

T6 – Análisis estadístico cuantitativo

Regressão linear Importando e manipulando dados

Importando dados

```
dap <- read.table("http://www.leg.ufpr.br/~walmes/cursoR/dap.txt", header=TRUE, sep="\t")
str(dap)
```

```
## 'data.frame': 991 obs. of 2 variables:
## $ DAP: num 4.87 7.38 5.95 5.73 6.4 ...
## $ HT : num 9.5 9.8 13 13.5 13.5 13.5 13.5 14.3 14.8 14.8 ...
```

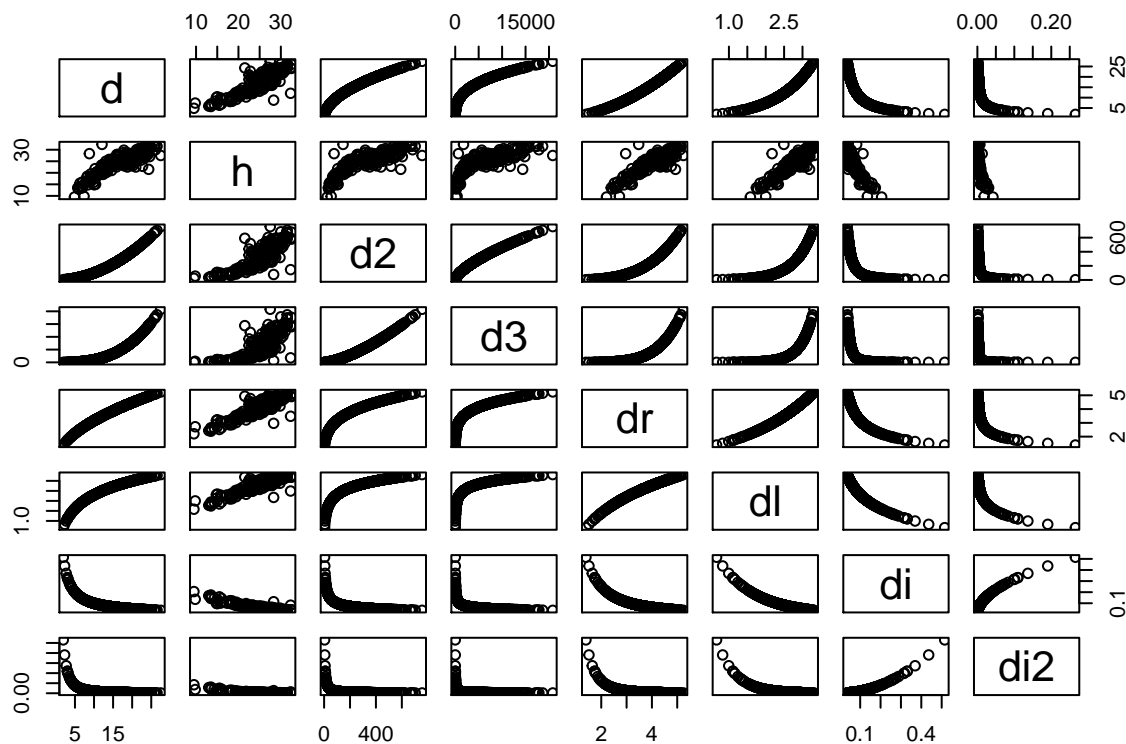
```
names(dap) <- c("d", "h")
```

Criando novas variáveis regressoras

```
dap$d2 <- dap$d^2
dap <- transform(dap, d2=d^2, d3=d^3, dr=sqrt(d), dl=log(d), di=1/d, di2=1/d^2)
str(dap)
```

```
## 'data.frame': 991 obs. of 8 variables:
## $ d : num 4.87 7.38 5.95 5.73 6.4 ...
## $ h : num 9.5 9.8 13 13.5 13.5 13.5 13.5 14.3 14.8 14.8 ...
## $ d2 : num 23.7 54.5 35.4 32.8 40.9 ...
## $ d3 : num 116 403 211 188 262 ...
## $ dr : num 2.21 2.72 2.44 2.39 2.53 ...
## $ dl : num 1.58 2 1.78 1.75 1.86 ...
## $ di : num 0.205 0.135 0.168 0.175 0.156 ...
## $ di2: num 0.0422 0.0183 0.0282 0.0305 0.0244 ...
```

```
pairs(dap)
```



```
dap <- dap[order(dap$d),]
dapcc <- dap[complete.cases(dap),]
rownames(dapcc) <- NULL
head(dapcc)

##           d      h      d2      d3      dr      dl      di      di2
## 1 4.8701  9.5 23.71787 115.5084 2.206830 1.583114 0.2053346 0.04216229
## 2 5.7296 13.5 32.82832 188.0931 2.393658 1.745646 0.1745323 0.03046151
## 3 5.9524 13.0 35.43107 210.8999 2.439754 1.783794 0.1679995 0.02822382
## 4 6.3344 15.5 40.12462 254.1654 2.516823 1.845995 0.1578681 0.02492235
## 5 6.3980 13.5 40.93440 261.8983 2.529427 1.855985 0.1562988 0.02442933
## 6 6.7482 13.5 45.53820 307.3009 2.597730 1.909276 0.1481877 0.02195958

str(dapcc)

## 'data.frame': 223 obs. of 8 variables:
## $ d : num 4.87 5.73 5.95 6.33 6.4 ...
## $ h : num 9.5 13.5 13 15.5 13.5 13.5 14.3 13.5 16 9.8 ...
## $ d2 : num 23.7 32.8 35.4 40.1 40.9 ...
## $ d3 : num 116 188 211 254 262 ...
## $ dr : num 2.21 2.39 2.44 2.52 2.53 ...
## $ dl : num 1.58 1.75 1.78 1.85 1.86 ...
## $ di : num 0.205 0.175 0.168 0.158 0.156 ...
## $ di2: num 0.0422 0.0305 0.0282 0.0249 0.0244 ...
```

