

Lista 5

Finanças Quantitativas

Diogo Wolff Surdi

May 24, 2020

Questão 4.1

4.1.1

O gráfico 1 não pode ser um gráfico de resíduos de uma regressão linear, pois quase todos os resíduos são positivos, exceto um ponto com resíduo próximo a zero. Como a regressão busca minimizar a soma dos quadrados dos resíduos, e a função quadrática é crescente em relação ao módulo de um número, então seria possível diminuir esse valor alterando (para cima) o eixo de resíduos iguais a 0. De modo formal, uma das C.P.O.s da regressão linear é de que a soma dos resíduos é 0, o que claramente não ocorre nesse gráfico.

4.1.2

A entrada diagonal de índice i da matriz H é medida da distância da observação i em x em relação à observação média em x (no caso multivariado, vetores de médias). Com isso, não é possível inferir informações sobre a diagonal a partir do gráfico dos resíduos, pois não há informações sobre as variáveis regressoras.

4.1.3

A regressão é uma LAD. Isso pode ser visto devido ao fato de que a linha de regressão se encaixa bem nos pontos "comuns" dos dados, sem aparente influência dos outliers vistos na região superior à direita. O argmin da LAD é a mediana, que é mais robusta do que a média (argmin da regressão de mínimos quadrados), logo o efeito dos outliers é muito menor.

Questão 2

Seja $y = X\hat{\beta}$. Nosso objetivo é encontrar

$$b = \underset{\hat{\beta}_0}{\operatorname{argmin}} (y - X\hat{\beta})'(y - X\hat{\beta})$$

Sabemos que $(y - X\hat{\beta})'(y - X\hat{\beta}) = y'y - \hat{\beta}'X'y - y'X\hat{\beta} + \hat{\beta}'X'X\hat{\beta} = y'y - 2y'X\hat{\beta} + \hat{\beta}'X'X\hat{\beta}$.
Derivando em $\hat{\beta}$ encontramos que:

$$\frac{\partial(y'X\hat{\beta})}{\partial\hat{\beta}} = X'y; \quad \frac{\partial(\hat{\beta}'X'X\hat{\beta})}{\partial\hat{\beta}} = 2X'X\hat{\beta}$$

Igualando a derivada a 0, encontramos que $2X'y = 2X'X\hat{\beta}$. Logo

$$X'y = X'X\hat{\beta} \implies \hat{\beta} = (X'X)^{-1}X'y$$

Questão 3

3.1

As ações são:

ITUB4 B3SA3 PETR4 BBDC4 PETR3 ABEV3 BBAS3 ITSA4 MGLU3 JBSS3 LREN3
VALE3 SUZB3 GNDI3 WEGE3 NTCO3 RADL3 RAIL3 BBDC3 RENT3 VIVT4 LAME4
UGPA3 EQTL3 BRFS3 BTOW3 BBSE3 CCRO3 BRDT3 SBSP3 GGBR4 KLBN11 SULA11
HYPE3 HAPV3 VVAR3 ENGI11 TIMP3 EGIE3 CSAN3 BPAC11 SANB11 CMIG4 PCAR3
BRML3 COGN3 ELET3 TOTS3 YDUQ3 BRAP4 BRKM5 CRFB3 FLRY3 IRBR3 ELET6
QUAL3 TAEE11 MULT3 CSNA3 MRFG3 CPFE3 AZUL4 EMBR3 ENBR3 MRVE3 CYRE3
CIEL3 GOAU4 BEEF3 IGTA3 USIM5 ECOR3 CVCB3 HGTX3 GOLL4

3.2

Farei a análise com dois anos de retornos diários, omitindo os dias que não possuem valores ou dos ativos ou da taxa livre de risco.

```
library(quantmod)
library(dplyr)
```

```
dados <- read.csv('(...) / carteira.csv', stringsAsFactors = FALSE)
codigos <- c(dados$Codigo)
codigos <- paste(codigos, ".SA", sep='')
codigos <- codigos[-77]
codigos <- codigos[-76]
getSymbols(codigos, src='yahoo', from = "2018-05-21",
           to = "2020-05-21", env = e)

getSymbols('^BVSP', src='yahoo', from = "2018-05-21",
           to = "2020-05-21", env = e)
```

