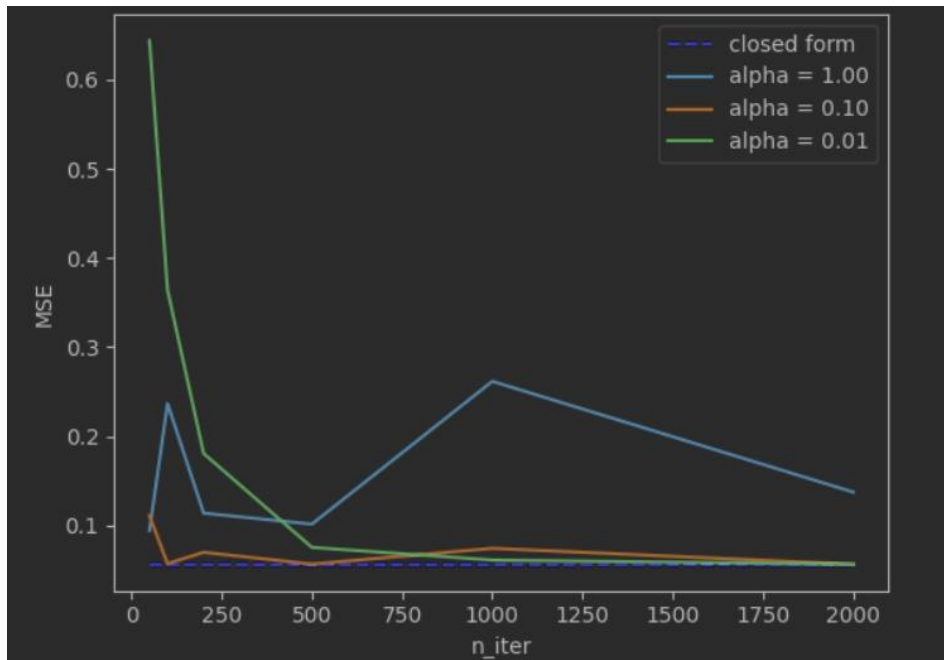


hertijon

Task 3:



Big alphas lead to bigger errors after a while, as the steps are bigger and can overstep the target. Small alphas start with big errors but will steadily improve as it gets closer to the target. Probably good to start with big alphas and then increasingly reduce the alpha.

Task 4a:

Linearity

- Somewhat fulfilled, the residuals in general go from mid left to mid right.

Normality

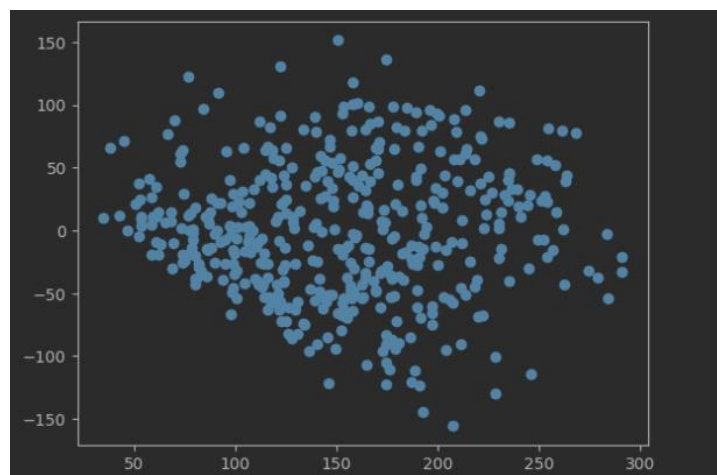
- Fulfilled, seems the errors are normally distributed.

Independence

- Somewhat fulfilled, the residuals seem to be independent, but tend to build a cone.

Equality of variance

- Not fulfilled, the residuals build a cone and are not equally distributed.



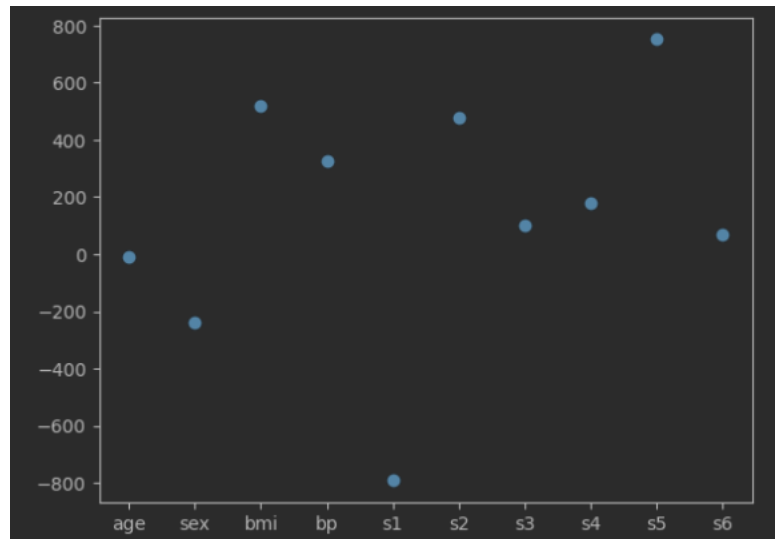
Task4b:

Which are the 3 most influential features?

1. **s1** -> -792.17563855
2. **s5** -> 751.27369956
3. **bmi** -> 519.84592005

How do you interpret the sign of the coefficients?

The sign of the coefficients is not relevant for determining the most influential features. It's the absolute value which matters the most. The bigger the value the more influential.



If you had to exclude 1 feature, which one would you select and why

Feature to be excluded would be **age** as it has the **smallest value**, -10.0098663. Meaning it will have the least amount of effect on the calculation.