309551101郭育麟

ML\_HW05

1 . Gaussian Process

1-1:

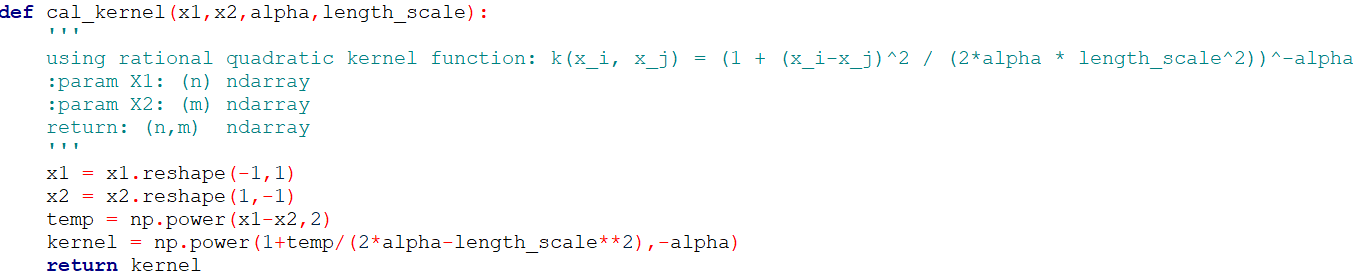
Kernel function : Rational quadratic kernel

Alpha : 1 , Length\_scale = 1

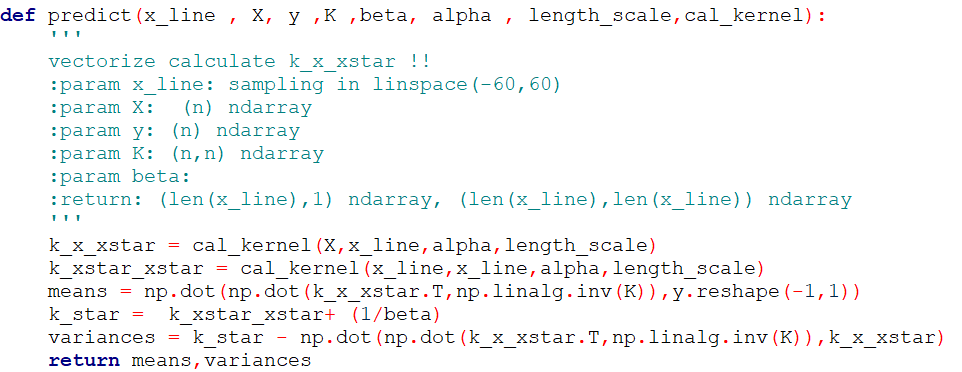
In this homework , first using cal\_kernel() to calculate data points’ similarity and return a kernel matrix . Then use predict() function to evaluate the new points (-60,60) each mean and variance . Finally use plot\_gaussian() function to draw the result .

Data points (raw data) are sample from

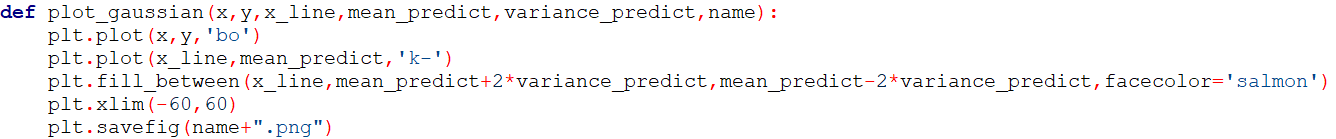
Cal\_kernel():



Predict():