3.3

Para o código tive que remover a empresa SULA11, pois os dados estavam incompletos (o na.omit acaba cortando informação demais). Ademais, estou utilizando como taxa livre de

risco os retornos diários de títulos americanos de 3 meses (não consegui dados bons para títulos brasileiros), acessando o site

<https://www.federalreserve.gov/datadownload/Choose.aspx?rel=H15>. Baixando o arquivo do site acima, o código abaixo deve rodar e retornar um dataframe com as regressões de todas as ações.

```
dados2 <- read.csv( ' (...) /FRB_H15.csv ', stringsAsFactors = FALSE)
dados2 <- dados2[ -c(1,2,3,4,5), ]
dados2$Series.Description <- as.Date(dados2$Series.Description)

riskfree <- with(dados2, dados2[(Series.Description >= "2019-05-21"
  & Series.Description <= "2020-05-21"), c(1,3)])
riskfree <- riskfree[ - grep( 'ND', riskfree$Market.yield.on.U.S...3.month... ),
riskfree$Market.yield.on.U.S...3.month... <-
as.numeric(riskfree$Market.yield.on.U.S...3.month...)

riskfree$Market.yield.on.U.S...3.month... <-
(1+riskfree$Market.yield.on.U.S...3.month...)^ (1/252)-1
riskfree <- xts(riskfree[, -1], order.by=riskfree$Series.Description)

pframe <- do.call(merge, as.list(e))

todos <- data.frame(merge(pframe, riskfree))
todos2 <- todos[, grepl("Close", names(todos))]
todos2 <- cbind(todos2, todos$riskfree)

retornos <- data.frame(matrix(ncol = 77, nrow = 259))
colnames(retornos) <- c( 'BVSP', codigos, 'riskfree' )
retornos$riskfree <- todos2$`todos$riskfree`

for (i in 1:76) {
for (j in 1:259) {
retornos[j,i] = log(todos2[j+1,i]/todos2[j,i])
}
}

retornos <- retornos[, -34]
retornos <- na.omit(retornos)

coefs <- data.frame(matrix(nrow = 74, ncol=3))
colnames(coefs) <- c( 'empresa', 'alpha', 'beta' )
coefs$empresa <- names(retornos)[2:75]

for (i in 1:74){
```

```

coefs[i,2] <- coef(lm(retornos[,i+1] ~ retornos[,1] -
  retornos[,76]))['(Intercept)']
coefs[i,3] <- coef(lm(retornos[,i+1] ~ retornos[,76] ~ retornos[,1] -
  retornos[,76]))['retornos[, 1]']
}

> head(coefs)
empresa      alpha      beta
1 ITUB4.SA -0.001890671 1.0250159
2 B3SA3.SA -0.001178668 0.9654480
3 PETR4.SA -0.006399874 1.7656781
4 BBDC4.SA -0.003531262 0.9705597
5 PETR3.SA -0.003861232 1.4758013
6 ABEV3.SA -0.002121174 0.8289363

> tail(coefs)
empresa      alpha      beta
69 IGTA3.SA -0.003501721 0.8349665
70 USIM5.SA -0.002212169 0.9357436
71 ECOR3.SA -0.003478726 0.8276687
72 CVCB3.SA -0.002883426 0.6259825
73 HGTX3.SA -0.004570013 0.8144341
74 GOLL4.SA -0.008074353 0.9891407

```

3.4

Para verificar a variação explicada, basta olhar para o R^2 . Tomando arbitrariamente que uma boa explicação é $R^2 > 0.6$, temos:

```

coefs$'R quadrado' <- NA
coefs$'P valor de alpha' <- NA

for (i in 1:74){
  coefs[i, 4] <- summary(lm(retornos[,i+1] ~ retornos[,76] ~
    retornos[,1] - retornos[,76]))$r.squared
  coefs[i, 5] <- coef(summary(lm(retornos[,i+1] ~ retornos[,76] ~
    retornos[,1] - retornos[,76])))[7]
}

> subset(coefs, coefs$'R quadrado'>0.6)
empresa      alpha      beta R quadrado P valor de alpha
3  PETR4.SA -0.006399874 1.7656781 0.6250968 2.061112e-04
5  PETR3.SA -0.003861232 1.4758013 0.6724419 2.766589e-03
7  BBAS3.SA -0.003676873 1.3266875 0.6726642 1.540962e-03
8  ITSA4.SA -0.006132765 1.7066715 0.7287029 3.817365e-06

