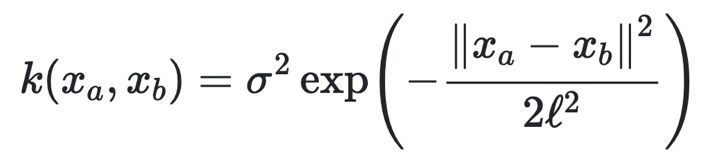
ML HW05 – Gaussian Processing

309553048\_多工所\_陳柏丞

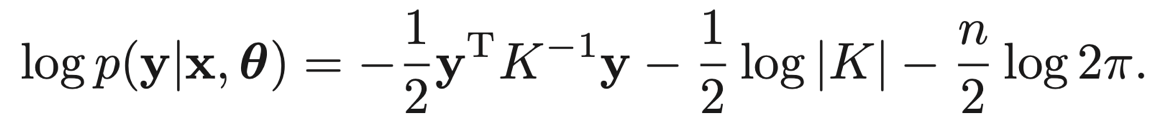
* **Code Explanation：**

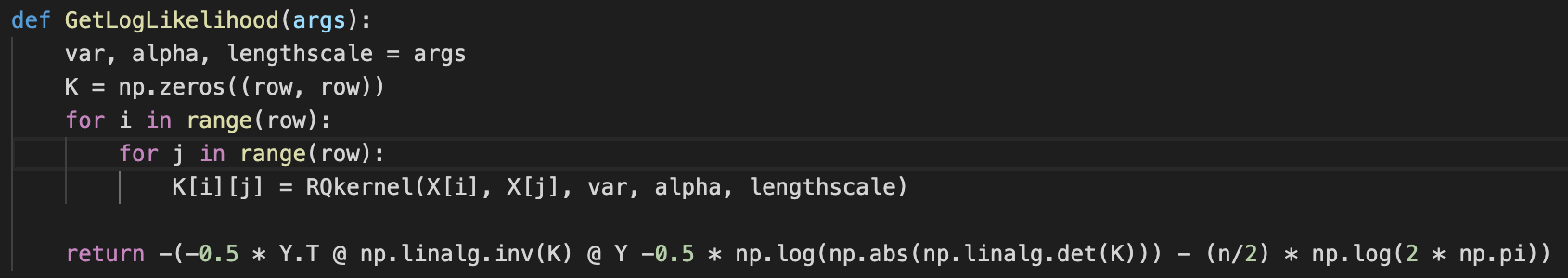
1. rational quadratic kernel (Xa, Xb, variance, alpha, lengthscale)



