

Assignment1

학번	2018023390	이름	이서연	제출일	2021.03.26
----	------------	----	-----	-----	------------

1. Time comparison (m, K) = (1000, 2000)

- Element wise : 21.58s
- Vectorized : 0.267s

2. Estimated unknown function parameter W, b

- $W = [1.29, 1.29]$
- $B = 0.17$

3. Empirically determined (best) hyperparameter α

- $1e-1$ 역시 testing accuracy 는 좋지만, $1e-2$ 에서 가장 적합해보인다. 추후 testing 을 더 늘리거나, 더 복잡한 문제를 풀거나, 더 많은 학습을 시킬 때를 고려해, 이후 과정은 모두 $1e-2$ 로 진행했다.



α	Training accuracy	Testing accuracy
1e-9	85	86
1e-8	79.4	79
1e-7	90.8	85
1e-6	86.6	90
1e-5	88	80
1e-4	92.1	85
1e-3	95.6	95
1e-2	100	99
1e-1	99.7	99

4. Accuracy

	m=10, n=100, K=2000	m = 100, n= 100, K = 2000	m = 1000, n = 100, K = 2000
Training accuracy	100.0	100.0	100.0
Testing accuracy	86.0	92.0	99.0

	m=1000, n=100, K = 20	m=1000, n=100, K = 200	m=1000, n=100, K = 2000
Training accuracy	95.30	98.6	100.0
Testing accuracy	94.00	98.0	99.0

5. Discussion

- 테스트를 반복적으로 해봤지만, element wise 일 때, vectorized 일 때 시간이 아닌 퍼포먼스에서 차이가 좀 있었다. 엄청 심각하게 나는 것은 아니지만 대략 0.5 – 1% 정도 accuracy 에서 차이가 있었다.
- 당연히, training sample 이 적을 때는 training accuracy 는 높고, testing accuracy 는 낮게 나왔으며 전체 학습 횟수가 적을 때는 전체적으로 accuracy 가 떨어지고, 전체 학습 횟수가 많을 때는 accuracy 가 높게 나왔다. 이를 통해, 이번 과제에서는 충분한 양의 training set 이 있어야만 더 다양한 범위의 데이터를 포괄적으로 다룰 수 있으므로, 학습하지 않은 데이터에서도 잘 작동함을 알 수 있었다. 또한, 충분한 수준의 학습을 통해 파라미터 W , b 가 충분히 업데이트 될 수 있도록 해야한다. W 와 b 는 (이번 예제에서는) 0 으로 초기화를 시켰다. (랜덤으로도 해봤다)
- W 가 특정 값으로 수렴되어야 한다고 생각했는데, 수렴되지 않고 계속 증가했다.