

1 Results

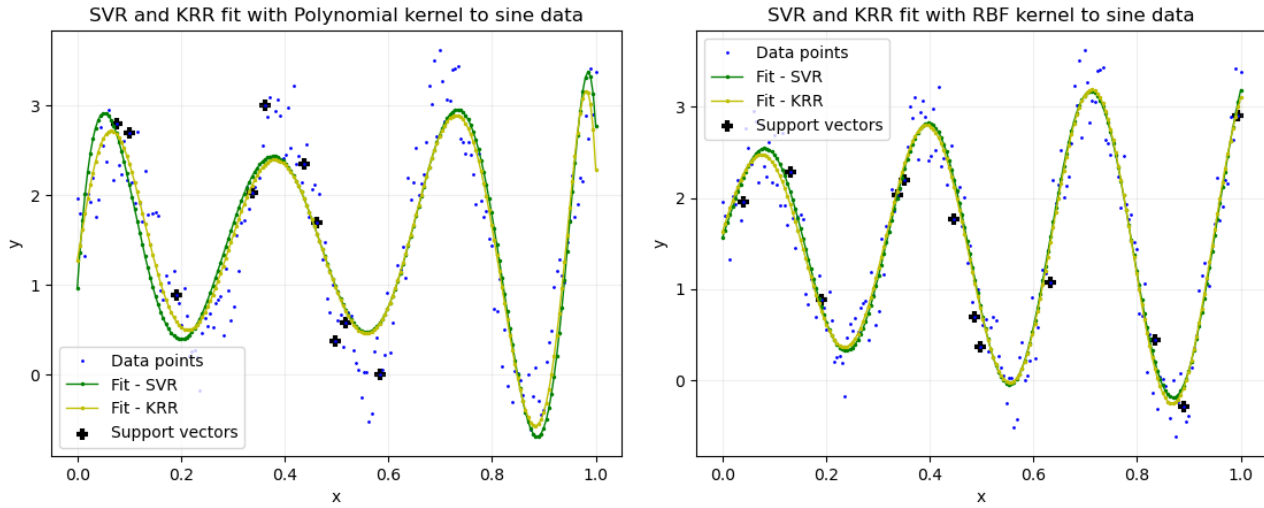
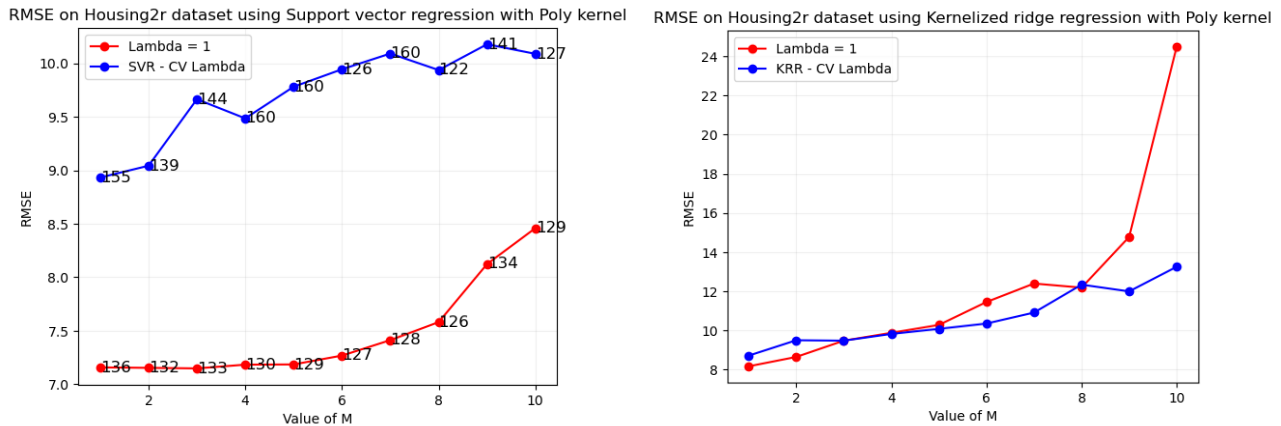


