ML Foundation HW3 PA Report

B04505036
李慕家

---

19. Implement the fixed learning rate gradient descent algorithm for logistic regression. Run the algorithm with $\eta = 0.01$ and $T = 2000$, what is $E_{out}(g)$ from your algorithm, evaluated using the $0/1$ error on the test set?

20. Implement the fixed learning rate stochastic gradient descent algorithm for logistic regression. Instead of randomly choosing n in each iteration, please simply pick the example with the cyclic order $n = 1, 2, \ldots, N, 1, 2, \ldots$.

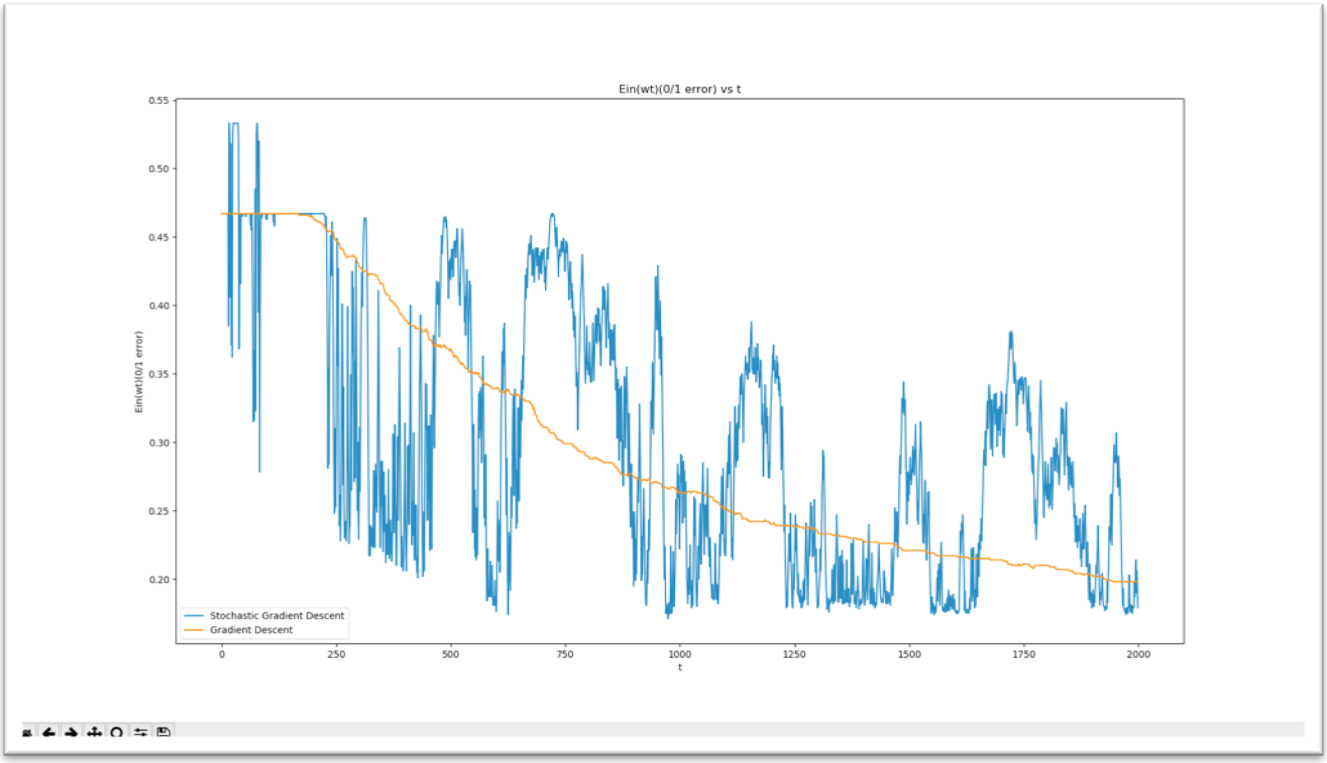
    Run the algorithm with $\eta = 0.001$ and $T = 2000$. What is $E_{out}(g)$ from your algorithm, evaluated using the $0/1$ error on the test set?

7. (20 points, *) For Questions 19 and 20 of Homework 3 on Coursera, plot a figure that shows $E_{in}(\mathbf{w}_t)$ as a function of t for both the gradient descent version and the stochastic gradient descent version on the same figure. Describe your findings. **Please print out the figure for grading.**

7.

