Report Template
FYS-STK4155 - Project 1

Bartosz Trela
University of Oslo
(Dated: September 26, 2025)

Regression is a fundamental tool in physics and data science for approximating functions and making predictions. In this project, we studied polynomial regression using Runge's function to illustrate issues of model complexity and overfitting. We compared Ordinary Least Squares (OLS), Ridge, and LASSO regressions, and evaluated their performance with resampling methods. We performed the optimization using both stochastic gradient descent (SGD) and its modern variants, such as Adam and RMSProp. Our results show that regularization improves the stability of the polynomial fits, while advanced optimizers improve convergence speed. The findings highlight the importance of balancing bias and variance to achieve better predictions.

I. INTRODUCTION

Regression methods are central in statistical learning, widely used to approximate functions, analyze data, and test theoretical models. Polynomial regression provides a simple but illustrative case, showing the bias–variance tradeoff, and how model complexity can lead to overfitting. Runge's function highlights these challenges, as higher-degree polynomials often produce oscillations instead of improving accuracy.

In this project, we applied Ordinary Least Squares (OLS), Ridge, and LASSO regression to polynomial features of Runge's function. We assessed their performance using resampling methods such as bootstrap and cross-validation, and compared analytical solutions with optimization techniques including stochastic gradient descent (SGD) and modern variants like momentum, Adagrad, RMSProp, and Adam.

Our aim was to study how regularization, optimization, and resampling affect regression performance and illustrate the balance between bias and variance. The report is organized as follows: Section II outlines the methods, Section III describes the implementation, Section IV presents results and discussion, and Section V gives our conclusions and perspectives.

II. METHODS

A. Method 1/X

- Space for methods

B. Implementation

- Space for implementation

C. Use of AI tools

- ChatGPT has been used to generate a text skeleton to quicken the process of writing, but also as a debugging tool to aid in solving some of the coding issues.

III. RESULTS AND DISCUSSION

- Here is the place for future results. So far we present 2 required figures along with one optional but important:

Figure 1: My dog.

Figure 1 displays our professor's dog. I decided to leave this figure for scientific purposes.

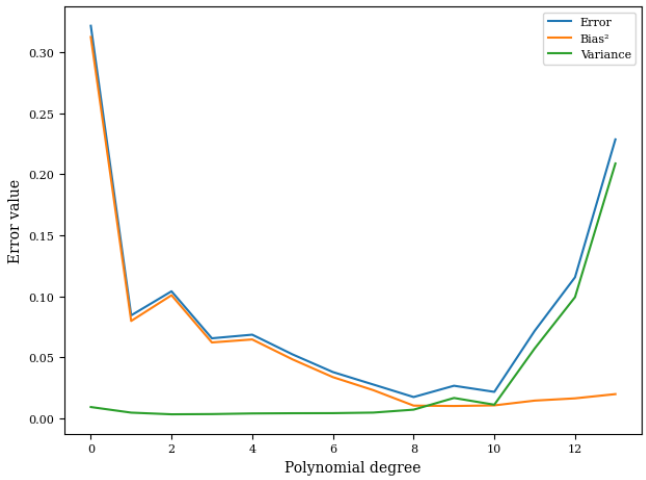

Figure 2: The bias-variance tradeoff.

Figure 2 [1]presents the trade-off between bias and variance.

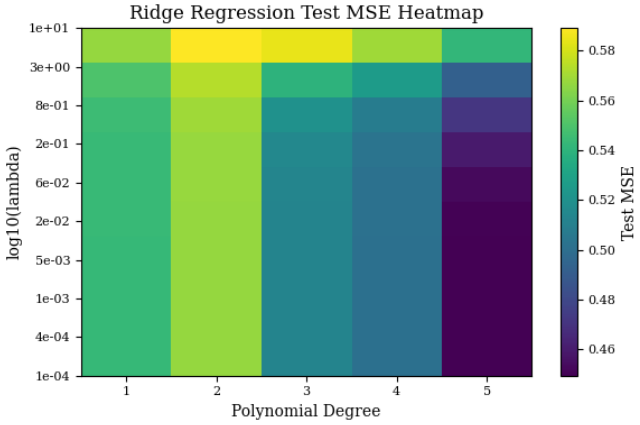

Figure 3: MSE values for Ridge Regression model vs. Polynomial degree and lambda.

Figure 3 presents the MSE of a Ridge regression model for various polynomial degrees and lambda values

IV. CONCLUSION

- Our main findings are to be determined. However, lecture notes and previous assignments give us some insight into what we can expect to achieve.

[1] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, et al., Journal of Machine Learning Research **12**, 2825 (2011).