de Matemática e Estatística
Departamento de Estatística



Anais

VII SEMANÍSTICA

VII Semana Acadêmica do Departamento de Estatística

da UFRGS

<http://www.ufrgs.br/semanistica>

Porto Alegre - 12 a 14 de setembro de 2016

Agrupamento de séries temporais financeiras aplicado a otimização de portfólios.

Cristiano Sulzbach^{1 2}

Marcio Valk³

Hudson Torrent⁴

Resumo: A construção de carteiras de ativos financeiros que tenham risco mínimo para um determinado nível de retorno esperado é objetivo na teoria de portfólios. Neste trabalho propomos uma estratégia para construção de portfólios baseada em agrupamento de séries temporais. A definição dos grupos é feita através do método hierárquico, juntamente com teste para homogeneidade de grupos baseado em U-estatísticas. A alocação de capital nos grupos é definida por um critério original que distribui os pesos de acordo com o desempenho de cada grupo, segundo alguma medida de desempenho. Os resultados mostram que o método de agrupamento, juntamente com o critério para alocação de capital, é superior à estratégia 1/N e ao método de mínima variância para a amostra e período considerados.

Palavras-chave: *Séries Temporais, Otimização de Portfólio, Agrupamento.*

1 Introdução

Em um cenário ideal um investidor deseja um alto grau de retorno com um baixo nível de risco. A teoria moderna de portfólios estipula que investidores construam carteiras otimizando o retorno esperado para um determinado nível de risco de mercado. De acordo com essa teoria, proposta por [3], busca-se construir uma *fronteira eficiente*, em que os portfólios tenham retorno esperado máximo para um dado nível de risco. Portfólios que proporcionam retorno máximo para um dado nível de risco são chamados de *portfólios eficientes*.

Recentemente, alguns trabalhos propõem técnicas que combinam métodos de agrupamento com métodos clássicos de alocação de ativos ([6, 1, 4]). A justificativa para a utilização de clusters é que na técnica de Markowitz os retornos esperados e a matriz de variâncias e covariâncias não podem ser estimadas com precisão a partir dos dados. Para evitar o viés de estimação, uma solução é adotar uma estratégia simples que não dependa da estimação de parâmetros, a qual consiste em alocar igualmente

¹UFRGS - Universidade Federal do Rio Grande do Sul. Email: cristiano.sulzbach@hotmail.com

²O autor agradece ao CNPQ pela bolsa de iniciação científica

³UFRGS - Universidade Federal do Rio Grande do Sul. Email: marciovalk@gmail.com

⁴UFRGS - Universidade Federal do Rio Grande do Sul. Email: hudsontorrent@gmail.com

o capital nos ativos. Essa estratégia chamada “*naive strategy*” ou 1/N, tem sido discutida na literatura e, embora não atenda aos pressupostos para alocação eficiente de portfólios, tem-se apresentado um *benchmark* difícil de ser superado.

Um método de classificação e agrupamento baseado em U-estatísticas é apresentado no trabalho de [5]. Esse método é baseado em um teste de homogeneidade para dois grupos de séries temporais e se mostra competitivo para encontrar a melhor divisão possível em dois grupos. No entanto, para o caso em que existem mais de dois grupos, esse método não foi estudado. Como todos os métodos de classificação e agrupamento, o método baseado em U-estatística depende da escolha da métrica. No contexto de volatilidade é mais adequado considerar medidas chamadas *Garch model based*.

Neste trabalho exploramos a construção de metodologias para otimização de portfólios de investimento baseada na utilização dessas técnicas de agrupamentos em séries temporais.

2 Medidas de Performance

Estamos interessados em medir a qualidade dos nossos portfólios, portanto é interessante conhecer índices que forneçam uma ideia da performance do portfólio. O índice mais conhecido é o *índice de Sharpe*. Este índice mede o excesso de retorno médio ganho em relação a uma taxa livre de risco (que no trabalho será a *Treasury bills*) por unidade de volatilidade. Numericamente temos que:

$$\widehat{SR}_k = \frac{\hat{\mu}_k - R_f}{\hat{\sigma}_k}$$

Em seguida, tem-se o *índice de Treynor*, que, analogamente ao índice de Sharpe, mensura o retorno ganho em excesso. A única diferença entre eles é que o índice de Treynor usa Beta para mensurar as incertezas. Também, pode-se dizer que quanto mais alto o valor deste índice, melhor a performance do portfólio. Em seguida, temos a fórmula:

$$\widehat{TR}_k = \frac{\hat{\mu}_k - R_f}{\beta}$$

Por último, tem-se *Certainty-Equivalent Return(CEQ)*. Este é um retorno garantido (livre de risco) que um determinado agente estaria disposto a aceitar; ao invés de correr risco para uma chance de maior de retorno. Em seguida, tem-se a expressão para calcular a taxa CEQ, na qual γ pode assumir diversos valores, neste trabalho será usado sempre com o valor 1.

