

# CO496 Coursework 3

Xiaonan Chong

CO496 - mathematics for machine learning

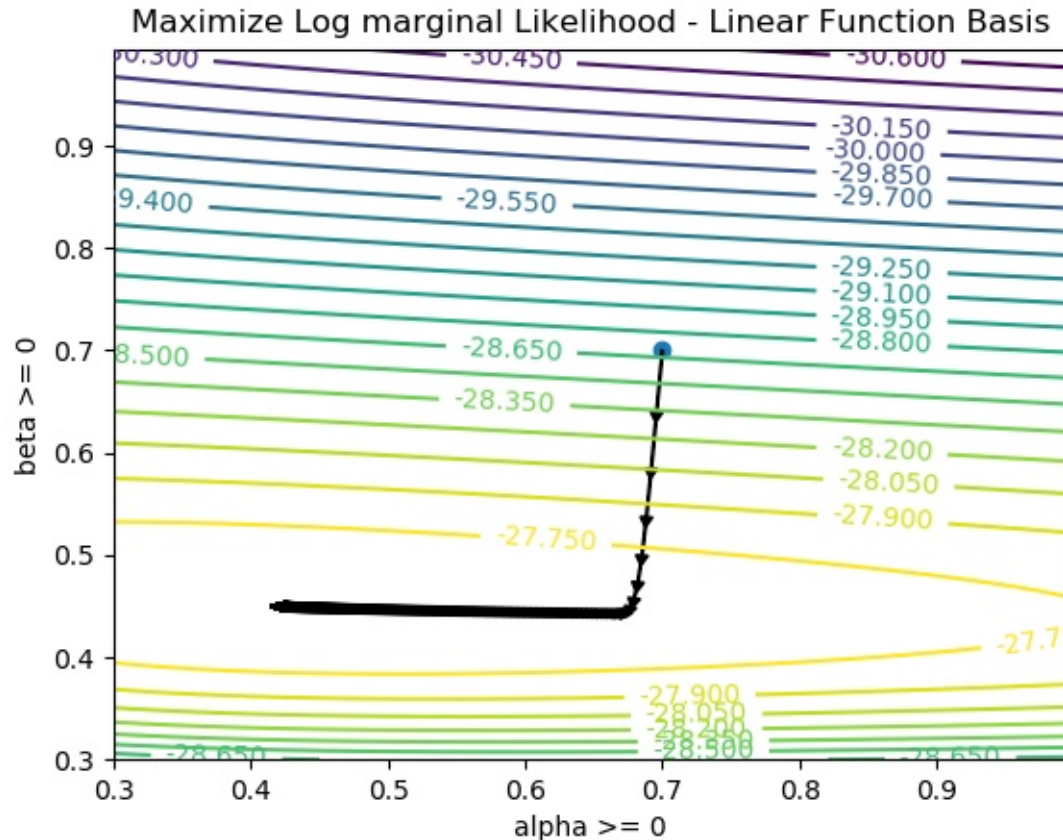
November 14, 2018

## Exercise 1. Bayesian Linear Regression

### b) Maximize log marginal likelihood - linear function basis

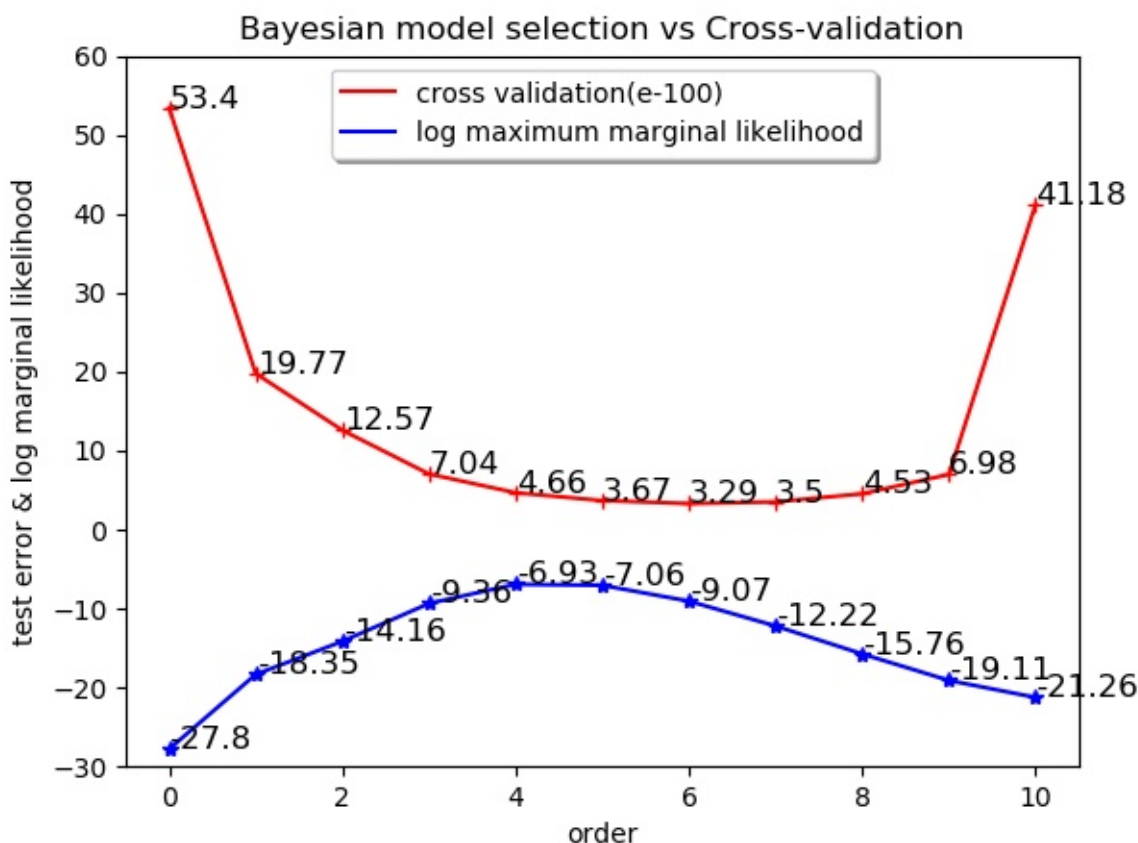
When  $\alpha = 0.42455539296074635$  and  $\beta = 0.44923131197151495$ , the maximum marginal likelihood is achieved, the log value is -27.608794629203047.

The step size is 0.005.



c) Maximize log marginal likelihood for order 0 to 10 - trigonometric function basis

By comparing both cross-validation for model selection and Bayesian model selection, we find that their conclusions slightly defer from each other. In the following picture, it is noted that cross validation favours model with order 6 but Bayesian model selection method will choose model with order 4. I believe it is because that cross-validation favors models that have high predictive ability but Bayesian model selection mean to find the model that explain the data well.



#### Merits of cross validation approach

Cross-validation targets on the practical issue that the amount of training data may be limited. By iteratively splitting the training data set and testing data set, cross-validation is able to compute the expected generalization error with low variance without cutting the amount of training data. Therefore tackles the mean issue in model selection.

One disadvantage of this approach is that the cost may be high because we need to train the model for multiple times.

#### Merits of Bayesian approach

The advantage of Bayesian model selection is that there is no need to place a prior that favours simpler models. By adopting the hierarchical generative process, given the data set, the posterior distribution of models can be computed by integrating out the model parameters. Therefore, unlike likelihood which is prone to overfitting, the marginal likelihood can well decide which model can generalize better.

#### d) Bayesian Linear Regression

