10-801: Advanced Optimization and Randomized Methods

Homework 1: Convex sets and functions

(Jan 22, 2014)

Instructor: *Suvrit Sra* **Due:** *Feb 03, 2014*

Visit: http://www.cs.cmu.edu/~suvrit/teach/ for academic rules for homeworks.

1. Prove that the functions in (a)–(b) below are convex, without resorting to second derivatives.

   (a) $f(x, y) = x^2/y$ for $y > 0$ on $\mathbb{R} \times \mathbb{R}_{++}$

   (b) $f(x) = \log(1 + e^{\sum_i a_i x_i})$ on $\mathbb{R}^n$ ($a_i \in \mathbb{R}$ for $1 \le i \le n$).

   (c) Using (b) show that $\det(X + Y)^{1/n} \ge \det(X)^{1/n} + \det(Y)^{1/n}$ for $X, Y \in \mathbf{S}^n_{++}$.

2. Let $f : \mathbb{R} \to \mathbb{R}_{++}$. Prove that f is log-convex *if and only if* $e^{cx} f(x)$ is convex for every $c \in \mathbb{R}$ (we assume f is continuous but not that it is differentiable).

3. Fenchel conjugates:

   (a) Derive the Fenchel conjugate for $x^T Ax + b^T x$ where $A \succeq 0$ may be rank-deficient

   (b) Consider the *quasi-norm* $f(x) := \|x\|_{1/2} := \left[\sum_{i=1}^n |x_i|^{1/2}\right]^2$. What is its *bi-conjugate* f^{**}?

4. Let a vector x be split into nonoverlapping subvectors $x_1, ..., x_G$, then we define its $\ell_{p,q}$*-mixed norm* as

$$\|x\|_{p,q} := \left(\sum_i^G \|x_i\|_q^p\right)^{1/p}, \quad p, q \ge 1.$$

   Derive the dual norm to this norm (Hint: it is another mixed-norm).

   (*Remark:* The norms $\ell_{1,2}, \ell_{1,\infty}$ and $\ell_{2,1}$ are perhaps the most interesting examples; they come up in multitask lasso and group lasso problems.)

5. Consider the normed metric space: $\mathbb{R}^n$. Define the function

$$d(x, y) := \frac{2\|x - y\|}{\|x\| + \|y\| + \|x - y\|}, \qquad \forall x, y \in \mathbb{R}^n.$$

   Prove that d is a metric on $\mathbb{R}^n \setminus \{0\}$.

6. Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is a symmetric function, (i.e., if $x = [x_1, x_2, \ldots, x_n]$ and $x_\sigma = [x_{\sigma(1)}, \ldots, x_{\sigma(n)}]$ for any permutation $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$, then $f(x_\sigma) = f(x)$). Let $S^{n\times n}$ be the set of $n \times n$ symmetric matrices, and $\lambda : S^{n\times n} \to \mathbb{R}^n$ the *eigenvalue map*, that maps a symmetric matrix to the sorted (↓) vector of its eigenvalues. Show that the Fenchel conjugate of the composite function

$$(f \circ \lambda)^* = f^* \circ \lambda.$$

   [*Hint:* This question is simpler than it appears. Use the fact that for any two matrices $X, Y \in S^{n\times n}$ we have the inequality

$$\mathrm{tr}(XY) \le \lambda(X)^T \lambda(Y).$$

   Also useful is to remember that $\lambda(\cdot)$ and tr enjoy the following invariance: $\lambda(QAQ^T) = \lambda(A)$ for orthogonal Q, and $\mathrm{tr}(QAQ^T) = \mathrm{tr}(A)$. To prove the claim, try showing $(f \circ \lambda)^* \le f^* \circ \lambda$ and $(f \circ \lambda)^* \ge f^* \circ \lambda$. It'll be helpful to consider $Y = U\mathrm{Diag}\lambda(Y)U^T$.]

7. [*Bonus*] Let x and y be vectors whose coordinates are in sorted order, so that

$$x_1 \ge x_2 \ge \ldots \ge x_n, \quad y_1 \ge y_2 \ge \ldots \ge y_n.$$

   Suppose now that x and y satisfy the following

$$\sum_{i=1}^k x_i \le \sum_{i=1}^k y_i \text{ for } 1 \le k < n$$
$$\sum_{i=1}^n x_i = \sum_{i=1}^n y_i.$$

   Prove that for convex function $f : \mathbb{R} \to \mathbb{R}$, it must hold that

$$\sum_{i=1}^n f(x_i) \le \sum_{i=1}^n f(y_i).$$