Homework 2

Li Ziniu

November 14, 2018

1 Proximity Rules

Proximity operator:

$$prox_f(x) = \arg\min_z \frac{1}{2}||z-x||^2 + f(z), \quad x, z \in R^n$$

Scaling and translation: if $f(x) = g(ax+b)$ with $a \neq 0$, $x, b \in R^n$ then

$$prox_f(x) = \frac{1}{a}\Big(prox_{a^2 g}(ax+b) - b\Big)$$

Proof:

$$\begin{aligned}
prox_f(x) &= \underset{z}{\text{argmin}} \frac{1}{2}||z-x||^2 + g(az+b) \\
&= \underset{z}{\text{argmin}} \frac{1}{2}||z||^2 + \langle z, x\rangle + g(az+b) \\
&= \underset{t=az+b}{\text{argmin}} \frac{1}{2}||\frac{t-b}{a}||^2 + \langle \frac{t-b}{a}, x\rangle + g(t) \\
&= \underset{t}{\text{argmin}} \frac{1}{a^2}\Big(\frac{1}{2}||t||^2 + a\langle t, x+b\rangle + a^2 g(t)\Big) \\
&= \underset{t}{\text{argmin}} \frac{1}{2}||t-(ax+b)||^2 + a^2 g(t) \\
&= prox_{a^2 g}(ax+b)_{|t=az+b} \\
&= \frac{1}{a}\Big(prox_{a^2 g}(ax+b) - b\Big)
\end{aligned}$$

2 Proximity Operator for $L_{2,1}$

The proximity operator for $L_{2,1}$ norm, one of the way to solve block-compress sensing is defined as

$$prox(x) = \min_z \frac{1}{2}||z-x||_2^2 + \lambda||z||_{2,1}$$

$$z = (z_1, z_2, \cdots, z_n)$$
$$z_i = \sum_k z_{ik}^2$$

First, decomposing z and x into individual components z_i, x_i, respectively,

$$L = \frac{1}{2}(\sum_k z_{ik}^2 - x_i)^2 + \lambda \sum_k z_{ik}^2$$

According to the soft-thresholding based on L_1 norm result and observing that $z_i \geq 0$,

$$z_i = \begin{cases} x_i - \lambda & if\ x_i \geq \lambda \\ 0 & if\ x_i < \lambda \end{cases}$$

It seems that we can not get to know the individual components of z_{ik} but only know the l_2 norm of z_i.

3 Convex function and its quadratic upper bound

3.1 Convex function

A function f is said to be convex when it satisfies:

$$\forall x, y \in C, \forall \alpha \in [0,1], f\Big(\alpha x + (1-\alpha)y)\Big) \leq \alpha f(x) + (1-\alpha) f(y)$$

where C is a convex set.

Proposition 1 *Suppose f is first-order differentiable. f is convex $\Leftrightarrow \forall x, y \in C, f(y) \geq f(x) + \nabla f(x)^T (y-x)$.* [1]

Proposition 2 *Suppose f is first-order differentiable. f is convex if and only if $dom f$ is convex and $\nabla f(x)$ is a monotone operator:* [2]

$$\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq 0$$

3.2 Quadratic upper bound

Proposition 3 *Suppose f is first-order differentiable. The following statements are equivalent.*[3]

(a) *Lipschitz continuity* of $\nabla f(x)$: there exists an $L > 0$ such that

$$\forall x, y \in dom f,\ ||\nabla f(x) - \nabla f(y)|| \leq L||x-y||$$

(b) $g(x) = \frac{L}{2}||x||^2 - f(x)$ is convex.

[1]See Appendix 4.1 for proof
[2]See Appendix 4.2 for proof
[3]This part refers to https://www.math.cuhk.edu.hk/course˙builder/1617/math6211a/cvxop.pdf

(c) *Quadratic upper bound*.

$$f(y) \leq f(x) + \nabla\langle f(x), y-x\rangle + \frac{L}{2}||y-x||^2$$

. *Proof*:

$(a \Rightarrow b)$

$$\begin{aligned}\langle \nabla g(x) - \nabla g(y), x-y\rangle &= \Big\langle L(x-y) - (\nabla f(x) - \nabla f(y)), x-y \Big\rangle \\ &= L||x-y||^2 - \langle \nabla f(x) - \nabla f(y), x-y\rangle \\ &\geq L||x-y||^2 - ||\nabla f(x) - \nabla f(y)|| \cdot ||x-y|| \\ &\geq 0\end{aligned}$$

Thus, we get the monotonicity of $\nabla g(x)$, which implies g(x) is convex.