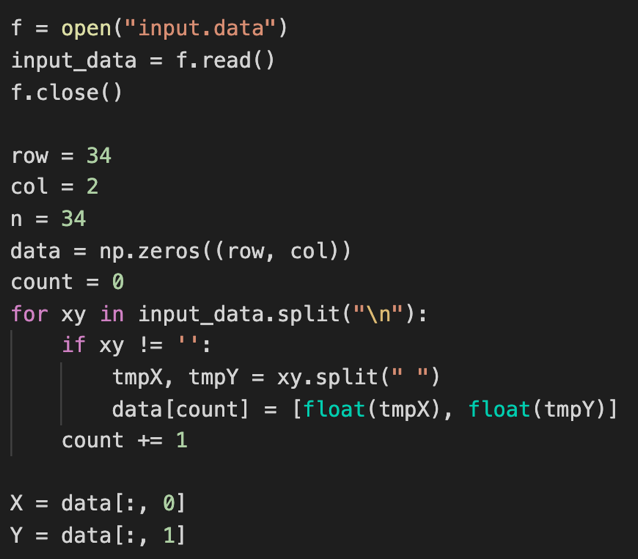


1. negative marginal log-likelihood

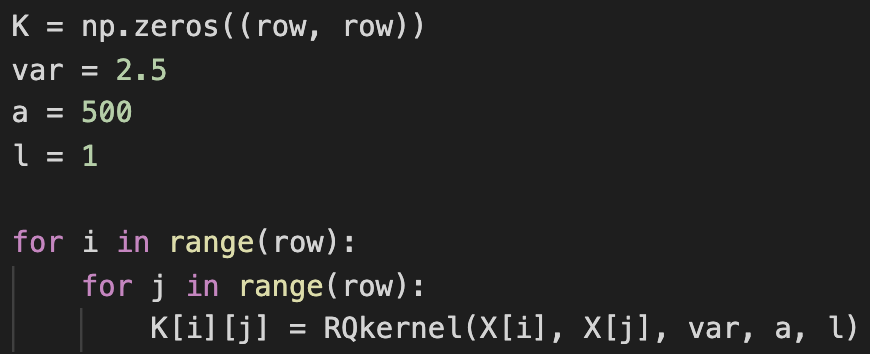




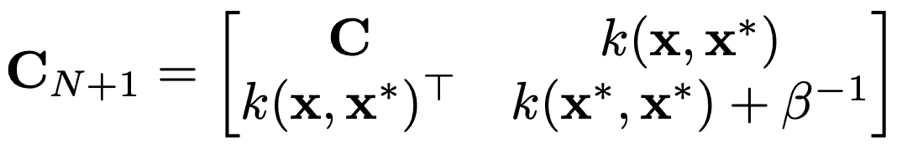
1. read data & separate data to X, Y



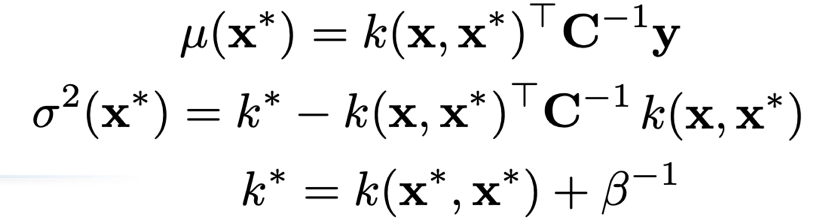
1. calculate covariance matrix by rational quadratic kernel

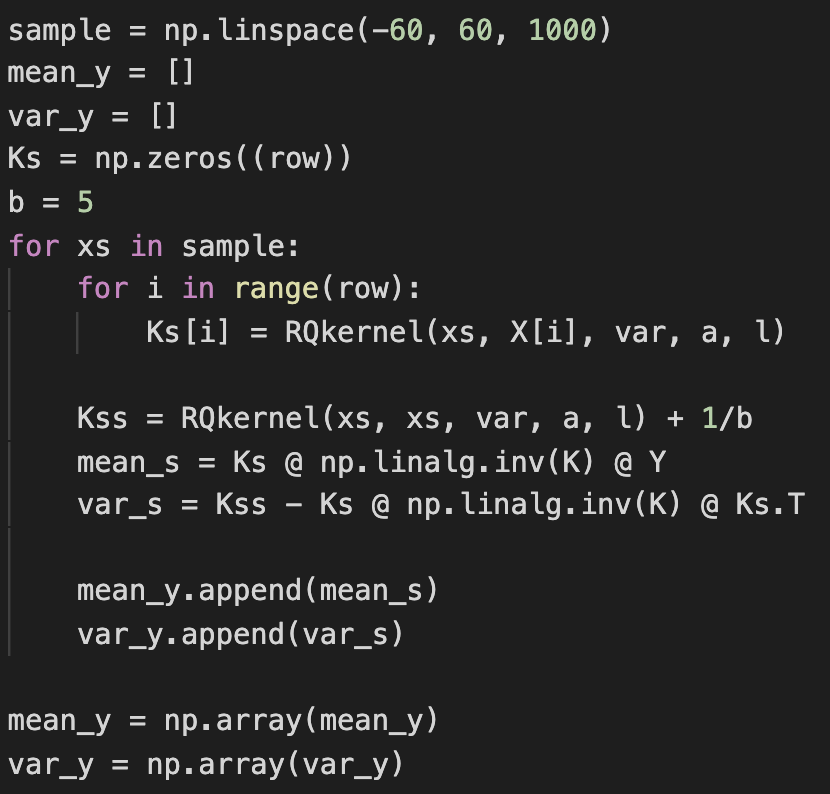


1. generate some sample & get new covariance matrix

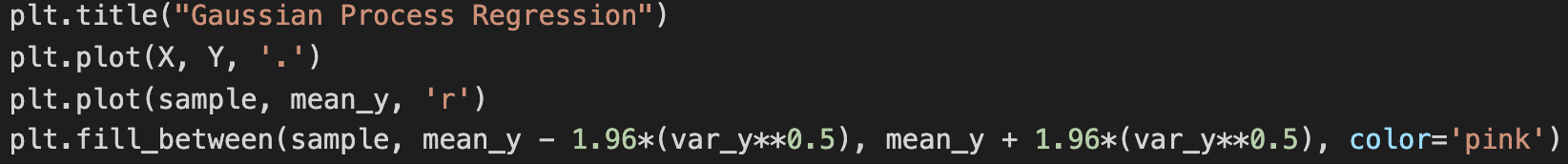


get prediction of mean and variance (conditional gaussian)





