

Análise de Carteiras usando o R - Parte 6

Bibliografia – BKM, cap. 8

Claudio Lucinda

FEA-RP/USP

Baixando Dados de Fundos

Dados de Fundos

- Nesta parte do curso, iremos utilizar o banco de dados de fundos de investimentos que eu utilizei no artigo com o João Nerasti:
- NERASTI, João Nascimento; LUCINDA, Claudio Ribeiro. Persistência de Desempenho em Fundos de Ações no Brasil (Persistence in Mutual Fund Performance in Brazil). Revista Brasileira de Finanças, v. 14, n. 2, p. 269, 2016.
- Esses dados estão na versão bruta no arquivo `Lucinda_Nerasti.xlsx`.
- O tratamento dos dados está no arquivo `Fund_Data.R`
- Puxamos informações sobre “Fatores de Risco” do site do NEFIN
- Os dados finais estão em `Fund_Data.RDS`.

Modelo de Índice Único

Carregando os dados e fazendo limpezas preliminares

```
load("Fund_Data.RDS")
```

```
# Dados diários de outubro de 2007 em diante
```

```
Fund>Returns_clean<-na.omit(Fund>Returns_clean)
```

```
fund_list_names<-colnames(Fund>Returns_clean)
```

```
Data_all<-merge(Fund>Returns_clean,Risk_Factors,  
                join="inner")
```

```
Data_all$Rm<-Data_all$Rm_minus_Rf+Data_all$Risk_free
```

Exemplo – Fundo 280

- Essa é uma amostra de 60 fundos em um período entre 2007 a 2014.
- Vou pegar esse fundo aqui para olhar as coisas – completamente aleatório:

```
Fund_Info_clean[14,]
```

```
##           Date           Fundo           CNPJ Classifica
## 280 Fundo_280 Nest Acoes FIC FIA 8.912577e+12 Ações IBOV
##           Empresa|gestora Data.do|Início.da.Série
## 280 Nest Investimentos           2007-07-12
```

Correlação entre os retornos do fundo e do índice

- É de se esperar que os retornos do fundo e um índice amplo de ações sejam bastante correlacionados.
- Estou seguindo a definição aqui do NEFIN para o índice amplo de mercado:
- The Market Factor is the difference between the value-weighted daily return of the market portfolio (using all the eligible stocks as defined in Section 2)

Índice de Mercado

- A stock traded in BOVESPA is considered “eligible” for year t if it meets 3 criteria:
- The stock is the most traded stock of the firm (the one with the highest traded volume during last year);
- The stock was traded in more than 80% of the days in year $t-1$ with volume greater than R\$ 500.000,00 per day. In case the stock was listed in year $t-1$, the period considered goes from the listing day to the last day of the year;
- The stock was initially listed prior to December of year $t-1$.

Correlação entre retorno do fundo e retorno de mercado

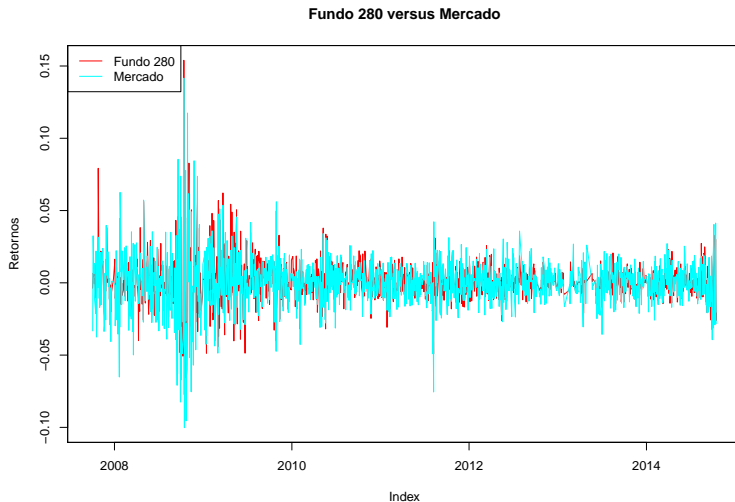
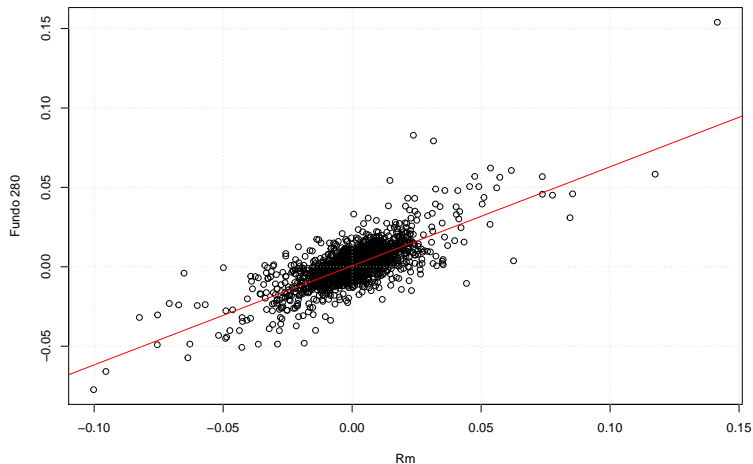


Diagrama de dispersão



Regressão

```
# Regressão  
model_280<-lm(Fundo_280~Rm, data=Data_fig)  
summary(model_280)  
confint(model_280)
```

