

# Linear Model

Julia Terhart & Philipp Kopper

02 June 2020

## Model Math

A linear model can be characterised by the following structural normal equation:

$$X^T X y = X^T y \beta \quad (1)$$

where  $X$  is the data or design matrix,  $y$  the dependent variable and  $\beta$  the coefficient vector.

## Fitting in R

A model can be fit with the following code (using the `iris` data):

```
linear_model <- lm(Petal.Width ~ ., data = iris)
```

A summary can be obtained via:

```
summary( linear_model)
```

```
##
## Call:
## lm(formula = Petal.Width ~ ., data = iris)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.59239 -0.08288 -0.01349  0.08773  0.45239
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -0.47314    0.17659   -2.679  0.00824 **
## Sepal.Length  -0.09293    0.04458   -2.084  0.03889 *
## Sepal.Width    0.24220    0.04776    5.072 1.20e-06 ***
## Petal.Length   0.24220    0.04884    4.959 1.97e-06 ***
## Speciesversicolor 0.64811    0.12314    5.263 5.04e-07 ***
## Speciesvirginica  1.04637    0.16548    6.323 3.03e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1666 on 144 degrees of freedom
## Multiple R-squared:  0.9538, Adjusted R-squared:  0.9522
## F-statistic: 594.9 on 5 and 144 DF, p-value: < 2.2e-16
```