EE 546

HW1

Qirui Sun

1. $x_{t+1} = x_t - \mu \nabla f(x_t)$

$$\|x_{t+1} - x^*\|_{\ell_2}^2 = \|x_t - \mu \nabla f(x_t) - x^*\|^2$$

$$= \|x_t - x^*\|^2 - 2\mu \langle \nabla f(x_t), x_t - x^* \rangle + \mu^2 \|\nabla f(x_t)\|_{\ell_2}^2$$

$$-2\mu \langle \nabla f(x), x - x^* \rangle \leq -2\mu \frac{1}{\alpha} \|x_t - x^*\|_{\ell_2}^2 - \frac{2\mu}{\beta} \|\nabla f(x)\|_{\ell_2}^2$$

$$\Rightarrow \|x_{t+1} - x^*\|_{\ell_2}^2 \leq \|x_t - x^*\|_{\ell_2}^2 - \frac{2\mu}{\alpha} \|x_t - x^*\|_{\ell_2}^2 - \frac{2\mu}{\beta} \|\nabla f(x)\|_{\ell_2}^2 + \mu \|\nabla f(x_t)\|^2$$

$$= (1 - \frac{2\mu}{\alpha}) \|x_t - x^*\|_{\ell_2}^2 + (\mu^2 - \frac{2\mu}{\beta}) \|\nabla f(x)\|^2$$

$\because 0 < \mu \leq \frac{2}{\beta} \Rightarrow \mu^2 - \frac{2\mu}{\beta} < 0$

$$\therefore \|x_{t+1} - x^*\|_{\ell_2}^2 \leq (1 - \frac{2\mu}{\alpha}) \|x_t - x^*\|_{\ell_2}^2$$

$$\leq (1 - \frac{2\mu}{\alpha})^t \|x_0 - x^*\|_{\ell_2}$$

2. $f(x-\mu\nabla f(x)) \le f(x) - \mu(1-\frac{\mu L}{2})\|\nabla f(x)\|_{\ell_2}^2$

$$x_{t+1} = x_t - \mu\nabla f(x_t)$$

$\therefore \quad f(x_{t+1}) \le f(x_t) - \mu(1-\frac{\mu L}{2})\|\nabla f(x_t)\|_{\ell_2}^2$

$$\mu(1-\frac{\mu L}{2})\|\nabla f(x_t)\|_{\ell_2}^2 \le f(x_t) - f(x_{t+1})$$

$\because \quad \mu \le 1/L \Rightarrow (1-\frac{\mu L}{2}) \ge 0$

$\therefore \quad \mu(1-\frac{\mu L}{2})\sum_{t=1}^{T}\|\nabla f(x_t)\|_{\ell_2}^2 \le f(x_1) - f(x_{T+1})$

$\because$ $f(x)$ is bounded from below, $\therefore$ $f(x) \ge b \quad \forall x$

$\therefore \quad \mu(1-\frac{\mu L}{2})\sum_{t=1}^{T}\|\nabla f(x_t)\|_{\ell_2}^2 \le f(x_1) - b$

$\therefore \lim_{T\to\infty}\sum_{t=1}^{T}\|\nabla f(x_t)\|_{\ell_2}^2 \le \frac{f(x_1)-b}{\mu(1-\frac{\mu L}{2})} << \infty$

$\therefore \lim_{T\to\infty}\|\nabla f(x_t)\|^2 \to 0 \Rightarrow \lim_{T\to\infty}\|\nabla f(x_t)\| \to 0$

(b) under the same assumption, x_t may not converge

for example $x_t = \frac{1}{t}$

$\lim_{t\to\infty} x_t \to 0 \qquad \lim_{t\to\infty}\sum_{t=1}^{t}\|x_t\|_{\ell_2}^2 \to \infty$

c)

$$f(x - \mu \nabla f(x)) \le f(x) - \mu\left(1 - \frac{\mu L}{2}\right) \|\nabla f(x)\|_{\ell_2}^2$$

$$x_{t+1} = x_t - \mu \nabla f(x)$$

$$f(x_t - \mu \nabla f(x_t)) \le f(x_t) - \mu\left(1 - \frac{\mu L}{2}\right) \|\nabla f(x_t)\|_{\ell}^2$$

$$f(x_{t+1}) \le f(x_t) - \mu\left(1 - \frac{\mu L}{2}\right) \|\nabla f(x_t)\|_{\ell}^2$$

$$\|\nabla f(x)\|_{\ell_2}^2 \ge \gamma (f(x) - f(x^*))$$

$$\Rightarrow f(x_{t+1}) \le f(x_t) - \mu\left(1 - \frac{\mu L}{2}\right) \gamma \cdot (f(x_t) - f(x^*))$$

$$f(x_{t+1}) - f(x^*) \le f(x_t) - f(x^*) - \mu\left(1 - \frac{\mu L}{2}\right) \gamma \left(f(x_t) - f(x^*)\right)$$

$$f(x_{t+1}) - f(x^*) \le \left(1 - \mu\left(1 - \frac{\mu L}{2}\right)\gamma\right) (f(x_t) - f(x^*))$$

$$\mu \le 1/L \Rightarrow 1 - \frac{\mu L}{2} \le \frac{1}{2}$$

$$\therefore f(x_{t+1}) - f(x^*) \le \left(1 - \frac{1}{2}\mu\gamma\right) (f(x_t) - f(x^*))$$