Report Template
FYS-STK3155 - Project X

Naden Jarel Anthony Koh
University of Oslo
(Dated: September 25, 2025)

In this project, we explore regression methods and resampling techniques as applied to polynomial fitting of the Runge function. We first implement Ordinary Least Square (OLS), Ridge and Lasso Regression for the Runge function. We then explore Gradient Descent Techniques and find more efficient techniques in ways such as using momentum, Stochastic Gradient Descent, as well as other methods and evaluate their Accuracy. Lastly, look at Bias-variance trade-off and the bootstrap resampling technique. These analyses are important because they illustrate how model complexity, regularization, and optimization interact—concepts that are central to building reliable and interpretable machine learning models.

I. INTRODUCTION

Traditional OLS regression relies on having $X^T X$ be invertible. This assumption tends to be violated when features are highly correlated. Regularization methods such as Ridge and Lasso regressions addresses this by introducing a penalty term to the cost function, albeit with certain limitations which will be explored in Results and Discussions. We also introduce a plain Gradient Descent Method. However, with this method, we have to be careful with our choice of learning rate. A high learning rate could lead to unstable convergence, while a low learning rate can lead to slow convergence. Furthermore, it is important to note that finding a local minima does not mean that we have found the best set of parameters. We first aim for faster convergence in Gradient Descent with Momentum, before exploring other optimization techniques as well as Stochastic Gradient Descent to help us find a better estimation of the optimal parameters. We then explore Resampling Methods which are useful when we have a finite dataset, and we don't know the model's true underlying distribution.

The report is structured as follows. Section I introduces the theoretical background for regression, bias-variance trade-off and resampling. Section II describes the methods and algorithms implemented. Section III presents results from OLS, Ridge and Lasso Regresion, with analyses of MSE and parameter behaviour. Finally, Section V concludes with reflections on the bias-variance trade-offs involved.

Our implementations make use of the `scikit-learn` library [2], which provides efficient tools for regression, resampling, and model evaluation. In addition, the theoretical foundations of our work are guided by insights from Hein [**?**] and Hastie et al. [1].

II. METHODS

A. Use of AI tools

B. Use of AI tools

In the course of this project, we made use of a large language model (LLM), ChatGPT, to support both the implementation and the reporting process. The LLM was used to debug LaTeX compilation errors for exercise 39, as well as in the debugging of code implementations in exercise 2.

III. RESULTS AND DISCUSSION

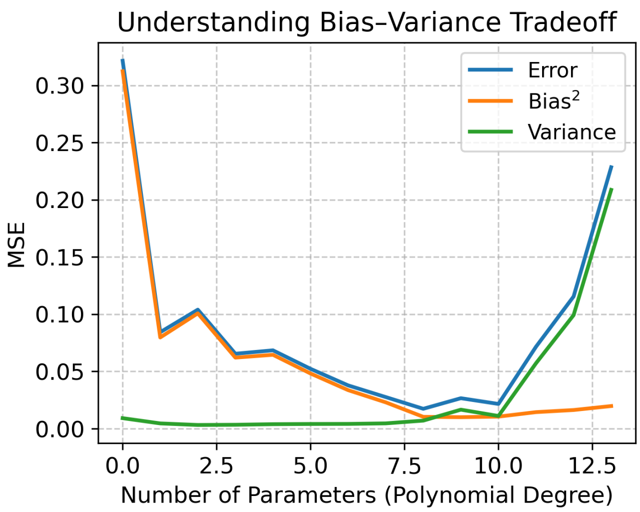

Figure 1: Illustration of the Bias–Variance trade-off.

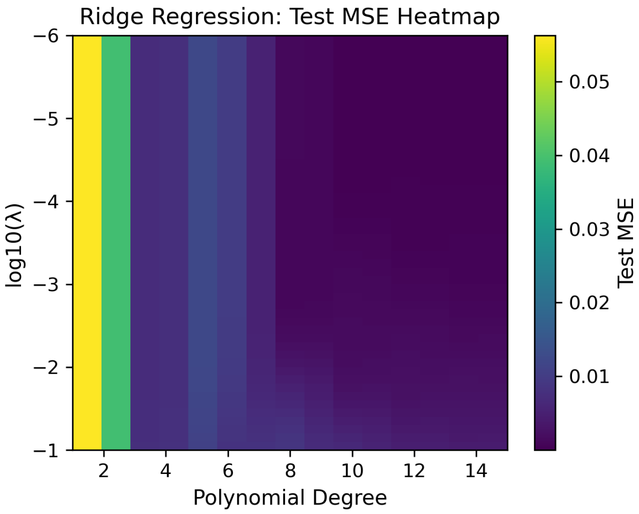

Figure 2: Illustration of how Lambda and Polynomial Degree affects MSE.

As shown in Figure 2, the test MSE depends strongly on both the polynomial degree and the regularization parameter λ. The Test MSE seems to decrease going up, which is when lambda decreases. At the same time, Test MSE also decrease as the polynomial degree increases, however it increases as the polynomial degree gets really high.

[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.

[2] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.