Optimization

Problem set 1

Due Monday, October 21st

1. Consider the closed convex set $B_1 = \{x \in \mathbb{R}^n | \|x\|_1 = \sum_i |x_i| \leq 1\}$. This is the unit ball of the ℓ_1 norm.

    (a) Show that B_1 is a polyhedron by explicitly expressing it as an intersection of halfspaces. How many halfspaces ("facets") are required in order to express B_1?

    (b) Explicitly express B_1 as a convex hull of a finite number of points. How many points ("vertices") are required in this characterization?

    (c) Contrast this with the ℓ_∞ unit ball, $B_\infty = \{x \in \mathbb{R}^n | \|x\|_\infty \leq 1\}$. How many halfspaces are required in order to express B_∞ as an intersection of halfspaces? How many points are required in order to express B_∞ as a convex hull?

    (d) For each point $\hat{x}$ on the boundary of B_1, identify the set of all supporting hyperplanes of B_1 at $\hat{x}$ explicitly. For each such $\hat{x}$, what is the dimensionality of this set?

2. Consider a polyhedron $C = \text{conv}\{v_1, \ldots, v_k\} \subset \mathbb{R}^n$ and a convex function $f : \mathbb{R}^n \to \mathbb{R}$.

    (a) Prove that a maximum of f over C is achieved at one of the vertices v_i. (Hint: assume the statement is false and use Jensen's inequality). Is it possible that the maximum is also achieved at an interior point?

    (A generalization of the above is that a maximum of a function over a closed and bounded convex set is achieved at an extreme point, i.e. a point which is not a convex combination of other points in the set).

    (b) Use the above to conclude that the *minimum* of a linear objective over the polyhedron C is always achieved at one of the vertices v_i.

3. In this problem we will define strong convexity more generally then it is defined by Boyd and Vandenberghe (Section 9.1.2). In particular, we will consider a definition that is valid also for non-differentiable functions.

   Definition: A function $f : \mathbb{R}^n \to \mathbb{R}$ is m-strongly convex if for every $x, y \in \mathbb{R}^n$ and every $\theta \in [0, 1]$:

$$f((1-\theta)x + \theta y) \leq (1-\theta)f(x) + \theta f(y) - \frac{m}{2}\theta(1-\theta)\left\|x-y\right\|_2^2$$

   (a) Prove that a continuously differentiable function f is m-strongly convex if and only if for every $x, y \in \mathbb{R}^n$,

$$f(y) \geq f(x) + \nabla f(x)^T(y-x) + \frac{m}{2}\left\|y-x\right\|_2^2.$$

   This generalizes the first order characterization of convexity (Section 3.1.3).

   (b) Prove that a twice continuously differentiable function f is m-strongly convex if and only if its domain and convex and for every $x \in \mathbb{R}^n$, all eigenvalues of the Hessian at x are greater or equal to m, i.e.:

$$\nabla^2 f(x) \succcurlyeq mI.$$

   This generalizes the second order characterization of convexity (Section 3.1.4) and is the definition used in Section 9.1.2.

   (c) Provide an example of a function that is strongly convex but not everywhere differentiable.

   (d) Let f be a m-strongly convex function, and x^* an optimum for $\min_{x\in\mathbb{R}^n} f(x)$. Prove that for any point $x \in \mathbb{R}^n$:

$$f(x) \geq f(x^*) + \frac{m}{2}\left\|x-x^*\right\|_2^2.$$

   Conclude that the optimum is unique and that any ϵ-suboptimal point must be close to the optimum. Provide an explicit upper bound on $\left\|x-x^*\right\|_2$ for an ϵ-suboptimal x. (Note that if f is convex but not strongly convex, ϵ-suboptimal points can be arbitrarily far away from the closest optimum).

Recommended review exercises from Boyd and Vandenberghe (please do not turn these in—they will *not* be graded): 2.12, 2.15, 3.6, 3.16, 3.18, 3.24, 3.26.