

Exercicios Aula 7

Affonso Amendola

NUSP 9301753

May 12, 2020

Exercício 1

Use os dados da tabela `tab_virgo.dat` para ajustar a relação $D_n - \sigma$. Use a função `lm()` para fazer uma regressão linear ordinária e a função `rlm()` para fazer uma regressão robusta. Compare os resultados.

Para ajustar usando `lm()`:

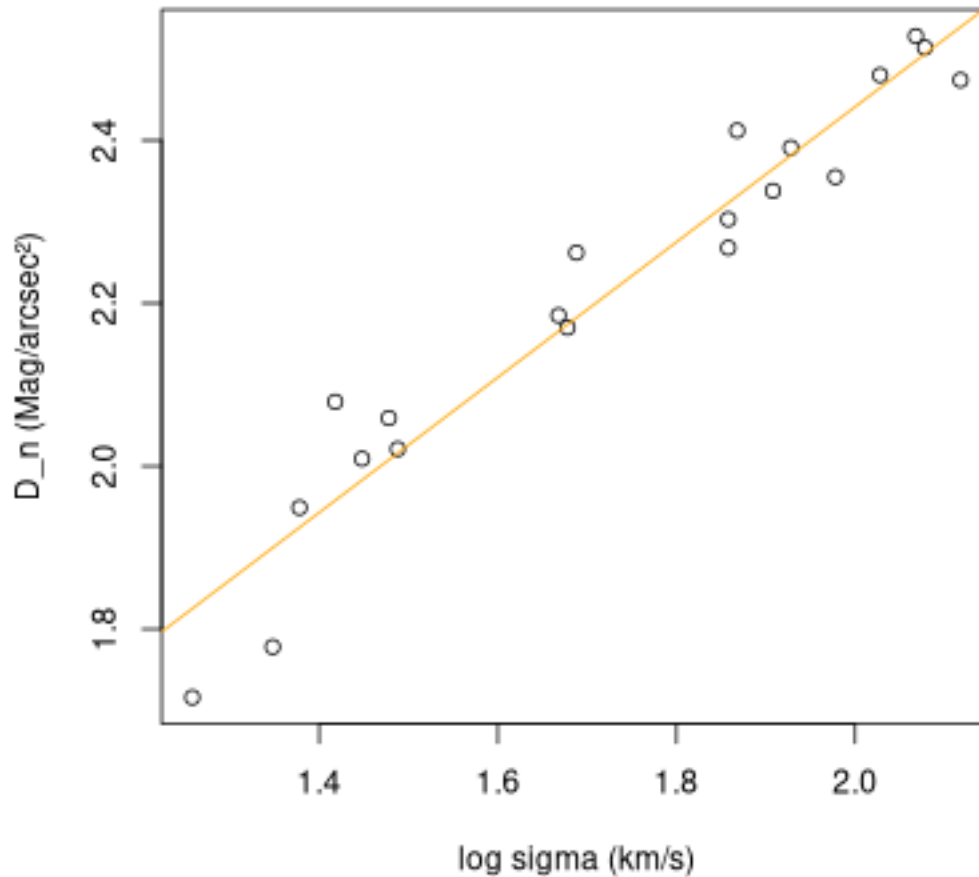
```
1 data <- read.table(file = "tab_virgo.dat", header = TRUE)
2
3 log_dn <- data[, 4]
4 log_sigma <- data[, 8]
5
6 x <- log_dn
7 y <- log_sigma
8
9 model <- lm(y~x)
10 plot(x,y, main="Aglomerado de Virgo", xlab="log sigma (km/s)", ylab= "D_n (Mag/arcsec^2)")
11
12 abline(model, col="orange")
13
14 print(model)
```

Os coeficientes obtidos na saída do programa foram:

```
1 Call:
2 lm(formula = y ~ x)
3
4 Coefficients:
5 (Intercept)          x
6      0.7806       0.8303
```

E obtive o seguinte gráfico

Aglomerado de Virgo



Para a regressão robusta `rlm()` o código a seguir foi utilizado:

```
1 data_robust <- data.frame(cbind(x,y))
2
3 fit <- lm(y~x, data=data_robust)
4
5 print(fit)
```

O output foi:

```
1 Call:
2 lm(formula = y ~ x, data = data_robust)
3
4 Coefficients:
5 (Intercept)          x
6    0.7806         0.8303
```

Introduzindo um outlier em $x, y = (1.2, 2.45)$

```
1 outlier <- data.frame(x= 1.2, y=2.45)
2
3 out_data <- rbind(data_robust, outlier)
```