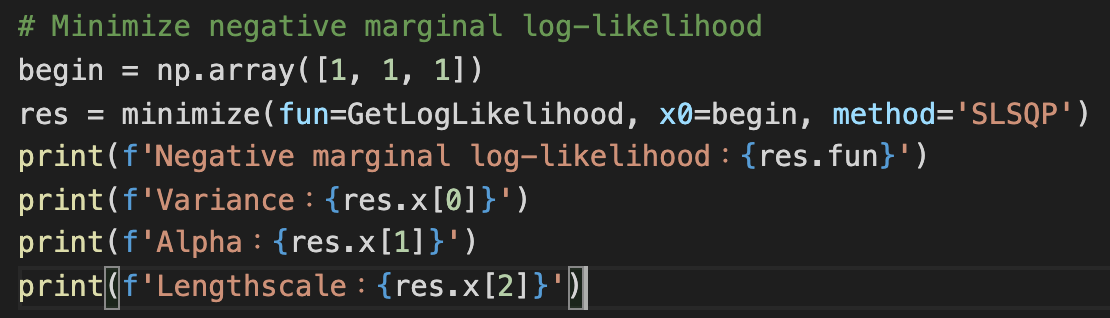
1. plot gaussian processing



Result (variance = 2.5, alpha = 500, lengthscale = 1)



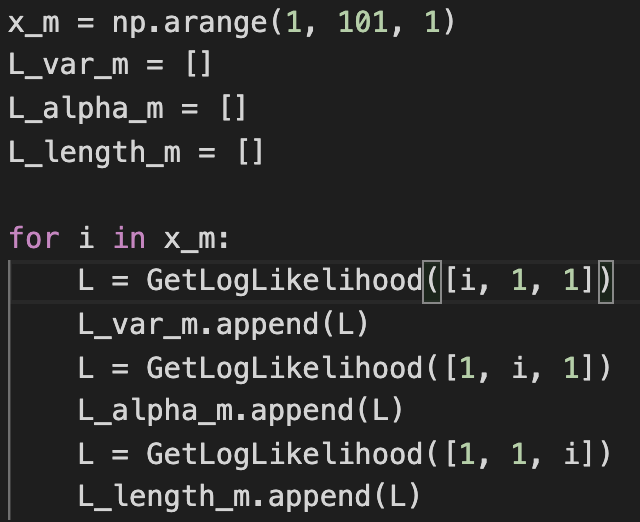
1. Minimize negative marginal likelihood



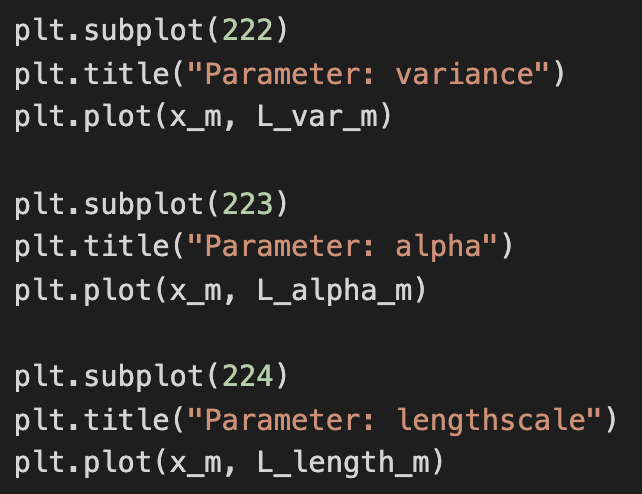
Optimize result (variance = 1.88, alpha = 8.51, lengthscale = 2.48)



