Solutions to hw2 homework on Convex Optimization https://web.stanford.edu/class/ee364b/homework.html

Andrei Keino

January 6, 2021

2.1 (8 points, 1 point per question)

Let f be a convex function with domain in R^n. We fix $x \in$ **int dom f** and $d \in R^n$. Recall the definition of the directional derivative of f at x along the direction d

$$f'(x, d) = \lim_{t \to 0} \frac{f(x + td) - f(x)}{t}$$

In this question we aim to show that $f'(x, d)$ exists and is finite, and that we have the following relationship between $\partial f(x)$ and $f'(x, d)$,

$$f'(x, d) = \sup_{g \in \partial f(x)} g^T d$$

(a) Show that the ratio $\frac{f(x+td)-f(x)}{t}$ is a nondecrasing function of $t > 0$. Deduce that $f'(x, d)$ exists and is either finite or equal to $-\infty$.
We know from the lectures that, since $x \in$ **int dom f**, the subdifferential set ∂f is non - empty, convex and compact.

Solution:

Proof of non - decreasing. Definition of subgradient is

$$f(z) \geq f(x) + g^T(z - x)$$

let $z = x + td$; then

$$f(x + td) \geq f(x) + g^T(x + td - x)$$

or

$$f(x + td) - f(x) \geq tg^T d$$

dividing both part of the inequality by t (as $t > 0$, we can do it) gives

$$\frac{f(x+td) - f(x)}{t} \geq g^T d$$

as the right - hand side of the equation is not depends of t, differentiating by t gives

$$\partial \frac{\frac{f(x+td)-f(x)}{t}}{\partial t} \geq 0$$

As the $\frac{\partial f'(x,d}{\partial t} \geq 0$, it means the function $f'(x,d)$ is nondecreasing by variable t.

Proof of possible equality to $-\infty$.
The definition of convexity:

$$f(\theta x + (1-\theta)y)) \leq \theta f(x) + (1-\theta) f(y)$$

where $0 < \theta < 1$.
let $t = 1 - \theta$, $0 < t < 1$. then

$$f((1-t)x + ty)) \leq (1-t) f(x) + t f(y)$$

or

$$f(x + t(y-x))) \leq f(x) + t(f(y) - f(x))$$

as we can choose y any of the point in domain f, we can set $d = y - x$. Then

$$f(x+td) \leq f(x) + t(f(y) - f(x))$$

or

$$f(x+td) - f(x) \leq t(f(y) - f(x))$$

or

$$\frac{f(x+td) - f(x)}{t} \leq f(y) - f(x)$$

As $f(x)$ can be equal to ∞ on the domain of f, so $f'(x,d) = \frac{f(x+td)-f(x)}{t}$ can be less or equal than (for the infinity with sign minus it means strictly equal) $-\infty$ on the domain of f. This means that $f'(x,d)$ can be equal to $-\infty$ on domain of f.

(b) Let $g \in \partial f(x)$. Show that $f'(x,d) \geq g^T d$. Deduce that $f'(x,d)$ is finite and $f'(x,d) \geq \sup_{g \in \partial f(x)} g^T d$.

Solution:
We already shown that

$$f'(x,d) \geq g^T d$$

in part (a). We also shown in part (a) that

$$\frac{f(x+td) - f(x)}{t} \leq f(y) - f(x)$$

Second upper inequality means that $f'(x,d)$ is bounded from upper side (i.e it can't be equal to ∞), it means its value is finite.
As the first of upper inequalities is correct $\forall$ subgradients in domain of f, it means, that it is correct for the supremum of these subgradients in domain f. It means that

$$f'(x,d) \geq \sup_{g \in \partial f(x)} g^T d.$$

In the remaining part of this question, we will establish the converse inequality $f'(x,d) \leq \sup_{g \in \partial f(x)} g^T d$, by showing the existence of a subgradient $g^* \in \partial f(x)$, such that $f'(x,d) \leq g^{*T} d$. We introduce two following sets

$$C_1 = \{(z,t) \mid z \in \mathbf{dom} f,\ f(z) < t\}$$
$$C_2 = \{(y,v) \mid y = x + \alpha d,\ v = f(x) + \alpha f'(x,d),\ \alpha \geq 0\}$$

(c) Prove that C_2 and C_2 are nonempty, convex and disjoint.

Solution:

C_1 is the epigraph of the convex function, therefore it is nonempty and convex.

C_2 is the nonempty set, because it have at least one point, which corresponds to $\alpha = 0$, $y = x$, $v = f(x)$. It is also a convex set, because $C_2^1 = \{y \mid y = x + \alpha d\}$ is a convex set as it is translated domain of f which is a convex set and $C_2^2 = \{v \mid v = f(x) + \alpha f'(x,d),\ \alpha \geq 0\}$ is either a straight line or a beam or a segment.