E refazer o ajuste:

```
1 fit_1 <- lm(y~x, data= out_data)
2 print(fit_1)
```

Tendo o seguinte ajuste:

```
1 Call:
2 lm(formula = y ~ x, data = out_data)
3
4 Coefficients:
5 (Intercept)          x
6      1.1463       0.6343
```

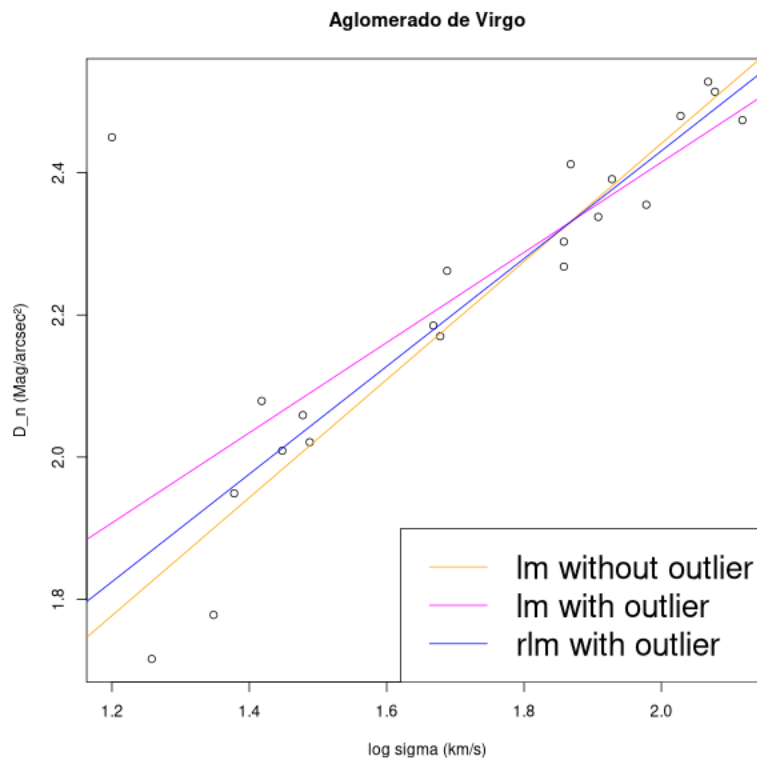
Após isso fiz o ajuste robusto usando o pacote MASS:

```
1 library(MASS)
2 robust_fit <- rlm(y~x, data = out_data)
3
4 print(robust_fit)
```

Obtendo o seguinte resultado final:

```
1 Call:
2 rlm(formula = y ~ x, data = out_data)
3 Converged in 8 iterations
4
5 Coefficients:
6 (Intercept)          x
7   0.9152115   0.7578167
8
9 Degrees of freedom: 21 total; 19 residual
10 Scale estimate: 0.0414
```

Com esses dados o seguinte gráfico foi elaborado para visualização e podermos comparar os resultados, com o outlier adicionado sendo claramente visível do topo superior esquerdo:



Exercício 2

A tabela `tab_virgo.dat` contém também um indicador de metalicidade denominado `Mg2`, associado ao magnésio. Usando mínimos quadrados linear, ajuste este índice em função do brilho superficial e da dispersão de velocidades (conjuntamente!).

Foi usado o seguinte código:

```
1 data <- read.table(file="tab_virgo.dat", header=TRUE)
2
3 sigmae <- data[,7]
4 log_sigma <- data[,8]
5 mg2 <- data[,9]
6
7 x <- log_sigma
8 y <- sigmae
9 z <- mg2
10
11 chi_2 = lm(z~x+y)
12 print(summary(chi_2))
```

Para obter o seguinte resultado:

```
1
2 Call:
3 lm(formula = z ~ x + y)
4
5 Residuals:
6      Min       1Q   Median       3Q      Max
7 -0.023324 -0.009331 -0.002051  0.011551  0.020893
8
9 Coefficients:
10              Estimate Std. Error t value Pr(>|t|)
11 (Intercept) -0.308387   0.093733  -3.290  0.00432 **
12 x            0.228062   0.013817  16.506 6.72e-12 ***
13 y            0.003663   0.003972   0.922  0.36931
14 ---
15 Signif. codes:  0   ***    0.001   **    0.01   *    0.05   .    0.1    1
16
17 Residual standard error: 0.01413 on 17 degrees of freedom
18 Multiple R-squared:  0.942, Adjusted R-squared:  0.9352
19 F-statistic: 138.1 on 2 and 17 DF, p-value: 3.068e-11
```

Ou seja, o resultado foi:

$$\begin{aligned}mg2 &= -0.308387 \\log_{\sigma} &= 0.228062 \\\sigma &= 0.003663\end{aligned}$$

Exercício 3