Regressão – Outputs

```
##  
## Call:  
## lm(formula = Fundo_280 ~ Rm, data = Data_fig)  
##  
## Residuals:  
##           Min           1Q       Median           3Q           Max  
## -0.038763 -0.005151 -0.000461  0.004704  0.067391  
##  
## Coefficients:  
##              Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 0.0006187  0.0002299   2.691  0.0072 **  
## Rm          0.6236462  0.0136305  45.754 <2e-16 ***  
## ---  
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
##  
## Residual standard error: 0.00918 on 1593 degrees of freedom  
## Multiple R-squared:  0.5679, Adjusted R-squared:  0.5676  
## F-statistic: 2093 on 1 and 1593 DF, p-value: < 2.2e-16
```

Calculando a matriz Variância-Covariância pelo modelo de índice único

- Desse jeito, a gente consegue obter os σ_e^2 e os β de uma vez só:

```
betas<-ldply(as.list(fund_list_names),  
             function(.x) coef(lm(get(.x)~Rm,  
                                  data=Data_all))[2])  
sigmas2<-ldply(as.list(fund_list_names),  
               function(.x) sigma(lm(get(.x)~Rm,  
                                      data=Data_all))^2)  
sigma2_Rm<-as.matrix(var(Data_all$Rm))  
betas_vec<-as.matrix(betas)  
sigmas2_vec<-as.matrix(sigmas2)
```

Calculando a matriz VC

```
cov.mkt<-sigma2_Rm[1,1]*(betas_vec%*%t(betas_vec))
D.matrix<-matrix(0,nrow=nrow(cov.mkt),ncol=ncol(cov.mkt))

for (i in 1:ncol(cov.mkt)) {
  D.matrix[i,i]<-sigmas2_vec[i,1]
}

cov.si<-cov.mkt+D.matrix
cov.hist<-cov(Fund>Returns_clean)
```

As duas matrizes

```
cov.si[1:4,1:4]
```

```
##           [,1]      [,2]      [,3]      [,4]
## [1,] 1.488538e-04 7.891651e-05 7.477208e-05 7.469266e-05
## [2,] 7.891651e-05 3.108718e-04 2.483075e-04 2.480438e-04
## [3,] 7.477208e-05 2.483075e-04 2.799540e-04 2.350173e-04
## [4,] 7.469266e-05 2.480438e-04 2.350173e-04 2.859826e-04
```

```
cov.hist[1:4,1:4]
```

```
##           Fundo_267  Fundo_268  Fundo_269  Fundo_270
## Fundo_267 1.470532e-04 8.945463e-05 8.776943e-05 8.828073e-05
## Fundo_268 8.945463e-05 3.054762e-04 2.867113e-04 2.878323e-04
## Fundo_269 8.776943e-05 2.867113e-04 2.750941e-04 2.758630e-04
## Fundo_270 8.828073e-05 2.878323e-04 2.758630e-04 2.810143e-04
```

```
is.positive.definite(cov.si)
```

```
## [1] TRUE
```

```
is.positive.definite(cov.hist)
```

```
## [1] FALSE
```

Os betas são estáveis?

- Como disse anteriormente, eu uso os dados históricos para fazer análises prospectivas.
- Para isso, estou implicitamente assumindo que a minha amostra de dados gera estimativas representativas do que vai acontecer no futuro.
- Vamos fazer uma estimativa dos betas em uma janela móvel de 252 observações.

Beta com janela móvel de 252 dias

```
rr <- rollapply(Data_fig, width = 252,  
                FUN = function(z) coef(lm(Fundo_280 ~ Rm, c  
                by.column = FALSE, align = "right"))
```

Gráfico

Estabilidade do beta



Modelo de Fator Único Estatístico

Modelo de Fator Único Estatístico

- Outra forma de determinar o tal do fator único é por meio da análise estatística.
- O instrumento utilizado é a chamada “Análise de Componentes Principais”.

Análise de Componentes Principais – Definição

- Análise de Componentes Principais (ACP) ou Principal Component Analysis (PCA) é um procedimento matemático que utiliza uma transformação ortogonal para converter um conjunto de observações de variáveis possivelmente correlacionadas num conjunto de valores de variáveis linearmente não correlacionadas chamadas de componentes principais.
- Esta transformação é definida de forma que o primeiro componente principal tem a maior variância possível (ou seja, é responsável pelo máximo de variabilidade nos dados), e cada componente seguinte, por sua vez, tem a máxima variância sob a restrição de ser ortogonal a (i.e., não correlacionado com) os componentes anteriores.

Extraindo o Primeiro Componente Principal da matriz de retornos

```
stat.fact<-statistical.factor.model(Fund>Returns_clean,  
                                   k=1)  
stat.fact.data<-stat.fact$factor_realizations  
stat.fact.betas<-stat.fact$factor_loadings  
cov.si.stat.fact<-extractCovariance(stat.fact)
```

Matriz VC

```
cov.si.stat.fact[1:4,1:4]
```

```
##           [,1]      [,2]      [,3]      [,4]
## [1,] 1.471256e-04 9.029562e-05 8.587855e-05 8.593852e-05
## [2,] 9.029562e-05 3.054951e-04 2.613879e-04 2.615704e-04
## [3,] 8.587855e-05 2.613879e-04 2.751104e-04 2.487749e-04
## [4,] 8.593852e-05 2.615704e-04 2.487749e-04 2.810341e-04
```

```
cov.hist[1:4,1:4]
```

```
##           Fundo_267  Fundo_268  Fundo_269  Fundo_270
## Fundo_267 1.470532e-04 8.945463e-05 8.776943e-05 8.828073e-05
## Fundo_268 8.945463e-05 3.054762e-04 2.867113e-04 2.878323e-04
## Fundo_269 8.776943e-05 2.867113e-04 2.750941e-04 2.758630e-04
## Fundo_270 8.828073e-05 2.878323e-04 2.758630e-04 2.810143e-04
```

```
is.positive.definite(cov.si.stat.fact)
```

```
## [1] TRUE
```