Exercise 2

Hou, Jue

014695647

1. Training and generalization error

Validation error for Regularization, (Train: Validation=10:90)

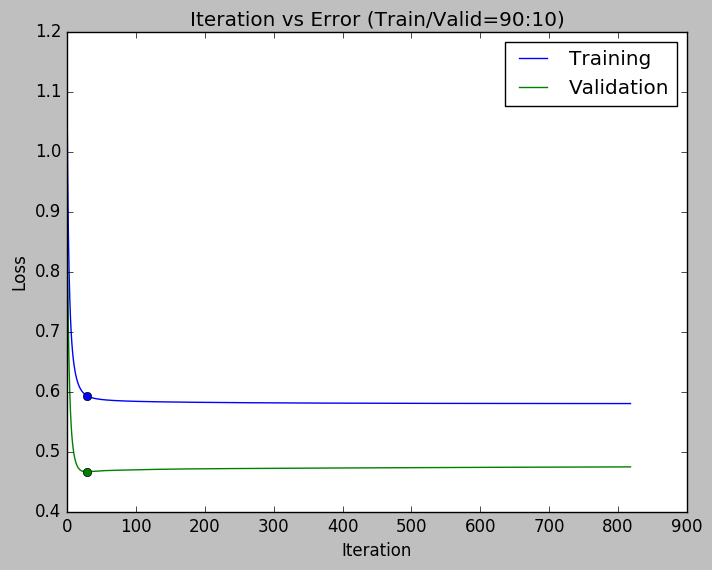
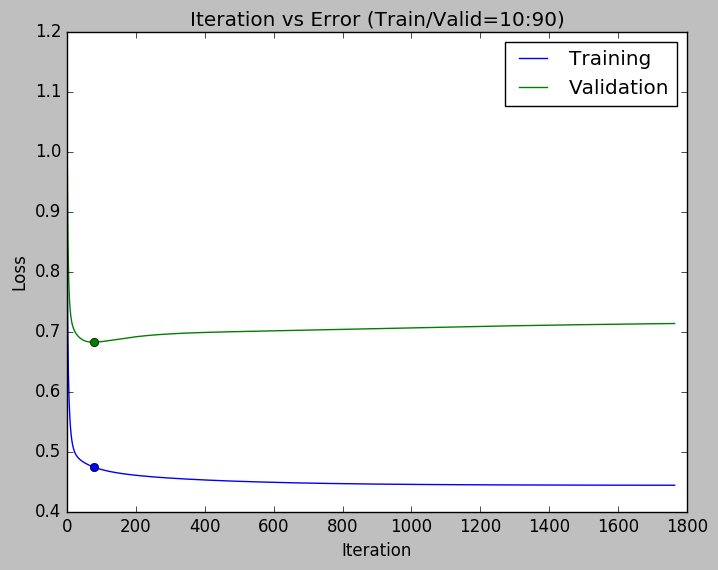
|  |  |  |
| --- | --- | --- |
| λ | Iteration | Validation Error |
| 0.001 | 1043 | 0.69577984 |
| 0.003 | 574 | 0.6896136 |
| 0.01 | 231 | 0.68939152 (min) |
| 0.03 | 96 | 0.70814333 |
| 0.1 | 34 | 0.77326792 |

|  |  |  |
| --- | --- | --- |
| Train: Validation=10:90 | Iteration | Test Error |
| ERM | 1764 | 0.82721003644915925 |
| Early-Stop | 78 | 0.79092831571209044 |
| Regularization (λ=0.01) | 231 | 0.79550146601373495 |

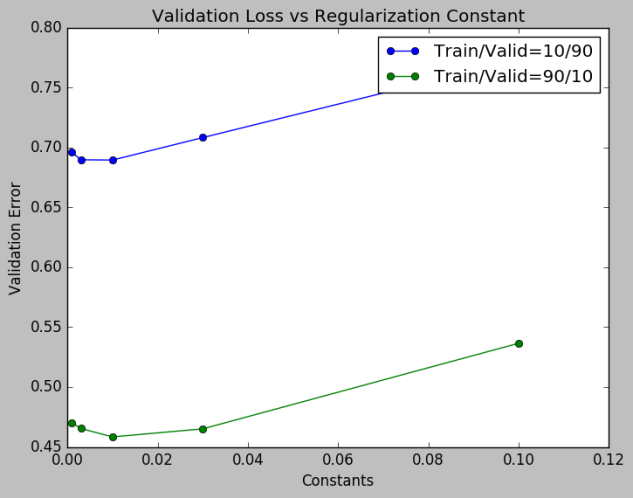
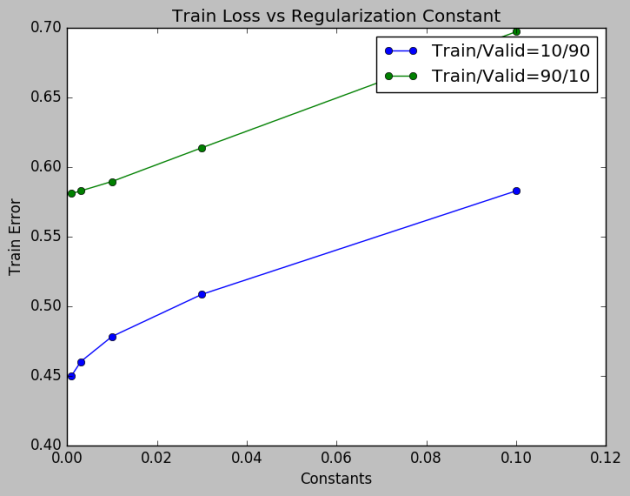
Validation error for Regularization (Train: Validation=90:10):

|  |  |  |
| --- | --- | --- |
| λ | Iteration | Validation Error |
| 0.001 | 576 | 0.4701488 |
| 0.003 | 364 | 0.46531423 |
| 0.01 | 167 | 0.45834984 (min) |
| 0.03 | 77 | 0.4650475 |
| 0.1 | 33 | 0.5364707 |