For T = 2000, η=0.01:

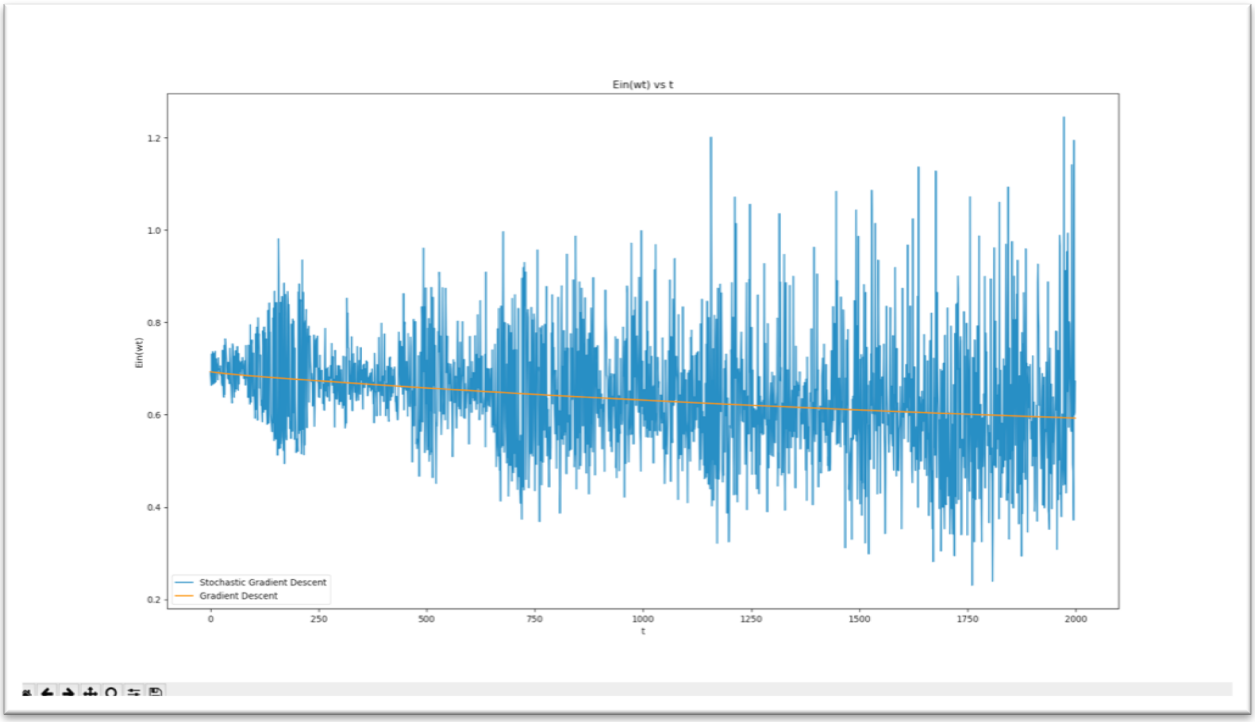
In Sample Error using 0/1 error:



```
update t: 2000, SGD error: 17.9%
```

