Assignment 3

CIS 621: Algorithm and Complexity

Problem 1 (4 points) Consider $f : \mathbb{R}^n \to \mathbb{R}$, $f(x) = \|x\|_2$. Show that, for $x \neq 0$, the subgradient is $\frac{x}{\|x\|_2}$; for $x = 0$, the subgradient is any element in the set of $\{y \mid \|y\|_2 \leq 1\}$.

Solution $x \neq 0$: For any z, we need to show $\|z\|_2 \geq \frac{x^T}{\|x\|_2}(z-x) + \|x\|_2$, which is equivalent to showing $\|z\|_2\|x\|_2 \geq x^T(z-x) + \|x\|_2^2 = x^T z$. This actually holds, because $x^T z \leq |x^T z| \leq \|x\|_2\|z\|_2$.

$x = 0$: For any z, we need to show $\|z\|_2 \geq y^T z$, where $\|y\|_2 \leq 1$. We note $y^T z \leq |y^T z| \leq \|y\|_2\|z\|_2 \leq \|z\|_2$.

(Hint: See "Cauchy-Schwarz Inequality.")

Problem 2 (6 points) Consider $f(x) = \max\{f_1(x), f_2(x)\}$, where $f_1, f_2 : \mathbb{R}^n \to \mathbb{R}$ are convex and differentiable. Show that, for $f_1(x) > f_2(x)$, the subgradient is $\nabla f_1(x)$; for $f_1(x) < f_2(x)$, the subgradient is $\nabla f_2(x)$; for $f_1(x) = f_2(x)$, the subgradient is any point on the line segment between $\nabla f_1(x)$ and $\nabla f_2(x)$.

Solution $f_1(x) > f_2(x)$: For any y, we have $f(y) = f_1(y)$ and $f(x) + \nabla f_1(x)^T(y-x) = f_1(x) + \nabla f_1(x)^T(y-x)$. We have $f(y) \geq f(x) + \nabla f_1(x)^T(y-x)$ due to the convexity of f_1. Thus, $\nabla f_1(x)$ is a subgradient.

$f_1(x) < f_2(x)$: Omitted. Similar to the above; due to the convexity of f_2.

$f_1(x) = f_2(x)$: Omitted. Similar to the above; due to the convexity of f_1 (or f_2).

Problem 3 (4 points) Use the definition of subdifferentials to show that the following two functions are *not* subdifferentiable at $x = 0$: (1) $f(x) = -x^{\frac{1}{2}}$; (2) $f(x)$ such that $f(0) = 1$ and $f(x) = 0$ for $x > 0$.

Solution Let's prove by contradiction. (1): Assuming $-x^{\frac{1}{2}}$ is subdifferentiable at $x = 0$, then there exists a subgradient at $x = 0$, denoted as g. For any y, we should have $f(y) \geq g(y-0) + f(0)$, i.e., $-y^{\frac{1}{2}} \geq gy$. However, we note that there does not actually exist such g that makes $-y^{\frac{1}{2}} \geq gy$ hold for every possible y. That is, $-x^{\frac{1}{2}}$ is not subdifferentiable at $x = 0$.

(2): Similarly, for any y, we should have $f(y) \geq g(y-0) + f(0)$, i.e., $0 \geq gy + 1$ if $y > 0$. There does not actually exist such g that makes $0 \geq gy + 1$ hold for every possible $y > 0$.

Problem 4 (2 points) Consider the subgradient method $x^+ = x - \alpha g$, where $g \in \partial f(x)$. Show that if $\alpha < \frac{2(f(x)-f(x^*))}{\|g\|_2^2}$, then we have $\|x^+ - x^*\|_2 < \|x - x^*\|_2$, i.e., every iteration moves closer to the optimal solution x^*.