$$CEQ_k = \hat{\mu}_k - \frac{\gamma \hat{\sigma}_k}{2}$$

2.1 Desempenho de alguns setores do mercado americano

Para compreender toda a tecnicidade inerente ao cálculo das medidas de desempenho, repetimos um estudo realizado por [2]. Os dados para esse exemplo são provenientes do *S&P500 sector data* e seguem o Padrão de Classificação Industrial Global. Fazem parte dele os 10 seguintes setores: 1, Consumo Facultativo; 2, Consumo Básico; 3, Energia; 4, Financeiro; 5, Saúde; 6, Industrial; 7, Tecnologia da Informação; 8, Materiais; 9 Telecomunicações e 10, Utilidades. Os dados foram obtidos semanalmente do período de 22 de setembro de 1989 a 28 de dezembro de 2012. Além destes, foi obtido o retorno das T-bills (usadas como taxa livre de risco). Os resultados podem ser vistos na seguinte tabela:

	SR	CEQ(%)	TR	Ann.Ret.(%)	Ann.risk(%)
Índice Obtido	0.0253	-1.5831	0.0095	0.4815	20.3208
Índice real	0.025	-1.59	0.010	0.51	20.50

Esse estudo foi fundamental para que pudéssemos verificar o cálculo dos índices e então ter maior confiabilidade na comparação de métodos de formação de portfólios.

3 Aplicação à dados do Ibovespa

Para essa aplicação consideramos dados diários do preço de fechamento dos ativos que compõem o índice Ibovespa, mas que tem histórico de comercialização no período de janeiro de 2010 a setembro de 2016. Dentre esses ativos, consideramos os mais líquidos, ou seja, desconsideramos os ativos em que não houve comercialização em mais de 5% dos dias. Assim, o banco de dados é composto por 63 ativos (ver Figura 1) com 1652 observações. Os dados foram obtidos da base de dados do google utilizando o pacote “*quantmod*” do R. Para a montagem do portfólio foram calculados os retornos diários e utilizadas 252 observações, o que é padrão para dados diários, representando o período de um ano. Os clusters são obtidos utilizando o método hierárquico (“*hclust*”) e a medida de distância utilizada é uma *Garch-Model-based*, apresentada [4], que mede a distância entre coeficientes garchs. Particularmente para um processo GARCH(1,1) a distância entre duas séries temporais é definida como

$$d_{garch} = \left[\frac{\alpha_{1,1}^2}{1 - \beta_{1,1}^2} + \frac{\alpha_{2,1}^2}{1 - \beta_{2,1}^2} - \frac{2\alpha_{1,1}\alpha_{2,1}}{1 - \beta_{1,1}\beta_{2,1}} \right]^{1/2}$$

em que α_i 's e β_i 's, $i=1,2$, são os coeficientes do modelo GARCH(1,1) da série i . Um dos problemas ao se utilizar o método hierárquico é que o número de grupos existente não é conhecido. Assim, é necessário que o usuário defina previamente o número de grupos ou use alguma técnica empírica para encontrar um número ótimo de grupos. É nesse ponto que utilizamos o teste de homogeneidade baseado em U-estatísticas proposto por [5]. Basicamente, seguimos as ramificações do método hierárquico e definimos

os grupos. Enquanto os grupos forem significativamente distintos, continuamos subdividindo. Se o teste atesta que os grupos são homogêneos, paramos de procurar mais grupos naquele ramo. Fazendo isso, encontramos 8 grupos significativos, os quais podem ser vistos em cada uma das caixas da Figura 1. Sendo homogêneo, é natural que dentro de cada grupo os ativos recebam a mesma carga de investimento, ou seja, os pesos dentro de cada grupo são alocados seguindo a estratégia 1/N. Para definir quanto vamos alocar em cada grupo, vamos adotar a estratégia de [6] e distribuí de acordo com o índice de Sharpe (ou Treynor) de cada grupo. Os grupos que apresentam maior índice de Sharpe recebem uma fatia maior do capital e os grupos com índices menores recebem uma penalização. Propomos o seguinte critério para obtenção dos pesos de cada grupo: Seja $SR_i, i = 1, \dots, ng$ o Sharpe Ratio de cada grupo, em que ng é o número de grupos (aqui $ng = 8$). Defina μ_{SR} como sendo a média dos valores absolutos de SR_i . Considere

