IE 539 Convex Optimization Assignment 1

Fall 2022

Out: 7th September 2022

Due: 23rd September 2022 at 11:59pm

Instructions

- Submit a PDF document with your solutions through the assignment portal on KLMS by the due date. Please ensure that your name and student ID are on the front page.
- Late assignments will **not** be accepted except in extenuating circumstances. Special consideration should be applied for in this case.
- It is **required** that you typeset your solutions in LaTeX. Handwritten solutions will not be accepted.
- Spend some time ensuring your arguments are **coherent** and your solutions **clearly** communicate your ideas.

Question:	1	2	3	4	5	6	Total
Points:	10	5	15	30	20	20	100

1. (10 points) Fix a point $y \in \mathbb{R}^d$ and a set $Z \subseteq \mathbb{R}^d$. Show that the set

$$X := \left\{x \in \mathbb{R}^d : \|x - y\|_2 \leq \operatorname{dist}(x, Z)\right\}$$

is convex, where $\operatorname{dist}(x, Z) := \min_{z \in Z} \|x - z\|_2$.

2. (5 points) Show that a quadratic function $f(x) = x^\top Qx + c^\top x$, where Q is symmetric positive semidefinite, is convex. Give an example of a quadratic (e.g., in $\mathbb{R}^2$) that is neither convex or concave.

3. This question asks you to prove that $f(x) = \|Ax - b\|_2^2/(1 - \|x\|_2^2)$ is convex on the open unit ball $X = \{x \in \mathbb{R}^d : \|x\|_2 < 1\}$. Answer the following sub-questions that lead to a proof.
   (a) (5 points) Show that $g(x, t) = \|x\|_2^2 + \|Ax - b\|_2^2/t$ over $(x, t) \in X \times \mathbb{R}_{++}$ is convex.
   (b) (5 points) Show that the epigraph of f is given by

$$\operatorname{epi}(f) = \{(x, t) : g(x, t) \leq 1, \|x\|_2 < 1, t > 0\} \cup \{(x, 0) : \|x\|_2 < 1, Ax = b\}.$$

   (c) (5 points) Complete the proof that f is convex on X.

4. Verify convexity/concavity of the following functions. You may use the first-order and second-order characterizations of convex functions, while there exists a direct proof based on the definition of convex functions.
   (a) (5 points) The *negative entropy function* is convex on $\mathbb{R}^d_{++}$:

$$f(x) := \sum_{i \in [d]} x_i \log(x_i).$$

   (b) (5 points) The *log-sum-exp function* is convex:

$$f(x) = \log\left(\sum_{i \in [d]} \exp(x_i)\right).$$

   [Hint: An elementary proof exists by showing $f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$. You may use (without proof) the inequality $\sum_{i \in [d]} |u_i|^\lambda |v_i|^{1-\lambda} \leq \left(\sum_{i \in [d]} |u_i|\right)^\lambda \left(\sum_{i \in [d]} |v_i|\right)^{1-\lambda}$.]

   (c) (5 points) The geometric mean is concave on $\mathbb{R}^d_{++}$:

$$f(x) = \left(\prod_{i \in [d]} x_i\right)^{1/d}.$$

   [Hint: compute $\frac{\lambda f(x) + (1-\lambda) f(y)}{f(\lambda x + (1-\lambda) y)}$. You may use (without proof) the arithmetic-geometric mean inequality: $\left(\prod_{i \in [d]} x_i\right)^{1/d} \leq \frac{1}{d} \sum_{i \in [d]} x_i$.]

   (d) (5 points) The log-determinant is concave on $\mathbb{S}^d_{++}$, the set of $d \times d$ positive "definite" matrices:

$$f(X) = \log \det(X).$$

   Note: the determinant of a matrix $X \in \mathbb{S}^d_{++}$ is simply the product of its eigenvalues (which are all positive by assumption). You will need the following properties of matrices/determinants which you can use without proof:
   - $\det(AB) = \det(BA) = \det(A)\det(B)$ for $A, B \in \mathbb{S}^d_{++}$.
   - If $A \in \mathbb{S}^d_{++}$, then we can write $A = PDP^\top$ where $P^\top P = PP^\top = I$ and D is diagonal with strictly positive entries. Let D^r be the diagonal matrix with all diagonal entries of D raised to the power $r \in \mathbb{R}$. We can define powers of A via $A^r = PD^rP^\top$, which have the same properties as the usual powers: $A^rA^s = A^sA^r = A^{r+s}$, $A^0 = I$. Furthermore, $\det(A^r) = \det(A)^r$.
   - $\alpha A + \beta B = A^{1/2}(\alpha I + \beta A^{-1/2}BA^{-1/2})A^{1/2}$ for any $A, B \in \mathbb{S}^d_{++}$ and $\alpha, \beta \geq 0$.

