HW6

2020年12月10日 5:55

MATH 3332 Data Analytic Tools
Homework 6

Due date: 14 December, 6pm, Monday

1. Find the sub-differentials.

   (a) $g_1(x) = \begin{cases} -x & \text{if } x \le 0, \\ x^2 & \text{if } x > 0, \end{cases}$ where $x \in \mathbb{R}$.

   (b) $g_2(\boldsymbol{x}) = \sqrt{x_1^2 + x_2^2} + \sqrt{x_3^2 + x_4^2}$, where $\boldsymbol{x} \in \mathbb{R}^4$.

   (c) $g_3(\boldsymbol{x}) = \|\boldsymbol{x}\|_\infty$, where $\boldsymbol{x} \in \mathbb{R}^2$.

   (d) $g_4(\boldsymbol{x}) = (|x_1| + |x_2|)^2$, where $\boldsymbol{x} \in \mathbb{R}^2$.

2. We consider $\min_{\boldsymbol{x}} g(\boldsymbol{x})$, where $g : \mathbb{R}^n \to \mathbb{R}$ is a convex function. In the backward sub-gradient algorithm (a.k.a. proximal algorithm), we used the iteration: given $\boldsymbol{x}^{(k)}$, generate

$$\boldsymbol{x}^{(k+1)} = \arg\min_{\boldsymbol{x} \in \mathbb{R}^n} \frac{1}{2}\|\boldsymbol{x} - \boldsymbol{x}^{(k)}\|_2^2 + \alpha_k g(\boldsymbol{x}).$$

   Prove that $\boldsymbol{x}^{(k+1)}$ is uniquely defined, i.e., prove that the solution of the minimization

$$\min_{\boldsymbol{x} \in \mathbb{R}^n} \frac{1}{2}\|\boldsymbol{x} - \boldsymbol{x}^{(k)}\|_2^2 + \alpha_k g(\boldsymbol{x})$$

   exists and is unique for any $\boldsymbol{x}^{(k)} \in \mathbb{R}^n$ and $\alpha_k > 0$.

3. Consider the following minimization

$$\min_{\boldsymbol{x}} \frac{1}{2}\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2^2 + \lambda \sum_{i=1}^{n} r(x_i)$$

   where $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, $\boldsymbol{b} \in \mathbb{R}^m$, $\lambda > 0$ are given, and $r : \mathbb{R} \to \mathbb{R}$ is defined by

$$r(t) = \begin{cases} 0, & \text{if } t \le 0, \\ t, & \text{if } t \ge 0. \end{cases}$$

   Find a forward-backward splitting algorithm with explicit formulas for solving this minimization problem. (*Hint: Use a forward step for the smooth convex term* $\frac{1}{2}\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2^2$ *and a backward step for the non-smooth convex term* $\lambda \sum_{i=1}^{n} r(x_i)$.)

1. Find the sub-differentials.

(a) $g_1(x) = \begin{cases} -x & \text{if } x \le 0, \\ x^2 & \text{if } x > 0, \end{cases}$ where $x \in \mathbb{R}$.

(b) $g_2(\boldsymbol{x}) = \sqrt{x_1^2 + x_2^2} + \sqrt{x_3^2 + x_4^2}$, where $\boldsymbol{x} \in \mathbb{R}^4$.

(c) $g_3(\boldsymbol{x}) = \|\boldsymbol{x}\|_\infty$, where $\boldsymbol{x} \in \mathbb{R}^2$.

(d) $g_4(\boldsymbol{x}) = (|x_1| + |x_2|)^2$, where $\boldsymbol{x} \in \mathbb{R}^2$.

1a) $\partial g_1(x) = \begin{cases} \partial(-x) & \text{if } x < 0 \\ \partial g_1(x) & \text{if } x = 0 \\ \partial(x^2) & \text{if } x > 0 \end{cases}$

if $x = 0$, $D_- g(x) = \lim_{y \to 0_-} \frac{g(y) - g(x)}{y - x}$

$= \lim_{y \to 0_-} \frac{-y}{y}$

$= -1$

$D_+ g(x) = \lim_{y \to 0_+} \frac{g(y) - g(x)}{y - x}$

$= \lim_{y \to 0^+} \frac{y^2}{y}$

$= 0$

$\partial g_1(x) = \begin{cases} \{-1\} & \text{if } x < 0 \\ [-1, 0] & \text{if } x = 0 \\ \{2x\} & \text{if } x > 0 \end{cases}$

b). $g_2(x) = \sqrt{x_1^2 + x_2^2} + \sqrt{x_3^2 + x_4^2}$ for $x \in \mathbb{R}^4$.

at $A = \{x \in \mathbb{R}^4 : x_1 = x_2 = 0\}$

$B = \{x \in \mathbb{R}^4 : x_3 = x_4 = 0\}$

gradient of $g_2(x)$, we need to

To compute subgradient of $g_2(x)$, we need to divide into 4 cases:

Case 1: if $x \notin A$ and $x \notin B$, then g_2 is differentiable because all the partial derivatives exists and continuous near that point.

So, $\nabla g_2(x) = \left(\frac{x_1}{\sqrt{x_1^2+x_2^2}}, \frac{x_2}{\sqrt{x_1^2+x_2^2}}, \frac{x_3}{\sqrt{x_3^2+x_4^2}}, \frac{x_4}{\sqrt{x_3^2+x_4^2}} \right)^T$

Case 2: if $x \in A$ but $x \notin B$, then we claim

$\partial g_2(x) = \{ (u, v, \frac{x_3}{\sqrt{x_3^2+x_4^2}}, \frac{x_4}{\sqrt{x_3^2+x_4^2}})^T \in \mathbb{R}^4 : u^2+v^2 \leq 1 \}$

Proving the claim,

$g_2(x) = h(x_1, x_2) + h(x_3, x_4)$

where $h(s,t) = \sqrt{s^2+t^2}$

Since $x \notin B$, $h(x_3, x_4)$ is differentiable

and $\nabla h_2(x_3, x_4) = \left(\frac{x_3}{\sqrt{x_3^2+x_4^2}}, \frac{x_4}{\sqrt{x_3^2+x_4^2}} \right)^T$

For $h(x_1, x_2) = \sqrt{x_1^2+x_2^2} = \|v\|_2$ where $v = \binom{x_1}{x_2}$

according to lecture notes, $\partial \|x\|_2$ at $x = (0,0) \in \mathbb{R}^2$,

we have:

$\partial h(0,0) = \{ u = \mathbb{R}^2 : \|u\|_2 \leq 1 \}$

$\rightarrow$ subgradient of g_2 for any points $x \in A$ is:

$\Rightarrow$ subgradient of g_2 for any point $x \in \ldots$

$$\left\{ \left(u, v, \frac{x_3}{\sqrt{x_3^2+x_4^2}}, \frac{x_4}{\sqrt{x_3^2 + x_4^2}}\right)^T \in \mathbb{R}^4 : u^2+v^2 \le 1 \right\}$$

Case 3: If $x \in B$ but $x \notin A$, similarly we have

$$\partial g_2(x) = \left\{ \left(\frac{x_1}{\sqrt{x_1^2+x_2^2}}, \frac{x_2}{\sqrt{x_1^2+x_2^2}}, u, v\right)^T \in \mathbb{R}^4 : u^2+v^2 \le 1 \right\}$$

Case 4: If $x \in A \cap B$, $x_1 = x_2 = x_3 = x_4 = 0$,

$$\partial g_2(0) = \left\{ (s,t,u,v)^T \in \mathbb{R}^4 : s^2+t^2 \le 1, u^2+v^2 \le 1 \right\}$$

Proving this, we rewrite $g_2(x)$ as:

$$g_2(x) = h(x_1, x_2) + h(x_3, x_4)$$

consider $\partial h(x_1, x_2)$ at $(x_1, x_2) = (0,0)$, we have

$$\partial h(0,0) = \{ u \in \mathbb{R}^2 : \|u\|_2 \le 1 \}$$

consider $\partial h(x_3, x_4)$ at $(x_3, x_4) = (0,0)$, we have

$$\partial h(0,0) = \{ u \in \mathbb{R}^2 : \|u\|_2 \le 1 \}.$$

$$\Rightarrow \partial g_2(0) = \{ (s,t,u,v)^T \in \mathbb{R}^4 : s^2+t^2 \le 1, u^2+v^2 \le 1 \}.$$

