Lab4

Simon Jonsson, Fanny Karelius

2017-10-02

This lab is for the purpose of implementing some of the simple features of the **lm** class. Methods will be declared below

General information

Package implements linear regression models, called **linreg** and is called as an example:

```
ex_form <- Petal.Length ~ Sepal.Width + Sepal.Length
linreg_ex <- linreg$new(formula=ex_form, data=iris)
```

The package calculates the following formulas: **Regression coefficients:**

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

The fitted values:

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

The residuals:

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

The degrees of freedom: where n is the number of observations and p is the number of parameters in the model.

$$df = n - p$$

The residual variance:

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

The variance of the regression coefficients:

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

The t-values for each coefficient:

$$t_\beta = \frac{\hat{\beta}}{\sqrt{\mathrm{Var}\left(\hat{\beta}\right)}}$$

Methods

Implemented methods are as follows:

print() prints out the coefficients and coefficient names, similar as done by the lm class
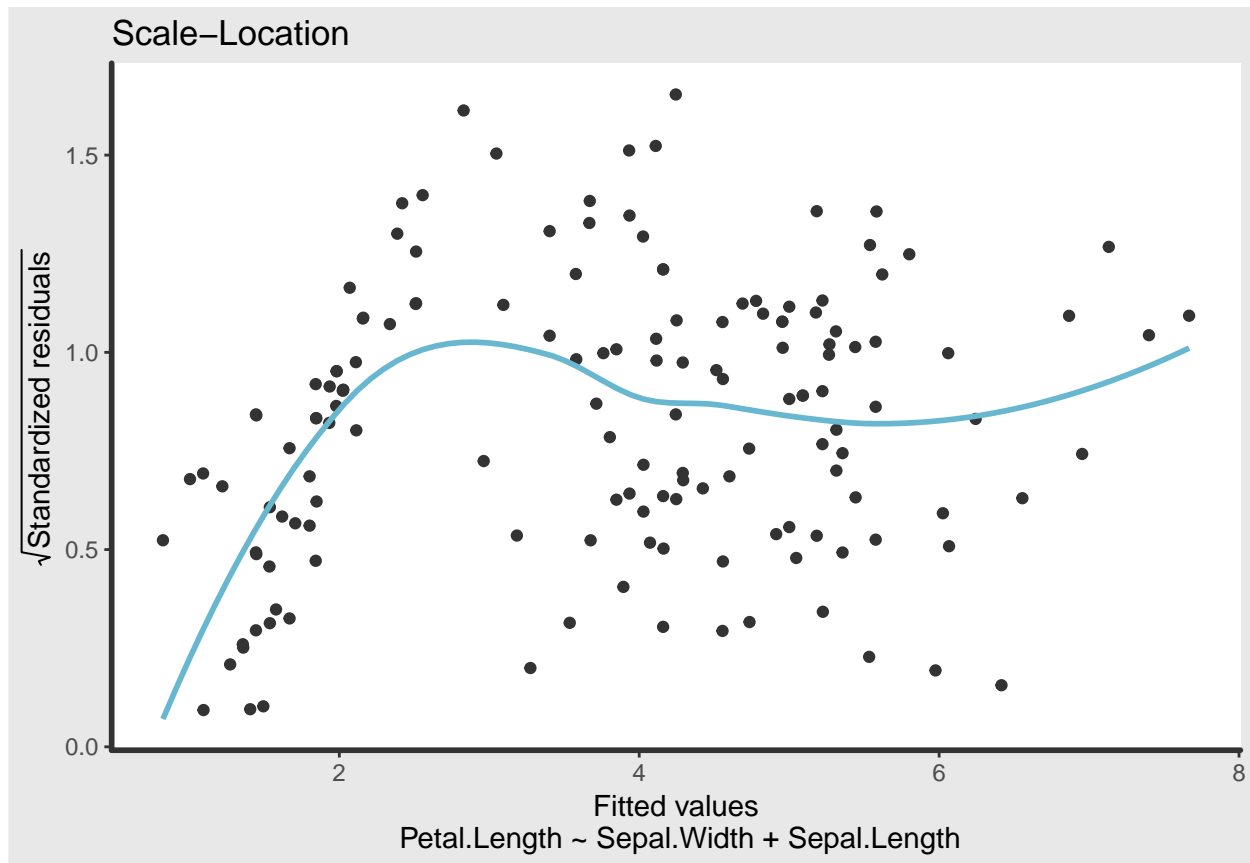
```
linreg_ex$print()
```

```
##
## Call:
## linreg(formula = Petal.Length ~ Sepal.Width + Sepal.Length, data = iris)
##
## Coefficients:
##  (Intercept)    Sepal.Width  Sepal.Length
##      -2.5248        -1.3386        1.7756
```

