Asset-allocation

Arthur Paolucci

02/10/2020

# Asset Allocation Problem

### 1. Problema

Este trabalho tem por objetivo estimar, dada uma janela de tempo, a alocação de um fundo de investimentos multimercado no Brasil, ou seja, determinar a exposição da sua carteira a diferentes fatores de risco. A motivação para tal, se dá pelo fato de que a Comissão de Valores Mobiliários (CVM) divulga a alocação dos fundos depois de um período de três meses, enquanto o mercado precisa saber como os fundos estão posicionados no momento da decisão de investimento.

### Revisão Bibliográfica

É amplamente conhecido que os retornos de uma carteira ou portfólio de mercado estão relacionados com diferentes fatores de risco. Observa-se na literatura a utilização de modelos de múltiplos fatores para medir e administrar a exposição a esses riscos. Esses fatores podem ser, por exemplo, indicadores macroeconômicos como taxa de juros, taxa de inflação, índices de preços de mercados internos ou externos, etc.

Existe um consenso de que a alocação em classes de ativos é a principal responsável pela variação nos retornos de uma carteira de investimentos. Isso é especialmente verdade se essa carteira é bem diversificada, ou seja, investe em vários fundos, que por sua vez investem em vários ativos.

Os fatores de risco utilizados neste trabalho foram obtidos a partir da “tropicalização” dos fatores utilizados por Sharpe na sua publicação “Asset Allocation: Management Style and Performance Measurement”. Em seu trabalho, o autor busca estimar o “estilo” (alocação em classes de ativos) de diferentes fundos a partir do seguinte modelo de fatores:

Neste modelo, cada fator representa o retorno de uma classe de ativos e interpreta-se os coeficientes como a sensibilidade da carteira aos fatores de risco escolhidos. É importante destacar que os coeficientes devem somar 1 (ou 100%), já que o retorno de cada ativo investido pelo fundo pode ser decomposto nos fatores de risco e o retorno do fundo deve ser explicado pela soma dos retornos de todos os ativos investidos por ele.

De fato, a soma dos termos em chaves na equação 1 pode ser interpretada como o retorno atribuído à escolha das classes de ativo a que o fundo está exposto, enquanto o resíduo pode ser explicado como o retorno atribuído à seleção de ativos. Apesar de somarem 100%, os coeficientes não devem ser confundidos com os pesos de uma carteira, já que os fatores representam diferentes riscos e não ativos.

**Sugestão AP:** **> Explicar a seleção (“tropicalização”) dos fatores**

**(Tabela tropicalização - sugestão: comparar escolhas do sharpe com as nossas)**

**> Explicar a escolha de janelas de 70 dias e dos períodos selecionados**

### Base de Dados

Todos os dados utilizados neste trabalho foram obtidos na base Economática, acessível para os alunos da FGV, ou calculados a partir da base. Na plataforma, é possível obter o histórico de preços da cota de fundos de investimento e ainda de diversos indicadores de mercado. Os dados obtidos na plataforma foram exportados para uma planilha e, em seguida, importados para o R. No programa em R, foram realizados procedimentos de limpeza e organização dos dados e foram calculados os retornos para cada variável utilizada no modelo pelo método dos retornos logarítmicos. A tabela abaixo descreve as variáveis presentes na base de dados utilizada:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Atributo | Descrição | Base/ Calculado | Fonte/ Cálculo | Papel na Regressão | Tipo | Domínio |  |
| Cota | Preço da cota do fundo “Verde…” em reais a.d. | Base | Economática | - | Numérico |  |  |

#### Importando os dados da Economatica

A FGV disponibiliza aos alunos o acesso à base de dados economatica. Para esse trabalho temos o interesse em obter as cotações do fundo analisado (JGP Strategy…) e dos índices que servirão como proxy para os fatores de risco.

Entretanto, para trabalhar esse conjunto de dados como cross-sectional devemos utilizar os retornos ao invés dos níveis de preço.