So,

$$\left\{ \left(u, v, \frac{x_3}{\sqrt{x_3^2+x_4^2}}, \frac{x_4}{\sqrt{x_3^2+x_4^2}}\right)^T \in \mathbb{R}^4 : u^2+v^2 \le 1 \right\} \quad \text{if } x_1 = x_2 = 0$$

$$\partial g_2(x) = \begin{cases} \{ \cdots \frac{\cdots}{\sqrt{x_3^2+x_4^2}}, \frac{\cdots}{\sqrt{x_3^2+x_4^2}} \cdots \\ \left\{ \left(\frac{x_1}{\sqrt{x_1^2+x_2^2}}, \frac{x_2}{\sqrt{x_1^2+x_2^2}}, u, v \right)^T \in \mathbb{R}^4 : u^2+v^2 \le 1 \right\} & \text{if } x_3 = x_4 = 0 \\ \left\{ (s,t,u,v)^T \in \mathbb{R}^4,\ s^2+t^2 \le 1,\ u^2+v^2 \le 1 \right\} & \text{if } x_1 = x_2 = x_3 = x_4 = 0 \\ \left\{ \left(\frac{x_1}{\sqrt{x_1^2+x_2^2}}, \frac{x_2}{\sqrt{x_1^2+x_2^2}}, \frac{x_3}{\sqrt{x_3^2+x_4^2}}, \frac{x_4}{\sqrt{x_3^2+x_4^2}} \right)^T \in \mathbb{R}^4 \right\} & \text{otherwise} \end{cases}$$

1(c). $g_3(x) = \|x\|_\infty$, $x \in \mathbb{R}^2$,

$g_3(x) = \max(|x_1|, |x_2|)$

Case 1: $x_1 > |x_2|$, then around (x_1, x_2), $\|x\|_\infty = x_1$

$\Rightarrow \nabla g_3(x) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

Case 2: $x_1 < -|x_2|$, then around (x_1, x_2), $\|x\|_\infty = -x_1$

$\Rightarrow \nabla g_3(x) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$

Case 3: $x_2 > |x_1|$, then around (x_1, x_2), $\|x\|_\infty = x_2$

$\Rightarrow \nabla g_3(x) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

Case 4: $x_2 < -|x_1|$, then around (x_1, x_2), $\|x\|_\infty = -x_2$

$\Rightarrow \nabla g_3(x) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$

Case 5: $x_2 = x_1 > 0$, $\partial g_3(x) = S : \left\{ \begin{pmatrix} u \\ v \end{pmatrix} : u, v \geq 0, u+v = 1 \right\}$

Derive the above statement:

Proving the above statement:

take $k \in S$, we have

$$\|x\|_\infty + \langle k, (y-x)\rangle = \max(|x_1|, |x_2|) + k_1(y_1 - x_1) + k_2(y_2 - x_2)$$

$$= x_1 + k_1(y_1 - x_1) + k_2(y_2 - x_2)$$

$$= x_1 + k_1 y_1 - k_1 x_1 + k_2 y_2 - k_2 x_2$$

$(\because k \in S, k_1 + k_2 = 1)$ $$= k_1 x_1 + k_2 x_1 + k_1(y_1 - x_1) + k_2(y_2 - x_2)$$

$$= k_1 y_1 + k_2 y_2$$

$$\leq (k_1 + k_2) \max\{|y_1|, |y_2|\}$$

$$= \|y\|_\infty$$

$\therefore S \subseteq \partial g_3(x)$, where $x_2 = x_1 > 0$.