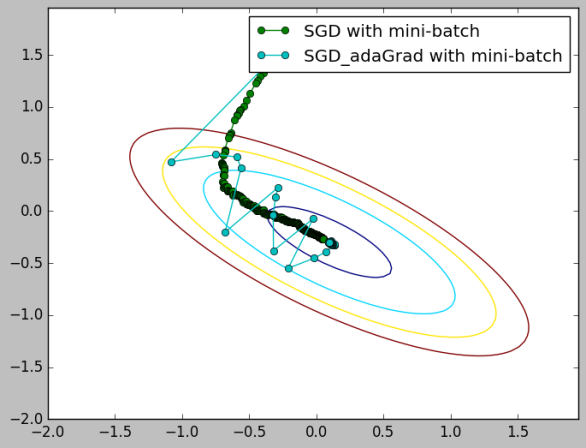
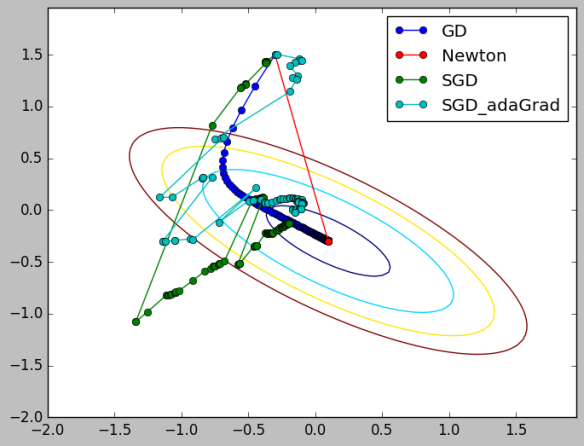
|  |  |  |
| --- | --- | --- |
| Train: Validation=90:90 | Iteration | Test error |
| ERM | 818 | 0.70310958205438279 |
| Early-Stop | 29 | 0.71095022358461124 |
| Regularization (λ=0.01) | 167 | 0.70638341531463922 |

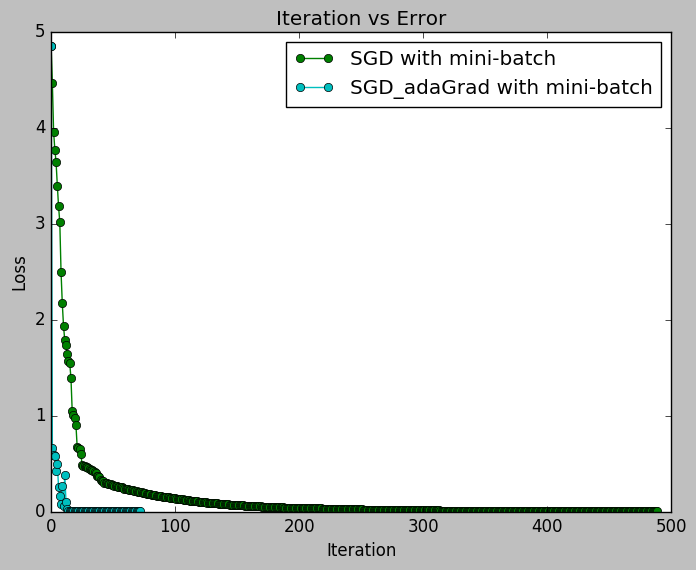
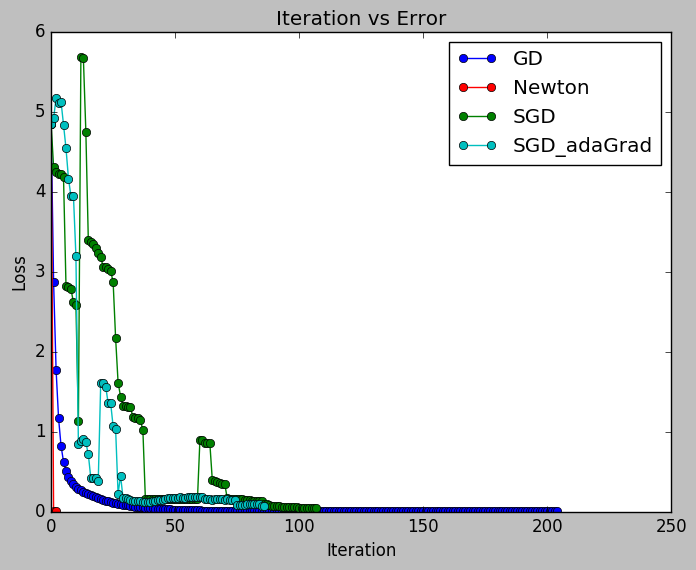


The above two plots show the error of training and validation in the function of iterations. The spots on both line is the point where Early-stopping stopped.



1. Comparison of optimization methods





The step size for GD and SGD are 0.05 and 0.01 respectively to achieve a better visual result and reach convergence with relatively small number of iterations.

The newton method will directly jump to the convergence because it solved the global optimal parameter for a convex function with the first and second derivative.

When only one sample is randomly selected, both SGDs fluctuate a lot to reach the convergence. When mini-batch is applied, SGD with adaptive gradient algorithm is not as smooth as standard SGD, but it obviously reaches convergence with fewer iterations.

1. Plate diagrams and independence

*K*

*N*

Estimated Complete Time: 6 hours