```

10	JBSS3.SA	-0.003421278	1.5183309	0.6063160	2.480604e-02
29	BRDT3.SA	-0.005286663	0.9207139	0.6264034	9.473536e-09
30	SBSP3.SA	-0.004344839	1.3931826	0.6659584	4.665828e-04
45	COGN3.SA	-0.006129905	1.7004298	0.7417666	1.719574e-06
52	FLRY3.SA	-0.004915613	1.0194310	0.6405466	5.075046e-07
53	IRBR3.SA	-0.003455530	0.9219265	0.6241863	1.302938e-04
58	CSNA3.SA	-0.003008471	1.2406074	0.6814542	4.582487e-03

Para testar a hipótese de que $\alpha = 0$, basta observar o p-valor do `summary(lm)`. Testando contra o nível de significância 0.025 temos:

```
> subset(coefs, coefs$'P valor de alpha'>0.025)
```

	empresa	alpha	beta	R quadrado	P valor de alpha
1	ITUB4.SA	-1.890671e-03	1.0250159	0.5968030	0.07098985
2	B3SA3.SA	-1.178668e-03	0.9654480	0.2154303	0.60527442
4	BBDC4.SA	-3.531262e-03	0.9705597	0.3063904	0.05187021
6	ABEV3.SA	-2.121174e-03	0.8289363	0.3586675	0.12318791
9	MGLU3.SA	4.393194e-05	1.2445474	0.4866509	0.97784039
14	GNDI3.SA	-1.967596e-03	1.2205791	0.3616754	0.32739923
19	BBDC3.SA	-2.310543e-03	0.8796261	0.4037981	0.08188959
20	RENT3.SA	-2.737210e-03	1.0979733	0.5393931	0.03028991
21	VIVT4.SA	1.911944e-03	1.6511584	0.4749589	0.37394520
22	LAME4.SA	-3.260190e-03	1.2646892	0.3359930	0.13952544
23	UGPA3.SA	-1.444341e-03	0.7598385	0.3281028	0.28379706
24	EQTL3.SA	-8.621131e-04	0.9927723	0.3653597	0.59452409
26	BTOW3.SA	4.046280e-04	0.8944781	0.4347814	0.74851594
28	CCRO3.SA	-2.496720e-03	1.1756139	0.5872175	0.04179526
32	KLBN11.SA	-2.590178e-03	0.7985194	0.3643142	0.04838198
34	HAPV3.SA	-1.706038e-03	0.5375073	0.2860372	0.10570440
36	ENGI11.SA	-2.426285e-03	0.9491199	0.4508906	0.06243007
43	PCAR3.SA	-3.683575e-03	1.2770299	0.4837428	0.02502130
44	BRML3.SA	-1.981059e-03	1.3484000	0.4154259	0.31751319
47	TOTS3.SA	-2.004734e-03	0.7724116	0.5060994	0.03482618
50	BRKM5.SA	-3.908970e-03	1.5341825	0.4789686	0.04954330
54	ELET6.SA	-3.250131e-03	0.7649283	0.2960232	0.02697724
59	MRFG3.SA	-1.114291e-03	1.1665254	0.5985997	0.34640098
61	AZUL4.SA	-1.111553e-03	0.4150506	0.1332584	0.39653016
63	ENBR3.SA	-1.229449e-03	1.2350516	0.5443774	0.37969559
64	MRVE3.SA	-2.991334e-03	0.9687313	0.2699026	0.13038512
70	USIM5.SA	-2.212169e-03	0.9357436	0.4103314	0.11222059
73	HGTX3.SA	-4.570013e-03	0.8144341	0.1961488	0.02608367

Logo apenas algumas empresas respeitam essa parte do modelo.