# Selecionar janelas de interesse  
index(dat) <- as.Date(index(dat))  
  
## Janela 1: janela de 70 dias até 2016-07-01  
end\_date <- index(dat[index(dat)>='2016-07-01',][1,])  
start\_date <- index(dat)[match(end\_date,index(dat))-69]  
jan1 <- window(dat, start = start\_date, end = end\_date)  
str(jan1)

## 'zoo' series from 2016-03-28 to 2016-07-01  
## Data: num [1:70, 1:22] 7.34 7.34 7.33 7.3 7.31 ...  
## - attr(\*, "dimnames")=List of 2  
## ..$ : NULL  
## ..$ : chr [1:22] "Cota" "Fechamento.ibov" "Fechamento.snp" "Fechamento.imab5" ...  
## Index: Date[1:70], format: "2016-03-28" "2016-03-29" "2016-03-30" "2016-03-31" "2016-04-01" ...

## Janela 2: janela de 70 dias até 2018-07-01  
end\_date <- index(dat[index(dat)>='2018-07-01',][1,])  
start\_date <- index(dat)[match(end\_date,index(dat))-69]  
jan2 <- window(dat, start = start\_date, end = end\_date)  
str(jan2)

## 'zoo' series from 2018-03-27 to 2018-07-02  
## Data: num [1:70, 1:22] 9.52 9.51 9.55 9.55 9.51 ...  
## - attr(\*, "dimnames")=List of 2  
## ..$ : NULL  
## ..$ : chr [1:22] "Cota" "Fechamento.ibov" "Fechamento.snp" "Fechamento.imab5" ...  
## Index: Date[1:70], format: "2018-03-27" "2018-03-28" "2018-03-29" "2018-03-30" "2018-04-02" ...

## Janela 3: janela de 70 dias até 2020-07-01  
end\_date <- index(dat[index(dat)>='2020-07-01',][1,])  
start\_date <- index(dat)[match(end\_date,index(dat))-69]  
jan3 <- window(dat, start = start\_date, end = end\_date)  
str(jan3)

## 'zoo' series from 2020-03-26 to 2020-07-01  
## Data: num [1:70, 1:22] 10.2 10 10.2 10 9.8 ...  
## - attr(\*, "dimnames")=List of 2  
## ..$ : NULL  
## ..$ : chr [1:22] "Cota" "Fechamento.ibov" "Fechamento.snp" "Fechamento.imab5" ...  
## Index: Date[1:70], format: "2020-03-26" "2020-03-27" "2020-03-30" "2020-03-31" "2020-04-01" ...

Os fatores “tropicalizados” utilizados no modelo são as taxas de retorno calculadas, dessa forma a hipótese de independência entre as observações da amostra aleatória do método de mínimos quadrados ordinários não é violada. No Item 4.3. a validade dessa hipótese é analisada em detalhe. As tabelas abaixo apresentam as estatísticas descritivas, para cada janela, de cada variável incluída no modelo de fatores:

**Estatisticas Janela 1:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Fator | Mean | Standard Deviation | Skewness | Kurtosis | Corr(Ret.verde) |
| Ret.verde | 7.2914810^{-4} | 0.0025 | 0.17683 | -0.11601 | 1 |
| Ret.ibov | 7.2250210^{-4} | 0.01675 | 0.08879 | -0.32057 | 0.30804 |
| Ret.snp | 4.6262110^{-4} | 0.00834 | -1.10564 | 4.33331 | 0.26019 |
| Ret.imab5 | 4.7373210^{-4} | 0.00153 | -0.33008 | 0.37232 | 0.51465 |
| Ret.imab5p | 0.00108 | 0.00638 | -0.13693 | 0.28799 | 0.46484 |
| Ret.dolar | -0.00237 | 0.01279 | -0.22905 | -0.3774 | -0.20607 |
| Ret.di17 | -4.6807710^{-4} | 0.00633 | -0.14933 | -0.17146 | -0.36593 |
| Ret.di21 | -0.00238 | 0.01497 | -1.30217 | 5.17812 | -0.06382 |