Regressão linear simples

Ajustando a equação da reta (regressão linear simples)

```
m0 <- lm(h~d, data=dapcc)
summary(m0)

##
## Call:
## lm(formula = h ~ d, data = dapcc)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -8.4544 -1.2683  0.0483  1.3332 11.4202
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 11.89977    0.53760   22.14  <2e-16 ***
## d           0.74241    0.02969   25.00  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.304 on 221 degrees of freedom
## Multiple R-squared:  0.7388, Adjusted R-squared:  0.7376
## F-statistic: 625.1 on 1 and 221 DF, p-value: < 2.2e-16

str(m0)

## List of 12
## $ coefficients : Named num [1:2] 11.9 0.742
```

```

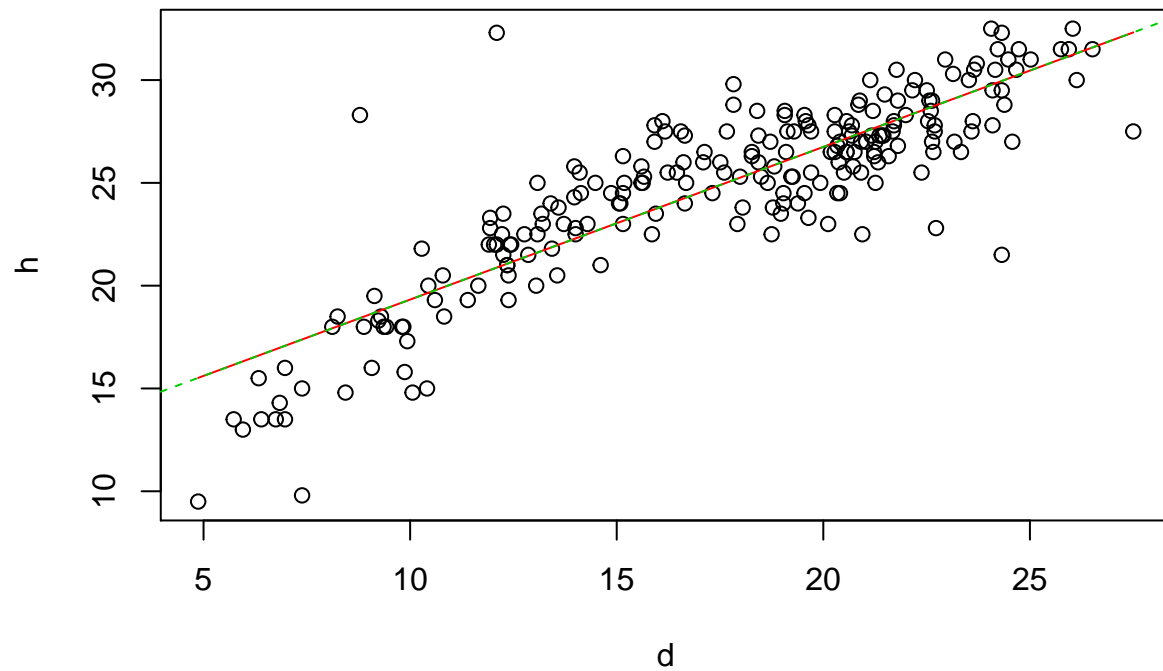
##   ..- attr(*, "names")= chr [1:2] "(Intercept)" "d"
##   $ residuals      : Named num [1:223] -6.02 -2.65 -3.32 -1.1 -3.15 ...
##   ..- attr(*, "names")= chr [1:223] "1" "2" "3" "4" ...
##   $ effects        : Named num [1:223] -369.975 57.617 -2.668 -0.462 -2.511 ...
##   ..- attr(*, "names")= chr [1:223] "(Intercept)" "d" "" "" ...
##   $ rank            : int 2
##   $ fitted.values: Named num [1:223] 15.5 16.2 16.3 16.6 16.6 ...
##   ..- attr(*, "names")= chr [1:223] "1" "2" "3" "4" ...
##   $ assign          : int [1:2] 0 1
##   $ qr              :List of 5
##   ..$ qr           : num [1:223, 1:2] -14.933 0.067 0.067 0.067 0.067 ...
##   .. ..- attr(*, "dimnames")=List of 2
##   .. .. ..$ : chr [1:223] "1" "2" "3" "4" ...
##   .. .. ..$ : chr [1:2] "(Intercept)" "d"
##   .. ..- attr(*, "assign")= int [1:2] 0 1
##   ..$ qraux: num [1:2] 1.07 1.14
##   ..$ pivot: int [1:2] 1 2
##   ..$ tol   : num 1e-07
##   ..$ rank  : int 2
##   ..- attr(*, "class")= chr "qr"
##   $ df.residual : int 221
##   $ xlevels      : Named list()
##   $ call         : language lm(formula = h ~ d, data = dapcc)
##   $ terms        :Classes 'terms', 'formula' language h ~ d
##   .. ..