March 8, 2023

# 1 Q2 (c and d part)

Write down the expression for the gradient of the squared loss function, $\nabla_w \mathcal{L}(X, w, y)$, using the given information.

To find the minimizer of the loss function, $\arg\min_w \mathcal{L}(X, w, y)$, set the gradient equal to zero and solve for $w$.

Use the given initialization $w_0$ and step size $\alpha$ to find expressions for the first two iterates $w_1$ and $w_2$, in terms of $w_0$, $X$, and $y$.

we can start like this for w1 and w2.
So, in the first iteration, we will have:

$$w_1 = w_0 - \alpha\left(X^\top (Xw_0 - y)\right) = w_0 - \alpha X^\top X w_0 + \alpha X^\top y$$

Then in the second iteration, we will have:

$w_2 = w_1 - \alpha X^\top (Xw_1 - y) = \left(1 - \alpha X^\top X\right) w_1 + \alpha X^\top y$

Replace w1 value from first iteration.

Generalize the expressions for $w_1$ and $w_2$ to find an expression for the $k$th iterate $w_k$ in terms of $w_0$, $X$, $y$, $\alpha$, and $k$.

You need to find pattern in w1 and w2. You can also check for w3 and w4(not compulsory). Based on which we could find generalised form for wk.

# 2 Q3

$z = w_1 x + b$

$\frac{\partial l}{\partial \mathbf{x}} = \frac{\partial l}{\partial y}\frac{\partial y}{\partial z}\frac{\partial z}{\partial x}$

$\frac{\partial l}{\partial \mathbf{b}} = \frac{\partial l}{\partial y}\frac{\partial y}{\partial z}\frac{\partial z}{\partial b}$

$\frac{\partial z}{\partial b} = 1$

$\frac{\partial z}{\partial x} = w_1$

$\frac{\partial y}{\partial z} = w_2^\top diag\left(sigmoid'\left(z\right)\right) = w_2^\top \odot sigmoid'\left(z\right)$

$\frac{\partial l}{\partial y} = (y - t)^2$

By using the above formula and the given values in the question, find the gradient values: