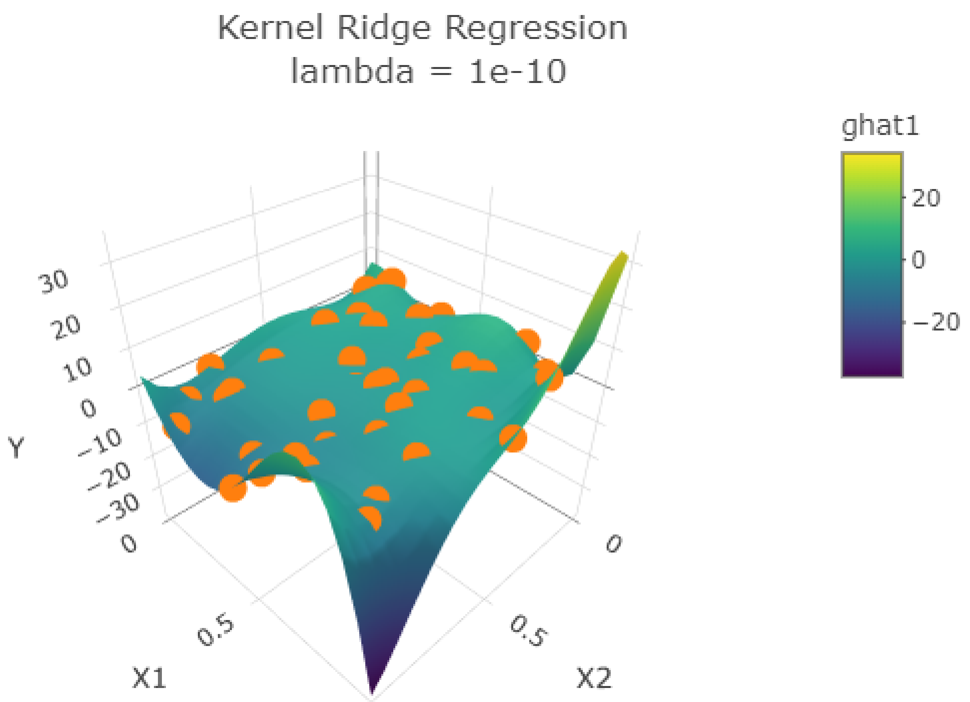
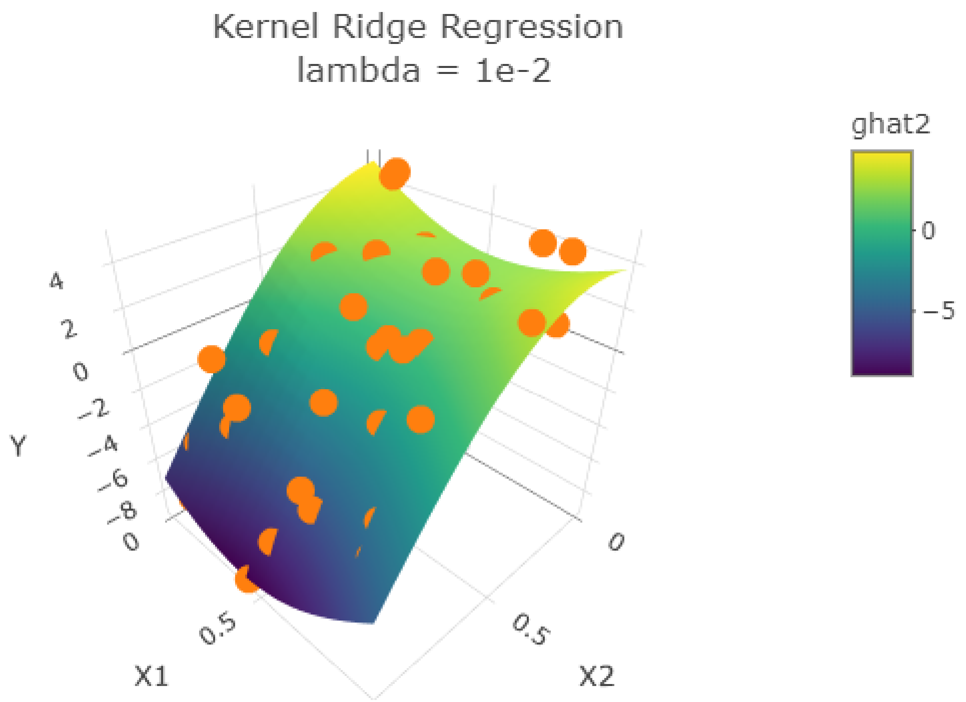
BIOST 527 - HOMEWORK 4

