Kyle Salitrik (kps168)
CMPSC 448 – HW2

1.1) $y_i = w^T x_i + \epsilon_i$; $\epsilon_i \overset{iid}{\sim} N(\emptyset, \sigma_\epsilon^2)$

$$y = w^T x + N(\emptyset, \sigma_\epsilon^2) \Rightarrow y_i \sim N(w^T x_i, \sigma^2) \Rightarrow y_i = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{\frac{(y_i - w^T x_i)^2}{2\sigma^2}\right\}$$

$$L(y_i; w^T x_i, \sigma^2) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \prod_{i=1}^{n} \left[\exp\left\{\frac{(y_i - w^T x_i)^2}{2\sigma^2}\right\}\right] = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{\sum_{i=1}^{n} \frac{(y_i - w^T x_i)^2}{2\sigma^2}\right\}$$

$$\ln(L) = n \cdot \ln\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) + \frac{1}{2\sigma^2} \sum_{i=1}^{n} (y_i - w^T x_i)^2$$

1.2) $f(w) = \frac{1}{n} \sum (w^T x_i - y_i)^2 = \frac{1}{n} (X w - Y)^2$

$$= \frac{1}{n} (Xw - Y)^T (Xw - Y) = \frac{1}{n} (X^T w^T - Y^T)(Xw - Y)$$

$$= \frac{1}{n} [X^T w^T X w - X^T w^T Y - \underbrace{Y^T X w}_{X^T w^T Y} + Y^T Y]$$

$$= \frac{1}{n} [X^T w^T X w - 2 X^T w^T Y - Y^T Y]$$

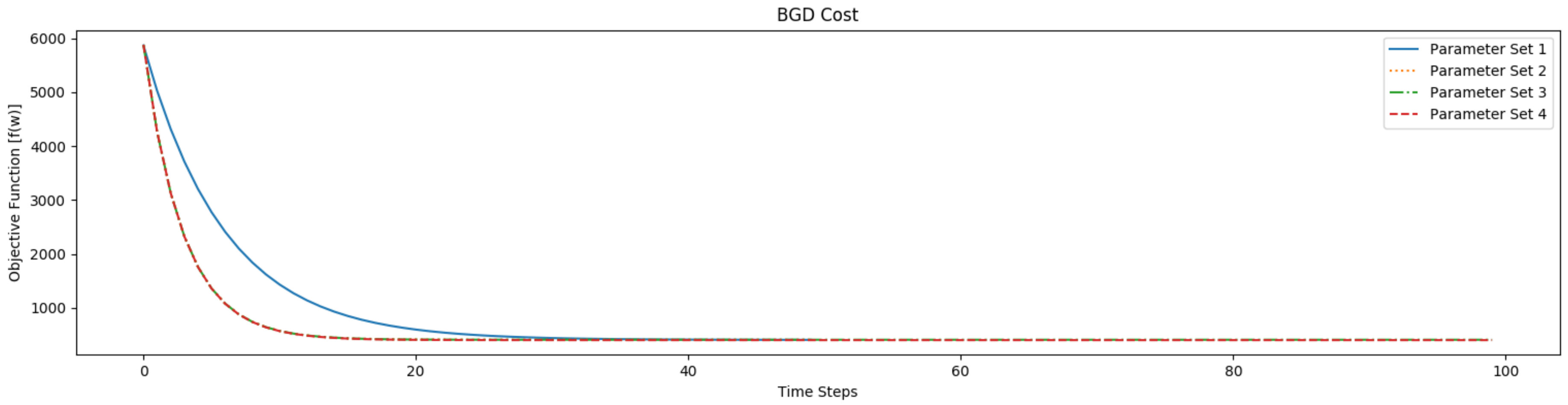
$$\frac{\partial f}{\partial w} = \frac{1}{\cancel{n}} \left[\frac{\partial}{\partial w} (X^T w^T X w) - 2 \frac{\partial}{\partial w} (X^T w^T Y) - \cancel{\frac{\partial}{\partial w} (Y^T Y)}^{\emptyset}\right] = \emptyset$$

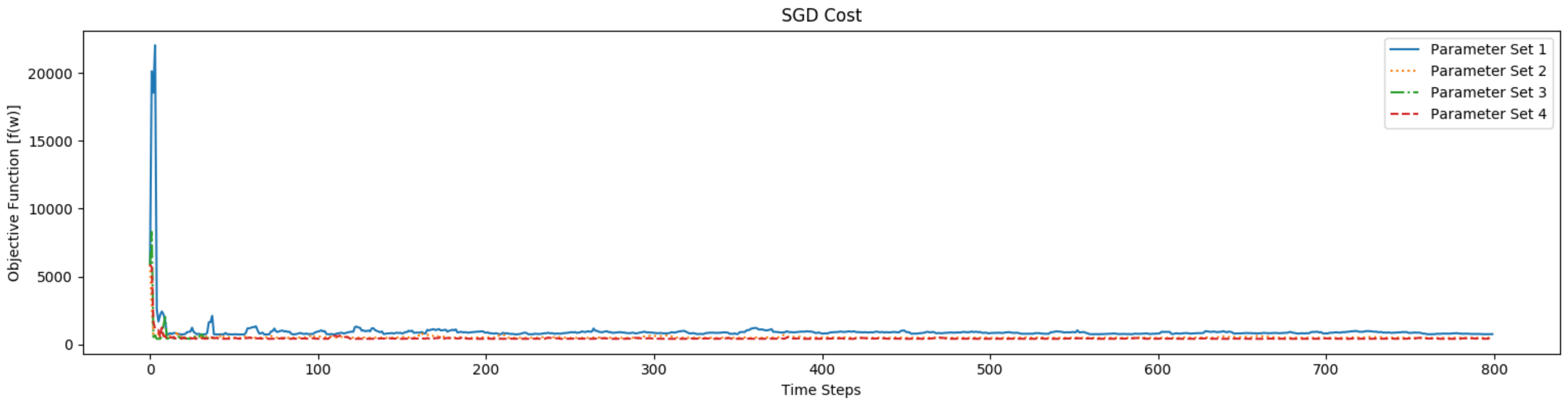
$$\Rightarrow X^T X \frac{\partial}{\partial w} [w^T w] - 2 X^T Y \frac{\partial}{\partial w} (w) = \emptyset$$

$$\Rightarrow \cancel{2} (X^T X) w - \cancel{2} X^T Y = \emptyset$$

$$\Rightarrow (X^T X)^{-1} (X^T X) w = (X^T X)^{-1} X^T Y$$

$$\Rightarrow \boxed{w = (X^T X)^{-1} X^T Y}$$


BGD Cost
Objective Function [f(w)]
Time Steps
Parameter Set 1
Parameter Set 2
Parameter Set 3
Parameter Set 4
6000
5000
4000
3000
2000
1000
0
20
40
60
80
100


SGD Cost
Objective Function [f(w)]
Time Steps
Parameter Set 1
Parameter Set 2
Parameter Set 3
Parameter Set 4
20000
15000
10000
5000
0
100
200
300
400
500
600
700
800