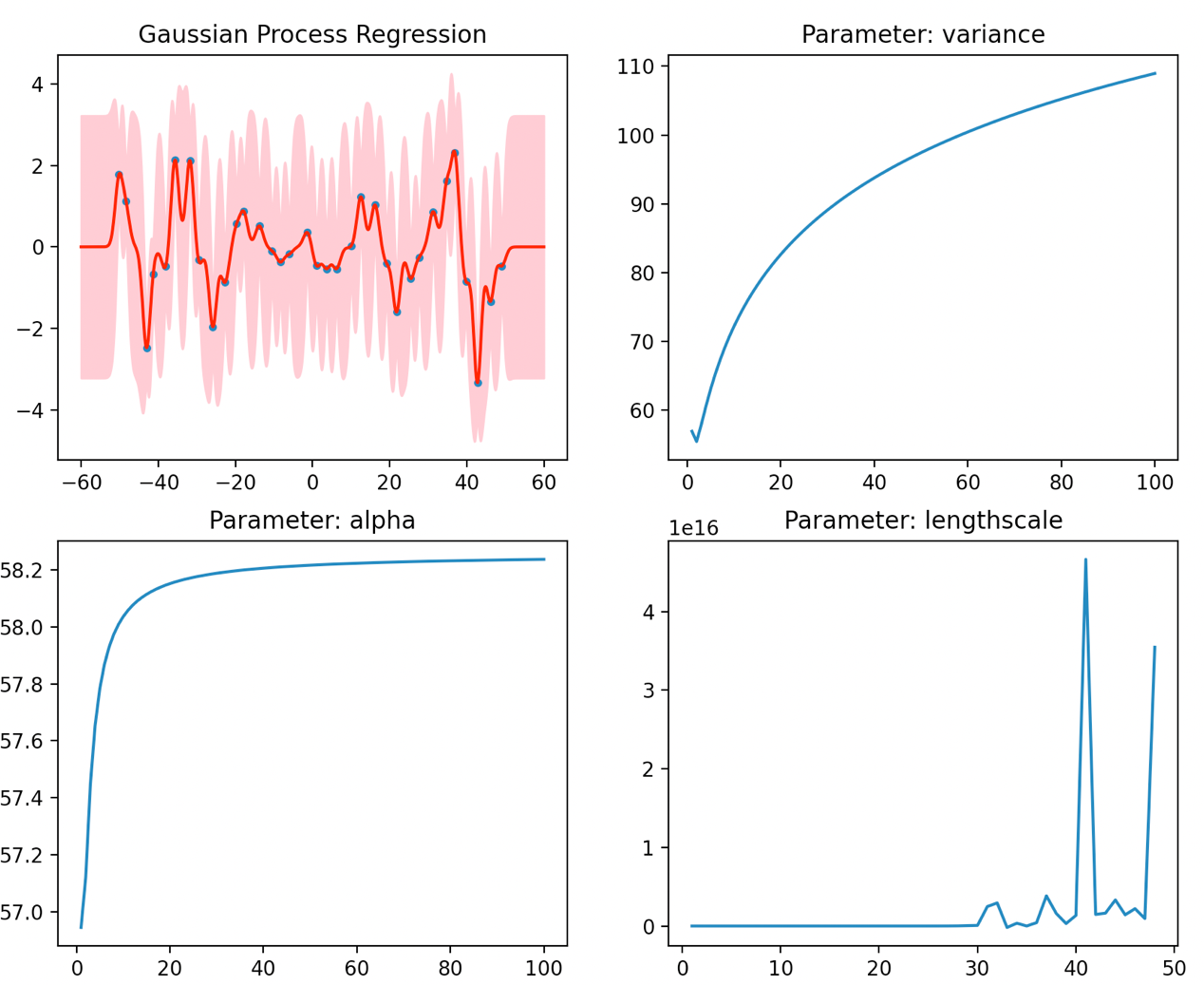
1. Testing other parameters; if a parameter set 1~100, others set 1. And get likelihood



Plot



result



* **Experiment settings and results：**

1. At first, I’m not sure which parameters is best, so I set parameters in kernel by variance = 2.5, alpha = 500, lengthscale = 1.

