FYS-STK4155: Project 1

Simen Nyhus Bastnes

5. October 2020

Abstract

The goal of this project is to do linear regression with bootstrapping and cross-validation on the Franke function and real data.

1 Introduction

With the emergence of more powerful computers, the field of machine learning has exploded in recent years, and is steadily becoming an integral part of many fields of science. While many of the concepts and algorithms used in machine learning today has been known for a long time, the emergence of more powerful dauowhiao h adwd One of the simplest and most studied methods is so-called linear regression.

The heart. [1]

2 Theory

2.1 Linear regression

2.1.1 Ordinary least squares

2.1.2 Ridge regression

2.1.3 LASSO regression

2.2 Resampling methods

2.2.1 Bootstrap

2.2.2 Cross-validation

2.3 Confidence intervals

Variance for β for Ridge regression can be found in [2]

$$\mathrm{Var}[\hat{\beta}(\lambda)] = \sigma^2(\mathbf{X}^\mathsf{T}\mathbf{X} + \lambda\mathbf{I}_{pp})^{-1}\mathbf{X}^\mathsf{T}\mathbf{X}[(\mathbf{X}^\mathsf{T}\mathbf{X} + \lambda\mathbf{I}_{pp})^{-1}]^\mathsf{T}$$

3 Implementation

4 Results

5 Discussion

6 Conclusion

7 Appendixes?

References

[1] Github repository, project 1. `https://github.com/simennb/fysstk4155-project1`.

[2] Wessel N. van Wieringen. Lecture notes on ridge regression, 2020.