[Special Topics] Homework 1
CS 4301.001
Sudarshana Jagadeeshi

Question 1-a:

Proof convex:

We want the second derivative to be greater or equal to 0. Differentiating w.r.t to w we get, $[\max(0, 1 - y_m w \ x_m)^2]'' = [\max(0, 2 - 2y_m w \ x_m) * (-y_m x_m))]' = [\max(0, -2y_m x_m)]$. This is positive everywhere, because the minimum value the function can take is 0.

Compute Gradient/Subgradient:

The ith entry of the gradient is: $\max(0, 2 - 2y_m \ w_i \ x_m) * (-y_m x_m))$. This is just the first derivative I put above, but w_i instead of w, because all other dimensions should differentiate out to 0.

Updates for Gradient w/ Fixed step size:

$(w_i)_{t+1} = (w_i)_t$ - gamma * gradient of $(w_i)_t$

⇒ $(w_i)_{t+1} = (w_i)_t$ - gamma * $\max(0, 2 - 2y_m \ w_i \ x_m) * (-y_m x_m))$.

Question 1-b:

Proof convex:

If we choose an x and y, we want to show $f(tx + (1-t)y) <= t f(x) + (1-t)f(y)$

$$\log \sum_{i=1}^{n} e^{a_i (t x_i + (1-t) y_i)}$$

$$\leq t \log \sum_{i=1}^{n} e^{a_i x_i} + (1-t) \log \sum_{i=1}^{n} e^{a_i y_i}$$

For a single point, assuming log is ln

$$a(tx + (1-t)y) \leq$$

$$tax + (1-t)ay$$ ✓ They are equal

Compute Gradient/Subgradient:
x is a vector(i.e. w_1 x_1+ w_2 x_2+...w_n x_n) The ith entry of the gradient is aw_i* exp(a^i x^i). The exponent stays the same, but we differentiate the vector x and get only a w_i.

Updates for Gradient w/ Fixed step size:
(w_i)_t+1 = (w_i)_t - gamma * gradient of (w_i)_t
⇒ (w_i)_t+1 = (w_i)_t - gamma * aw_i* exp(a^i x^i)

Question 1-c:
Proof convex:

Compute Gradient/Subgradient:

Updates for Gradient w/ Fixed step size:

Question 2-1-a:
Refer to gd_userinput.py. I've hosted the linreg.data file via url. Input for X and y must be separated by spaces.

Question 2-1-b:
Refer to problem2-1b.xlsx. The generating code is in gd_final.py.

Question 2-2-a
Refer to file subgradient_final.py

Question 2-2b
Refer to file problem2-2b.xlsx

Question 3-1-a:
Yes. We know that the infimum preserves convexity, so we want to just know if the inside is a convex function. Deriving once we get:
[e^a (ln(e^a / e^a + 1) + ln(e^a +1))] / (e^a + 1)^2.

Deriving again we get,

$$-\frac{e^{a}\left(\left(e^{a}-1\right)\ln\left(\frac{e^{a}}{e^{a}+1}\right)+\left(e^{a}-1\right)\ln\left(e^{a}+1\right)-e^{a}-1\right)}{\left(e^{a}+1\right)^{3}}$$

Which should be positive everywhere. Something went wrong with my math, but the original function is convex.

Question 3-1-b:

Question 3-1-c:

Question 3-2-a:
Refer to file cvxnum.py
min of -132.66535571301728 @ approximately x=-4.75