   (e) (5 points) The *conjugate* of a function $f : \mathbb{R}^d \to \mathbb{R}$:

$$f^*(x) = \sup_{y \in \mathbb{R}^d} \{\langle y, x\rangle - f(y)\}$$

(f) (5 points) The sum of k largest components of $x \in \mathbb{R}^d$:

$$f(x) = x_{\sigma(1)} + \cdots + x_{\sigma(k)}$$

where $x_{\sigma(1)} \geq \cdots \geq x_{\sigma(d)}$ are the rearrangement of $x_1, \ldots, x_d$ in nonincreasing order.

5. Let $C \subseteq \mathbb{R}^d$ be a set. The *dual cone* of C is defined as

$$C^* = \left\{ y \in \mathbb{R}^d : y^\top x \geq 0 \ \forall x \in C \right\}.$$

(a) (5 points) Show that C^* is a convex cone.

(b) (5 points) Let $K = \{(x, y, z) \in \mathbb{R} \times \mathbb{R}_{++} \times \mathbb{R} : y \exp(x/y) \leq z\}$. Show that K is a convex cone.

(c) (2 points) The dual cone of K is given by

$$K^* = \left\{ (u, v, w) \in \mathbb{R}^3 : ux + vy + wz \geq 0 \ \forall (x, y, z) \in K \right\}.$$

Show that for any $(u, v, w) \in K^*$ with $w = 0$, the point satisfies $u = 0$ and $v \geq 0$. Moreover, argue that $\{(0, v, 0) : v \geq 0\} \subset K^*$.

(d) (2 points) Show that for any $(u, v, w) \in K^*$ with $w > 0$ and $u = 0$, the point satisfies $v \geq 0$. Moreover, argue that $\{(0, v, 0) : v \geq 0\} \subset K^*$. Moreover, argue that $\{(0, v, w) : v \geq 0, w > 0\} \subset K^*$.

(e) (2 points) Show that for any $(u, v, w) \in K^*$ with $w > 0$ and $u \neq 0$, the point satisfies $u < 0$.

(f) (4 points) Finish the proof that the dual cone of K from part (b) is given by

$$K^* = \{(u, v, w) : u < 0, -u \exp(v/u) \leq \exp(1) w\} \cup \{(0, v, w) : v, w \geq 0\}.$$

[Hint: argue that if $w > 0$ and $u < 0$, then $-u \exp v/u \leq \exp(1) w$ holds. You can use the fact (without proof) that minimizers of a differentiable convex function $f : \mathbb{R} \to \mathbb{R}$ occur at points where $f'(x) = 0$. We prove a more general version of this in lectures.]

6. This question asks you to show that the following matrix functions are norms. Note that matrices A we take are not necessarily square matrices.

(a) (5 points) The spectral norm of a (real) matrix:

$$\|A\|_2 = \sqrt{\lambda_{\max}(A^\top A)}$$

where $\lambda_{\max}(A^\top A)$ is the largest eigenvalue of $A^\top A$.

(b) (5 points) The Frobenius norm of a matrix:

$$\|A\|_F = \sqrt{\operatorname{tr}\left\{A^\top A\right\}}$$

where $\operatorname{tr}\{A^\top A\}$ is the trace of $A^\top A$, defined by, the sum of diagonal entries of $A^\top A$.

(c) (10 points) The nuclear norm of a symmetric matrix $S \in \mathbb{S}^{d \times d}$

$$\|S\|_{\text{nuc}} = \sum_{i=1}^{d} \sqrt{\lambda_i(S^\top S)}$$

where $\lambda_1(S^\top S), \ldots, \lambda_d(S^\top S)$ are the eigenvalues of $S^\top S$.