Proof of disjointedness:
We should show that there is exists a nonzero vector $(a, \beta) \in R^n \times R$ such as

$$a^T(x + \alpha d) + \beta(f(x) + \alpha f'(x,d)) \leq a^T z + \beta w$$

for all α $geq0$, $z \in \mathbf{dom} f$, and $f(z) < w$.
Solution: As we shown earlier,

$$f'(x,d) \leq f(y) - f(x)$$

where $x, y \in \mathbf{dom} f$. As x, y can be any points in domain f, it follows that

$$f'(x,d) \leq \min_{z \in \mathbf{dom} f}(f(z)) - \max_{z \in \mathbf{dom} f}(f(z))$$

Lets just derive equation for β.

$$\beta(f(x) + \alpha f'(x,d) - w) \leq a^T(z - x - \alpha d)$$

or

$$\beta \leq \frac{a^T(z - x - \alpha d)}{f(x) + \alpha f'(x,d) - w}$$

I don't know how to solve items (e) - (g)

(h) Let $A \in R^{m \times n}$, $b \in R^m$, $\lambda > 0$, and fix a direction $d \in R^n$. Consider the function $\frac{1}{2}||Ax - b||_2^2 + \lambda||x||_1$. Compute $f'(0, d)$. Remark: you can either compute it directly by using the definition of the directional derivative, or, use the variational formula $f'(0, d) = \sup_{g \in \partial f(0)} g^T d$.

Solution:

$\nabla||x||_1 = sign(x)$

$\nabla||Ax - b||_2^2 = \nabla((Ax - b)^T(Ax - b)) = 2(Ax - b)^T A$

see`https://math.stackexchange.com/questions/606646/matrix-derivative-ax-btax-b`
and `http://www.math.uwaterloo.ca/~hwolkowi//matrixcookbook.pdf`

So,
$\nabla(\frac{1}{2}||Ax - b||_2^2 + \lambda||x||_1) = 2(Ax - b)^T A + \lambda sign(x)$

Then

$$f'(0, d) = d^T((Ax - b)^T A + \lambda[-1, 1]_n)$$

where $[-1, 1]_n$ is a vector in R^n with component values in range
$-1 \leq x_i \leq 1,\ i \in 1, \ldots, n$.

2.2 (4 Points)

In this question, we will show that a subgradient of the function
$h(x) = \min_{z \in C} ||x - z||_2$ is

$$g = \frac{x - z^*}{||x - z^*||_2}$$

where C is a compact set in R^n, x is a given point in R^n, which does not belong to C, and
$z^* = P_C(x) := \arg\min_{z \in C} ||x - z||_2$ denotes the Euclidean projection of x onto C (which exists and is unique).

(a) (0.5 point) Use the fact that $||x - z||_2 = \max_{u:||u||_2 \leq 1} u^T(x - z)$ to transform the minimization problem] $h(x) = \min_{z \in C} ||x - z||_2$ into the following saddle point problem

$$\min_{z \in C} \max_{u:||u||_2 \leq 1} u^T(x - z)$$

Solution:

We get it by substituting expression $\max_{u:||u||_2\leq 1} u^T(x-z)$ instead the expression $||x-z||_2$.

(b) (2 points) Now, we will use (a simple version of) the Sion's minimax theorem, which can be stated as follows.
Let $X \subseteq R^n$, $Y \subseteq R^n$, be compact and convex sets. Let f be a real valued function on $X \times Y$ such that
- $f(x,\cdot)$ is continuous and concave on Y, $\forall x \in X$.
- $f(\cdot,y)$ is continuous and convex on X, $\forall y \in Y$.

Then, we have
$\min_{x\in X} \max_{y\in Y} f(x,y) = \max_{y\in Y} \min_{x\in X} f(x,y)$

Further, there exists a (saddle) point $(x^*, y^*) \in X \times Y$ such that

$$f(x*, y*) = \min_{x\in X} f(x, y^*) = \max_{y\in Y} f(x*, y) = \min_{x\in X} \max_{y\in Y} f(x,y) = \max_{y\in Y} \min_{x\in X} f(x,y)$$

Apply Sion's minimax theorem to conclude that

$$\min_{z\in C} \max_{u:||u||_2\leq 1} u^T(x-z) = \max_{u:||u||_2\leq 1} \min_{z\in C} u^T(x-z)$$

Define $u^* = \frac{x-z^*}{||x-z^*||_2}$. Show that (z^*, u^*) is a saddle point of the above minimax problem.

Solution:
C is compact and convex. u defined on the closed sphere S of unity radius, therefore its domain is compact and convex also. The function $f(u,z) = u^T(x-z)$ is linear in sense of both $f(z,\cdot)$ on C and $f(\cdot,u)$ on S. It means that it is concave and convex in both cases. So, applying Sion's minimax theorem we have:

$$\min_{z\in C} \max_{u:||u||_2\leq 1} u^T(x-z) = \max_{u:||u||_2\leq 1} \min_{z\in C} u^T(x-z)$$

where $f(z,u) = u^T(x-z)$ and there exist a saddle point (z^*, u^*) in $C \times S$ such that

$$f(z^*, u^*) = \min_{z\in C} f(z, u^*) = \max_{u\in S} f(z^*, y) = \min_{z\in C} \max_{u\in S} f(z,u) = \max_{z\in C} \min_{u\in S} f(z,u)$$

define $u^* = \frac{x-z^*}{||x-z^*||_2}$. It is evident, that u^* is the solution of the problem $\max_{u:||u||_2\leq 1} u^T(x-z)$. Also, this is evident, that the point z^* is the solution of the problem $z^* = P_C(x) := \arg\min_{z\in C} u^T(x-z)_2$. Then, by Sion's theorem the point $(z^*, u^*) : f(z^*, u^*) = \min_{z\ inC} f(z, u^*)$, where $f(z, u*) = u^{*T}(x-z)$ is a saddle point of the problem

$$\min_{z\in C} \max_{u\in S} u^T(x-z)_2.$$