Marcin Pietrasik
Sean Scheideman
Why2

1. The before value of the fourth point (4.0, 2.1) is nonzero because the weights of six of its indices have been updated when the second point (4.0,2.0) was being learned. The before value of point four is the sum of the six weights of the tile indices that are shared by points two and four. These indices are:

[39, 160, 281, 403, 777, 898]

2. Since the target is non-stationary, the training function cannot account for the variability in the target caused by the addition of a random variable taken from a distribution with mean zero and standard deviation 0.1. The mean error will equal the variability, 0.1, and the MSE=(0.1)^2=0.01, and therefore the MSE will always be about 0.01 off.

```
Example ( 0.1 , 0.1 , 3.0 ):     f before learning:  0.0     f after learning :  0.3
Example ( 4.0 , 2.0 , -1.0 ):     f before learning:  0.0     f after learning :  -0.1
Example ( 5.99 , 5.99 , 2.0 ):     f before learning:  0.0     f after learning :  0.2
Example ( 4.0 , 2.1 , -1.0 ):     f before learning:  -0.075     f after learning :  -0.1675
The estimated MSE:  0.249130196078
The estimated MSE:  0.0584510915818
The estimated MSE:  0.022491343592
The estimated MSE:  0.0144300625331
The estimated MSE:  0.0126172422618
The estimated MSE:  0.0116093976956
The estimated MSE:  0.0115199205562
The estimated MSE:  0.0112872432055
The estimated MSE:  0.0113983357389
The estimated MSE:  0.0112387558105
The estimated MSE:  0.0111245712785
```