EECS 127/227AT **Fall 2020**

Homework 4

Homework 4 is due on Gradescope by Friday 10/02 at 11.59 p.m.

1 Convexity of Sets and Functions

This is a warm up problem to study when sets and functions are convex and to understand how to form the convex hull of a set. The relevant parts of the textbooks are Secs. 8.2.1 and 8.2.2. of the textbook of Calafiore and El Ghaoui and Secs. 3.1 and 3.2 of the textbook of Boyd and Vandenberghe.

(a) For the each of the sets given below, determine if it is a convex set or not with a brief justification for your answer.

   i. The empty set ϕ in $\mathbb{R}^n$.

   ii. A singleton set $\{x_0\}$, where $x_0 \in \mathbb{R}^n$.

   iii. $\mathbb{R}^n$.

   iv. $\mathcal{P} = \{z \in \mathbb{R}^n : \|z - z_0\|_2 = \epsilon\}$, where $z_0 \in \mathbb{R}^n$, $\epsilon > 0$.

   v. $\mathcal{P} = \{z \in \mathbb{R}^n : \|z - z_0\|_2 \leq \epsilon\}$, where $z_0 \in \mathbb{R}^n$, $\epsilon > 0$.

   vi. (**Optional**)
   $\mathcal{P} = Q \bigcap R$, where Q and R are convex sets in $\mathbb{R}^n$.

   vii. $\mathcal{P}$ = *Minkowski sum* of sets Q and R, where Q and R are convex sets in $\mathbb{R}^n$, where the Minkowski sum of two sets Q and R is defined as $Q + R := \{q + r \ : \ q \in Q, \ r \in R\}$.

   viii. $\mathcal{P} = \{(x_1, x_2) : (x_1 \geq x_2 - 1 \ \text{ and } \ x_2 \geq 0) \ \text{ OR } \ (x_1 \leq x_2 - 1 \ \text{ and } \ x_2 \leq 0)\}$

(b) For each of the functions given below, determine if it is a convex function or not, with a proper justification for your answer.

   i. $f(x) = a^\top x + b$, where $a \in \mathbb{R}^n$ and $b \in \mathbb{R}$.

   ii. $f(x) = \cos(x), x \in [\frac{\pi}{2}, \frac{3\pi}{2}]$.

   iii. $f(x) = x^\top Q x + a^\top x + b$, where $Q \in \mathbb{S}^n_+$, $a \in \mathbb{R}^n$ and $b \in \mathbb{R}$.

   iv. $\sum_{i=1}^m w_i f_i(x)$, $w_i \in \mathbb{R}$ and each f_i convex with domain $\mathbb{R}^n$.

   v. $g(f(x))$, where $f : \mathbb{R}^n \to \mathbb{R}$ is convex with domain $\mathrm{dom}(f)$ and $g : \mathbb{R} \to \mathbb{R}$ is convex and *nondecreasing* with domain $\mathbb{R}$.
   (Recall that a function $g : \mathbb{R} \to \mathbb{R}$ is called nondecreasing on its domain if, for every $x, y \in \mathrm{dom}(g)$, when we have $x < y$ we also have $g(x) \leq g(y)$.)

(c) For each of the sets given in the following question, find the convex hull of the set. Namely, state what the convex hull of the set is and justify why it is the convex hull of the set.

   i. $\mathcal{P} = \{(1,1), (2,3), (5,2)\}$, viewed as a subset of $\mathbb{R}^2$ (here (x, y) stands for the vector $\begin{bmatrix} x & y \end{bmatrix}^T$ as usual).

   ii. $\mathcal{P} = \{(\theta, y) \ : \ y = \sin(\theta), \theta \in \mathbb{R}\}$, viewed as a subset of $\mathbb{R}^2$.

   iii. $\mathcal{P} = \{z : \|z - z_0\|_2 = \epsilon\}$, where $z_0 \in \mathbb{R}^n$ and $\epsilon > 0$.

2 Visualizing Rank 1 Matrices

In this problem, we explore the effect of rank constraints on the convexity of sets of matrices. The relevant parts of the textbooks are Secs. 8.2.1, 8.2.2. and the beginning of Sec. 8.3 of the textbook of Calafiore and El Ghaoui and Secs. 3.1 and 3.2 of the textbook of Boyd and Vandenberghe.

First, consider the set of all 2×2 matrices with diagonal elements $(1, 2)$, which we can write as

$$\mathcal{R} := \left\{ \begin{bmatrix} 1 & x \\ y & 2 \end{bmatrix} \ : \ x, y \in \mathbb{R} \right\}.$$

(a) Is the set $\mathcal{R}$ convex? If so, provide a proof, and if not, provide a counterexample.

