Oliver Brown
Mentor: Antonio Olivas
Project: Function Estimation Using Reproducing Kernel Hilbert Spaces

This project investigated the theory and application of Reproducing Kernel Hilbert Spaces (RKHS) in the context of nonparametric regression. The work began with an overview of Hilbert spaces, emphasizing their completeness and inner product structure, which allow for the representation of infinite-dimensional function spaces. A key theoretical result at this stage was the Riesz Representation Theorem, which guarantees that every bounded linear functional on a Hilbert space can be uniquely represented as an inner product with a fixed element of that space. A linear functional is simply a function that takes a function as input and returns a real number, while preserving linearity. This theorem is fundamental to RKHS theory, as it underpins the existence of a reproducing kernel: for each point x, the evaluation functional $f \to f(x)$ is linear and bounded, and thus corresponds to an inner product with a unique function K_x in the space.

The concept of RKHS was then introduced as a Hilbert space of functions in which evaluation at any point can be written as an inner product with a corresponding kernel function. A central technique explored was the kernel trick. Rather than explicitly mapping data into high-dimensional feature spaces, kernel functions are used to compute inner products directly. Several standard kernels were considered, including linear, polynomial, and Gaussian, each inducing different structures in the associated function space.

Mercer's Theorem provided an additional theoretical foundation. It states that any continuous, symmetric, positive semidefinite kernel can be expressed through an expansion involving eigenfunctions in L^2, the space of square-integrable functions. That is, L^2 contains all functions f such that the integral of $|f(x)|^2$ over a given domain is finite. This result validates the construction of RKHS from such kernels and offers insight into how the kernel function shapes the complexity of the function class.

The next stage of the project focused on implementing kernel ridge regression (KRR), a regularized least-squares method within the RKHS framework. This method introduces a tuning parameter λ to control the trade-off between fidelity to the data and model complexity. KRR is particularly well-suited to problems involving high-dimensional predictors, where linear models may be insufficient.

Empirical evaluation was conducted on simulated data and the Boston Housing dataset, which includes 13 predictors and a response variable for median home value (in thousands of dollars). Five-fold cross-validation on the housing dataset showed that the Gaussian KRR model performed best (MSE 14.25), using bandwidth set to the median pairwise distance—a common method for controlling the smoothness of the function estimation. The 2nd-order polynomial KRR model also performed well (MSE 16.57), outperforming quadratic ridge regression (MSE 19.04). The Sobolev KRR model yielded an MSE of 18.38. Linear KRR (MSE 24.52) closely matched linear ridge regression (MSE 24.14). The cosine kernel performed poorly (MSE 33.80). Overall, nonlinear kernels—particularly Gaussian and polynomial—captured the data's structure most effectively.

This study illustrates the theoretical elegance and practical strength of RKHS methods. By selecting appropriate kernels and tuning parameters, kernel-based models provide a flexible and effective approach to high-dimensional regression.