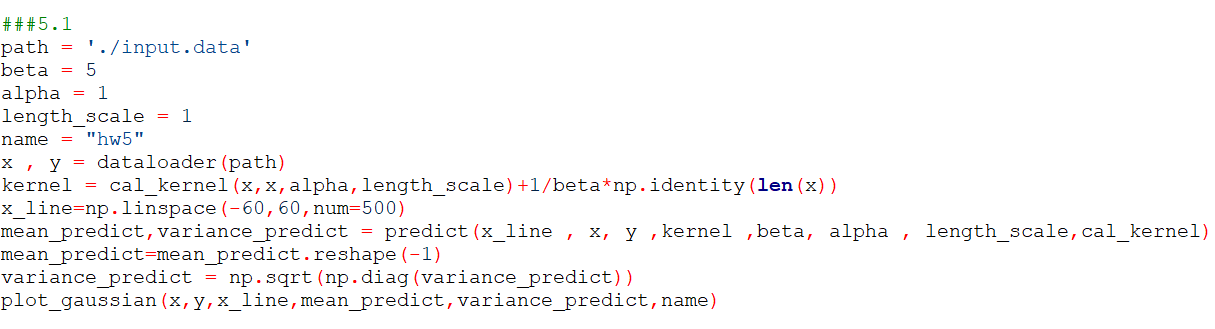


plot\_gaussian():

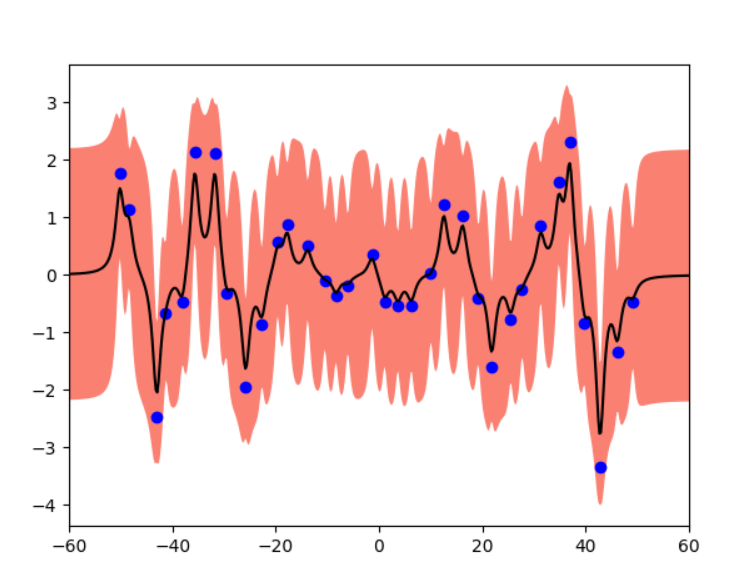


Given raw data ,mean and variance can plot the graph , 95% confidence interval means that .

Main():



Result:



In the blue dots (raw data) interval , the variance would smaller than new points (-60~60) . Raw data has more specific range than new points .

1-2:

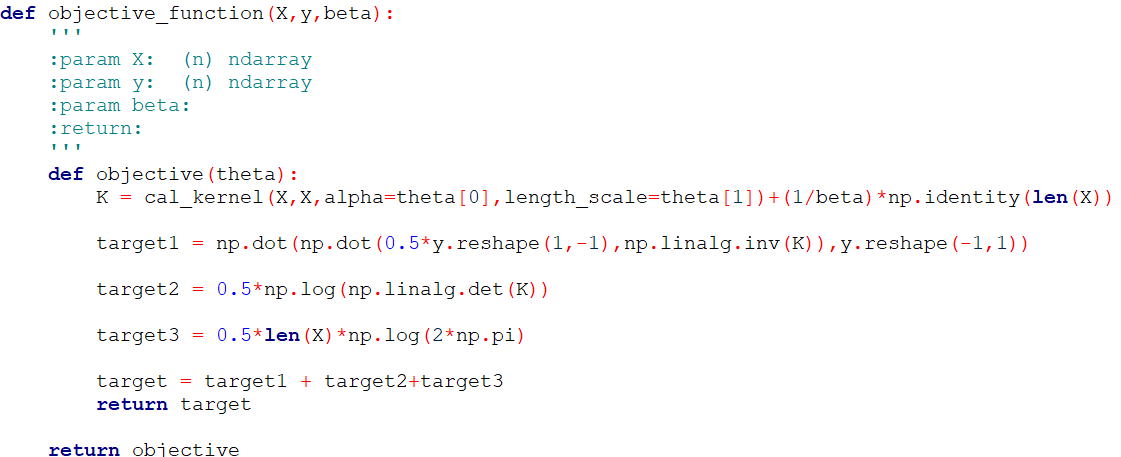
In this exercise , we need to optimize our parameter(alpha and lengh\_scale) first . I use scipy.optimize minimize the negative likelihood function . After minimize I can get the better parameters so it can predict result same as 1-1 .