Figure 1: Fitted data on 1D Sine dataset

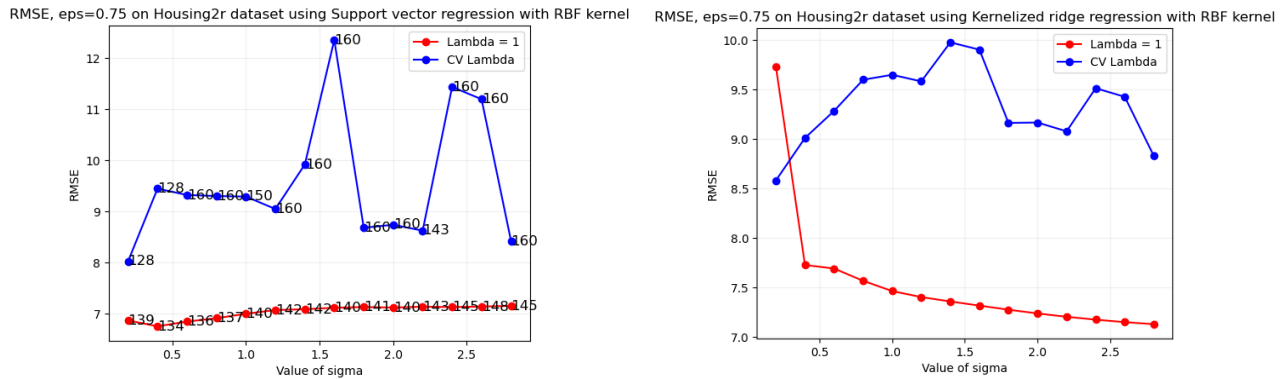
In Figure 1 SVR (*Support Vector Regression*) and KRR (*Kernelized Ridge Regression*) fit on 1D Sine data set is presented. Two kernels are used: Polynomial and RBF. All kernels are fitted with 5-fold cross validation (details in Table 1). We can see that polynomial kernel has high value of parameter M , while RBF kernel is fitted with higher λ penalizing regularization parameter.

Model	Kernel	λ	(σ/M)	RMSE
SVR	Polynomial	e^{-25}	12	1.4624
SVR	RBF	0.1	0.1	1.4703
KRR	Polynomial	e^{-20}	10	0.4337
KRR	RBF	0.1	0.1	0.2768

Table 1: Parameters for 1D Sine

Figure 2: RMSE vs. Kernel parameter value (Polynomial degree M)

In the Figures 2 and 3, performance of models is presented on Housing2r dataset. KRR and SVR is used with best parameters, which were evaluated via 5-fold CV. Each kernel is fitted with both models and compared with default parameters (e.g. $\lambda = 1$).

Figure 3: RMSE vs. Kernel parameter value (RBF parameter σ)

2 Conclusion

From the plots, we can see that the performance of the model is slightly worse on this data set compared to the kernelized ridge regression. Nevertheless, I would prefer SVR to be my learning algorithm, due to having one more hyperparameter to tune in ϵ , making it more flexible than kernelized ridge regression. Also, with these small sets of data we can use more effectively SVR, which doesn't have closed form solution as KRR and we can also benefit from faster prediction time in SVR. Besides mentioned, mean values of all λ for KRR is 8.02, while SVR has significantly smaller value 0.024. This fact indicates that housing model is sparse model, where SVR is more adequate as sparse model solution from KRR, which is trying to use regularization to achieve sparsity. Using bias-variance decomposition (for polynomial kernel, $M = 1, \lambda = 1$), estimated values of SVR variance is 0.299 with bias of 62.896, while KRR decomposition estimated values for bias of 103.248 and significant bigger variance of 1.537.