Emma Twitmyer, Deme Apostolou, Varun Chitturi

Lab 1B Write Up

Task 1

Show that the linear MMSE problem is a particular case of the ERM problem in (2).

given ERM:

$$w^* = \underset{w}{\operatorname{argmin}}\ r(w)$$

$$= \underset{w}{\operatorname{argmin}} \frac{1}{N} \sum_{i=1}^{N} \ell(y_i, \phi(x_i; w))$$

def: $\ell(y, \hat{y}) = (y - \hat{y})^2$

$\phi(x; w) = w^T x$

substitute: $\underset{w}{\operatorname{argmin}} \frac{1}{N} \sum_{i=1}^{N} (y_i - w^T x_i)^2$

this considers linear parametrizations and quadratic losses

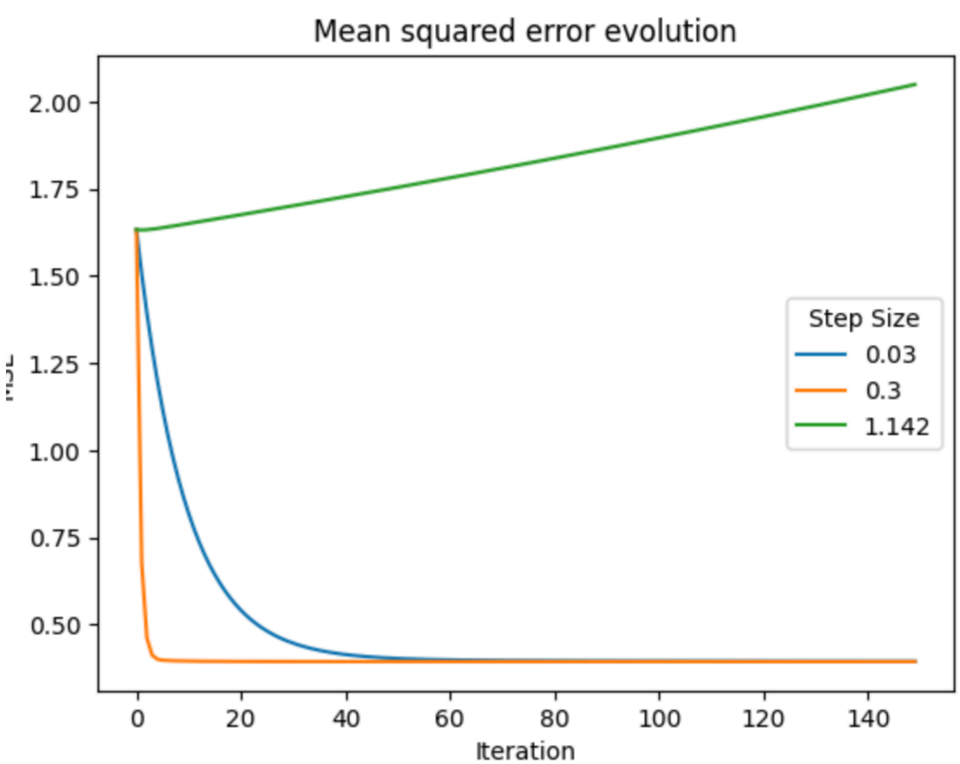
-

Task 2

A) Plot of r(w(k)) versus k.:

