The University of Texas at Austin
Optimization

Homework 8

Constantine Caramanis, Sujay Sanghavi

Submitting solutions: *Please submit your solutions as a single pdf file. If you have code or figures, please include these in the pdf.*

1. **Subgradients**. Show the following for sub-gradients

   (a) If $g(x) = f(Ax + b)$, then $\partial g(x) \supseteq A^T \partial f(Ax + b)$.

   (b) If $f(x) = \max_{1 \le i \le m} f_i(x)$, then $\partial f(x) \supseteq \text{conv}\left(\bigcup_{i: f_i(x) = f(x)} \partial f_i(x)\right)$

   In fact, the containments above are equalities, but the reverse inclusions are more delicate, so we are only asking you to show one inclusion. For your interest, we will detail the argument for both directions on the solution.

2. **More Subgradients**. Show that for any convex function f, and for any points x, y, and $u \in \partial f(x)$, $v \in \partial f(y)$,

   $$\langle u - v, x - y \rangle \ge 0.$$

   This shows that the subdifferential is a monotone operator for any convex function. We will not use this specific property in this class, but it is heavily used for more advanced aspects of convex optimization and convex analysis.

3. **Coordinate Descent**

   (a) Give an example that shows that coordinate descent may not find the optimum of a convex function. That is, provide a simple function f and a point x such that coordinate descent starting from x will *not* get to the global minimum of f.

   (b) Let $f(x, y) = x^2 + y^2 + 3xy$, where x, y are scalars. Note that f is not convex. Would coordinate descent with exact line search always converge to a stationary point ?

4. **Frank Wolfe and PGD**. In the lectures, we showed an example where Frank-Wolge and projected gradient descent (PGD) behave very differently. Replicate that here, and show your plots of the trajectory obtained.

   $$\begin{aligned} \text{min}: \quad & 100x_1^2 + x_2^2 + (x_3 - 20)^2 \\ \text{s.t.}: \quad & x_1 + x_2 + x_3/20 = 1 \\ & x_1, x_2, x_3 \ge 0. \end{aligned}$$