(b) Let us now define $\mathcal{R}^{(1)} \subset \mathcal{R}$ as the set of all rank-1 matrices in $\mathcal{R}$. Write out conditions on x and y which determine when a matrix $\begin{bmatrix} 1 & x \\ y & 2 \end{bmatrix} \in \mathcal{R}$ lies in $\mathcal{R}^{(1)}$.

(c) Is the set $\mathcal{R}^{(1)}$ convex? If so, provide a proof, and if not, provide a counterexample.
Hint: The image of a convex set under any affine mapping is a convex set. Here the function that maps $\begin{bmatrix} a & x \\ y & b \end{bmatrix}$ to $\begin{bmatrix} x \\ y \end{bmatrix}$ is a linear map from $\mathbb{R}^{2\times 2}$ to $\mathbb{R}^2$ and thus affine. This suggests that we could consider the image of $\mathcal{R}^{(1)}$ under this map and check if it is convex.

(d) In this class, we will sometimes pose optimization problems in which we optimize over sets of matrices. Since low-dimensional models are often easier to interpret, it would be nice to impose constraints that force the solutions to have low rank. Suppose we wish to solve the optimization problem

$$\min_{A \in \mathcal{R}^{(1)}} \|A\|_F^2,$$

which is equivalent to

$$\begin{aligned} &\min_{A \in \mathcal{R}} \|A\|_F^2 \\ \text{s.t.} \quad &\text{rk}(A) = 1. \end{aligned}$$

Is this optimization problem convex?

3 Further characterizations of convexity

The purpose of this problem is to develop some additional intuition about convex functions. The relevant parts of the textbooks are Secs. 8.2.1 and 8.2.2. of the textbook of Calafiore and El Ghaoui and Secs. 3.1 and 3.2 of the textbook of Boyd and Vandenberghe.

(a) Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function with dom$(f) = \mathbb{R}^n$. Suppose that f is bounded from above, i.e. there is a finite constant $K < \infty$ such that $f(x) \leq K$ for all $x \in$ dom(f). Show that f must be a constant function.

Hint: You can try to argue by contradiction, starting with the assumption that f is not constant and then seeing what the convexity of f would then imply.

(b) Let $\phi : \mathbb{R}^n \to \mathbb{R}^m$ be an affine function, i.e.

$$\phi(x) = Ax + b \text{ for all } x \in \mathbb{R}^n,$$

for some $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$.

Let $f : \mathbb{R}^m \to \mathbb{R}$ with domain dom(f) be a convex function. Let $g : \mathbb{R}^n \to \mathbb{R}$ be defined by $g(x) = f(\phi(x))$ for $x \in \phi^{-1}(\mathrm{dom}(f))$, i.e. g is the composition of f and ϕ. Here $\phi^{-1}(\mathrm{dom}(f))$ denotes the inverse image of $\mathrm{dom}(f)$ under ϕ, i.e.

$$\phi^{-1}(\mathrm{dom}(f)) := \{x \ : \ \phi(x) \in \mathrm{dom}(f)\}.$$

Show that g is a convex function on $\mathbb{R}^n$, with $\mathrm{dom}(g) = \phi^{-1}(\mathrm{dom}(f))$.

(c) Recall that $\mathbb{R}^{m \times n}$ can be viewed as a vector space, and that every matrix $A \in \mathbb{R}^{m \times n}$ has a well-defined largest singular value, denoted $\sigma_1(A)$, which is a nonnegative number.

Show that $\sigma_1 : \mathbb{R}^{m \times n} \to \mathbb{R}_+$, the function that maps a matrix to its largest singular value, is a convex function, with domain $\mathbb{R}^{m \times n}$.

Hint: You can use the characterization of the largest singular value of a matrix as its induced ℓ_2 norm, together with the result of the preceding part of this problem.

4 Convex Matrix Functions

This problem is related to Problem 5 of Homework 1. Relevant sections in the textbook include Chapter 4 and Secs. 8.2.1 and 8.2.2. of the textbook of Calafiore and El Ghaoui and Secs. 3.1 and 3.2 of the textbook of Boyd and Vandenberghe.

Suppose we are given a function $\varphi : \mathbb{R} \longrightarrow \mathbb{R}$ such that the Taylor series expansion

$$\varphi(x) = \sum_{k=0}^{\infty} \frac{\varphi^{(k)}(0)}{k!} x^k,$$

is valid for all $x \in \mathbb{R}$. Then we can extend φ to symmetric matrices $A \in \mathbb{S}^n$ via

$$\varphi(A) := \sum_{k=0}^{\infty} \frac{\varphi^{(k)}(0)}{k!} A^k,$$

where by convention $A^0 := I$. By the spectral theorem for symmetric matrices, we can write $A = \sum_{i=1}^{n} \lambda_i u_i u_i^\top$, where λ_i and u_i are the eigenvalues and a respective choice of corresponding eigenvectors of A. We can then show that

$$\varphi(A) = \sum_{i=1}^{n} \varphi(\lambda_i) u_i u_i^\top, \tag{1}$$

and so we can take $\sum_{i=1}^{n} \varphi(\lambda_i) u_i u_i^\top$ to be the definition of $\varphi(A)$ even if the Taylor series expansion of φ is not valid, as long as the domain of φ includes all the eigenvalues of A.

