CSE 203B W21 Homework 2

Due Time : 11:50pm, Monday Feb. 1, 2021 Submit to Gradescope
Gradescope: `https://gradescope.com/`

In this homework, we work on exercies from text book including level sets of convex, concave, quasi-convex, quasi-concave functions (3.2), second-order conditions for convexity on affine sets (3.9), Kullback-Leibler divergence (3.13), determination of convex, concave, quasi-convex, quasi-concave functions (3.16), conjugate functions (3.36), and gradient and Hessian of conjugate functions (3.40). Extra assignments are given on entropy, Kullback-Leibler divergence, and conjugate function.

Total points: 30. Exercises are graded by completion, assignments are graded by content.

I. Exercises from textbook chapter 3 (6 pts, 1pt for each problem)

3.2, 3.9, 3.13, 3.16, 3.36, 3.40.

II. Assignments (24 pts)

II. 1. Entropy. (8 pts)

In information theory, entropy is defined as

$$H(X) = -\sum_{i=1}^{n} p_i \log p_i,$$

where $P(X = x_i) = p_i$ for $i \in \{1, 2, ..., n\}$, $p_i > 0 \, \forall i$ and $\sum_{i=1}^{n} p_i = 1$.

(1) Show that the entropy $H(X)$ is concave.

(2) Derive the conjugate function of the negative entropy $(-H(X))$.

II. 2. Properties and application of Kullback-Leibler divergence. (8 pts): It is often useful to measure the difference between two probability distributions over the same random variable. Here, we define the Kullback-Leibler (KL) divergence as $KL(p, q) = \sum_i p_i \log(\frac{p_i}{q_i})$, where $p_i > 0, \; q_i > 0, \; \forall i \in \{1, 2, ..., n\}$, and $\sum_{i=1}^{n} p_i = \sum_{i=1}^{n} q_i = 1$.

(1) Assume we have two six-sided dice with the probability distributions as shown in the given table. Please calculate the KL divergences $KL(P, Q)$ and $KL(Q, P)$ of the given example.

Event	P	Q
1	1/6	1/12
2	1/6	1/6
3	1/6	1/4
4	1/6	1/4
5	1/6	1/6
6	1/6	1/12

(2) Use the example in (2) to illustrate that the Hellinger distance can serve as a lower bound:

$$KL(P, Q) \geq \sum_i (\sqrt{p_i} - \sqrt{q_i})^2$$

.

(3) Create a simple but nontrivial numerical example to demonstrate that $KL(P, Q) \geq 0$ for all possible P, Q and the equality holds only when $P = Q$.

II. 3. Further practice on conjugate functions. (4 pts)
Consider the function

$$f(x) = \begin{cases} \frac{1}{2}x^2, & |x| < 1, \\ |x| - \frac{1}{2}, & |x| > 1. \end{cases}$$

Find the conjugate function f^*.

II. 4. Show that the dual of the l_p norm is the l_q norm, where $\frac{1}{p} + \frac{1}{q} = 1$. (4 pts)