$(b \Leftrightarrow c)$

$$\begin{aligned}g(x)\,\text{is convex} &\Leftrightarrow g(y) \geq g(x) + \nabla g(x)^T(y-x) \\ &\Leftrightarrow \frac{L}{2}||y||^2 - f(y) \geq \frac{L}{2}||x||^2 - f(x) + \langle Lx - \nabla f(x), y-x\rangle \\ &\Leftrightarrow f(y) \leq f(x) + \nabla\langle f(x), y-x\rangle + \frac{L}{2}||y-x||^2\end{aligned}$$

$(c \Rightarrow a)$

$$f(y) \leq f(x) + \nabla\langle f(x), y-x\rangle + \frac{L}{2}||y-x||^2$$

$$f(x) \leq f(y) + \nabla\langle f(y), x-y\rangle + \frac{L}{2}||x-y||^2$$

Add these two inequalities,

$$\langle \nabla f(x) - f(y), x-y\rangle \leq L||x-y||^2$$

By Cauchy-inequality,

$$\langle \nabla f(x) - f(y), x-y\rangle = ||\nabla f(x) - \nabla f(y)|| \cdot ||x-y|| \leq L||x-y||^2$$

$$||\nabla f(x) - \nabla f(y)|| \leq L||x-y||$$

Thus, we get the *Lipschitz continuity* of $\nabla f(x)$.

4 Appendix

4.1 Proposition 1

Suppose f is first-order differentiable. f is convex $\Leftrightarrow \forall x, y \in C, f(y) \geq f(x) + \nabla f(x)^T(y-x)$.
Proof:

1) $\forall x, y \in C, f(y) \geq f(x) + \nabla f(x)^T (y - x) \Rightarrow f$ is convex.

$\forall x, y \in C, \forall \alpha \in [0,1]$, let $z = \alpha x + (1 - \alpha y) \in C$, according to the inequality:

$$f(x) \geq f(z) + \nabla f(z)^T (x - z) \tag{1}$$

$$f(y) \geq f(z) + \nabla f(y)^T (y - z) \tag{2}$$

(1) $\times$ α + (2) $\times$ $(1 - \alpha)$, we get that

$$\alpha f(x) + (1 - \alpha) f(y) \geq f(z) + \nabla f(z)^T (\alpha x + (1 - \alpha) y - z) = f(z)$$

Thus, f is convex.

2) f is convex $\Rightarrow \forall x, y \in C, f(y) \geq f(x) + \nabla f(x)^T (y - x)$.

First, we construct such a function $g : (0, 1] \mapsto R$, and $x, z \in C$, as well as $x \neq z$,

$$g(\alpha) = \frac{f\Big(x + \alpha(z - x)\Big) - f(x)}{\alpha}$$

$$\lim_{a \to 0^+} g(\alpha) = \nabla f(x)^T (z - x), g(1) = f(z) - f(x)$$

To prove 2), we just need to prove $\lim_{a \to 0^+} g(\alpha) \leq g(1)$, which implies $g(\alpha)$ is monotonic between (0, 1]. Next, we are going to prove $g(\alpha)$ is monotonic.

$$\forall\, 0 < \alpha_1 < \alpha_2 < 1, \quad \text{let } \bar{\alpha} = \frac{\alpha_1}{\alpha_2} < 1, \bar{z} = x + \alpha_2 (z - x) \in C$$

It is easy to find that

$$f\Big(x + \bar{\alpha}(\bar{z} - x)\Big) \leq \bar{\alpha} f(\bar{z}) + (1 - \bar{\alpha}) f(x)$$

$$\frac{f\Big(x + \bar{\alpha}(\bar{z} - x)\Big) - f(x)}{\bar{\alpha}} \leq f(\bar{z}) - f(x)$$

Using the definition of $\bar{a}$ and $\bar{z}$, by algebraic calculation,

$$\frac{f\Big(x + \alpha_1 (z - x)\Big) - f(x)}{\alpha_1} \leq \frac{f\Big(x + \alpha_2 (z - x)\Big) - f(x)}{\alpha_2}$$

$$g(\alpha_1) \leq g(\alpha_2)$$

Thus, $g(\alpha)$ is monotonic. By above equivalent analysis, we complete the proof.

4.2 Proposition 2

Suppose f is first-order differentiable. f is convex if and only if $dom f$ is convex and $\nabla f(x)$ is a monotone operator:

$$\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq 0$$

Proof:

1) ($\Rightarrow$)

From proposition 1, we have that

$$f(x) \geq f(y) + \nabla f(y)^T (x - y) \tag{3}$$

$$f(y) \geq f(x) + \nabla f(x)^T (y - x) \tag{4}$$

By algebraic calculation of (3) + (4),

$$f(x) + f(y) \geq f(y) + f(x) + (\nabla f(y)^T - \nabla f(x)^T)(x - y)$$

$$\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq 0$$

Thus, we get the monotonicity of $\nabla f(x)$.

2) ($\Leftarrow$) Let $g(\alpha) = f(x+\alpha(y-x))$. Then $g'(\alpha) = \Big\langle \nabla f \Big(x+\alpha(y-x) \Big), y-x \Big\rangle \geq 0$, which means $g(\alpha)$ is monotonic. Hence,

$$\begin{aligned} f(y) = g(1) = g(0) &+ \int_0^1 g'(\alpha) d\alpha \\ &\geq g(0) + \int_0^1 g'(0) d\alpha \\ &\geq f(x) + \langle \nabla f(x), y - x \rangle \end{aligned}$$

Thus, we get the convexity of $f(x)$.