[Special Topics] Homework 4
CS 4301.001
Sudarshana Jagadeeshi

1-1, 1-2.

1-1.
$$\min_{B\in\mathbb{R}^{n\times n}} \sum_{i,j} (A_{ij}-B_{ij})^2$$
subject to $B \succeq 0$
$Q_{ij}=0$

1-2.
$$\mathcal{L}(B,\Lambda,\mu) = \langle A-B, A-B\rangle - \langle B,\Lambda\rangle + \sum_{i,j}\mu_{ij}Q_{ij} \quad \text{s.t. } \Lambda \succeq 0$$
$$\frac{\partial \mathcal{L}}{\partial B} = \overset{0}{\cancel{\langle A,A\rangle}} - 2\overset{A}{\langle A,B\rangle} + \overset{2B}{\cancel{\langle B,B\rangle}} \; \overset{A}{\cancel{\langle B,\Lambda\rangle}} = 0$$
$$-2A + 2B - \Lambda = 0$$
$$B = \frac{2A+\Lambda}{2}$$
The dual is:
$$g(\Lambda,\mu) = \frac{1}{4}\langle \Lambda - A, \Lambda - A\rangle - \frac{1}{2}\langle 2A+\Lambda, \Lambda\rangle + \langle \mu, Q\rangle$$
subject to $\Lambda \succeq 0$, $Q_{ij}=0$

1-3.
Notes:
- The gradient is 2 * (B-A)
- The projection is performed by taking any negative eigenvalue of B and zeroing it out.

https://colab.research.google.com/drive/1rx7wJrPcS1XtFknVlYgF6pTLwunt1Qtg?usp=sharing

2-1.
The objective function is convex when fixing one of C,U.

2-1 $$\min_{C \in \mathbb{R}^{m \times k},\, U \in \mathbb{R}^{k \times n}} \sum_{i,j} \left[A_{ij} \log\left(\frac{A_{ij}}{\sum_k C_{ik} U_{kj}}\right) - A_{ij} + \sum_k C_{ik} U_{kj} \right]$$

subject to $C \geq 0, U \geq 0$

2-2.
Take partials to find the argmin. I tried to solve them for 0, but it doesn't make sense here.

$$\frac{\partial}{\partial C_{ab}} = \sum_{i,j} \left[A_{ij} \cancel{\log(A_{ij})}^{0} - A_{ij} \log(CU) - \cancel{A_{ij}}^{0} + CU \right]'$$

$$= \sum_{i,j} (CU)' \left(1 - \frac{A_{ij}}{CU}\right)$$

$$= \sum_{i,j} \sum_{l} U_{bl} \left(1 - \frac{A_{ij}}{CU}\right)$$

$$\frac{\partial}{\partial U_{ab}} = \sum_{i,j} \sum_{l} C_{la} \left(1 - \frac{A_{ij}}{CU}\right)$$

https://colab.research.google.com/drive/1EY790Bm7R9ERLp3Iud7MrTtnGkZsNzZd?usp=sharing

2-3.
The algorithm won't always converge if the function isn't smooth, like the |x+y| - 3|y-x| function we saw in class. However, this function is smooth and doesn't need subgradients, so it should converge.