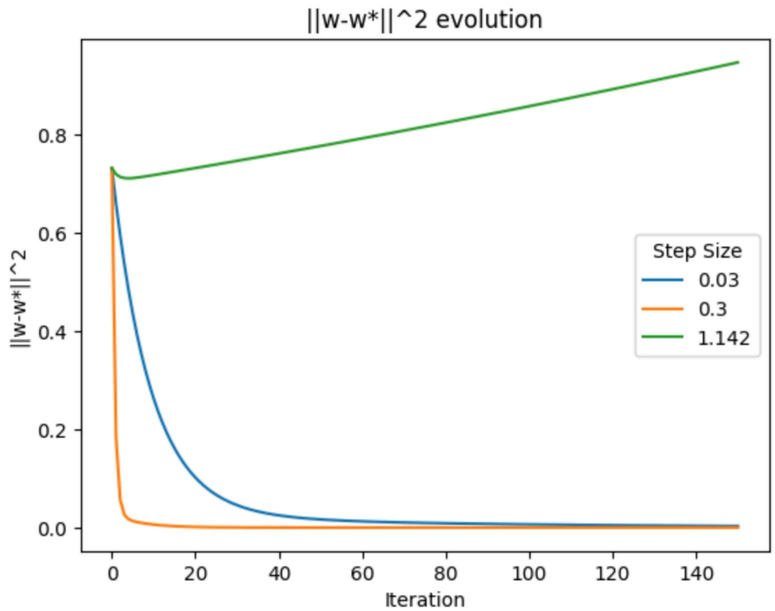
- step size = epsilon (denoted in legend)

- Plot: step size = epsilon (legend)