To show $\partial g_3(x) \subseteq S$, assume $k = \begin{bmatrix} k_1 \\ k_2 \end{bmatrix}$, $k_1 + k_2 > 1$,

then take $y = 2\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, we have

$$\|y\|_\infty - \|x\|_\infty - \langle k, y - x\rangle = 2\max\{|x_1, x_2|\} - \max\{|x_1|, |x_2|\} - k_1(2x_1 - x_1) - k_2(2x_2 - x_2)$$

$(\because x_2 = x_1 > 0)$ $$= 2x_1 - x_1 - k_1 x_1 - k_2 x_2$$

$$= (1 - k_1 - k_2) x_1 < 0$$

Similarly, we assume $k = \begin{bmatrix} k_1 \\ k_2 \end{bmatrix}$ with $k_1 + k_2 < 1$,

then $y = \frac{1}{2}\begin{bmatrix} x_1 \\ \end{bmatrix}$, we have then $y = \frac{1}{2}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, we have

$$\|y\|_\infty - \|x\|_\infty - \langle k, y-x \rangle = \frac{1}{2}\max\{|x_1|, |x_2|\} - \max\{|x_1|, |x_2|\}$$
$$- k_1(\tfrac{1}{2}x_1 - x_1) - k_2(\tfrac{1}{2}x_2 - x_2)$$
$$= \frac{1}{2}x_1 - x_1 + \frac{1}{2}k_1x_1 + \frac{1}{2}k_2x_2$$
$$(\because x_2 = x_1 > 0) \quad = \frac{1}{2}(k_1 + k_2 - 1)x_1 < 0$$

Similarly, if $k = \begin{pmatrix} k_1 \\ k_2 \end{pmatrix}$ with $k_1 < 0$ or $k_2 < 0$, we could use similar approach for taking $y = \begin{pmatrix} \frac{1}{2}x_1 \\ x_2 \end{pmatrix}$ or $y = \begin{pmatrix} x_1 \\ \frac{1}{2}x_2 \end{pmatrix}$ to get contradiction:

$k_1 < 0$: $$\|y\|_\infty - \|x\|_\infty - \langle k, y-x \rangle = \max\{|\tfrac{1}{2}x_1|, |x_2|\} - \max\{|x_1|, |x_2|\}$$
$$- k_1(\tfrac{1}{2}x_1 - x_1) - k_2(x_2 - x_2)$$
$$= x_1 - x_1 + \frac{1}{2}k_1x_1 - 0$$
$$(\because k_1 < 0, x_1 > 0) \quad = \frac{1}{2}k_1x_1 < 0$$

$k_2 < 0$: $$\|y\|_\infty - \|x\|_\infty - \langle k, y-x \rangle = \max\{|x_1|, |\tfrac{1}{2}x_2|\} - \max\{|x_1|, |x_2|\} -$$
$$k_1(x_1 - x_1) - k_2(\tfrac{1}{2}x_2 - x_2)$$
$$= x_2 - x_2 + \frac{1}{2}k_2x_2 - 0$$

$= x_2 - x_2 + \frac{1}{2}k_2 x_2 - 0$

$= \frac{1}{2}k_2 x_2 < 0$

$\Rightarrow \partial g_3(x) \subseteq S$, where $x_1 = x_2 > 0$

$\Rightarrow$ we have $\partial g_3(x) = S$ where $x_1 = x_2 > 0$.

same as:

$$\partial g_3(x) = \left\{ \begin{pmatrix} r \\ 1-r \end{pmatrix} \in \mathbb{R}^2 : 0 \leq r \leq 1 \right\}$$

So, similarly, we have case 6-8 as the following:

Case 6 $x_2 = x_1 < 0$,

$$\partial g_3(x) = \left\{ \begin{pmatrix} -r \\ -1+r \end{pmatrix} \in \mathbb{R}^2 : 0 \leq r \leq 1 \right\}$$

Case 7: $x_1 = -x_2$, $x_1 < 0$

$$\partial g_3(x) = \left\{ \begin{pmatrix} -r \\ 1-r \end{pmatrix} \in \mathbb{R}^2 : 0 \leq r \leq 1 \right\}$$

case 8. $x_1 = -x_2$, $x_1 > 0$

$$\partial g_3(x) = \left\{ \begin{pmatrix} r \\ -1+r \end{pmatrix} \in \mathbb{R}^2 : 0 \leq r \leq 1 \right\}$$

case 9: If $x_1 = x_2 = 0$, we claim that

$$\partial g_3(x) = S_2 := \{ x : \|x\|_1 < 1 \}$$

$$\partial g_3(x) = S_2 := \{x : \|x\|_1 < 1\}$$

Take $k = \begin{pmatrix} k_1 \\ k_2 \end{pmatrix} \in S_2$. i.e. $|k_1| + |k_2| \le 1$.

