SOLUTIONS HW 7

1 Problem 1

(a)

$$\begin{aligned}
& g \text{ is a subgradient of } f \text{ at } x \\
\iff & f(y) \geq f(x) + g^\mathsf{T}(y-x), \;\; \forall\, y \in \mathbb{R}^n \\
\iff & af(y) \geq af(x) + ag^\mathsf{T}(y-x), \;\; \forall\, y \in \mathbb{R}^n, \;\; a > 0 \\
\iff & ag \text{ is a subgradient of } af \text{ at } x.
\end{aligned}$$

(b) If g_1 is a subgradient of f_1 and g_2 is a subgradient of f_2 at x, then

$$\begin{aligned}
& f_1(y) \geq f_1(x) + g_1^\mathsf{T}(y-x), \;\; \forall\, y \in \mathbb{R}^n \\
& f_2(y) \geq f_2(x) + g_2^\mathsf{T}(y-x), \;\; \forall\, y \in \mathbb{R}^n \\
\Rightarrow \;\; & f_1(y) + f_2(y) \geq f_1(x) + f_2(x) + (g_1+g_2)^\mathsf{T}(y-x), \;\; \forall\, y \in \mathbb{R}^n,
\end{aligned}$$

which implies that $g_1 + g_2$ is a subgradient of $f_1 + f_2$ at x.

(c)

$$\begin{aligned}
& g \text{ is a subgradient of } f \text{ at } Ax+b \\
\iff & f(y) \geq f(Ax+b) + g^\mathsf{T}(y-(Ax+b)), \;\; \forall\, y \in \mathbb{R}^n \\
\iff & f(Ay+b) \geq f(Ax+b) + g^\mathsf{T}(Ay+b-(Ax+b)), \;\; \forall\, y \in \mathbb{R}^n, \;\; \text{(since } A \text{ is invertible)} \\
\iff & h(y) \geq h(x) + (A^\mathsf{T}g)^\mathsf{T}(y-x), \;\; \forall\, y \in \mathbb{R}^n, \;\; \text{(here } h(y) = f(Ay+b)\text{)} \\
\iff & A^\mathsf{T}g \text{ is a subgradient of } h \text{ at } x.
\end{aligned}$$

2 Problem 2

Inspired by the 1-D case, we can conjecture that the subdifferential of $f(x) = |x_1|+|x_2|+|x_3|$ at $(x_1, x_2, x_3) = (0,0,0)$ is: $\partial f(0,0,0) = \{(s_1, s_2, s_3) \in \mathbb{R}^3 : |s_1| \leq 1, |s_2| \leq 1, |s_3| \leq 1\}$.

Now we provide a proof. First, let's prove that for any (s_1, s_2, s_3) satisfying $|s_1| \leq 1, |s_2| \leq 1, |s_3| \leq 1$, the following inequality holds for all $(x_1, x_2, x_3) \in \mathbb{R}^3$:

$$|x_1| + |x_2| + |x_3| \geq s_1x_1 + s_2x_2 + s_3x_3.$$

For any $(x_1, x_2, x_3) \in \mathbb{R}^3$, we have $|x_1| + |x_2| + |x_3| \geq |s_1x_1| + |s_2x_2| + |s_3x_3| \geq s_1x_1 + s_2x_2 + s_3x_3$ due to the fact $|s_1| \leq 1$, $|s_2| \leq 1$, and $|s_3| \leq 1$. Thus, any (s_1, s_2, s_3) with $|s_1| \leq 1, |s_2| \leq 1, |s_3| \leq 1$ is a subgradient and hence belongs to $\partial f(0,0,0)$.

Now, let's prove the converse statement that for any $(s_1, s_2, s_3) \in \partial f(0,0,0)$, we must have $|s_1| \leq 1$, $|s_2| \leq 1$, and $|s_3| \leq 1$. By the definition of the subdifferential, for all $(x_1, x_2, x_3) \in \mathbb{R}^3$, we have $|x_1|+|x_2|+|x_3| \geq s_1x_1+s_2x_2+s_3x_3$. Taking $(x_1, x_2, x_3) = (1,0,0)$, we get $1 \geq s_1$, and hence $s_1 \leq 1$. Taking $(x_1, x_2, x_3) = (-1,0,0)$, we get $1 \geq -s_1$ and hence $s_1 \geq -1$ Therefore, $|s_1| \leq 1$. Similarly, we can show that $|s_2| \leq 1$ by choosing $(x_1, x_2, x_3) = (0,1,0)$ and $(x_1, x_2, x_3) = (0,-1,0)$. We can further show $|s_3| \leq 1$ by choosing $(x_1, x_2, x_3) = (0,0,1)$ and $(x_1, x_2, x_3) = (0,0,-1)$. Therefore, for any $(s_1, s_2, s_3) \in \partial f(0,0,0)$, we must have $|s_1| \leq 1$, $|s_2| \leq 1$, and $|s_3| \leq 1$.

Finally, we can conclude $\partial f(0,0,0) = \{(s_1, s_2, s_3) \in \mathbb{R}^3 : |s_1| \leq 1, |s_2| \leq 1, |s_3| \leq 1\}$.

3 Problem 3

Yes, the method always converges to the global minimum solution. We only need to verify three assumptions. First, the function $f(x) = |x_1| + |x_2| + |x_3|$ is a convex function (by the triangle inequality property). Second, the global minimum of f exists and can be attained by the point $(0, 0, 0)$. Finally, the norm of the subgradient of f is always upper bounded by 1. Therefore, all the three assumptions in our lecture note are satisfied, and the subgradient method with $\alpha_k = \frac{1}{\sqrt{k+1}}$ always converges to the global minimum at the rate $O\left(\frac{\log(N)}{\sqrt{N}}\right)$.