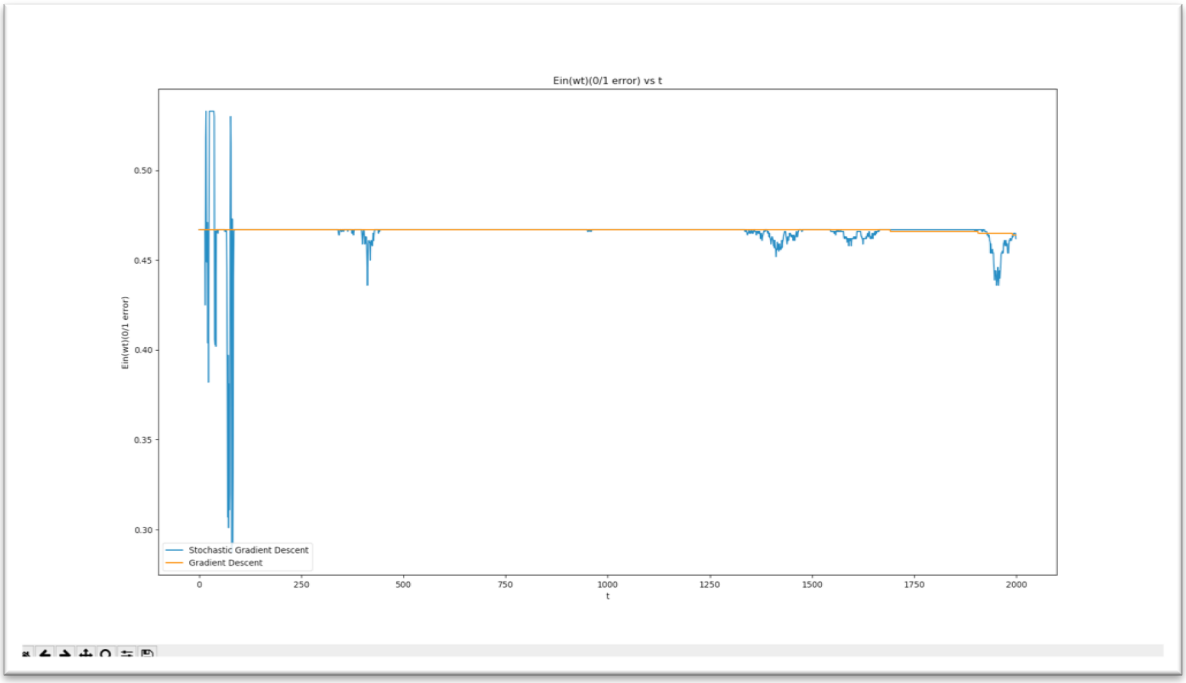
```
update t: 2000, GD error: 19.8%
```

Cross-Entropy Error:



For T = 2000, η=0.001:

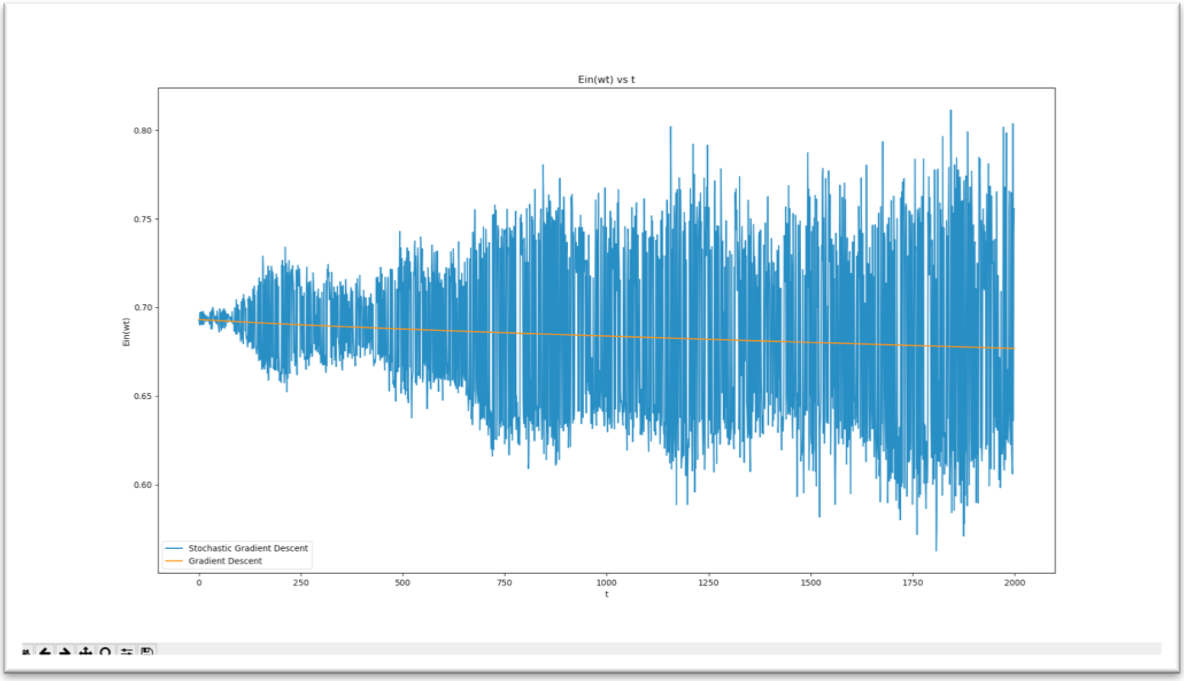
In Sample Error using 0/1 error:




```
update t: 2000, SGD error: 46.2%
```

```
update t: 2000, GD error: 46.4%
```

