Comparision between

element-wise learning

and

vectorized learning

2016025423 박주언

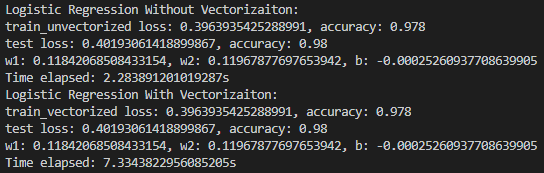
이번 과제를 수행하기 위해서 다음과 같이 두 파일을 만들었다.

* LogisticRegressionWithoutVectorization
* LogisticRegressionWithVectorization

말 그대로 전자는 vectorization없이 학습을 하는 것이고, 후자는 vectorization을 하고 학습을 수행하는 것이다.

둘 다 로직 자체는 같게 설정하였다. 같은 sigmoid 함수와 같은 loss function을 사용했으며, learning rate의 경우에도 똑같이 설정하였다. 다만, 둘을 비교하는 데에 있어서 같은 data를 가지고 비교를 해야 올바른 비교결과를 얻을 수 있기 때문에 random data를 만드는 과정에 있어서 같은 random value를 만든 후, unvectorized version과 vectorized version에 같이 넣어주는 식으로 같은 data를 부여했다.

실험에 앞서서 실험 환경은 Windows 10 64 bit, python version 3.7.3 (64bit)을 사용하였다.



# Time Comparison

-tested with train num = 1000, test num = 100-

|  |  |
| --- | --- |
|  | Time(s) |
| Element-wise version | 2.283891201019287s |
| Vectorized version | 7.3343822956085205s |

시간의 경우에는 예상했던 것과 다르게 Vectorized version이 더 오래 걸렸다.

# Estimated unknown function parameters W & b

-tested with train num = 1000, test num = 100-

|  |  |  |
| --- | --- | --- |
|  | Unvectorized | Vectorized |
| w1 | 0.11842068508433154 | 0.11842068508433154 |
| w2 | 0.11967877697653942 | 0.11967877697653942 |
| b | -0.00025260937708639905 | -0.00025260937708639905 |

같은 data로 실험을 하다 보니 unvectorized version과 vectorized version 모두 다 같은 값의 w1, w2, 그리고 b를 가졌다.

# Empirically determined (best) hyper parameter 𝜶

# Accuracy

For unvectorized version:

|  |  |  |  |
| --- | --- | --- | --- |
|  | m=10, k=100 | m=100, k=100 | m=1000, k=100 |
| With train set |  |  | 0.974 |
| With test set |  |  | 0.98 |
|  | m=100, k=10 | m=100, k=100 | m=100, k=1000 |
| With train set |  |  |  |
| With test set |  |  |  |

For vectorized version:

|  |  |  |  |
| --- | --- | --- | --- |
|  | m=10, k=100 | m=100, k=100 | m=1000, k=100 |
| With train set |  |  | 1.0 |
| With test set |  |  | 1.0 |
|  | m=100, k=10 | m=100, k=100 | m=100, k=1000 |
| With train set |  |  |  |
| With test set |  |  |  |