KATIE WOOD

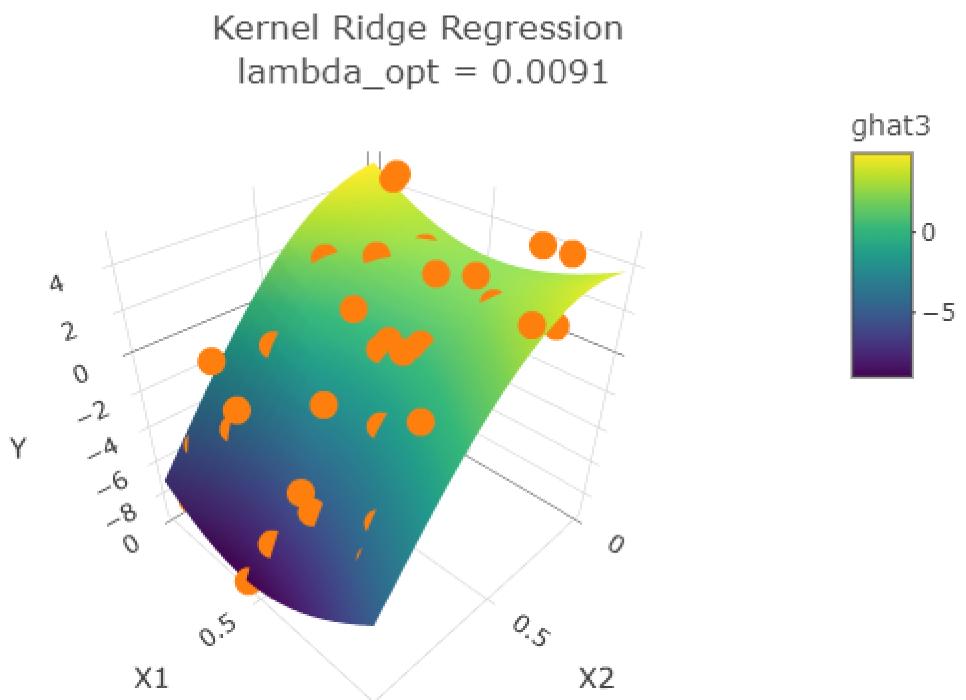
Problem 1



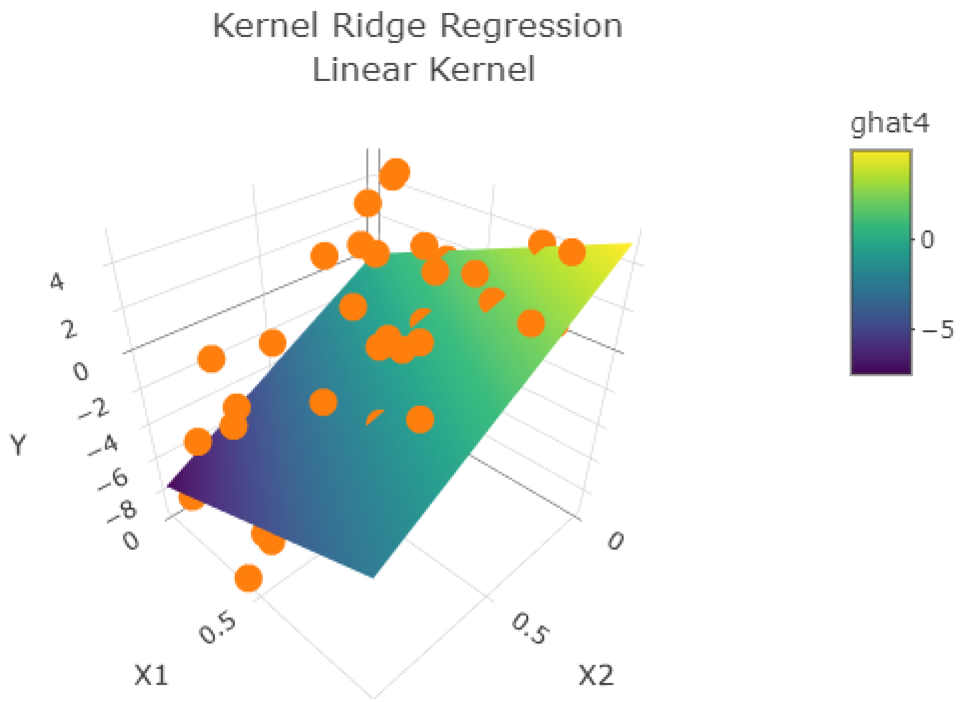
Above we show the result of Kernel Ridge Regression using a Gaussian kernel with the penalty $\lambda = 10^{-10}$. As we expect from a small penalty, $\hat{g}_1$ provides an excellent fit to the training data. However, we suspect that $\hat{g}_1$ may be over-fitting to the noise in the data since the true g^* is a hyperbolic paraboloid (degree 2 in each independent variable), meanwhile $\hat{g}_1$ shows much more texture than degree 2 would allow. Particularly near the boundaries, $\hat{g}_1$ changes rapidly, which is not reflected in the data.



Above we show the result of Kernel Ridge Regression using a Gaussian kernel with the penalty $\lambda = 10^{-2}$. As we expect from a larger penalty, $\hat{g}_2$ does not fit the training data as seamlessly as $\hat{g}_1$, but may nonetheless provide a more accurate representation of the true underlying g^*. Indeed, $\hat{g}_2$ displays what looks like degree 2 curvature.



Above we show the result of Kernel Ridge Regression using a Gaussian kernel with the optimal penalty $\lambda^* \approx 0.0091$. According to Generalized Cross-Validation, $\hat{g}_3$ should provide the most accurate representation of the true underlying g^*. We note that this value of λ^* is only slightly smaller than the $\lambda = 10^{-2}$ we used for $\hat{g}_2$. Thus, similar to our remarks for $\hat{g}_2$, we notice that $\hat{g}_3$ displays curvature reminiscient of a hyperbolic paraboloid (which is degree 2 in each independent variable), which fits the training data well without over-fitting to noise.



Lastly, we show the result of Kernel Ridge Regression using a linear kernel with the optimal penalty $\lambda^* \approx 0.086$. Here, $\hat{g}_4$ under-fits the training data. This makes sense because the linear kernel can only capture behavior up to degree 1 in each independent variable, meanwhile the true g^* is degree 2.