

CIS 621 Assignment 2

Steven Walton

February 18, 2019

Problem 1

Consider $f : \mathbb{R}^n \rightarrow \mathbb{R}, f(x) = \|x\|_2$. Show that

$$g = \begin{cases} \frac{x}{\|x\|_2} & x \neq 0 \\ \{y \mid \|y\|_2 \leq 1\} & x = 0 \end{cases}$$

As usual, we will just take the derivative of f . We know this is what to do because of the definition of subgradients. Again, we use Einsteinian notation.

$$\begin{aligned} f(x) &= \|x\|_2 \\ &= \sqrt{x_i^2} \\ \frac{\partial f}{\partial x_i} &= (2x_i) \left(\frac{1}{2} \frac{1}{\sqrt{x_i^2}} \right) \\ &= \frac{x_i}{\sqrt{x_i^2}} \\ \nabla f &= \frac{x}{\|x\|_2} \end{aligned}$$

Therefore we can see that any gradient of the function is a subgradient. We should also recognize that this is the same as $\text{norm}\{x\}$.

Dealing with the case where $x = 0$ we can see that the derivative is $\{g \mid \|g\|_2 \leq 1\}$. \therefore we have our solution.

Problem 2

Consider $f(x) = \max\{f_1(x), f_2(x)\}$ where $f_1, f_2 : \mathbb{R}^n \rightarrow \mathbb{R}$ 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 $\nabla f_2(x)$, and for $f_1(x) = f_2(x)$ the subgradient is any point on the line segment between $\nabla f_1(x), \nabla f_2(x)$.

This problem is fairly straight forward. Remembering back to the definition of subgradients we may recall, again, that when f is differentiable then ∇f is the only

possible choice for g . Clearly from there we can see

$$\begin{aligned}
 f_1(x) &> f_2(x) \\
 \max\{f_1(x), f_2(x)\} &= f_1(x) \\
 \Rightarrow g(x) &= \nabla f_1(x) \\
 f_1(x) &< f_2(x) \\
 \max\{f_1(x), f_2(x)\} &= f_2(x) \\
 \Rightarrow g(x) &= \nabla f_2(x)
 \end{aligned}$$

The trick is that when $f_1 = f_2$ then both are acceptable solutions to the max function. Knowing this we can see that $g = \nabla f_1$ and $g = \nabla f_2$ are acceptable solutions. \therefore we can conclude that any solution between these two solutions is also true. Or more explicitly

$$g(x) = \alpha \nabla f_1(x) + \beta \nabla f_2(x)$$

where $\alpha + \beta = 1$

Problem 3

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) $f(x)$ s.t. $f(0) = 1$ and $f(x) = 0, x > 0$

Part 1)

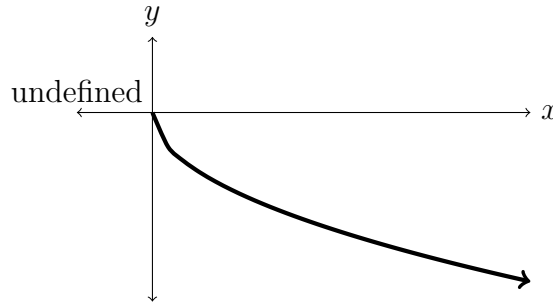


Figure 1: $-x^{\frac{1}{2}}$

$$\begin{aligned}
 \nabla f &= -\frac{1}{2}x^{-\frac{1}{2}} \\
 &= -\frac{1}{2} \frac{1}{\sqrt{x}}
 \end{aligned}$$

Trivially we can see that when $x = 0$ this function is not defined.

To show that it has no subdifferential we need to note that $\lim_{x \rightarrow 0^-} f(x) \neq \lim_{x \rightarrow 0^+} f(x)$ because $\lim_{x \rightarrow 0^-} f(x)$ is undefined and $\notin \mathbb{R}$. \therefore there is no subdifferential. (While a

subdifferential does not need to be differentiable at a point, its limit must exist and be well defined)

Part 2)

This functions is not differentiable, because this function has is discontinuous at

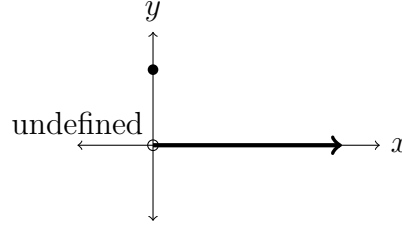


Figure 2: $f(0) = 1, f(x) = 0 \forall x > 0$

$x = 0$. Similar to part 1 we may notice that the left limit is undefined, and in this case doesn't exist in any space. \therefore there is no subdifferential.