$\forall\, y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in \mathbb{R}^2$,

$$\|y\|_\infty - \|\vec{0}\|_\infty - \langle k, y - \vec{0} \rangle = \|y\|_\infty - \langle k, y \rangle$$

$$= \max\{|y_1|, |y_2|\} - k_1 y_1 - k_2 y_2$$

$$\geq (|k_1| + |k_2|) \max\{|y_1|, |y_2|\} - k_1 y_1 - k_2 y_2$$

$$\geq |k_1||y_1| + |k_2||y_2| - k_1 y_1 - k_2 y_2 \geq 0$$

$\therefore S_2 \subseteq \partial g_3(x)$.

If $k \notin S_2$, then $|k_1| + |k_2| > 1$. Let $y = \begin{pmatrix} \text{sign}(k_1) \max\{|k_1|, |k_2|\} \\ \text{sign}(k_2) \max\{|k_1|, |k_2|\} \end{pmatrix}$

$$\|y\|_\infty - \|\vec{0}\|_\infty - \langle k, y - \vec{0} \rangle = \|y\|_\infty - \langle k, y \rangle$$

$$= \max\{|k_1|, |k_2|\} - |k_1| \max\{|k_1|, |k_2|\} - |k_2| \max\{|k_1|, |k_2|\}$$

$-|k_2| \max\{|k_1|, |k_2|\}$

$= (1-|k_1|-|k_2|)\max\{|k_1|, |k_2|\} < 0$

$\Rightarrow \partial g_3(x) \subseteq S_2, \quad \partial g_3(x) = \{x : \|x\|_1 < 1\}.$

Conclusion:

$$\partial g_3(x) = \begin{cases} \left\{\begin{pmatrix}1\\0\end{pmatrix}\right\} & \text{if } x_1 > |x_2| \geq 0 \\ \left\{\begin{pmatrix}0\\1\end{pmatrix}\right\} & \text{if } x_2 > |x_1| \geq 0 \\ \left\{\begin{pmatrix}-1\\0\end{pmatrix}\right\} & \text{if } x_1 < -|x_2| \leq 0 \\ \left\{\begin{pmatrix}0\\-1\end{pmatrix}\right\} & \text{if } x_2 < -|x_1| \leq 0 \\ \left\{\begin{pmatrix}r\\1-r\end{pmatrix} \in \mathbb{R}^2, 0 \leq r \leq 1\right\} & \text{if } x_1 = x_2 > 0 \\ \left\{\begin{pmatrix}-r\\1-r\end{pmatrix} \in \mathbb{R}^2, 0 \leq r \leq 1\right\} & \text{if } x_1 = -x_2 < 0 \\ \left\{\begin{pmatrix}-r\\-1+r\end{pmatrix} \in \mathbb{R}^2, 0 \leq r \leq 1\right\} & \text{if } x_1 = x_2 < 0 \\ \left\{\begin{pmatrix}r\\-1+r\end{pmatrix} \in \mathbb{R}^2, 0 \leq r \leq 1\right\} & \text{if } x_1 = -x_2 > 0 \\ \left\{x \in \mathbb{R}^2, \|x\|_1 \leq 1\right\} & \text{if } x_1 = x_2 = 0. \end{cases}$$

1d). $g_4(x) = (|x_1| + |x_2|)^2$ where $x \in \mathbb{R}^2$.