- attr(*, "variables")= language list(h, d)
##   .. ..- attr(*, "factors")= int [1:2, 1] 0 1
##   .. .. ..- attr(*, "dimnames")=List of 2
##   .. .. .. ..$ : chr [1:2] "h" "d"
##   .. .. .. ..$ : chr "d"
##   .. ..- attr(*, "term.labels")= chr "d"
##   .. ..- attr(*, "order")= int 1
##   .. ..- attr(*, "intercept")= int 1
##   .. ..- attr(*, "response")= int 1
##   .. ..- attr(*, ".Environment")=<environment: R_GlobalEnv>
##   .. ..- attr(*, "predvars")= language list(h, d)
##   .. ..- attr(*, "dataClasses")= Named chr [1:2] "numeric" "numeric"
##   .. .. ..- attr(*, "names")= chr [1:2] "h" "d"
##   $ model         :'data.frame': 223 obs. of 2 variables:
##   ..$ h: num [1:223] 9.5 13.5 13 15.5 13.5 13.5 14.3 13.5 16 9.8 ...
##   ..$ d: num [1:223] 4.87 5.73 5.95 6.33 6.4 ...
##   ..- attr(*, "terms")=Classes 'terms', 'formula' language h ~ d
##   .. .. ..- attr(*, "variables")= language list(h, d)
##   .. .. ..- attr(*, "factors")= int [1:2, 1] 0 1
##   .. .. .. ..- attr(*, "dimnames")=List of 2
##   .. .. .. .. ..$ : chr [1:2] "h" "d"
##   .. .. .. .. ..$ : chr "d"
##   .. .. ..- attr(*, "term.labels")= chr "d"
##   .. .. ..- attr(*, "order")= int 1
##   .. .. ..- attr(*, "intercept")= int 1
##   .. .. ..- attr(*, "response")= int 1
##   .. .. ..- attr(*, ".Environment")=<environment: R_GlobalEnv>
##   .. .. ..- attr(*, "predvars")= language list(h, d)
##   .. .. ..- attr(*, "dataClasses")= Named chr [1:2] "numeric" "numeric"
##   .. .. .. ..- attr(*, "names")= chr [1:2] "h" "d"

```

```
## - attr(*, "class")= chr "lm"
```

Verificando o ajuste

```
plot(h~d, dapcc) # xlab=, ylab=  
lines(fitted(m0)~d, dapcc, col="red")  
abline(m0, col=3, lty=2)
```



Análise de resíduos

```
par(mfrow=c(2,2))  
plot(m0)
```