Cross-Entropy Error:



在 eta=0.001 時，learning 的效果非常差，無論是用 gradient descent 或是 stochastic gradient descent 都無法有效降低 0/1 error。而當 eta=0.01，兩種算法都能有效降低 0/1 error, gradient descent 無論在 0/1 error 或 cross-entropy error 上都能穩定的下降，而 stochastic gradient descent 在計算速度上比 gradient descent 有很大的優勢，雖然 0/1 error 的曲線抖動很大，但整體而言 0/1 error 也能夠快速下降。

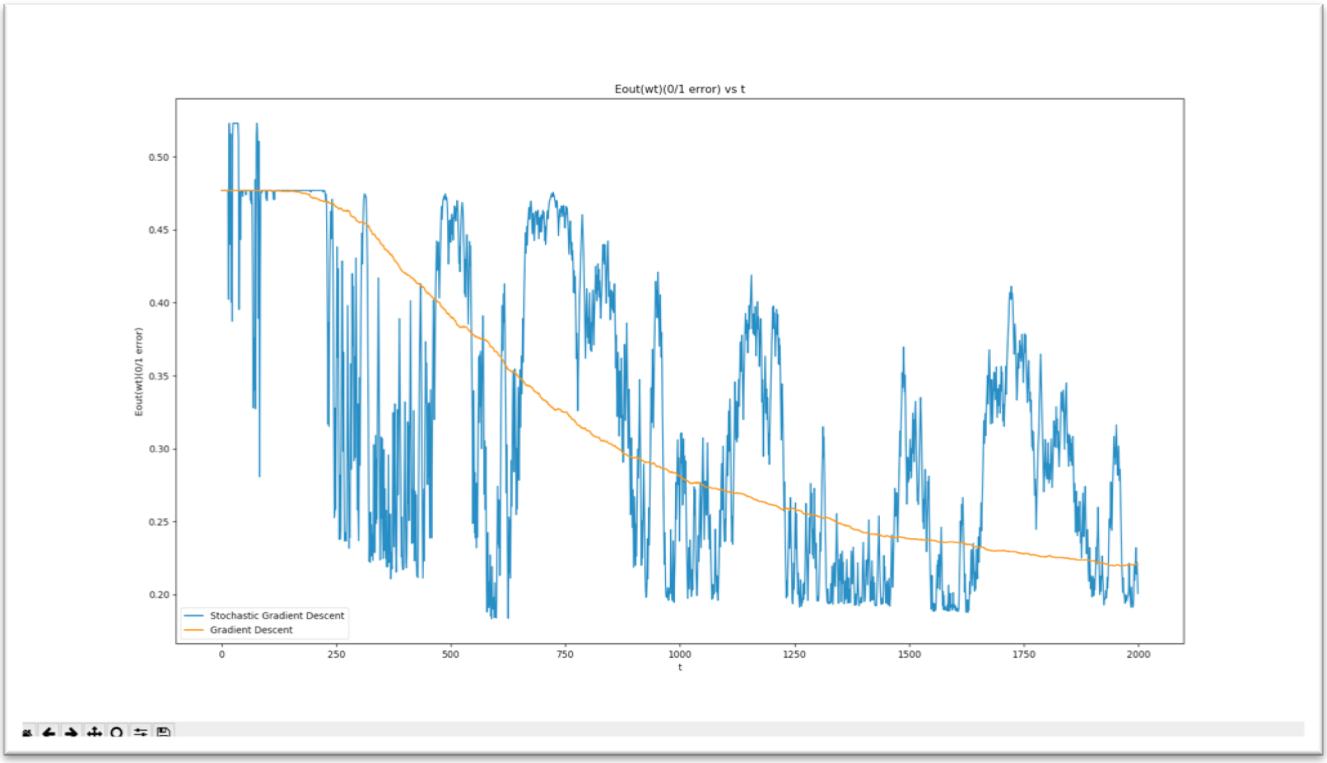
---

8. (20 points, *) For Questions 19 and 20 of Homework 3 on Coursera, plot a figure that shows $E_{out}(\mathbf{w}_t)$ as a function of t for both the gradient descent version and the stochastic gradient descent version on the same figure. Describe your findings. **Please print out the figure for grading.**

8.