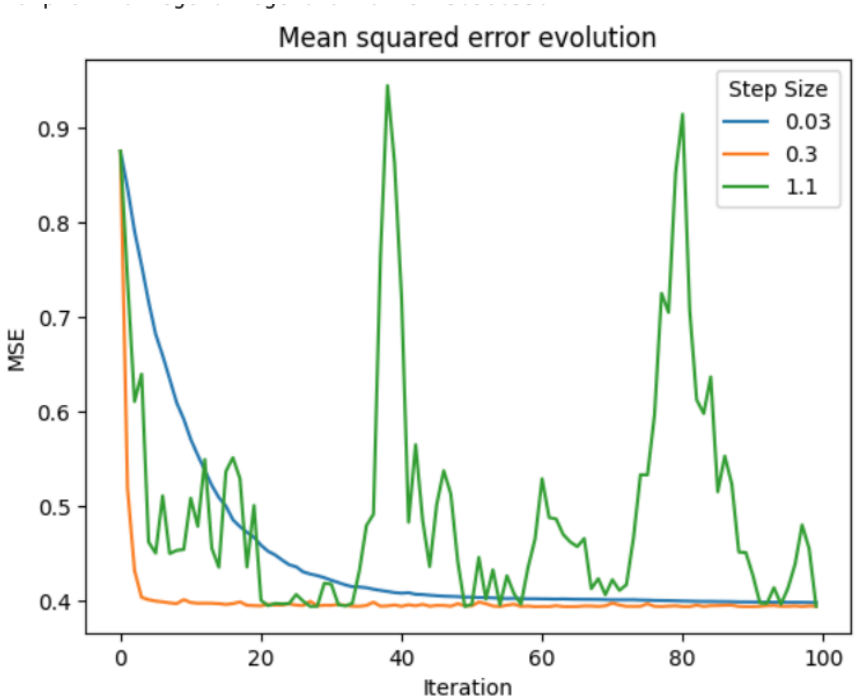
B) Plot of ||w-w*||^2

- Epsilon reported in legend

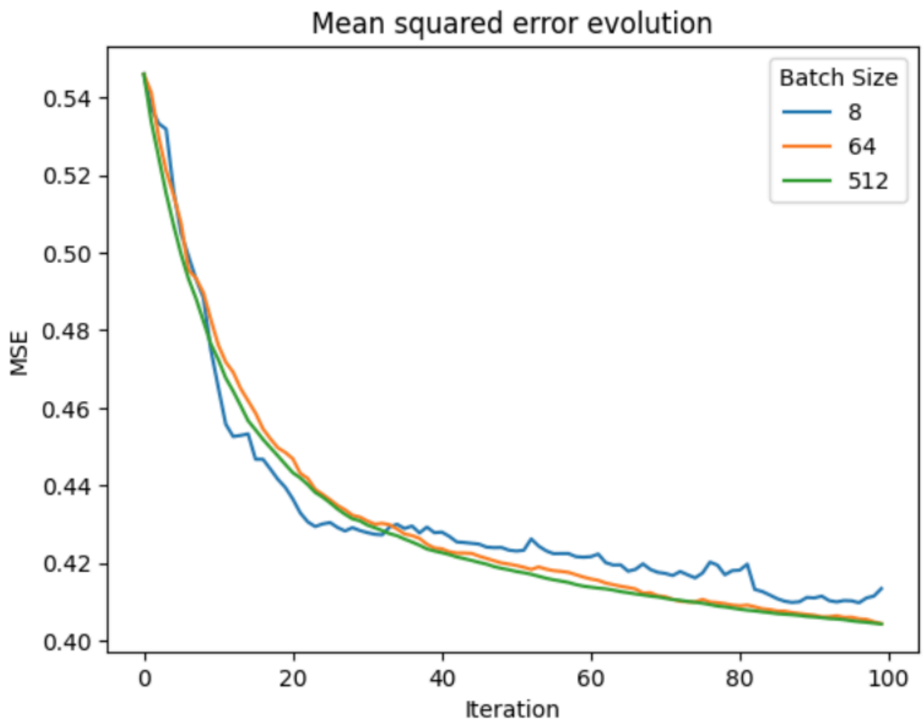


Task 3

A) Plot of r(w(k)) versus k. The epsilon values are .03, .3 and 1.1 and are denoted as step size in the legend

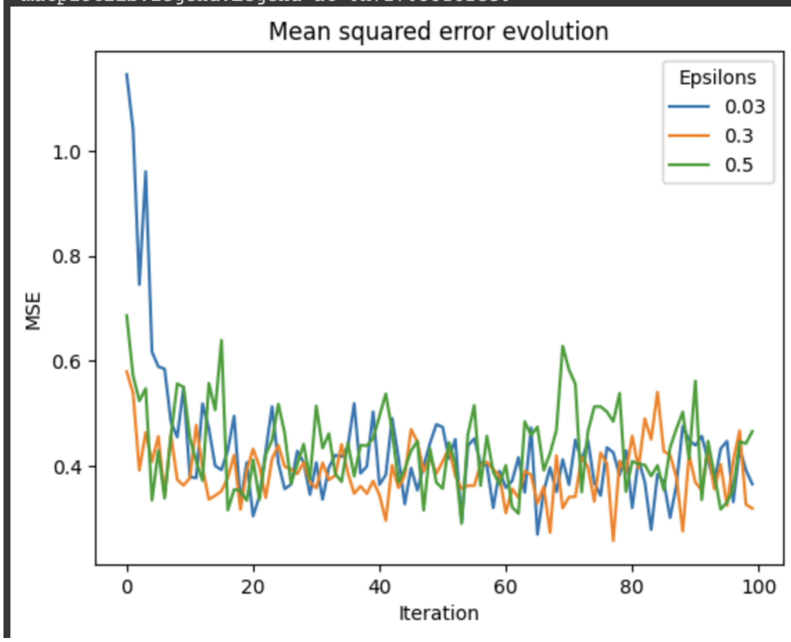


B) Plot of r(w(k)) versus k. The epsilon value (also known as learning rate) is constant at .03 but the batch sizes (B) are 8, 64 and 512, as demonstrated in the legend.



Task 4

```
Learning rate: 0.03, MSE: 0.366
Learning rate: 0.30, MSE: 0.319
Learning rate: 0.50, MSE: 0.466
<matplotlib.legend.Legend at 0x7f703818fee0>
```




```
Batch Size: 8.00, MSE: 0.313
Batch Size: 64.00, MSE: 0.434
Batch Size: 512.00, MSE: 0.361
<matplotlib.legend.Legend at 0x7f703800b730>
```