Practice#1 Report

2016024893 오 성준

* Time comparison (element-wise version vs. vectorized version, (m, K) = (1000, 2000))

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | 1회차 | 2회차 | 3회차 | 4회차 | Average |
| Vectorized | 0.421 | 0.463 | 0.404 | 0.442 | 0.4325 |
| Element Wise | 45.361 | 45.082 | 44.470 | 43.470 | 44.596 |

* Vectorized Version이 Element Wise Version보다 평균적으로 103배 더 빨랐습니다.
* Estimated unknown function parameters W & b
* Weight 와 bias는 상황마다 다릅니다. 그래서 지금은 alpha가 default인 0.01일 때의 weight와 bias의 평균을 구했습니다.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | 1회차 | 2회차 | 3회차 | 4회차 | 5회차 | Average |
| Weight | (1.30,1.30) | (1.30,1.28) | (1.32,1.32) | (1.30,1.29) | (1.26,1.25) | (1.30,1.29) |
| Bias | 0.0475 | -0.0164 | 0.0505 | -0.0255 | -0.059 | 0.001 |

Weight: (1.30, 1.29)

Bias: 0.001

* Empirically determined (best) hyper parameter, a

<About Accuracy>

|  |  |  |  |
| --- | --- | --- | --- |
|  | 1회차 | 2회차 | 3회차 |
| A=0.0001 | 98.8/97.0 | 97.5/96.0 | 99.1/100 |
| A=0.001 | 99.6/99.0 | 99.6/99.0 | 99.9/100 |
| A=0.01 | 99.9/100 | 99.7/98.0 | 99.8/100 |
| A=0.1 | 99.6/100 | 99.7/98.0 | 99.8/100 |
| A=1 | 99.6/100 | 99.7/98.0 | 99.9/100 |
| A=10 | 100/100 | 99.9/99.0 | 100/100 |

* 위의 표에서 보듯이 alpha가 클 수록 미세하나마 accuracy가 커진다는 것을 알 수 있으며 표에 표기하지 않았지만 cost도 가장 작았다 그러므로 이번 과제에선 alpha=10이 best hyper parameter입니다.
* Accuracy (fill in the blanks in the tables below and add them to the report)

|  |  |  |  |
| --- | --- | --- | --- |
|  | m= 10, n= 100, K = 2000 | m=100, n=100, K = 2000 | m=1000, n=100, K=2000 |
| Accuracy (m train set) | 100 | 100 | 99.2 |
| Accuracy (n test set) | 95.0 | 99.0 | 100 |

|  |  |  |  |
| --- | --- | --- | --- |
|  | m= 10, n= 100, K = 20 | m= 10, n= 100, K = 200 | m= 10, n= 100, K = 2000 |
| Accuracy (m train set) | 100 | 100 | 100 |
| Accuracy (n train set) | 89.0 | 99.0 | 95.0 |

* Discussion (what you learned in this experiment)

1. Vectorization version이 Element Wise version보다 압도적으로 빠릅니다.
2. Training set이나 여러가지 환경에 따라서 가장 좋은 learning rate 값이 다르고 다른 hyper parameter 값도 다르다.
3. Training set의 수가 작으면 train set에 대한 정확도는 증가하지만 Test set에 대한 정확도는 감소한다. 이는 Iteration에서도 마찬가지로, iteration이 감소하면 test set에 대한 정확도도 감소한다.

* 제출물에 대하여

1. Practice1.py는 요구하신 부분을 전부 구현한 코드입니다.

* Iteration이 10이 될 때마다 cost, weight, bias를 순서대로 print합니다.

1. Alpha.py는 alpha값이 0.0001, 0.001, 0.01, 0.1, 1, 10 일 때 결과를 print합니다. Best alpha를 구하기 위해서 만든 프로그램입니다.