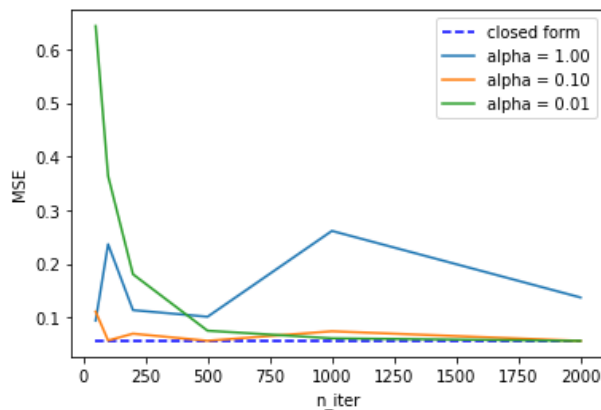


### Task 3c)

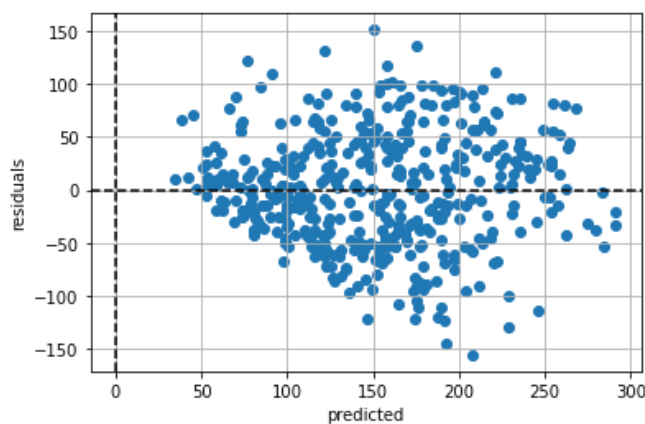


The alpha 0.01 is very small and thus leads to smaller steps. Because of that it needs a lot of iteration until the MSE is good enough and is similar to the true values.

The alpha 1.00 on the other hand is too large and leads to a behaviour which “over-shoots” the target and at high iterations can be bad.

The optimum alpha value for our test is 0.10 because already at the iterations under 250 the MSE is already quite small.

### Task 4a)



- 1) The relationship seems linear
- 2) The residuals seem to be independent from the predicted values
- 3) It seems to be normal because there are not that many outliers
- 4) The error distribution seems to go up the higher the predicted value, so it is not equally variant

### Task 4b)

All coefficients:

```
[ -10.01219782 -239.81908937 519.83978679 324.39042769 -792.18416163 476.74583782 101.04457032  
177.06417623 751.27932109 67.62538639]
```

If you take the absolute value of every coefficient, then the bigger the value the more influential it is to the cost function. So, the three most influential values and features are:

- 792 -> s1
- 751 -> s5
- 519 -> bmi

If I had to exclude one feature, it would be the age because the absolute coefficient of it is 10 and thus doesn't change much.