plot() returns **Scale-Location** and **Residuals vs Fitted** plots p1 and p2 respectively
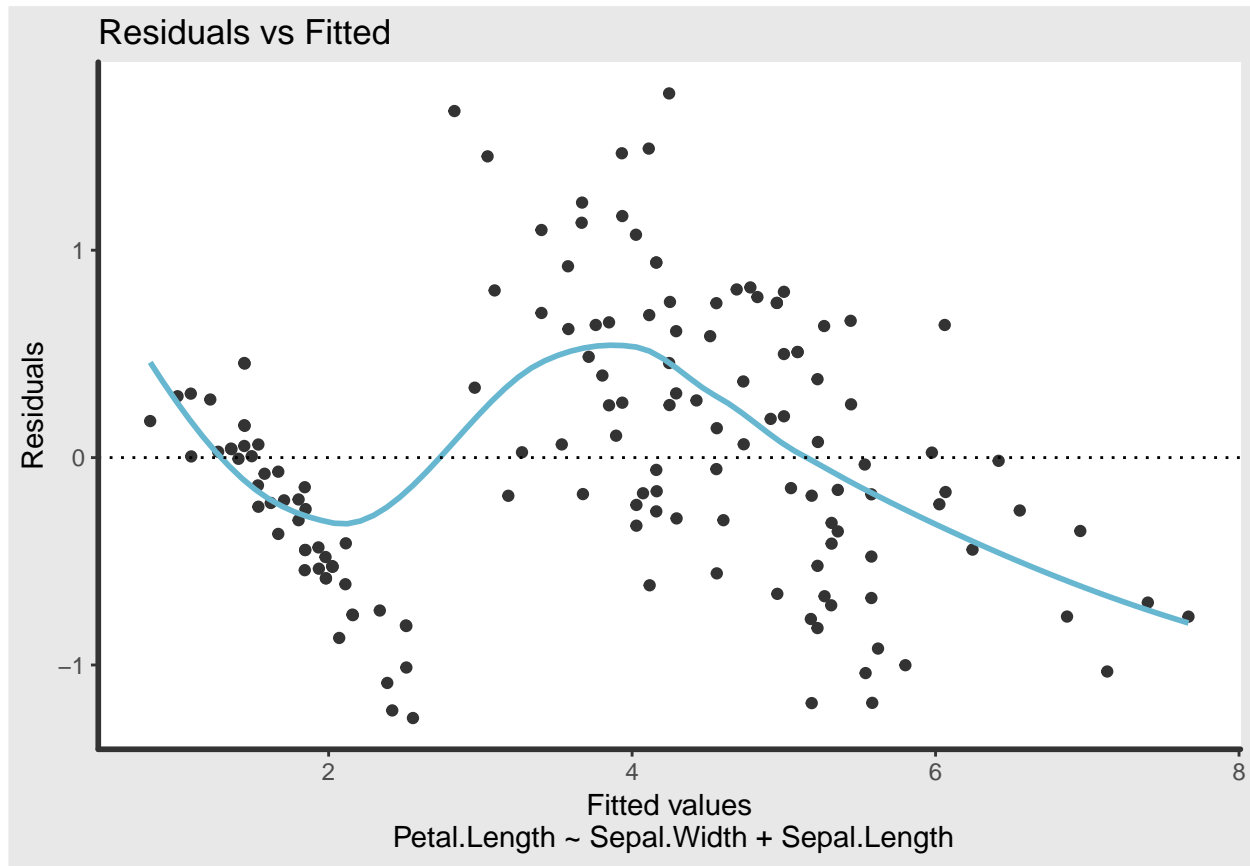
```
linreg_ex$plot()
```

```
## $p1
```

```
## `geom_smooth()` using method = 'loess'
```



```
##
## $p2
```

```
## `geom_smooth()` using method = 'loess'
```

resid() returns a vector of residuals $\hat{\mathbf{e}}$

```
head(linreg_ex$resid())
```

```
##             [,1]
## 1 -0.445578965
## 2 -0.759772100
## 3 -0.236928933
## 4  0.006767993
## 5 -0.134157381
## 6 -0.142807413
```

pred() returns the predicted values $\hat{\mathbf{y}}$

```
head(linreg_ex$pred())
```

```
##       [,1]
## 1 1.845579
## 2 2.159772
## 3 1.536929
## 4 1.493232
## 5 1.534157
## 6 1.842807
```

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

```
head(linreg_ex$coef())
```

```
##  (Intercept)  Sepal.Width Sepal.Length
```

```
##    -2.524762    -1.338623    1.775593
```

summary() prints out the coefficients with their **standard error**, **t-value**, and **p-value** as well as the estimate of $\hat{\sigma}$ and the degrees of freedom in the model

```
linreg_ex$summary()
```

```
## Call:
## linreg(formula = Petal.Length ~ Sepal.Width + Sepal.Length, data = iris)
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    -2.525      0.563  -4.481        1
## Sepal.Width    -1.339      0.122 -10.940        1
## Sepal.Length    1.776      0.064  27.569        1
## ---
##
## Residual standard error: 0.6465 on 147 degrees of freedom
```