Problem 4

Consider the subgradient method $x^+ = x - \alpha g$ where $g \in \nabla f(x)$. Show that if $\alpha < \frac{2(f(x)-f(x^*))}{\|g\|_2^2}$, then $\|x^+ - x^*\|_2 < \|x - x^*\|_2$. (Every iteration moves closer to the optimal x^*)

$$\begin{aligned}
 \|x^+ - x^*\|_2^2 &= \|x - \alpha g - x^*\|_2^2 \\
 &= \|x - x^*\|_2^2 - 2\alpha g(x - x^*) + \alpha^2 \|g\|_2^2 \\
 &< \|x - x^*\|_2^2 - 2\alpha(f(x) - f(x^*)) + 2 \left(\frac{(2(f(x) - f(x^*)))^2}{(\|g\|_2^2)^2} \right) \|g\|_2^2 \\
 &< \|x - x^*\|_2^2 - 2\alpha(f(x) - f(x^*)) + 4(f(x) - f(x^*))\alpha \\
 &< \|x - x^*\|_2^2 - 6\alpha(f(x) - f(x^*))
 \end{aligned}$$

From here we can see that because of the extra term and the upper bound provided to α we can conclude that $\|x^+ - x^*\| < \|x - x^*\|$

Problem 5

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 non-negative constants

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

We will rewrite this as

$$\begin{aligned}
& \min \sum_{s \in \mathcal{S}, u \in \mathcal{U}} a_{s,u} x_{s,u} + \sum_{s \in \mathcal{S}, u \in \mathcal{U}} b_{s,u} ((x_{s,u} + 1) \log(x_{s,u} + 1) - x_{s,u}) \\
& s.t. \quad c_u - \sum_{s \in \mathcal{S}} x_{s,u} \leq 0, \forall u \in \mathcal{U} \\
& \quad d_s - \sum_{u \in \mathcal{U}} x_{s,u} \leq 0, \forall s \in \mathcal{S} \\
& \quad -x_{s,u} \leq 0, \forall s \in \mathcal{S} \\
& \quad -x_{s,u} \leq 0, \forall u \in \mathcal{U}
\end{aligned}$$

From here we can see that we clearly have two separable problems, one over the domain for \mathcal{S} and another over \mathcal{U} .

$$\begin{aligned}
& \min \sum_{s \in \mathcal{S}} a_{s,u} x_{s,u} + \sum_{s \in \mathcal{S}} b_{s,u} ((x_{s,u} + 1) \log(x_{s,u} + 1) - x_{s,u}) \\
& \quad + \lambda_2 \left(d_s - \sum_{u \in \mathcal{U}} x_{s,u} - x_{s,u} \right) \\
& s.t. \quad c_u - \sum_{u \in \mathcal{U}} x_{s,u} \leq 0, \forall u \in \mathcal{U} \\
& \quad -x_{s,u} \leq 0, \forall u \in \mathcal{U}
\end{aligned}$$

and

$$\begin{aligned}
& \min \sum_{u \in \mathcal{U}} a_{s,u} x_{s,u} + \sum_{u \in \mathcal{U}} b_{s,u} ((x_{s,u} + 1) \log(x_{s,u} + 1) - x_{s,u}) \\
& \quad + \lambda_1 \left(c_u - \sum_{s \in \mathcal{S}} x_{s,u} - x_{s,u} \right) \\
& s.t. \quad d_s - \sum_{u \in \mathcal{U}} x_{s,u} \leq 0, \forall s \in \mathcal{S} \\
& \quad -x_{s,u} \leq 0, \forall s \in \mathcal{S}
\end{aligned}$$

Where $\lambda_1, \lambda_2 \geq 0$

\therefore Lagrangian Relaxation, we are able to solve these equations faster and they are separable over the two different domains and \therefore parallelizeable.

To find an optimal value we will need to use an iterative method, such as gradient decent, steepest decent, or Newton's method. To create an effective algorithm we would need to pass back data per iteration. Each domain would be calculated separately, passed back, and then be added together. Then the next step can take place in the same manner.