## SGD

| Epoches | w | t | beta | overall | training time |
|---|---|---|---|---|---|
| 1 | 0.60614 | 0.728225 | 0.273687291 | 0.53602 | 9.3333252 |
| 2 | 0.537158 | 0.010215 | 0.195189647 | 0.24752 | 19.6473019 |
| 5 | 0.521522 | 0.676156 | 0.136404612 | 0.44469 | 50.1611658 |
| 10 | 0.546745 | 0.153906 | 0.095304414 | 0.26532 | 98.454093 |
| 20 | 0.499547 | 0.178716 | 0.241286876 | 0.30652 | 196.0749034 |
| 30 | 0.565303 | 0.481736 | 0.16680693 | 0.40462 | 284.3656852 |
| 50 | 0.468067 | 0.744258 | 0.163351239 | 0.45856 | 452.7361426 |

## SGD with gamma_n decayed

| Epoches | w | t | beta | overall | training time |
|---|---|---|---|---|---|
| 1 | 0.060286 | 0.028369 | 0.040626915 | 0.04309 | 9.6457014 |
| 2 | 0.075889 | 0.05648 | 0.011416728 | 0.04793 | 18.857489 |
| 5 | 0.083291 | 0.048317 | 0.044151757 | 0.05859 | 45.8980195 |
| 10 | 0.075628 | 0.123348 | 0.028876164 | 0.07595 | 95.1072255 |
| 20 | 0.082017 | 0.127219 | 0.025648754 | 0.0783 | 232.7508287 |
| 30 | 0.064784 | 0.053344 | 0.03296617 | 0.05036 | 324.9480153 |
| 50 | 0.064327 | 0.014579 | 0.038502974 | 0.03914 | 521.0642649 |

## ASGD with gamma_n decayed

| Epoches | w | t | beta | overall | training time |
|---|---|---|---|---|---|
| 1 | 0.045245 | 0.064369 | 0.009784338 | 0.0398 | 9.6457014 |
| 2 | 0.044765 | 0.058129 | 0.019710774 | 0.04087 | 18.857489 |
| 5 | 0.044513 | 0.077618 | 0.09482906 | 0.07232 | 45.8980195 |
| 10 | 0.04803 | 0.057314 | 0.020287092 | 0.04188 | 95.1072255 |
| 20 | 0.050978 | 0.061093 | 0.019816488 | 0.04396 | 232.7508287 |
| 30 | 0.049563 | 0.060916 | 0.019148668 | 0.04321 | 324.9480153 |
| 50 | 0.047128 | 0.057966 | 0.018818047 | 0.0413 | 521.0642649 |

## Batch SGD

| Epoches | w | t | beta | overall | training time |
| --- | --- | --- | --- | --- | --- |
| 1 | 0.05355577 | 0.045678081 | 0.028518 | 0.042584 | 0.1416787 |
| 2 | 0.041311624 | 0.012815092 | 0.021274 | 0.025134 | 0.2503323 |
| 5 | 0.051235597 | 0.02765044 | 0.013338 | 0.030741 | 0.6718789 |
| 10 | 0.064952874 | 0.039133281 | 0.026665 | 0.043584 | 1.2208895 |
| 20 | 0.047311392 | 0.055365312 | 0.04015 | 0.047609 | 2.4829876 |
| 30 | 0.053150234 | 0.010168958 | 0.038961 | 0.034093 | 3.7565146 |
| 50 | 0.046644138 | 0.038369516 | 0.014064 | 0.033026 | 6.5684521 |

## Batch SGD with gamma_n decayed

| Epoches | w | t | beta | overall | training time |
| --- | --- | --- | --- | --- | --- |
| 1 | 0.325615414 | 0.042762602 | 0.390445 | 0.252941 | 0.1520523 |
| 2 | 0.161053346 | 0.061418642 | 0.15053 | 0.124334 | 0.3063951 |
| 5 | 0.030722476 | 0.013924609 | 0.025862 | 0.023503 | 0.7744487 |
| 10 | 0.015212755 | 0.009314168 | 0.006667 | 0.010398 | 1.5824177 |
| 20 | 0.016251554 | 0.005469714 | 0.01053 | 0.01075 | 3.1191398 |
| 30 | 0.016920479 | 0.010562736 | 0.008106 | 0.011863 | 3.73341 |
| 50 | 0.016478227 | 0.005390627 | 0.005485 | 0.009118 | 6.2214178 |

## Batch ASGD with gamma_n decayed

| Epoches | w | t | beta | overall | training time |
| --- | --- | --- | --- | --- | --- |
| 1 | 0.440955496 | 0.03045396 | 0.571561 | 0.347657 | 0.1520523 |
| 2 | 0.317493132 | 0.149950352 | 0.309234 | 0.258893 | 0.3063951 |
| 5 | 0.173151489 | 0.068903273 | 0.236488 | 0.159514 | 0.7744487 |
| 10 | 0.070297864 | 0.036918572 | 0.064212 | 0.057143 | 1.5824177 |
| 20 | 0.039935517 | 0.016705293 | 0.026844 | 0.027828 | 3.1191398 |
| 30 | 0.022734145 | 0.00980124 | 0.018848 | 0.017128 | 3.73341 |
| 50 | 0.017021331 | 0.003461593 | 0.012884 | 0.011122 | 6.2214178 |