

-

Tutorial

R Tutorial (R-Tutorial.html)

ggplot2

ggplot2 Short Tutorial (ggplot2-Tutorial-With-R.html)

ggplot2 Tutorial 1 - Intro (Complete-Ggplot2-Tutorial-Part1-With-R-Code.html)

ggplot2 Tutorial 2 - Theme (Complete-Ggplot2-Tutorial-Part2-Customizing-Theme-With-R-Code.html)

ggplot2 Tutorial 3 - Masterlist (Top50-Ggplot2-Visualizations-MasterList-R-Code.html)

ggplot2 Quickref (ggplot2-cheatsheet.html)

Foundations

Linear Regression (Linear-Regression.html)

Statistical Tests (Statistical-Tests-in-R.html)

Missing Value Treatment (Missing-Value-Treatment-With-R.html)

Outlier Analysis (Outlier-Treatment-With-R.html)

Feature Selection (Variable-Selection-and-Importance-With-R.html)

Model Selection (Model-Selection-in-R.html)

Logistic Regression (Logistic-Regression-With-R.html)

Advanced Linear Regression (Environments.html)

Advanced Regression Models

Advanced Regression Models (adv-regression-models.html)

Time Series

Time Series Analysis (Time-Series-Analysis-With-R.html)

Time Series Forecasting (Time-Series-Forecasting-With-R.html)

More Time Series Forecasting ([Time-Series-Forecasting-With-R-part2.html](#))

High Performance Computing

Parallel computing ([Parallel-Computing-With-R.html](#))

Strategies to Speedup R code ([Strategies-To-Improve-And-Speedup-R-Code.html](#))

Useful Techniques

Association Mining ([Association-Mining-With-R.html](#))

Multi Dimensional Scaling ([Multi-Dimensional-Scaling-With-R.html](#))

Optimization ([Profiling.html](#))

InformationValue package ([Information-Value-With-R.html](#))

Stay up-to-date. Subscribe!

(https://docs.google.com/forms/d/1xkMYkLNFU9U39Dd8S_2JC0p8B5t6_Yq6zUQjanQQJpY/viewform)

Chat! (<https://docs.google.com/forms/d/13GrkCFcNa-TOIIIQghsz2SIEbc-YqY9eJX02B19I5Ow/viewform>)

Contents

How To Specify A Robust Regression Model

Compare Performance of `rlm()` with `lm()`

Robust Regression

Robust regression can be used in any situation where OLS regression can be applied. It generally gives better accuracies over OLS because it uses a weighting mechanism to weigh down the influential observations. It is particularly resourceful when there are no compelling reasons to exclude outliers in your data.

Robust regression can be implemented using the `rlm()` function in MASS package. The outliers can be weighted down differently based on `psi.huber`, `psi.hampel` and `psi.bisquare` methods specified by the `psi` argument.

How To Specify A Robust Regression Model

```
library(MASS)
rlm_mod <- rlm(stack.loss ~ ., stackloss, psi = psi.bisquare) # robust reg model
summary(rlm_mod)
#> Call: rlm(formula = stack.loss ~ ., data = stackloss)
#> Residuals:
#>      Min       1Q   Median       3Q      Max
#> -8.91753 -1.73127  0.06187  1.54306  6.50163
#>
#> Coefficients:
#>              Value      Std. Error t value
#> (Intercept) -41.0265      9.8073    -4.1832
#> Air.Flow      0.8294      0.1112     7.4597
#> Water.Temp    0.9261      0.3034     3.0524
#> Acid.Conc.   -0.1278      0.1289    -0.9922
#>
#> Residual standard error: 2.441 on 17 degrees of freedom
```

Compare Performance of rlm() with lm()

Lets build the equivalent lm() model so we can compare the errors against the respective fitted values.

```
lm_mod <- lm(stack.loss ~ ., stackloss) # lm reg model
```

Calculate the Errors

```
# Errors from lm() model
DMwR::regr.eval(stackloss$stack.loss, lm_mod$fitted.values)
#>      mae      mse      rmse      mape
#> 2.3666202 8.5157125 2.9181694 0.1458878

# Errors from rlm() model
DMwR::regr.eval(stackloss$stack.loss, rlm_mod$fitted.values)
#>      mae      mse      rmse      mape
#> 2.1952232 9.0735283 3.0122298 0.1317191
```

As expected, the errors from the robust regression model is lesser than the linear regression model.

© 2016-17 Selva Prabhakaran. Powered by jekyll (<http://jekyllrb.com/>), knitr (<http://yihui.name/knitr/>), and pandoc (<http://johnmacfarlane.net/pandoc/>). This work is licensed under the Creative Commons License. (<http://creativecommons.org/licenses/by-nc/3.0/>)