Solution

$$
\begin{aligned}
& \|x^+ - x^*\|_2^2 \\
= & \|x - \alpha g - x^*\|_2^2 \\
= & \|x - x^*\|_2^2 + \alpha^2\|g\|_2^2 - 2\alpha g^T(x - x^*) \\
\leq & \|x - x^*\|_2^2 + \alpha^2\|g\|_2^2 - 2\alpha(f(x) - f(x^*)) \\
< & \|x - x^*\|_2^2 + \alpha\frac{2(f(x)-f(x^*))}{\|g\|_2^2}\|g\|_2^2 - 2\alpha(f(x) - f(x^*)) \\
= & \|x - x^*\|_2^2
\end{aligned}
$$

That is, $\|x^+ - x^*\|_2 < \|x - x^*\|_2$.

Problem 5 (4 points) Use "Lagrange relaxation" and "dual decomposition" to design a distributed algorithm to find the optimal value of the objective function of the following problem, and describe your distributed algorithm elaborately. $a_{s,u}$, $b_{s,u}$, c_u, d_s, $\forall s \in \mathcal{S}$, $\forall u \in \mathcal{U}$ are all nonnegative constants.

$$
\begin{aligned}
\min \quad & \sum_{s\in\mathcal{S}}\sum_{u\in\mathcal{U}} a_{s,u}x_{s,u} + \sum_{s\in\mathcal{S}}\sum_{u\in\mathcal{U}} b_{s,u}\left((x_{s,u}+1)\ln(x_{s,u}+1) - x_{s,u}\right) \\
s.t. \quad & \sum_{s\in\mathcal{S}} x_{s,u} \geq c_u,\ \forall u \in \mathcal{U},
\end{aligned}
$$

$$\sum_{u \in \mathcal{U}} x_{s,u} \geq d_s, \ \forall s \in \mathcal{S}\,,$$
$$x_{s,u} \geq 0, \ \forall s \in \mathcal{S}, \ \forall u \in \mathcal{U}\,.$$

Solution If we split the problem across u, we need to put the second constraint into the objective; if we split the problem across s, we need to put the first constraint into the objective. Below, we split across u. Let's use $\alpha_s \geq 0$ as the dual variable associated to the second constraint.

We do dual decomposition. Let's set $\alpha_s^{(0)} = 0$, $\forall s$. Then we do the following iteration, $\forall k \geq 0$:

$$\alpha_s^{(k+1)} = \max\{\alpha_s^{(k)} + t(d_s - \sum_{u \in \mathcal{U}} x^*_{s,u}), 0\}\,, \forall s,$$

where t is the step size and $x^*_{s,u}$ is the optimal solution to the following problem (i.e., Lagrange relaxation):

$$\begin{aligned} \min \quad & \sum_{s \in \mathcal{S}} \sum_{u \in \mathcal{U}} a_{s,u} x_{s,u} + \sum_{s \in \mathcal{S}} \sum_{u \in \mathcal{U}} b_{s,u} \left((x_{s,u}+1)\ln(x_{s,u}+1) - x_{s,u}\right) + \sum_{s \in \mathcal{S}} \alpha_s^{(k)} (d_s - \sum_{u \in \mathcal{U}} x_{s,u}) \\ s.\,t. \quad & \sum_{s \in \mathcal{S}} x_{s,u} \geq c_u, \ \forall u \in \mathcal{U}\,, \\ & x_{s,u} \geq 0, \ \forall s \in \mathcal{S}, \ \forall u \in \mathcal{U}\,. \end{aligned}$$

Note that the above problem can be solved in a distributed manner, i.e., each u solves the following problem by any convex program solver:

$$\begin{aligned} \min \quad & \sum_{s \in \mathcal{S}} a_{s,u} x_{s,u} + \sum_{s \in \mathcal{S}} \left((x_{s,u}+1)\ln(x_{s,u}+1) - x_{s,u}\right) - \sum_{s \in \mathcal{S}} \alpha_s^{(k)} x_{s,u} \\ s.\,t. \quad & \sum_{s \in \mathcal{S}} x_{s,u} \geq c_u\,, \\ & x_{s,u} \geq 0, \ \forall s \in \mathcal{S}\,. \end{aligned}$$