For T = 2000, η=0.01:

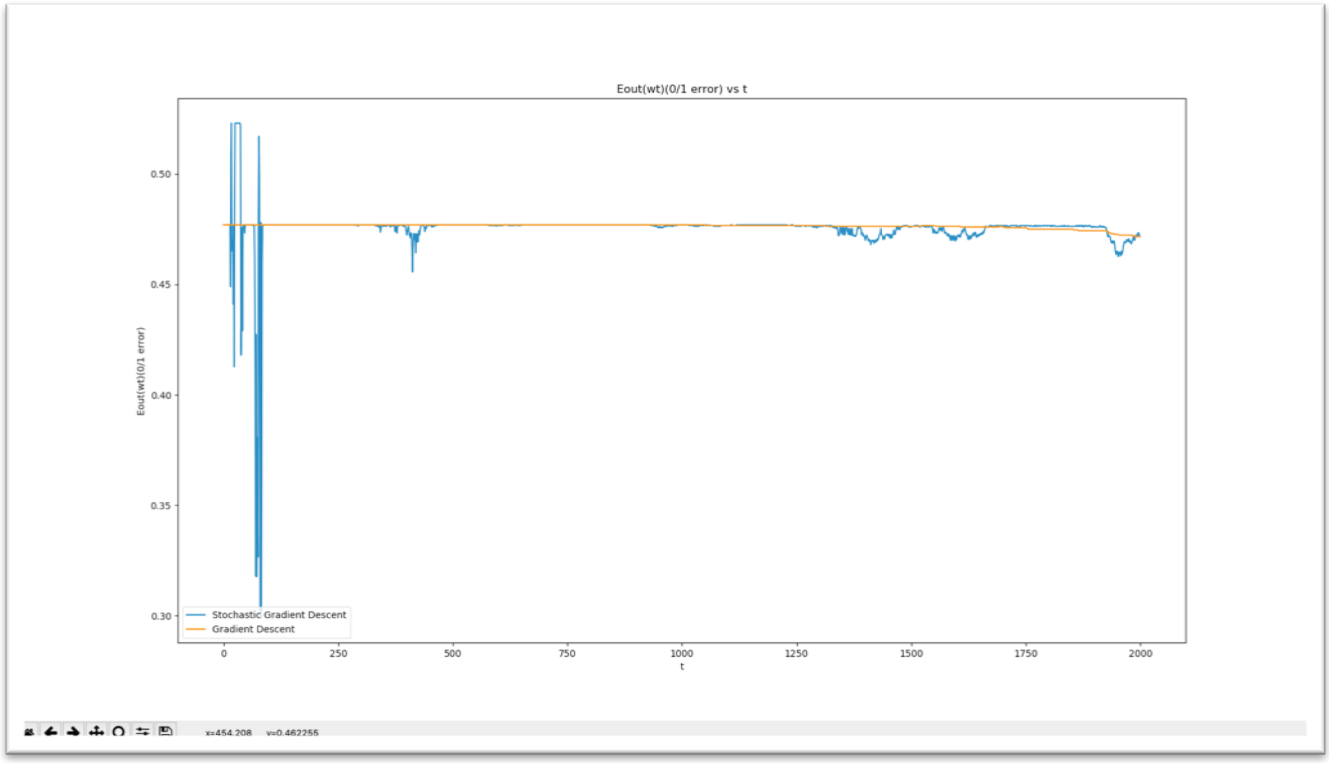
Out of Sample Error using 0/1 error:




```
--------------------------
update t: 2000, SGD error: 20.1%
--------------------------
update t: 2000, GD error: 22.1%
```

For T = 2000, η=0.001:

Out of Sample Error using 0/1 error:




```
update t: 2000, SGD error: 47.2%
update t: 2000, GD error: 47.2%
```

與上一題有非常相似的發現，代表 logistic regression 在 in sample 或 out of sample 的表現非常相似。