if $x_1 \neq 0$ and $x_2 \neq 0$,

if $x_1 \neq 0$ and $x_2 \neq 0$,

$$\frac{\partial g_4(x)}{\partial x_1} = 2(|x_1|+|x_2|) \cdot \frac{x_1}{|x_1|}$$

$$\frac{\partial g_4(x)}{\partial x_2} = 2(|x_1|+|x_2|) \cdot \frac{x_2}{|x_2|}$$

$$\therefore \nabla g_4(x) = 2\begin{pmatrix} x_1 + \frac{x_1 |x_2|}{|x_1|} \\ x_2 + \frac{x_2 |x_1|}{|x_2|} \end{pmatrix}$$

if $x_1 = 0$ and $x_2 \neq 0$,

$$\frac{\partial g_4(x)}{\partial x_1} = 2(|x_1|+|x_2|)\alpha_1 \quad \text{, where } |\alpha_1|<1$$

$$\frac{\partial g_4(x)}{\partial x_2} = 2(|x_1|+|x_2|) \cdot \frac{x_2}{|x_2|} = 2x_2$$

if $x_1 \neq 0$ and $x_2 = 0$, similarly,

$$\frac{\partial g_4(x)}{\partial x_1} = 2x_1, \quad \frac{\partial g_4(x)}{\partial x_2} = 2(|x_1|+|x_2|) \cdot \alpha_2,$$

where $|\alpha_2|<1$.

if $x_1 = x_2 = 0$,

$\partial g_4(x)$ $\quad (|x_1|+|x_2|)\alpha_1$ where $|\alpha_1|<1$

$$\frac{\partial g_4(x)}{\partial x_1} = 2(|x_1| + |x_2|)\alpha_1 \quad \text{where } |\alpha_1| < 1$$

$$\frac{\partial g_4(x)}{\partial x_2} = 2(|x_1| + |x_2|)\alpha_2 \quad \text{where } |\alpha_2| < 1$$

$$\Rightarrow \partial g_4(x) = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}, \text{ where } |\alpha_1| < 1 \text{ and } |\alpha_2| < 1.$$

In conclusion:

$$\partial g_4(x) = \begin{cases} \left\{ 2\begin{pmatrix} x_1 + \frac{x_1|x_2|}{|x_1|} \\ x_2 + \frac{x_2|x_1|}{|x_2|} \end{pmatrix} \right\} & \text{if } x_1 \neq 0 \text{ and } x_2 \neq 0. \\ \left\{ 2\begin{pmatrix} (|x_1| + |x_2|)\alpha_1 \\ x_2 \end{pmatrix}, |\alpha_1| < 1 \right\} & \text{if } x_1 = 0 \text{ and } x_2 \neq 0. \\ \left\{ 2\begin{pmatrix} x_1 \\ (|x_1| + |x_2|)\alpha_2 \end{pmatrix}, |\alpha_2| < 1 \right\} & \text{if } x_1 \neq 0 \text{ and } x_2 = 0 \\ \left\{ \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}, |\alpha_1| < 1, |\alpha_2| < 1 \right\} & \text{if } x_1 = x_2 = 0 \end{cases}$$

Q2. Let $F_{(k+1)}(x) = \frac{1}{2}\|x - x^{(k)}\|_2^2 + \alpha_k g(x)$

Let $f(x) = \frac{1}{2}\|x - x^{(k)}\|_2^2$

$F_{(k+1)}(x) = f(x) + \alpha_k g(x)$

1. Obviously, $F_{(k+1)}(x)$ is continuous as $\frac{1}{2}\|x-x^{(k)}\|_2^2$ is continuous and $g(x)$ is continuous.

2. $F_{(k+1)}(x)$ is coercive as $\|x\| \to +\infty$,

$f(x) \to +\infty$
And $\|x\| \to +\infty$, $g(x) \to +\infty$. since $g(x)$ is convex function.

3. $(1)+(2) \Rightarrow \min_{x \in \mathbb{R}^n} F(x)$ exists at least 1 solution.

4. As $f(x)$ is strictly convex and $g(x)$ is convex
$\Rightarrow F(x)$ is strictly convex.

5. conclusion: $\min_{x \in \mathbb{R}^n} F(x)$ exists and its unique.

