Question 2

Ian Dover

January 2023

1 Problem Description

Prove that:
$CV_i = y_i - f_i(x_i) = \frac{y_i - f(x_i)}{1 - \frac{K(x_i, x_i)}{\sum_{j=1}^{n} K(x_i, x_j)}}$

2 Part 1

$CV_i = y_i - f_i(x_i)$

$CV_i = y_i - \frac{\sum_{j=1}^{n} K(x_i, x_j) y_j - K(x_i, x_i) y_i}{\sum_{j=1}^{n} K(x_i, x_j) - K(x_i, x_i)}$

$CV_i = \frac{\sum_{j=1}^{n} K(x_i, x_j) y_i - \sum_{j=1}^{n} K(x_i, x_j) y_j}{\sum_{j=1}^{n} K(x_i, x_j) - K(x_i, x_i)}$

3 Part 2

$CV_i = \frac{y_i - f(x_i)}{1 - \frac{K(x_i, x_i)}{\sum_{j=1}^{n} K(x_i, x_j)}}$

$CV_i = \frac{y_i - \frac{\sum_{j=1}^{n} K(x_i, x_j) y_j}{\sum_{j=1}^{n} K(x_i, x_j)}}{1 - \frac{K(x_i, x_i)}{\sum_{j=1}^{n} K(x_i, x_j)}}$

$CV_i = \frac{y_i - \frac{\sum_{j=1}^{n} K(x_i, x_j) y_j}{\sum_{j=1}^{n} K(x_i, x_j)}}{\frac{\sum_{j=1}^{n} K(x_i, x_j) - K(x_i, x_i)}{\sum_{j=1}^{n} K(x_i, x_j)}}$

$CV_i = \frac{\sum_{j=1}^{n} K(x_i, x_j) y_i - \sum_{j=1}^{n} K(x_i, x_j) y_j}{\sum_{j=1}^{n} K(x_i, x_j) - K(x_i, x_i)}$

4 Part 3

The results of Part 1 and Part 2 are the same. This means that both equations are equivalent, thus proving:
$CV_i = y_i - f_i(x_i) = \frac{y_i - f(x_i)}{1 - \frac{K(x_i, x_i)}{\sum_{j=1}^{n} K(x_i, x_j)}}$

5 Part 4

Doing LOOCV, only a single model will need to be created, but it would need to be evaluated at every point in the dataset.