**Estatisticas Janela 2:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Fator | Mean | Standard Deviation | Skewness | Kurtosis | Corr(Ret.verde) |
| Ret.verde | -1.530810^{-4} | 0.00261 | 0.56725 | 0.74125 | 1 |
| Ret.ibov | -0.00222 | 0.01394 | -0.55792 | 0.33693 | 0.57737 |
| Ret.snp | 3.6164110^{-4} | 0.008 | -0.56696 | 0.6379 | 0.5139 |
| Ret.imab5 | -6.0018910^{-5} | 0.0025 | -0.80765 | 4.94592 | 0.44209 |
| Ret.imab5p | -9.4961310^{-4} | 0.00501 | 0.11619 | 0.56857 | 0.64233 |
| Ret.dolar | 0.0022 | 0.01026 | -1.92269 | 8.88002 | -0.21302 |
| Ret.di19 | 0.00214 | 0.01822 | 1.26345 | 3.67099 | -0.3581 |
| Ret.di23 | 0.00245 | 0.01381 | 0.55181 | 1.68525 | -0.44113 |

**Estatisticas Janela 3:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Fator | Mean | Standard Deviation | Skewness | Kurtosis | Corr(Ret.verde) |
| Ret.verde | 0.00196 | 0.00946 | 0.69417 | 2.96592 | 1 |
| Ret.ibov | 0.00357 | 0.02285 | -0.15094 | 0.23776 | 0.77174 |
| Ret.snp | 0.00329 | 0.02117 | -8.3419210^{-4} | 1.53933 | 0.81666 |
| Ret.imab5 | 7.5527110^{-4} | 0.00263 | -3.06574 | 17.47893 | 0.20285 |
| Ret.imab5p | 0.0012 | 0.01172 | -1.69967 | 10.64168 | 0.40333 |
| Ret.dolar | 8.1872610^{-4} | 0.0169 | -0.23034 | -0.84309 | -0.41428 |
| Ret.di21 | -0.00747 | 0.03095 | 0.76854 | 4.96674 | -0.08984 |
| Ret.di25 | -0.00443 | 0.03091 | 1.88079 | 11.25917 | -0.26715 |

### Metodologia (OLS)

Cochrane (p. 435, 2005) explica que os retornos podem ser observados como um conjunto de dados “cross section” ou dados em séries temporais, dependendo da pergunta que se queira responder. Neste caso, estuda-se o nível de retorno dado certos fatores de risco num instante t, dito isso é adequado o uso do método clássico de OLS para obter os coeficientes dos fatores do modelo.

A seguir, vamos escrever o modelo a estimar pelo método de OLS, calculamos a regressão linear para cada janela de tempo e vemos se são cumpridas as propriedades algébricas dos estimadores de OLS para que sejam aceitas como certas as hipóteses de Gauss-Markov.

##### 1. Relação Linear (nos coeficientes)

O modelo populacional é descrito como:

O modelo é linear nos parâmetros, vamos estimar os coeficientes do modelo utilizando os estimadores de OLS.

# Regressão linear OLS  
## Regressão da janela 1 (2016)  
regres1 <- lm(Ret.verde~Ret.ibov+Ret.snp+Ret.imab5+Ret.imab5p+  
 Ret.dolar+Ret.di17+Ret.di21, data=jan1)  
b1.hat<-coef(regres1) #coeficientes  
regres1.hat<-fitted(regres1) #retorno previsto  
u1.hat<-resid(regres1) #residuos  
summary(regres1)

