Management Sciences Topics: Convex Optimization
Homework 3

1. Problem 1

   We may change the step size to
$$\eta_k = \frac{2}{\mu(k+2)}$$
   as when $\mu > 0$ in SSG. Now we show that

   First, by three term lemma,
$$\langle x^{k+1} - x^k, g_x^k \rangle + \frac{1}{2\eta_k}\|x^{k+1} - x^k\|_2^2 \le \langle x - x^k, g_x^k \rangle + \frac{1}{2\eta_k}\|x - x^k\|_2^2 - \frac{1}{2\eta_k}\|x - x^{k+1}\|_2^2.$$
   By strong convexity of f w.r.t. x,
$$f(x^k, y) - f(x^*, y) \le \langle g_x^k, x^k - x^* \rangle - \frac{\mu}{2}\|x^k - x^*\|_2^2.$$
   Then by mean inequality,
$$\begin{aligned}
f(x^k, y) - f(x^*, y) &\le \langle x^* - x^k, g_x^k \rangle + \frac{1}{2\eta_k}\|x^* - x^k\|_2^2 - \frac{1}{2\eta_k}\|x^* - x^{k+1}\|_2^2 - \frac{1}{2\eta_k}\|x^{k+1} - x^k\|_2^2 - \frac{\mu}{2}\|x^k - x^*\|_2^2 \\
&\le \frac{1}{2}\eta_k\|g_x^k\|_2^2 + (\frac{1}{2\eta_k} - \frac{\mu}{2})\|x^k - x^*\|_2^2 - \frac{1}{2\eta_k}\|x^* - x^{k+1}\|_2^2 \\
&\le \frac{1}{2}\eta_k M^2 + (\frac{1}{2\eta_k} - \frac{\mu}{2})\|x^k - x^*\|_2^2 - \frac{1}{2\eta_k}\|x^* - x^{k+1}\|_2^2 \\
&= \frac{1}{\mu(k+2)}M^2 + \frac{\mu k}{4}\|x^k - x^*\|_2^2 - \frac{\mu(k+2)}{4}\|x^* - x^{k+1}\|_2^2.
\end{aligned}$$
   Then
$$\sum_{k=0}^{K-1}(k+1)(f(x^k, y) - f(x^*, y)) \le \sum_{k=0}^{K-1}\frac{k+1}{\mu(k+2)}M^2 + \sum_{k=0}^{K-1}\frac{\mu k(k+1)}{4}\|x^k - x^*\|_2^2 - \sum_{k=0}^{K-1}\frac{\mu(k+1)(k+2)}{4}\|x^* - x^{k+1}\|_2^2,$$
   and
$$\frac{k(k+1)}{2}(f(\bar{x}, y) - f(x^*, y)) \le \frac{KM^2}{\mu},$$
   so
$$f(\bar{x}, y) - f(x^*, y) \le \frac{2M^2}{(k+1)\mu}.$$
   Similarly, we have
$$f(x, \bar{y}) - f(x, y^*) \le \frac{2M^2}{(k+1)\mu}.$$
   Lets $y = \bar{y}$ in the first equation and $x = x^*$ in the second equation, by adding these two equations, we get
$$f(\bar{x}, \bar{y}) - f(x^*, y^*) \le \frac{4M}{(K+1)\mu}.$$
   Then the algorithm has a convergence rate of $O(\frac{1}{k})$.

2. Problem 2.