Algorithms for Convex Optimization

Assignment 4

Please note that all statements are based on the vectorial l_2-norm without special instructions.

1. Let Ω be a nonempty, convex set and let α, $\beta \geq 0$. Does the equation $\alpha\Omega + \beta\Omega = (\alpha+\beta)\Omega$ hold?

2. Show that if $f : \mathbb{R} \to [0, \infty)$ is a convex function, then its $q-$power f^q is also convex for any $q > 1$.

3. We say that $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ is positively homogeneous if $f(\alpha x) = \alpha f(x)$ for all $\alpha > 0$ and that f is subadditive if $f(x+y) \leq f(x) + f(y)$ for all $x, y \in \mathbb{R}^n$. Show that a proper positively homogeneous function is subadditive if it is convex.

4. Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is twice continuously differentiable convex, then f is $L-$smooth w.r.t. l_2-norm if and only if $\lambda_{\max}(\bigtriangledown^2 f(x)) \leq L$ for any $x \in \mathbb{R}$.

5. Find the proximal mapping of $g(x) = \delta_{[0,\eta]\cap\mathbb{R}}(x)$, where $\eta > 0$.

6. Calculate the subdifferentials.

   (1). $f(x) = e^{|x|}$, $x \in \mathbb{R}$.

   (2). $f(x_1, x_2) = |x_1| + |x_2|$, $(x_1, x_2) \in \mathbb{R}^2$.

7. Find the conjugate for each of the following functions.

   $(1) f(x) = ax^2 + bx + c, \ where\ a > 0.$

   $(2) f(x) = \delta_{\mathbb{B}}(x), \ where\ \mathbb{B} = [-1, 1].$

   $$(3) f(x) = \ln\left(\sum_{i=1}^{n} e^{x_i}\right) \text{ where } x := (x_1, \cdots, x_n)^T \in \mathbb{R}^n.$$

8. Let $f : \mathbb{R}^n \to (-\infty, +\infty]$ be a proper function and let $\alpha > 0$. Define $g(x) = \alpha f(\frac{x}{\alpha})$.

   (a) Then $g^*(y) = \alpha f^*(y)$, $y \in \mathbb{R}^n$

   (b) Then $\text{prox}_g(x) = \alpha \text{prox}_{\frac{f}{\alpha}}(\frac{x}{\alpha})$

9. Prove the Fenchel's dual of

$$\min_{x \in \mathbb{R}^n} \; f(x) + g(x)$$

has the form:

$$\max_{y \in \mathbb{R}^n} -f^*(y) - g^*(y).$$

10. Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be proper convex and let $x \in \operatorname{int}(\operatorname{dom}(f))$. Then the function $d \mapsto f'(x; d)$ is convex.

11. Let $f : \mathbb{R} \to \mathbb{R}$ be convex. Prove that f is nondecreasing if and only if $\partial f(x) \subset [0, \infty)$ for all $x \in \mathbb{R}$.

12. Give an example of two lower semicontinuous real-valued functions whose product is not lower semicontinuous.

13. Let $\Omega_i := [a_i, b_i] \subset \mathbb{R}$, $i = 1, \cdots, m$ be m disjoint intervals. Find the optimality condition for

$$\min f(x) = \sum_{i=1}^{m} d(x; \Omega_i), \quad x \in \mathbb{R}$$

14. Show that

$$\max\{a_1, a_2, \cdots, a_k\} = \max_{\lambda \in \Delta_k} \sum_{i=1}^{k} \lambda_i a_i,$$

where $\Delta_k = \{\lambda \in \mathbb{R}^k \mid \lambda \geq 0,\, e^T\lambda = 1\}$ with $e = (1, \cdots, 1)^T$.

15. Suppose C is a closed, nonempty, convex set, show that

$$f(x) = \frac{1}{2}\|x\|^2 - \frac{1}{2}d_C^2(x)$$

is $1-$smooth.

16. Review the theorem on $L-$smoothness and boundedness of the Hessian.

17. Justify the $\mathbb{O}(\frac{1}{k})$ rate of convergence of cyclic block proximal gradient method and randomized block proximal gradient method (The Convex Case).