Result：



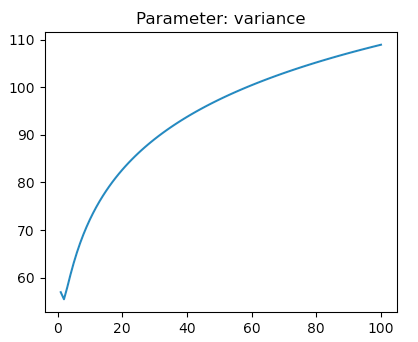
1. Then, I use the function, minimize(), and plotting with better parameter (variance = 1.88, alpha = 8.51, lengthscale = 2.48)

Optimized result：



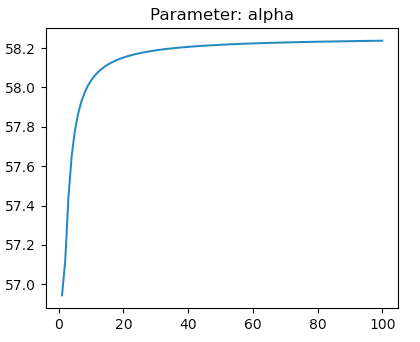
1. Furthermore, I want to know if I change one parameter in kernel and others are fixed, how likelihood could change.

Variance line graph



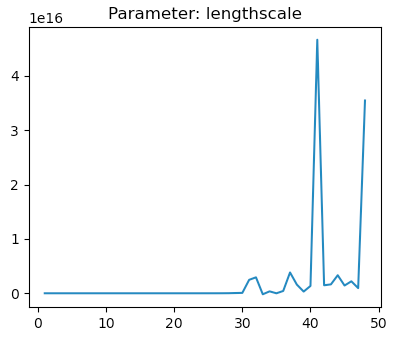
Minimal likelihood is happened by variance around 2, and others are fixed to 1.

Alpha line graph



Obviously, likelihood in bigger alpha has ceiling when others are fixed to 1.

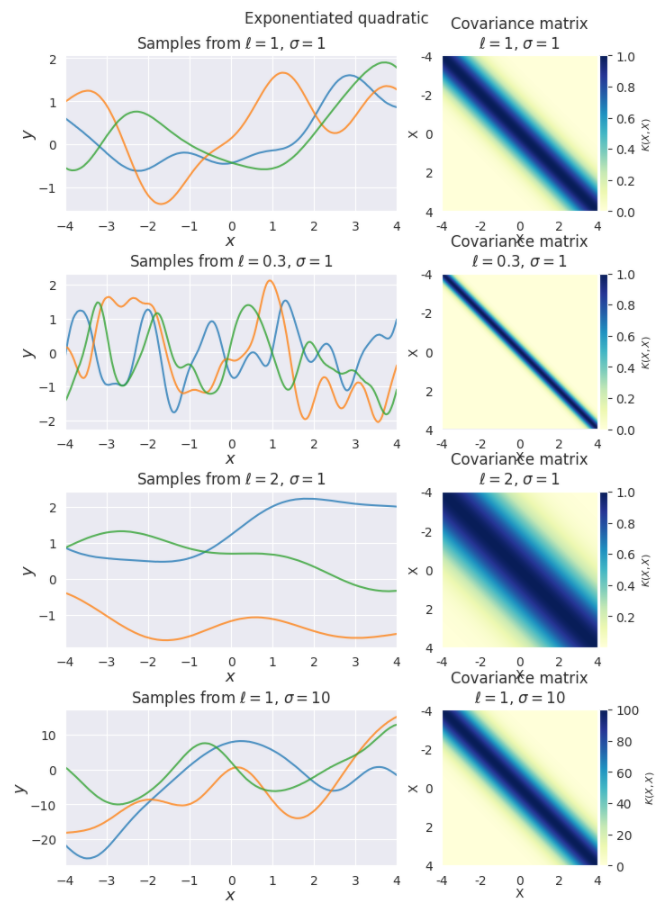
Lengthscale line graph



We couldn’t calculate likelihood when lengthscale over 50. It means the covariance matrix is singular, that is divide by zero.

* **Observations and discussion：**

We could find that which parameter in kernel is, determining the width of confidence interval.



Summary,

1. The larger variance is, the larger likelihood is. Then, the minimal likelihood happened by variance around 2, and others are fixed to 1.
2. The larger alpha is, the larger likelihood is. the likelihood has ceiling around 58.3.
3. With larger lengthscale, likelihood couldn’t be calculated because the covariance matrix is singular, that is, divided by zero.
4. If use minimize() to get minimal likelihood, our parameters are,   
   Negative marginal log-likelihood：51.989679759470846

Variance：1.8833709914285772

Alpha：8.509966881390344

Lengthscale：2.4762273397925236