##   
## Call:  
## lm(formula = Ret.verde ~ Ret.ibov + Ret.snp + Ret.imab5 + Ret.imab5p +   
## Ret.dolar + Ret.di17 + Ret.di21, data = jan1)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.006060 -0.001411 0.000269 0.001361 0.004376   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 9.17e-05 2.72e-04 0.34 0.73719   
## Ret.ibov -3.24e-03 2.03e-02 -0.16 0.87382   
## Ret.snp 1.06e-01 3.68e-02 2.87 0.00562 \*\*   
## Ret.imab5 1.19e+00 3.33e-01 3.58 0.00069 \*\*\*  
## Ret.imab5p -2.33e-02 7.52e-02 -0.31 0.75744   
## Ret.dolar -2.71e-02 2.13e-02 -1.27 0.20830   
## Ret.di17 5.06e-02 6.26e-02 0.81 0.42134   
## Ret.di21 -5.14e-03 1.69e-02 -0.30 0.76251   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.00202 on 62 degrees of freedom  
## Multiple R-squared: 0.415, Adjusted R-squared: 0.349   
## F-statistic: 6.28 on 7 and 62 DF, p-value: 1.39e-05

## Regressão da janela 2 (2018)  
regres2 <- lm(Ret.verde~Ret.ibov+Ret.snp+Ret.imab5+Ret.imab5p+  
 Ret.dolar+Ret.di19+Ret.di23, data=jan2)  
b2.hat<-coef(regres2) #coeficientes  
regres2.hat<-fitted(regres2) #retorno previsto  
u2.hat<-resid(regres2) #residuos  
summary(regres2)

##   
## Call:  
## lm(formula = Ret.verde ~ Ret.ibov + Ret.snp + Ret.imab5 + Ret.imab5p +   
## Ret.dolar + Ret.di19 + Ret.di23, data = jan2)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.002635 -0.000916 -0.000093 0.000549 0.006030   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 9.51e-05 1.98e-04 0.48 0.63260   
## Ret.ibov 5.58e-02 1.55e-02 3.60 0.00063 \*\*\*  
## Ret.snp 1.07e-01 2.49e-02 4.29 6.4e-05 \*\*\*  
## Ret.imab5 1.18e-01 1.96e-01 0.60 0.55022   
## Ret.imab5p 3.25e-01 7.37e-02 4.41 4.1e-05 \*\*\*  
## Ret.dolar 4.76e-02 2.34e-02 2.04 0.04580 \*   
## Ret.di19 4.26e-02 2.63e-02 1.62 0.11069   
## Ret.di23 -1.76e-02 2.03e-02 -0.86 0.39105   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.00154 on 62 degrees of freedom  
## Multiple R-squared: 0.688, Adjusted R-squared: 0.652   
## F-statistic: 19.5 on 7 and 62 DF, p-value: 1.61e-13

## Regressão da janela 3 (2020)  
regres3 <- lm(Ret.verde~Ret.ibov+Ret.snp+Ret.imab5+Ret.imab5p+  
 Ret.dolar+Ret.di21+Ret.di25, data=jan3)  
b3.hat<-coef(regres3) #coeficientes  
regres3.hat<-fitted(regres3) #retorno previsto  
u3.hat<-resid(regres3) #residuos  
summary(regres3)

##   
## Call:  
## lm(formula = Ret.verde ~ Ret.ibov + Ret.snp + Ret.imab5 + Ret.imab5p +   
## Ret.dolar + Ret.di21 + Ret.di25, data = jan3)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.008934 -0.002115 -0.000218 0.001651 0.014358   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.000711 0.000551 1.29 0.20   
## Ret.ibov 0.199721 0.036716 5.44 9.6e-07 \*\*\*  
## Ret.snp 0.248032 0.029575 8.39 8.5e-12 \*\*\*  
## Ret.imab5 -0.471024 0.459079 -1.03 0.31   
## Ret.imab5p 0.056476 0.106238 0.53 0.60   
## Ret.dolar 0.029659 0.037031 0.80 0.43   
## Ret.di21 0.016981 0.026720 0.64 0.53   
## Ret.di25 -0.026302 0.040191 -0.65 0.52   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.0042 on 62 degrees of freedom  
## Multiple R-squared: 0.823, Adjusted R-squared: 0.803   
## F-statistic: 41.2 on 7 and 62 DF, p-value: <2e-16

As hipóteses de Gauss-Markov para que o método de OLS seja não viesado são:

##### 2. Média Condicional Zero

O erro tem um valor esperado igual a zero, dados quaisquer valores das variáveis independentes. Em formula .