18. Justify the conjugate subgradient theorem.

19. Let $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ be a proper convex function. Let $\Omega \in \bar{\mathbb{R}}$ be a convex set. Is it true in general that the inverse image $f^{-1}(\Omega)$ is a convex subset of $\mathbb{R}^n$.

20. Use the definition. Let Ω be a nonempty, convex and bounded subset of $\mathbb{R}^n$. Define the function

$$\mu_\Omega(x) := \sup\{\|x - \omega\| \mid \omega \in \Omega\}, \; x \in \mathbb{R}^n.$$

Show μ_Ω is convex.

21. Let $f : \mathbb{R}^n \to$ be proper closed convex, and let $\mu > 0$, the Moreau envelope of f is

$$M_f^\mu(x) = \min_{u \in \mathbb{R}^n} \left\{ f(u) + \frac{1}{2\mu} \|x - u\|^2 \right\}.$$

If $\arg\min_{x \in \mathbb{R}^n} f(x) \neq \varnothing$, show that

(a). $M_f^\mu < \infty$.

(b). $M_f^\mu \leq f$.

(c). M_f^μ is convex.

(d). $\arg\min_{x \in \mathbb{R}^n} f(x) = \arg\min_{x \in \mathbb{R}^n} M_f^\mu(x)$.

(e). $\min_{x \in \mathbb{R}^n} M_f^\mu(x) = \min_{x \in \mathbb{R}^n} f(x)$.

22. Consider convex optimization in the form:

$$\min_{x \in \mathbb{R}^n} F(x) := f(x) + g(x), \tag{1}$$

where $f, g : \mathbb{R}^n \to \mathbb{R}$ are two proper closed convex functions. Suppose that f is L_f-smooth and $X^* = \arg\min_{x \in \mathbb{R}^n} F(x) \neq \varnothing$, and the optimal value of problem (1) is denoted by F_{opt}.

Let $\{x^k\}_{k \geq 0}$ be the sequence generated by the proximal gradient method for problem (1) with a constant stepsize $t_k = \frac{1}{L_f}$.

(a). Write down the scheme of the proximal gradient method for problem (1) with the constant stepsize $t_k = \frac{1}{L_f}$.

(b). Show that for any $x^* \in X^*$ and $k \geq 0$,

$$F(x^*) - F\left(x^{k+1}\right) \geq \frac{L_f}{2} \left\|x^* - x^{k+1}\right\|^2 - \frac{L_f}{2} \left\|x^* - x^k\right\|^2.$$

(c). Show that for any $x^* \in X^*$ and $k \geq 0$,

$$F(x^k) - F_{\text{opt}} \leq \frac{L_f \left\|x^0 - x^*\right\|^2}{2k}.$$

(d). Show that for any $x^* \in X^*$ and $k \geq 0$,

$$\left\|x^{k+1} - x^*\right\| \leq \left\|x^k - x^*\right\|.$$

(e). Show that the sequence $\{x^k\}_{k \geq 0}$ converges to an optimal solution of problem (1).

23. Define the perspective function $P : \mathbb{R}^{n+1} \to \mathbb{R}^n$ with domain $\operatorname{dom}(P) = \mathbb{R}^n \times \mathbb{R}_{++}$, as $P(z, t) = \frac{z}{t}$. Show that the inverse image of any convex set $C \subset \mathbb{R}^n$ under the perspective function

$$P^{-1}(C) = \left\{ (x, t) \in \mathbb{R}^{n+1} \,\middle|\, \frac{x}{t} \in C, t > 0 \right\}$$

is convex.

24. Express the conjugate of the perspective of a convex function f in terms of f^*, where the perspective of f is the function $g : \mathbb{R}^{n+1} \to \mathbb{R}$ defined by

$$g(x,t) := tf\left(\frac{x}{t}\right),$$

with domain

$$\mathrm{dom}(g) = \left\{(x,t) \,\middle|\, \frac{x}{t} \in \mathrm{dom}(f), t > 0\right\}.$$

25. Consider the model

$$\min_{x \in \mathbb{R}^n} f(x) + g(x)$$

where f, g satisfy the standing assumption 1 on the slides. If $g = \delta_C$ associated with a convex set C. The PGM (proximal gradient method) for the model reduces to the classical projected gradient method. Consider the convergence analysis for the projected gradient method.