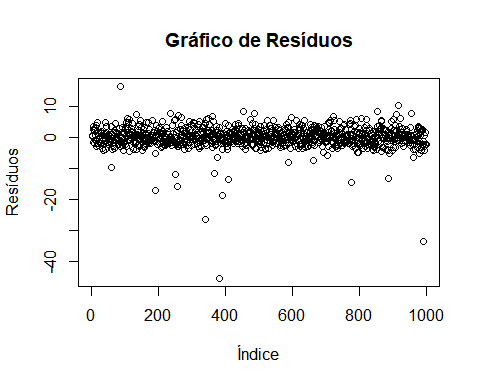
Exercício 3

Gustavo Beretta Gonçalves

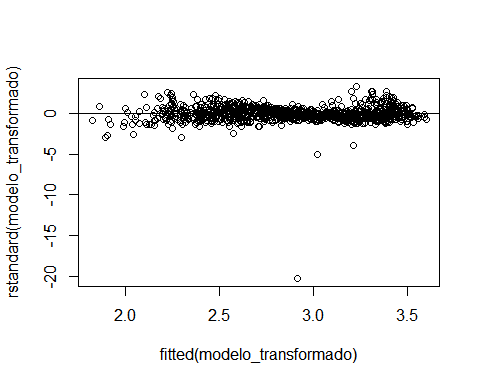
#leitura da base de dados  
base <- read.csv2("selecao.csv", dec=".")  
  
#a)  
set.seed(09092003)  
amostra = base[sample(nrow(base), 1000),]  
  
#b)  
modelo.null <- lm(formula = y ~ 1, data = amostra)  
modelo.forward <- step(object = modelo.null,  
 scope = list(lower = ~1, upper = ~ x1 + x2 + x3 + x4 + x5 + x6 + x7 + x8 + x9 + x10),  
 direction = "forward",  
 trace = 0)  
  
#c)  
residuos <- residuals(modelo.forward)  
plot(residuos, main = "Gráfico de Resíduos", ylab = "Resíduos", xlab = "Índice")



#podemos verificar um número consideravel de "outliers", isso pode ser melhorado para deixar o modelo mais ajustado.  
  
valores\_unicos\_x10 <- unique(amostra$x10)  
print(valores\_unicos\_x10)

## [1] 1 0

#podemos aplicar uma transformação logarítmica na variável resposta e na variável independente x9  
  
modelo\_transformado <- lm(log(y) ~ x7 + x10 + log(x9) + x6 + x4 + x5 + x8, data = amostra)  
  
#d)  
plot(fitted(modelo\_transformado), rstandard(modelo\_transformado))  
abline(0, 0)



#essa transformação se mostrou eficaz, pois a presença de outliers foi fortemente reduzida, resultando em um modelo mais ajustado