##### 3. Amostra Aleatória (iid)

Assumimos que nossa amostra é aleatória e independente. Segundo Cochrane, o preço de um ativo é dado por p\_t+1 =p\_t +ε\_t+1, ou seja, se σ2(ε )=constante os preços seguem um “random walk”. Essa definição é amplamente utilizada na teoria de finanças tradicionais e quer dizer que, se os mercados financeiros são eficientes, o preço de amanhã deve ser explicado pelo preço de hoje (que reflete toda a informação disponível sobre esse ativo) mais um choque imprevisível. Intuitivamente, se o preço hoje está muito menor que o preço de amanhã, os agentes vão antecipar esse aumento de preço e tentar comprar esse ativo, levando a um aumento de preço hoje e não amanhã. Sendo assim, os retornos de um ativo devem ser tão imprevisíveis como jogar uma moeda.

##### 4. Multicolinearidade não-perfeita

A hipóteses RLM.3 Colinearidade não perfeita diz: “Na amostra (e, portanto, na população) nenhuma das variáveis independentes é constante, e não há relações lineares exatas entre as variáveis independentes”

A continuação vemos, para cada janela de tempo (1, 2 e 3), que as correlações entre as variáveis independentes não são perfeitas, por tanto não há relações lineares entre elas.

$\underline{Testando\space as\space propiedades\space algebraicas:}$

* Propriedade 1: Os valores esperados dos resíduos da Janela 1, 2 e 3 são:
* *= 4.8792110^{-20}*
* *= 1.2151110^{-20}*
* *= -2.5550510^{-19}*

Podemos observar que os valores esperados dos resíduos para as três janelas de tempo estão muito perto de zero e por tanto a hipóteses de média condicional zero é cumprida.

* Propriedade 2:

**Janela 1:**

|  |  |
| --- | --- |
| **Fator j** |  |
| ret.ibov | 2.8591710^{-22} |
| ret.snp | 4.1426510^{-23} |
| ret.imab5 | 5.851110^{-23} |
| ret.imab5p | 9.4614810^{-22} |
| ret.dolar | -5.2705810^{-22} |
| ret.di17 | -1.4263210^{-21} |
| ret.di21 | -2.8112110^{-22} |

**Janela 2:**

|  |  |
| --- | --- |
| **Fator j** |  |
| ret.ibov | 2.6646610^{-22} |
| ret.snp | -2.7463310^{-22} |
| ret.imab5 | 1.5059910^{-22} |
| ret.imab5p | 2.8380710^{-22} |
| ret.dolar | 1.9500210^{-22} |
| ret.di17 | -3.9508710^{-22} |
| ret.di21 | -2.6315110^{-22} |

**Janela 3:**

|  |  |
| --- | --- |
| **Fator j** |  |
| ret.ibov | -7.0422610^{-21} |
| ret.snp | -1.6538710^{-21} |
| ret.imab5 | -8.0805710^{-22} |
| ret.imab5p | -3.0590210^{-21} |
| ret.dolar | 5.4870110^{-21} |
| ret.di17 | 8.4579210^{-21} |
| ret.di21 | 9.9421610^{-21} |

Podemos observar que todos os valores estão perto do zero, pelo qual assumimos que as hipóteses são verdadeiras.

