IE 539 Convex Optimization Assignment 1

Fall 2024

Out: 23rd September 2024

Due: 6th October 2024 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	Total
Points:	10	10	35	25	20	100

1. (10 points) Prove that a function is covex if and only if its epigraph is a convex set.
2. (10 points) Suppose that $g_1, \dots, g_p : \mathbb{R}^d \to \mathbb{R}$ are convex and $h_1, \dots, h_q : \mathbb{R}^d \to \mathbb{R}$ are affine functions. Prove that
$$C := \{x \in \mathbb{R}^d : \ g_i(x) \le 0 \text{ for } i = 1, \dots, p, \ h_i(x) = 0 \text{ for } j = 1, \dots, q\}$$
is a convex set.
3. 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 $f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda) f(y)$. You may use (without proof) the inequality $\sum_{i \in [d]} |u_i|^\lambda |v_i|^{1-\lambda} \le \left(\sum_{i \in [d]} |u_i|\right)^\lambda \left(\sum_{i \in [d]} |v_i|\right)^{1-\lambda}$.]
   (c) (10 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: $\left(\prod_{i \in [d]} x_i\right)^{1/d} \le \frac{1}{d} \sum_{i \in [d]} x_i$.]
   (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, $\det(A^r) = \det(A)^r$.
   - $\alpha A + \beta B = A^{1/2}(\alpha I + \beta A^{-1/2} B A^{-1/2}) A^{1/2}$ for any $A, B \in \mathbb{S}^d_{++}$ and $\alpha, \beta \ge 0$.

   (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 $x_{\sigma(1)} \ge \cdots \ge x_{\sigma(d)}$ are the rearrangement of $x_1, \dots, x_d$ in nonincreasing order.
4. This question asks you to show that the following formulation for uncertainty quantification is a convex optimization problem.
$$\begin{aligned} \text{maximize} \quad & x^\top (\bar{\Sigma} + S) x \\ \text{subject to} \quad & \bar{\Sigma} + S \succeq 0, \\ & \|S\|_{\text{nuc}} \le \epsilon, \\ & S \in \mathbb{R}^{d \times d} \end{aligned}$$
where $\bar{\Sigma}$ is an empirical covariance matrix and $A \succeq 0$ means matrix A being positive semidefinite.

(a) (10 points) The *nuclear norm* of a square matrix $S \in \mathbb{R}^{d\times d}$ is defined as

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

where $\lambda_1(S^\top S), \ldots, \lambda_d(S^\top S)$ are the eigenvalues of $S^\top S$. Prove that the nuclear norm is a norm.

(b) (10 points) Prove that

$$\left\{S \in \mathbb{R}^{d\times d} : \ \bar{\Sigma} + S \succeq 0\right\}$$

is a convex set.

(c) (5 points) Prove that the formulation is a convex optimization problem.

5. 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) (10 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) (10 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 $A^\top A$.