DS 595/ MA 590

Optimization for Deep learning & ML

Homework 1

1. Suppose that $f : R^d \to R$ and $g : R^d \to R$ and $h : R \to R$ are convex functions. Which of the following functions must be convex? Please explain why it is convex (if yes), or provide a counterexample (if no)
   (a) $h(f(x))$
   (b) $f^2(x)$
2. Consider the following optimization problem

$$\min_{x \in R^2} \|Qx - c\|$$

$$\text{subject to} \qquad a^\top x = b$$

Where $Q = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, $c = \begin{pmatrix} 6 \\ 7 \end{pmatrix}$, $a = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$, $b = 5$

(a) Formulate this as a quadratic program.

(b) Use a computer to find the optimal solution to this quadratic program (for example, you could use the CVXopt software we will see in class)

3. Consider the constraint set $\Omega = \{Ax \leq b\}$ for some matrix $A \in R^m \times R^n$ and some vector $b \in R^n$. Give a method of computing the following:

   (a) A "membership oracle" for Ω. That is, design an algorithm which, given a point $x \in R^n$, determines whether the point x is in Ω.

   (b) A "projection oracle" for Ω. That is, given a point $x \in R^n$, find a point y such that $\|y - x\| = \min_{z \in \Omega} \|z - x\|$ (in other words, find an algorithm which can project any point x onto the set Ω). (Hint: you may want to solve an optimization problem to find $\min_{z \in \Omega} \|z - x\|$.)

   (c) Implement your algorithm in part (b) on a computer using a convex optimization solver (for example, you could use the CVXopt software we will see in class). In this part, you can assume that $n = 2$, and that $A = \begin{pmatrix} 4 & 2 \\ 1 & 3 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$, and $b = \begin{pmatrix} 5 \\ 4 \\ 0 \\ 0 \end{pmatrix}$.

4. Let $\mathcal{S}$ be the set of symmetric positive definite matrices.

   (a) Give an example of a subset of $\mathcal{S}$ which is both compact (i.e., closed and bounded) and convex. How do you know your example is convex? How do you know it is a subset of $\mathcal{S}$?

   (b) Give an example of a convex function $f : \mathcal{S} \to \mathbb{R}$.

   (c) Let $\Omega \subseteq \mathcal{S}$ be a convex subset of $\mathcal{S}$.
   Is the following optimization problem convex?

$$\min_{A \in \Omega} \quad -\log(\det(A)) \tag{1}$$

   Please explain why it is convex (if yes), or provide a counterexample (if no) (to make this problem a bit easier, you may assume that Ω contains only *diagonal* matrices)

5. A small airline flies between three cities: Ithaca, Newark, and Boston. They offer several flights but, for this problem, we focus on the Friday afternoon flight that departs from Ithaca, stops in Newark, and continues to Boston. There are three types of passengers:

   (i) Those traveling from Ithaca to Newark.

   (ii) Those traveling from Newark to Boston.

   (iii) Those traveling from Ithaca to Boston.

   The aircraft is a small commuter plane that seats 30 passengers. The airline offers three fare classes:

   (i) Y class: full coach.

   (ii) B class: nonrefundable.

   (iii) M class: nonrefundable, 3-week advanced purchase.

   Ticket prices have been set and advertised as follows:

	Ithaca-Newark	Newark-Boston	Ithaca-Boston
Y	300	160	360
B	220	130	280
M	100	80	140

   Based on past experience, demand forecasters at the airline have determined the following upper bounds on the number of potential customers in each of the 9 possible origin-destination/fare-class combinations:

	Ithaca-Newark	Newark-Boston	Ithaca-Boston
Y	4	8	3
B	8	13	10
M	22	20	18

   The goal is to decide how many tickets from each of the 9 origin-destination/fare-class combinations to sell. The constraints are that the plane cannot be overbooked on either of the two legs of the flight and that the number of tickets made available cannot exceed the forecasted maximum demand. The objective is to maximize revenue.

   (a) Formulate this problem as a linear programming problem.

   (b) Find the optimal solution to this linear program using a computer (for example, you could use the CVXopt software we will see in class).

6. Consider the following simple model for how the coronavirus spreads from city to city. There are n cities. Each city is a vertex on a graph, and the cities are connected by edges (which represent flight connections between the cities). The weight w_{ij} of each edge i, j is the transmission rate from city i to city j. In other words, if at time t there are $x_i(t)$ infected people in each city i, then at time $t+1$ there will be $x_j(t+1) = \sum_{i=1}^{n} w_{ij} x_i(t)$ infected people in city j.

   For simplicity, we assume that there is a very large number of people in each city (so in our model there is no upper bound on the number of people who can be infected).

   Let $x(t) = (x_1(t), \ldots, x_n(t))^\top$ be the vector whose entries give the number of infected people in each city on day t. The number of infected people in each city on day t is given by the formula

$$x(t) = W^t x(0) \tag{2}$$

   where W is the matrix with (i, j)'th entry w_{ij}.

   (a) Is the matrix W symmetric (i.e., does it make sense to assume that $w_{ij} = w_{ji}$) for each i, j? Why or why not?

   (b) After a very long time (in the limit as $t \to \infty$), we expect the total number of infected people to be roughly e^{rt} for some number r. What is r equal to?

   (c) Now suppose that we have a way of reducing the transmission, by testing people before they get onto a flight (you get the answer to the test right away, so there is no delay, and the test is 100% accurate). So if we had enough tests for everybody, we could completely stop city-to-city transmission. But the factory can only make 10,000 tests per day. Formulate an optimization problem which minimizes the growth rate r. (hint: you may need to make some additional assumptions here, e.g., about how many people travel between the different cities)

   (d) Is your optimization problem in part (c) convex? Please explain why or why not.