* Propriedade 3: O ponto esta sempre sobre a reta de regressão de OLS

|  |  |  |  |
| --- | --- | --- | --- |
| **Janela 1** | **Janela 2** | **Janela 3** |  |
| $Fitted\space Value =$ | 7.2914810^{-4} | -1.530810^{-4} | 0.00196 |
|  | 7.2914810^{-4} | -1.530810^{-4} | 0.00196 |

Também podemos observar que as correlações entre as variáveis independentes não são perfeitas:

**Matriz de correlacoes entre os fatores Janela 1:**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Fator | Ret.ibov | Ret.snp | Ret.imab5 | Ret.imab5p | Ret.dolar | Ret.di17 | Ret.di21 |
| Ret.ibov | 1 | 0.50746 | 0.24343 | 0.44399 | -0.33297 | -0.37188 | -0.07136 |
| Ret.snp | 0.50746 | 1 | -0.16683 | 0.01784 | -0.30857 | 0.01393 | 0.08797 |
| Ret.imab5 | 0.24343 | -0.16683 | 1 | 0.80795 | 0.04429 | -0.75792 | -0.09329 |
| Ret.imab5p | 0.44399 | 0.01784 | 0.80795 | 1 | -0.15847 | -0.68497 | -0.19831 |
| Ret.dolar | -0.33297 | -0.30857 | 0.04429 | -0.15847 | 1 | -0.03542 | 0.11565 |
| Ret.di17 | -0.37188 | 0.01393 | -0.75792 | -0.68497 | -0.03542 | 1 | 0.05042 |
| Ret.di21 | -0.07136 | 0.08797 | -0.09329 | -0.19831 | 0.11565 | 0.05042 | 1 |

**Matriz de correlacoes entre os fatores Janela 2:**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Fator | Ret.ibov | Ret.snp | Ret.imab5 | Ret.imab5p | Ret.dolar | Ret.di19 | Ret.di23 |
| Ret.ibov | 1 | 0.27338 | 0.38453 | 0.44793 | -0.23299 | -0.40104 | -0.30494 |
| Ret.snp | 0.27338 | 1 | 0.09631 | 0.18585 | -0.09519 | -0.02541 | -0.02289 |
| Ret.imab5 | 0.38453 | 0.09631 | 1 | 0.79508 | -0.56285 | -0.91074 | -0.66136 |
| Ret.imab5p | 0.44793 | 0.18585 | 0.79508 | 1 | -0.59193 | -0.7545 | -0.72291 |
| Ret.dolar | -0.23299 | -0.09519 | -0.56285 | -0.59193 | 1 | 0.5842 | 0.42331 |
| Ret.di19 | -0.40104 | -0.02541 | -0.91074 | -0.7545 | 0.5842 | 1 | 0.66667 |
| Ret.di23 | -0.30494 | -0.02289 | -0.66136 | -0.72291 | 0.42331 | 0.66667 | 1 |

**Matriz de correlacoes entre os fatores Janela 3:**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Fator | Ret.ibov | Ret.snp | Ret.imab5 | Ret.imab5p | Ret.dolar | Ret.di21 | Ret.di25 |
| Ret.ibov | 1 | 0.5552 | 0.41902 | 0.63003 | -0.56217 | -0.30168 | -0.45199 |
| Ret.snp | 0.5552 | 1 | 0.10655 | 0.21534 | -0.32727 | -0.01258 | -0.12843 |
| Ret.imab5 | 0.41902 | 0.10655 | 1 | 0.80462 | -0.27185 | -0.78478 | -0.86603 |
| Ret.imab5p | 0.63003 | 0.21534 | 0.80462 | 1 | -0.44815 | -0.62987 | -0.86414 |
| Ret.dolar | -0.56217 | -0.32727 | -0.27185 | -0.44815 | 1 | 0.14479 | 0.31295 |
| Ret.di21 | -0.30168 | -0.01258 | -0.78478 | -0.62987 | 0.14479 | 1 | 0.69089 |
| Ret.di25 | -0.45199 | -0.12843 | -0.86603 | -0.86414 | 0.31295 | 0.69089 | 1 |

##### 5. Homocedasticidade

### Report Results

### Conclusão