Optimization and Machine Learning, Spring 2020
Homework 6

(Due Tuesday, June 16 at 11:59pm (CST))

1. Which of the following sets are convex?
   (a) A *wedge*, *i.e.*, $\{x \in \mathbb{R}^n \mid a_1^T x \leq b_1, a_2^T x \leq b_2\}$. (5 points)
   (b) The set of points closer to a given point than a given set, *i.e.*,
   $$\{x \mid ||x - x_0||_2 \leq ||x - y||_2 \text{ for all } y \in S\}$$
   where $S \subseteq \mathbb{R}^n$. (5 points)
   (c) The set of points closer to one set than another, *i.e.*,
   $$\{x \mid \mathbf{dist}(x, S) \leq \mathbf{dist}(x, T)\}$$
   where $S, T \subseteq \mathbb{R}^n$, and
   $$\mathbf{dist}(x, S) = \inf\{||x - z||_2 \mid z \in S\}.$$
   (5 points)
   (d) The set $\{x \mid x + S_2 \subseteq S_1\}$, where $S_1, S_2 \subseteq \mathbb{R}^n$ with S_1 convex. (5 points)
   (e) The set of multiplication
   $$\{x \in \mathbb{R}_+^n \mid \prod_{i=1}^n x_i \geq 1\}.$$
   (5 points)

2. Determine whether the following functions are convex, strictly convex, concave, strictly concave, both or neither.
   (a) $f(x) = e^x - 1$ on $\mathbb{R}$. (5 points)
   (b) $f(x_1, x_2) = x_1 x_2$ on $\mathbb{R}_{++}^2$. (5 points)
   (c) $f(x) = \log(\sum_{i=1}^n \exp(x_i))$ on $\mathbb{R}^n$, use the second-order condition. (5 points)
   (d) $f(w) = ||Xw - y||_2^2 + \lambda ||w||_2^2$ for $\lambda > 0$. (5 points)
   (e) The log-likelihood of a set of points $\{x_1, \cdots, x_n\}$ that are normally distributed with mean μ and finite variance $\sigma > 0$ is given by:
   $$f(\mu, \sigma) = n \log(\frac{1}{\sqrt{2\pi}\sigma}) - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2$$
   Show that if we view the log likelihood for fixed σ as a function of the mean, *i.e.*,
   $$g(\mu) = n \log(\frac{1}{\sqrt{2\pi}\sigma}) - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2$$
   then g is strictly concave.
   Show that if we view the log likelihood for fixed μ as a function of z, *i.e.*,
   $$h(z) = n \log(\frac{\sqrt{z}}{\sqrt{2\pi}}) - \frac{z}{2} \sum_{i=1}^n (x_i - \mu)^2$$
   then h is strictly concave (equivalently, we say f is strictly concave in $z = \frac{1}{\sigma^2}$).
   We say $f(x, y)$ with $x \in \mathbb{R}^m$ and $y \in \mathbb{R}^n$ is jointly convex if
   $$f\big(\lambda(x_1, y_1) + (1 - \lambda)(x_2, y_2)\big) \leq \lambda f\big((x_1, y_1)\big) + (1 - \lambda) f\big((x_2, y_2)\big).$$
   Show that f is not jointly concave in $\mu, \frac{1}{\sigma^2}$. (5 points)

3. Consider the problem

$$\begin{array}{ll} \text{minimize} & \|Ax-b\|_1 / \left(c^T x + d\right) \\ \text{subject to} & \|x\|_\infty \leq 1, \end{array}$$

where $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m, c \in \mathbb{R}^n$ and $d \in \mathbb{R}$. We assume that $d > \|c\|_1$, which implies that $c^{\mathrm{T}}x + d > 0$ for all feasible x.

(a) Show that this is a quasiconvex optimization problem. (5 points)

(b) Show that it is equivalent to the convex optimization problem

$$\begin{array}{ll} \text{minimize} & \|Ay-bt\|_1 \\ \text{subject to} & \|y\|_\infty \leq t, \\ & c^T y + dt = 1, \end{array}$$

with variables $y \in \mathbb{R}^n, t \in \mathbb{R}$. (10 points)

4. Consider the QCQP

$$\begin{array}{ll} \text{minimize} & (1/2)x^T P x + q^T x + r \\ \text{subject to} & x^T x \leq 1, \end{array}$$

with $P \in \mathbf{S}^n_{++}$. Show that $x^\star = -(P+\lambda I)^{-1} q$ where $\lambda = \max\{0, \bar{\lambda}\}$ and $\bar{\lambda}$ is the largest solution of the nonlinear equation

$$q^{\mathrm{T}}(P+\lambda I)^{-2} q = 1.$$

(15 points)

5. Consider the inequality form LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \preceq b, \end{array}$$

with $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m$. Let $w \in \mathbb{R}^m_+$. If x is feasible for the LP, *i.e.*, satisfies $Ax \preceq b$, then it also satisfies the inequality

$$w^T A x \leq w^T b.$$

Geometrically, for any $w \succeq 0$, the halfspace $H_w = \{x \mid w^T A x \leq w^T b\}$ contains the feasible set for the LP. Therefore if we minimize the objective $c^T x$ over the halfspace Hw we get a lower bound on p^*.

(a) Derive an expression for the minimum value of $c^T x$ over the halfspace H_w (which will depend on the choice of $w \succeq 0$). (5 points)

(b) Formulate the problem of finding the best such bound, by maximizing the lower bound over $w \succeq 0$. (5 points)

(c) Relate the results of (a) and (b) to the Lagrange dual of the LP. (10 points)