Advanced Statistical Modeling

Non-parametric models - Iteratively Re-Weighted Least Squares

Haoran Mo, Alexandra Yamaui

November 22, 2017

In this task we are going to implement the Iteratively Re-Weighted Least Squares algorithm (IRWLS), which is the most frequently used method to solve the maximization problem of the log-likelihood function. This function is at the same time used to estimate the coefficients of the Logistic Regression model.

We will used zero as the initial value of the coefficients β_0 (beta_0) and β_1 (beta_1) and we will build a new response variable z, which is a linear combination of the points x. The formula is presented below

$$z_i = \beta_0 + \beta_1 x_i + \frac{y_i - p_i}{p_i(1-p_i)},\ i = 1, ..., n$$

where y_i is the original response variable and p_i is defined as below, which comes from the logistic function for the conditional distribution of the response variable y:

$$p_i = \frac{e^{\beta_0 + \beta_1 x_i}}{1 + e^{\beta_0 + \beta_1 x_i}}$$

```r
IRWLS <- function(x,y) {
  n <- length(x)
  beta_0 <- 0
  beta_1 <- 0
  s <- 0
  p <- c()
  v <- c()
  z <- c()
  convergence = 1
  #convergence != TRUE

  while (convergence > 0.0001) {  # we set 0.0001 instead of 0 due to computing cost concerned.
    for (i in 1:n) {
      p[i] <- exp(beta_0 + beta_1*x[i])/(1 + exp(beta_0 + beta_1*x[i]))
      z[i] <- beta_0 + beta_1*x[i] + ((y[i]-p[i])/p[i]*(1-p[i]))
      v[i] <- p[i]*(1-p[i])
    }

    lr <- lm(z ~ x, weights = v)


    beta_0_pre <- beta_0
    beta_1_pre <- beta_1
    beta_0 <- lr$coefficients[1]
    beta_1 <- lr$coefficients[2]

    convergence <- (abs(beta_0-beta_0_pre) + abs(beta_1-beta_1_pre))/2
    s <- s + 1
  }
```

```r
  return(c(beta_0,beta_1))
}

# IRWLS(x,y)
```