Note that we have seen the same sort of construction in Homework 1, in the special case where φ is the exponential function.

As usual, we will denote by $a_{ij} \in \mathbb{R}$ the element of A at the ith row and jth column.

(a) Let $A \in \mathbb{S}^n$ and $\varphi : \mathbb{R} \to \mathbb{R}$ be a function whose domain includes the eigenvalues of A. Define $\varphi(A)$ as in equation (1). Suppose P is an orthogonal matrix. Show that $\varphi(PAP^\top)$ is well-defined and in fact $\varphi(PAP^\top) = P\varphi(A)P^\top$.

(b) Suppose that $\psi : \mathbb{R} \longrightarrow \mathbb{R}$ is affine. Show that

$$\sum_{i=1}^{n} \psi(a_{ii}) = \operatorname{tr}(\psi(A)).$$

(c) Suppose now that $\varphi : \mathbb{R} \longrightarrow \mathbb{R}$ is convex, with $\operatorname{dom}(\varphi) = \mathbb{R}$. Show that

$$\sum_{i=1}^{n} \varphi(a_{ii}) \leq \operatorname{tr}(\varphi(A)).$$

Hint: Start by writing the eigendecomposition $A = U\Lambda U^\top$ and finding an explicit expression for the a_{ii} in terms of the entries of the orthogonal matrix U and the diagonal matrix Λ.

(d) **(Optional)**

Suppose $\varphi : \mathbb{R} \to \mathbb{R}$ is convex with domain $\mathbb{R}$. Show for $\theta \in [0, 1]$ that

$$\operatorname{tr}(\varphi(\theta B + (1-\theta)C)) \leq \theta \operatorname{tr}(\varphi(B)) + (1-\theta) \operatorname{tr}(\varphi(C)).$$

Hint: Try diagonalizing $\theta B + (1-\theta)C$ and using what was proved in the first part of this problem.

(e) Suppose now that $A \in \mathbb{S}^n_+$, i.e. A is a positive semidefinite matrix. Deduce that

$$\det(A) \leq \prod_{i=1}^{n} a_{ii}.$$

This inequality is called *Hadamard's inequality*.

Note that this inequality can be false if A is symmetric but not positive semidefinite, for example $A = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}$.

Hint: Separately consider the case where A is singular. When A is positive definite, consider the function $-\log : \mathbb{R} \to \mathbb{R}$, which is a convex function with domain $\mathbb{R}_{++}$ (i.e. the set of strictly positive real numbers). For any positive definite matrix $A \in \mathbb{S}^n_{++}$ we can define $-\log(A)$ using equation (1), because all the eigenvalues of A are strictly positive. We can then prove the claim in part (c) of this problem replacing ϕ by $-\log$, even though the domain of $-\log$ is only $\mathbb{R}_{++}$, as long as the matrix A is positive definite.

5 Subdifferentials

The purpose of this problem is to develop some intuition about subgradients and subdifferentials of convex functions that may not be differentiable. The relevant section is Sec. 8.2.3 of the textbook of Calafiore and El Ghaoui.

Consider the function $f : \mathbb{R} \to \mathbb{R}$ given by

$$f(x) = \begin{cases} 0 & \text{if } -1 \leq x \leq 0, \\ \frac{3}{2}x & \text{if } 0 < x \leq 2, \\ x^2 - 1 & \text{otherwise.} \end{cases} \tag{2}$$

Verify for yourself that f is continuous and convex with $\mathrm{dom}(f) = \mathbb{R}$. You do not need to provide a justification for these facts. A graph of f is provided in Figure 1.

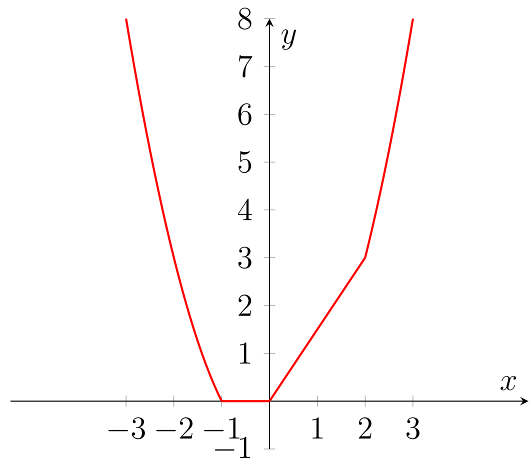

Figure 1: The graph of the function in equation (2).

Find the subdifferential $\partial f(x)$ at all $x \in \mathbb{R}$.