Testarei a normalidade do erro usando o teste K-S. Como são cerca de 400 observações, o ponto crítico ao nível de significância 0.05 é em torno de 0.07

```
coefs$'K-S test' <- NA
```

```

for (i in 1:74){
  coefs[i, 6] <- ks.test(resid(lm(retornos[,i+1] ~ retornos[,1]
})

> subset(coefs, coefs$'K-S test' < 0.07)
[1] empresa      alpha      beta      R quadrado
P valor de alpha K-S test
<0 rows> (or 0-length row.names)

```

Logo os erros não são distribuídos normalmente.

Para verificar a estabilidade do β , vou fazer as regressões para o período anterior, entre 2018 e 2019. Devo notar que o quantmod não conseguiu baixar os dados da Natura desse período, e vou cortar essa ação para fazer a comparação.

```

coefs2 <- coefs2[-16,]

codigos <- c(dados$Codigo)
codigos <- paste(codigos, ".SA", sep='')
codigos <- codigos[-77]
codigos <- codigos[-76]
codigos <- codigos[-16]
getSymbols(codigos, src='yahoo', from = "2018-05-21", to = "2019-05-21", env = )

getSymbols('^BVSP', src='yahoo', from = "2018-05-21", to = "2019-05-21", env = )

dados2 <- read.csv('(...) /FRB_H15.csv', stringsAsFactors = FALSE)
dados2 <- dados2[-c(1,2,3,4,5),]
dados2$Series.Description <- as.Date(dados2$Series.Description)

riskfree2 <- with(dados2, dados2[(Series.Description >= "2018-05-21" & Series
riskfree2 <- riskfree2[- grep('ND', riskfree2$Market.yield.on.U.S..Treasury.s
riskfree2$Market.yield.on.U.S..Treasury.securities.at.3.month...constant.matu
as.numeric(riskfree2$Market.yield.on.U.S..Treasury.securities.at.3.month...co

riskfree2$Market.yield.on.U.S..Treasury.securities.at.3.month...constant.matu
(1+riskfree2$Market.yield.on.U.S..Treasury.securities.at.3.month...constant.m
riskfree2 <- xts(riskfree2[, -1], order.by=riskfree2$Series.Description)

pframe <- do.call(merge, as.list(f))

todos <- data.frame(merge(pframe, riskfree2))
todos2 <- todos[, grepl("Close", names(todos))]
todos2 <- cbind(todos2, todos$riskfree2)

```

```

retornos <- data.frame(matrix(ncol = 76, nrow = 259))
colnames(retornos) <- c('BVSP', codigos, 'riskfree2')
retornos$riskfree2 <- todos2$todos$riskfree2

for (i in 1:75) {
  for (j in 1:259) {
    retornos[j,i] = log(todos2[j+1,i]/todos2[j,i])
  }
}

retornos <- retornos[, -7]
retornos <- na.omit(retornos)

coefs2 <- data.frame(matrix(nrow = 74, ncol=3))
colnames(coefs2) <- c('empresa', 'alpha', 'beta')
coefs2$empresa <- names(retornos)[2:75]

for (i in 1:73){
  coefs2[i,2] <- coef(lm(retornos[,i+1] ~ retornos[,75] ~ retornos[,1] - retornos[,2]))
  coefs2[i,3] <- coef(lm(retornos[,i+1] ~ retornos[,75] ~ retornos[,1] - retornos[,2]))
}

> head(coefs2)
  empresa      alpha      beta
1 ITUB4.SA -0.005382903 0.8896331
2 B3SA3.SA -0.005822410 1.1001610
3 PETR4.SA -0.002843259 1.2858247
4 BBDC4.SA -0.004185322 1.0461081
5 PETR3.SA -0.006150539 1.0956626
6 BBAS3.SA -0.004283986 0.6353060

> tail(coefs2)
  empresa      alpha      beta
69 USIM5.SA -0.005882214 1.2188308
70 ECOR3.SA -0.004451138 0.9900901
71 CVCB3.SA -0.003573700 0.2644555
72 HGTX3.SA -0.009525867 0.7613893
73 GOLL4.SA -0.005646744 1.2751726

```

Como pode-se ver, os valores de β são bem diferentes dos valores de 2019-2020, logo não vale a consistência de β .