**COMMENTS**

Next you should fit a linear model to your training data using a linear least squares approach (hint torch.linalg.lstsq might be useful, although the same solution can be computed in multiple ways). Comment on the model you have produced: what is its geometrical meaning, could it be regarded as a neural network, and how well does it model the test data?

Linear least square method helps to find the best fit line that separates the data. Here in our current study we have to find a solution for the interpolation problem. The ground surface will be having different latitudes and longitudes. In order to map the variations in the ground levels the error rate is calculated using the metrics Mean square error and root mean square error. It finds the linear relationship. RBF can be used as a function to solve the interpolation problem with different basis function and by choosing the centre.

It could not be regarded as a neural network. Least square with RBF tries to model the variations in the earth surface. But for complex datasets the model is unsuitable.

For test data according based on my analysis the error rate is bit high comparing with RBF based least square method .Hence trying with random forest, decision tree would be more better than RBF based least square method.

To extend this model you can select a number of data points as the centres for a set of basis functions. Each basis function should return the exponential of minus the squared distance to the basis function centre. Each 2D input data point can be replaced with the values of the basis functions. Start with 10 basis functions which will transform the 2D data into 10D data. Fit another linear model to the new data and measure how well it models the test data. Comment on how it compares to the purely linear model.

When we compare with pure linear model, the linear model outperformed well than RBF based model. The MSE rate of linear model is 0.0108 and for RBF based least square method it is 0.0107.To transform 2D into 10 D data we tried with a list of basis functions [10, 15, 20, 25, 30,35,40].The centres are also selected for RBF using Index based centre selection. The width parameter is also tuned to reduce the error rate.

How does varying the number of basis functions affect the accuracy of the final model, and what do you think would be the impact of using all of the training data as basis function center? Is there a good way to select which training data points might make the best basis function centres?

Varying the basis functions in RBF may increase the efficiency of the model. Centre are computed by using Index based centre selection method in our approach. But K means clustering method also can be used to find the basis function centre.

Can you think of other ways that you could model this kind of data? What if the amount of data was very large - say the grid was 1,000,000 x 1,000,000 instead of 50 x 50 - what difficulties might this kind of data introduce and how might you go about solving this problem?

For very large datasets deep learning algorithms can be used. In addition to that dimensionality reduction like PCA and clustering methods like k means also can be employed.