Exercises week 39
FYS-STK3155 - Project 1

Stan Daniels, Teresa Ghirlandi, Francesco Minisini, Carolina Ceccaci
University of Oslo
(Dated: September 26, 2025)

Regression in machine learning is a fundamental technique for predicting outcomes based on input features. It finds relationships between variables so that predictions on unseen data can be made. A major challenge arises as model complexity increases: low-degree models may underfit, while high-degree polynomial regression can become unstable and overfit the data. In this project, we study the Runge function, a well-known function that highlights the difficulties of high-degree polynomial interpolation. We apply Ordinary Least Squares, Ridge, and Lasso regression, complemented by gradient descent and its variants, including momentum, Adagrad, RMSprop, and ADAM. Resampling techniques such as bootstrap and cross-validation are used to evaluate model generalization and analyze the bias-variance trade-off. The results show that OLS fits become highly unstable for high polynomial degrees, while Ridge and Lasso regularization significantly improve stability and predictive accuracy. Gradient descent methods reproduce the analytical results, though their performance depends strongly on learning-rate strategies. Overall, the study highlights the importance of regularization and resampling for controlling overfitting and improving the reliability of regression models.

I. INTRODUCTION

The aim of this project is to study various regression methods, such as Ordinary Least Squares, Ridge Regression, and Lasso Regression. It focuses on fitting polynomials to a specific one-dimensional function, the Runge function:

$$\frac{1}{1+25x^2}$$

The Runge function shows the difficulties of high-degree polynomial interpolation and this makes it an ideal test case to compare the performances of the different methods. First, an OLS regression analysis is performed, exploring the dependence on the number of data points and the degree of polynomial. The analysis is then extended to Ridge and Lasso regressions, which add a regularization parameter λ. Gradient descent methods are implemented. The analysis starts with the standard gradient descent method, but then, to improve efficiency and convergence, several variants of the gradient descent have been developed, such as momentum, stochastic gradient descent and adaptive methods, including Adagrad, RMSprop, ADAM. The performance of OLS, Ridge and Lasso is then compared with the gradient descent-based optimization methods. In order to evaluate model performance and investigate bias-variance trade-off, resampling techniques such as bootstrap and cross-validation are applied, highlighting how different choices of model complexity and regularization affect the trade-off between bias and variance. These techniques provide insight into the stability of the models and the reliability of their predictions. Overall, this project aims to illustrate the strengths and the limitations of each method.
The structure of this project is as follows:

- Section II "Methods", describes the regression techniques and optimization algorithms, as well as the resampling methods.
- Section III "Results and Discussion", presents the numerical results, compares the performance of the different methods and discusses their implications in terms of bias-variance trade-off
- section IV "Conclusion", summarizes the main results and the insights gained from the methods studied.

II. METHODS

A. Method ..

[1]

B. Implementation

[2]

C. Use of AI tools

A large language model (LLM) is used to assist in generating Python code snippets for plotting and analysis. The LLM was used only for code suggestions, explanations, debug while all code execution, results, and interpretation were performed independently.

III. RESULTS AND DISCUSSION

As shown in Figure 1, ...
As shown in Figure 2, ...

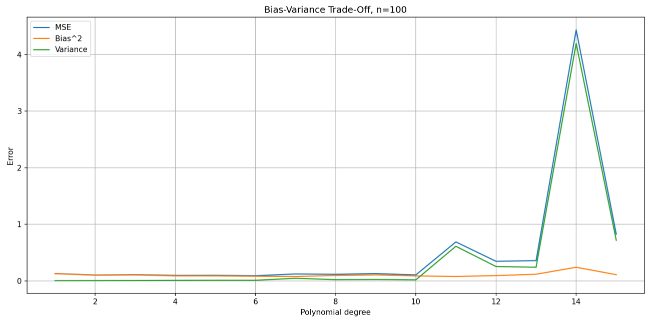

Figure 1: Bias-variance trade-off for polynomial regression on the Runge function. The blue curve shows the MSE, the orange curve shows the squared bias and the green curve represents the variance.

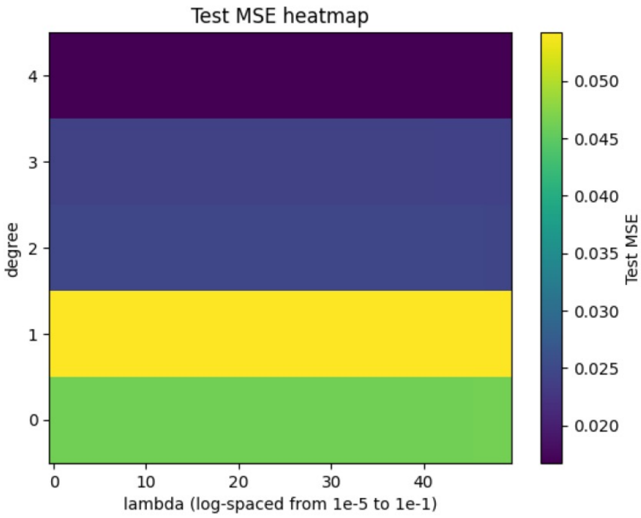

Figure 2: Heat map showing the Ridge regression error as a function of polynomial degree d and the parameter λ. Colors indicate the magnitude of the error, with darker shades corresponding to lower errors.

IV. CONCLUSION

[1] T. Hastie, R.Tibshirani, and J.Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition. Springer Series in Statistics* (Springer, New York, 2009), URL `https://link.springer.com/book/10.1007%2F978-0-387-84858-7`.

[2] 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).