Ridge Regression

If we use a 10th order polynomial function to fit the data, we observe from the expression for the estimated function that the ***estimated polynomial coefficients have a very large magnitude***. This is especially evident for the **higher order polynomials**. **Ridge regression controls the magnitude of these polynomial coefficients by introducing the parameter alpha**. *Alpha is a parameter we select before fitting or training the model*. Each row in the following table represents an increasing value of alpha.

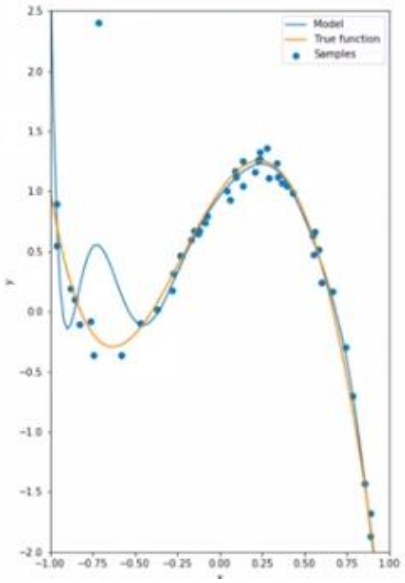
Ridge Regression

$$\hat{y} = 1 + 2\,x - 3\,x^2 - 2x^3 - 12\,x^4 - 40x^5 + 80x^6 + 71x^7 - 141x^8 - 38x^9 + 75x^{10}$$

Alpha	x	x^2	x^3	x^4	x^5	x^6	x^7	x^8	x^9	x^{10}
0	2	-3	-2	-12	-40	80	71	-141	-38	75
0.001	2	-3	-7	5	4	-6	4	-4	4	6
0.01	1	-2	-5	-0.04	0.15	-1	1	-0.5	0.3	1
1	0.5	-1	-1	-0.614	0.70	-0.38	-0. 56	-0.21	-0.5	-0.1
10	0	-0.5	-0.3	-0.37	-0.30	-0.30	-0.22	-0.22	-0.22	-0.17

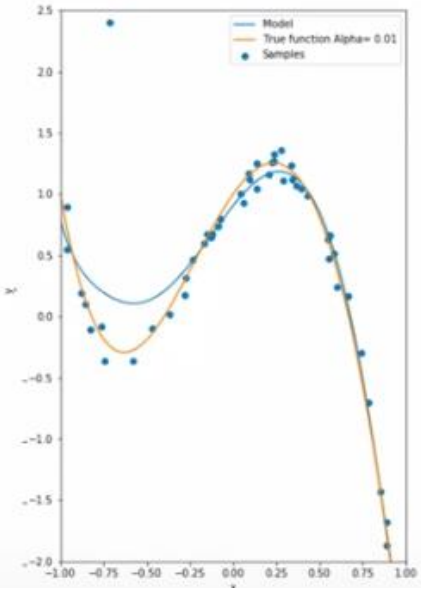
Ridge Regression

alpha
0
0.001
0.01
1
10



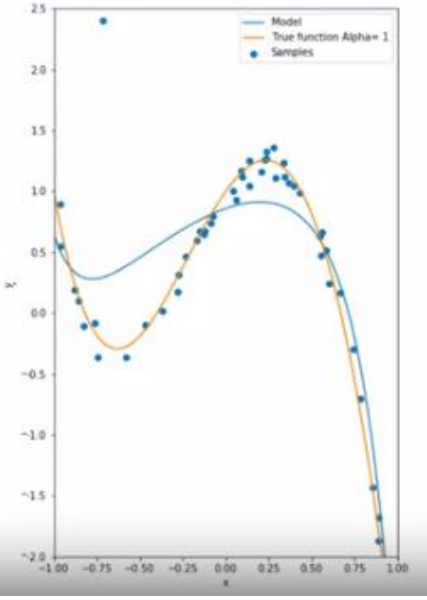
Ridge Regression

alpha
0
0.001
0.01
1
10



Ridge Regression

alpha
0
0.001
0.01
1
10



We use cross-validation to select Alpha!

Ridge Regression