Linreg specification

Milda Poceviciute, Henrik Karlsson

2017-09-27

This vignette summarises the *Least-squares estimation and related techniques* of linear regression that are available in linreg package.The examples are provided based on Iris data set and linear regression formula:

```
example_formula <- Petal.Length~Sepal.Width+Sepal.Length
```

General information

The package contains *linreg* RC object, hence the syntax of calling it is:

```
linreg_example <- linreg$new(example_formula, data=iris)
```

The *linreg* includes calculations of these values:

- *Estimate of regression coefficients:*

$$\hat{\beta} = (\mathbf{X^T X})^{-1} \mathbf{X^T y}$$

- *The fitted values:*

$$\hat{\mathbf{y}} = \mathbf{X}\hat{\beta}$$

- *The residuals:*

$$\hat{\mathbf{e}} = \mathbf{y} - \hat{\mathbf{y}}$$

- *The degrees of freedom:*

$$df = n - p$$

  where n is the number of observations and p is the number of parameters in the model.
- *The residual variance:*

$$\hat{\sigma}^2 = \frac{\mathbf{e^T e}}{\mathbf{df}}$$

- *The variance of the regression coefficients:*

$$\widehat{\mathrm{Var}(\hat{\beta})} = \hat{\sigma}^2 (\mathbf{X^T X})^{-1}$$

Methods

Below are listed available methods with examples:

- **print()**, returns the coefficients and coefficient names.

```
linreg_example$print()
```

```
##
## Call:
## linreg(formula = Petal.Length ~ Sepal.Width + Sepal.Length, data = iris)
##
## Coefficients:
## (Intercept) Sepal.Width Sepal.Length
## -2.524762 -1.338623 1.775593
```
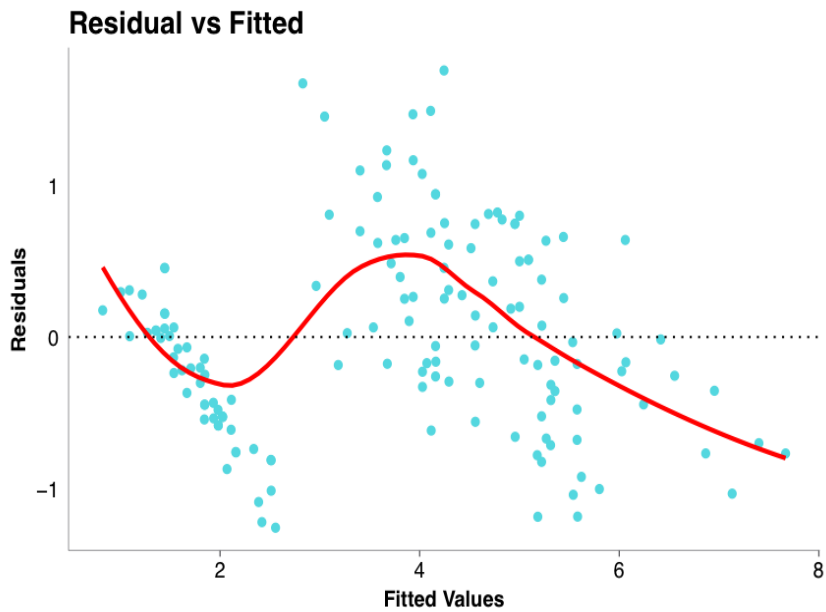
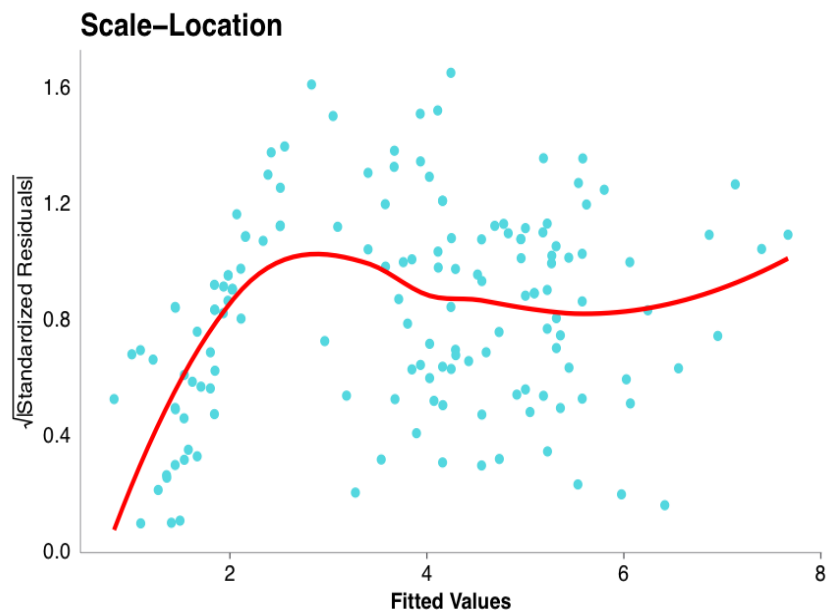- **plot()**, returns *Residuals vs Fitted* and *Scale Location* plots

```
linreg_example$plot()
```

$Residual_vs_Fitted

Residual vs Fitted

##
$Scale_Location

Scale–Location

- **resid()**, returns return the vector of residuals **e.**(in the example only first 5 values are displayed)

```
residuals <- linreg_example$resid()
residuals[1:5]
```

```
## [1] -0.45 -0.76 -0.24  0.01 -0.13
```

- **pred()**, returns the predicted values $\hat{y}$ (in the example only first 5 values are displayed)

```
predicted <- linreg_example$pred()
predicted[1:5]
```

```
## [1] 1.85 2.16 1.54 1.49 1.53
```

- **coef()**, returns the coefficients as a named vector

```
linreg_example$coef()
```

```
##  (Intercept)  Sepal.Width Sepal.Length
##    -2.524762    -1.338623     1.775593
```

- **summary()**, returns the coefficients, their standard error, $t-value$, $p-value$, and the degrees of freedom in the model

```
linreg_example$summary()
```

```
##
## Call:
## linreg(formula = Petal.Length ~ Sepal.Width + Sepal.Length, data = iris)
##
## Coefficients:
##
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)    -2.52       0.56   -4.48      ***
## Sepal.Width    -1.34       0.12  -10.94      ***
## Sepal.Length    1.78       0.06   27.57      ***
##
## Residual standard error: 0.65 on 147 degrees of freedom
```

```
## R version 3.3.1 (2016-06-21)
## Platform: x86_64-apple-darwin13.4.0 (64-bit)
## Running under: OS X 10.12.6 (Sierra)
##
## locale:
## [1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8
##
## attached base packages:
## [1] stats     graphics  grDevices utils     datasets  methods   base
##
## other attached packages:
## [1] linreg_1.0   ggplot2_2.2.1
##
## loaded via a namespace (and not attached):
##  [1] Rcpp_0.12.12     digest_0.6.12    rprojroot_1.2    plyr_1.8.4
##  [5] grid_3.3.1       gtable_0.2.0     backports_1.1.0  magrittr_1.5
##  [9] evaluate_0.10.1  scales_0.5.0     rlang_0.1.2      stringi_1.1.5
## [13] lazyeval_0.2.0   rmarkdown_1.6    labeling_0.3     tools_3.3.1
## [17] stringr_1.2.0    munsell_0.4.3    yaml_2.1.14      colorspace_1.3-2
## [21] htmltools_0.3.6  knitr_1.17       tibble_1.3.4
```