Q3. Let $f(x) = \frac{1}{2}\|Ax-b\|_2^2$ and

$$g(x) = \sum_{i=1}^{n} r(x_i)$$

The FBS algorithm is:

$$x^{(k+1)} = x^{(k)} - \alpha_k (\nabla f(x^{(k)}) + \lambda u^{(k+1)}), \text{ where } u^{(k+1)} \in \partial g(x^{(k+1)})$$

$$\Updownarrow$$

$$0 \in x^{(k+1)} - (x^{(k)} - \alpha_k \nabla f(x^{(k)})) + \alpha_k \lambda u^{(k+1)}$$

$$\Updownarrow$$

$$0 \in \partial\left(\tfrac{1}{2}\| x^{(k+1)} - x^{(k)} - \alpha_k \nabla f(x^{(k)})\|^2 + \alpha_k \lambda g(x^{(k+1)})\right)$$

$$\Rightarrow \begin{cases} y^{(k+1)} = x^{(k)} - \alpha_k \nabla f(x^{(k)}) \\ x^{(k+1)} = \underset{x \in \mathbb{R}^n}{\operatorname{argmin}} \tfrac{1}{2}\|x - y^{(k+1)}\|^2 + \alpha_k \lambda \sum_{i=1}^{n} r(x_i) \end{cases}$$

$$\Leftrightarrow 0 \in x - y^{(k+1)} + \alpha_k \lambda u^{(k+1)}$$

$$\Leftrightarrow y_i^{(k+1)} \in x_i + \alpha_k \lambda u_i^{(k+1)}, \quad u_i \in \partial(r(x_i))$$

$$\partial(r(x_i)) = \begin{cases} 0 & \text{if } x_i < 0 \\ [0,1] & \text{if } x_i = 0 \\ 1 & \text{if } x_i > 0 \end{cases}$$

$$\partial(\ldots) = \begin{cases} [0,1] & \text{if } x_i = 0 \\ 1 & \text{if } x_i > 0 \end{cases}$$

$$\Leftrightarrow\ y_i^{(k+1)} = \begin{cases} x_i + \alpha_k \lambda & \text{if } x_i > 0 \\ [0, \alpha_k \lambda] & \text{if } x_i = 0 \\ x_i & \text{if } x_i < 0 \end{cases}$$

$$\Leftrightarrow\ x_i = \begin{cases} y_i^{(k+1)} - \alpha_k \lambda & \text{if } y_i^{(k+1)} > \alpha_k \lambda \\ 0 & \text{if } y_i^{(k+1)} \in [0, \alpha_k \lambda] \\ y_i^{(k+1)} & \text{if } y_i^{(k+1)} < 0 \end{cases}$$

let $T_{\alpha\lambda}(y_i^{(k+1)}) = \begin{cases} y_i^{(k+1)} - \alpha_k \lambda & \text{if } y_i^{(k+1)} > \alpha_k \lambda \\ 0 & \text{if } y_i^{(k+1)} \in [0, \alpha_k \lambda] \\ y_i^{(k+1)} & \text{if } y_i^{(k+1)} < 0 \end{cases}$

$$\Rightarrow \begin{cases} y^{(k+1)} = x^{(k)} - \alpha_k \nabla f(x^{(k)}) \\ x^{(k+1)} = \begin{pmatrix} T_{\alpha\lambda}(y_1^{(k+1)}) \\ \vdots \\ T_{\alpha\lambda}(y_n^{(k+1)}) \end{pmatrix} \end{cases}$$

where $\nabla f(x^{(k)}) = A^T(Ax^{(k)} - b)$

$$\Rightarrow \begin{cases} y^{(k+1)} = x^{(k)} - \alpha_k \left(A^T(Ax^{(k)} - b)\right) \\ x^{(k+1)} = \begin{pmatrix} T_{\alpha\lambda}(y_1^{(k+1)}) \end{pmatrix} \end{cases}$$

$$c/\ \left\{\ x^{(k+1)} = \begin{pmatrix} T_{\alpha\lambda}\left(y_1^{(k+1)}\right) \\ \vdots \\ T_{\alpha\lambda}\left(y_n^{(k+1)}\right) \end{pmatrix}\right.$$