$$SRC_i = \frac{SR_i}{cp * \mu_{SR}},$$

uma correção do Sharpe Ratio de cada grupo, em que $cp > 0$ é o *coeficiente de penalização* que permitirá valorizar grupos com maior Sharpe e penalizar os grupos com Índice baixo ou negativo. Agora defina os pesos como sendo

$$w_i = \frac{e^{SRC_i}}{\sum_{i=1}^{ng} (e^{SRC_i})}.$$

Nessa aplicação os grupos de 1 a 8 estão representados pelas cores, *preto, magenta, vermelho, verde, cyan, amarelo, azul e cinza*. Os índices de Sharpe para cada grupo, calculado para 252 observações foi $SR = \{-0.018690361, 0.009399394, 0.034921843, -0.040598077, 0.039889066, 0.044910269, 0.076644713, -0.015720382\}$. Usando a estratégia acima com $cp = 0.7$, os pesos para cada um dos 8 grupos ficaram definidos como $w = \{0.011473638, 0.035996137, 0.101724901, 0.004703584, 0.124518831, 0.152755559, 0.555879349, 0.012948001\}$. Com essa estratégia, obtemos um Sharpe ratio para o portfólio via agrupamento de 0.002416785, contra -0.01989971 do método 1/N e -0.001533676 do método de mínima variância.

4 Conclusão

Neste trabalho propusemos uma estratégia para construção de portfólios baseada em agrupamento de séries temporais através do método hierárquico, juntamente com teste para homogeneidade de grupos baseado em U-estatísticas. A alocação de capital nos grupos foi definida de uma maneira diferente da normamente adotada na literatura. Optamos por um critério que distribui os pesos de acordo com o desempenho de cada grupo, segundo alguma medida de desempenho. Os resultados mostram que o

Cluster dos 63 ativos mais líquidos do Ibovespa, 2010–2016

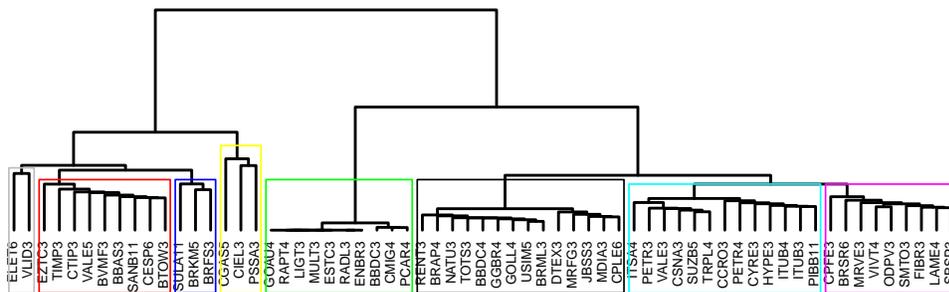


Figura 1: Agrupamento dos 63 ativos mais líquidos do Ibovespa no período de 01/2010 a 09/2016. A métrica utilizada é baseada na estimativa de parâmetros do modelo GARCH para a volatilidade dos retornos.

método de agrupamento, juntamente com o critério para alocação de capital, é superior à estratégia 1/N e ao método de mínima variância. Considerar rebalanceamento da carteira e custos de transação são pontos a serem abordados em pesquisa futura.

Referências

- [1] D’Urso, P., Cappelli, C., Lallo, D. D., and Massari, R. (2013). Clustering of financial time series. *Physica A: Statistical Mechanics and its Applications*, 392(9):2114 – 2129.
- [2] Koundouri, P., Kourogenis, N., Pittis, N., and Samartzis, P. (2016). Factor models of stock returns: Garch errors versus time-varying betas. *Journal of Forecasting*, 35(5):445–461. for.2387.
- [3] Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1):77–91.
- [4] Otranto, E. (2008). Clustering heteroskedastic time series by model-based procedures. *Computational Statistics and Data Analysis*, 52(10):4685 – 4698.
- [5] Valk, M. and Pinheiro, A. (2012). Time-series clustering via quasi U -statistics. *J. Time Ser. Anal.*, 33(4):608–619.
- [6] Zhang, J. and Maringer, D. (2010). *A Clustering Application in Portfolio Management*, pages 309–321. Springer Netherlands, Dordrecht.