Initially function:

After differential could get objective function .

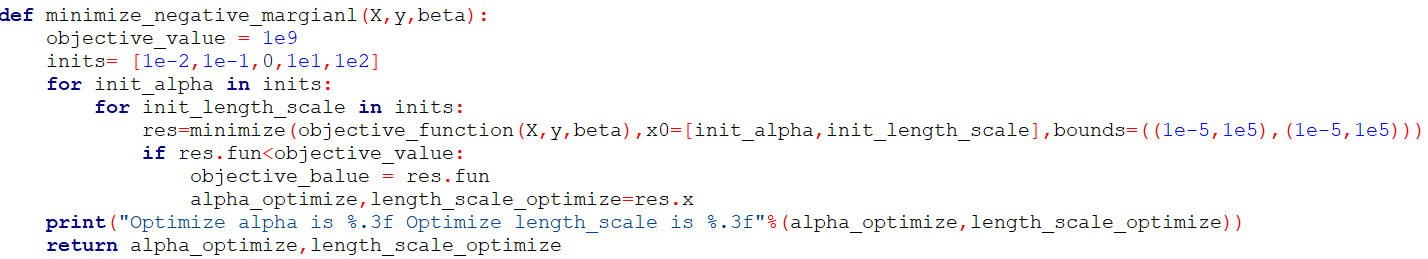
Objective function :

Objective\_function():



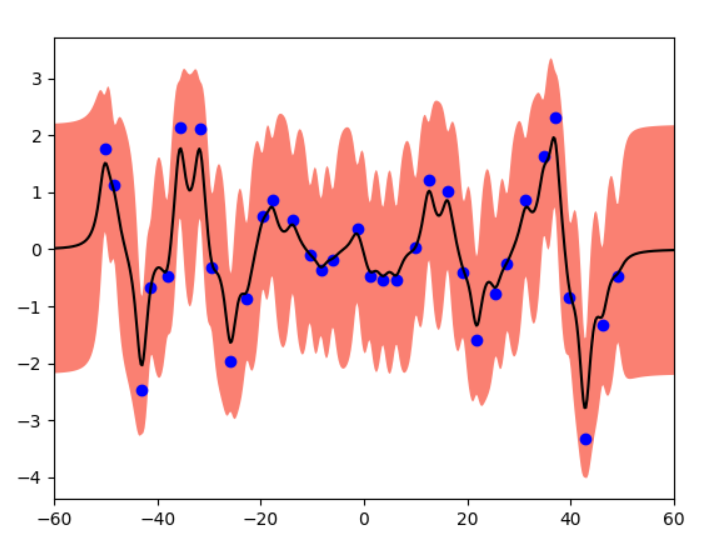
In this function , calculate the objective function value .

Minimize\_negative\_marginal():



In this function , initial sets parameters from [1e-2,1e-1,1e1,1e2] and want to minimize the objective function . The reason why choose [1e-2 , 1e-1, 1e1 , 1e2] is avoid the minimize value is local minimum not global minimum . Finally , return alpha and length\_scale .

Result:



Compare:

In result 1-2 , the variance in blue dot (raw data) becomes smaller . The parameters which has been optimized would let the likelihood be maximum . So it can has better performance than 1-1