1.1) 1) Use strong-convexity to go from function values to iterates.

2) follow the same proof as GD

3) Note that

$$f\left(x^t - \frac{1}{L}\nabla f(x^t)\right) \leq f(x^t) - \frac{1}{2L}\|\nabla f(x^t)\|_2^2$$

find relation between L^t & L

1.2 1) $\alpha = \|\alpha\|_1 \, \text{sign}(\alpha)$ ①

if $\alpha \in \mathbb{R}^d$; $\|\alpha\|_\infty = \max_i |\alpha_i|$ ②

Use 1, 2 with proof for GD.

2) Use norm inequalities

if $\alpha \in \mathbb{R}^d$ then,

$$\|\alpha\|_\infty \leq \|\alpha\| \leq \sqrt{d}\,\|\alpha\|_\infty$$

1.3) same proof as with d blocks

coordinate descent $\Rightarrow$ 1 coordinate/block

$$f(x^{t+1}) \leq f(x^t) - \frac{1}{2L}\|\nabla_{j_t} f(x^t)\|_2^2$$

Lipschitz-continuity wrt to each co-ordinate

coordinate chosen at iteration t

$P(j_t) = \frac{1}{d}$ [random selection of coordinate]

$$E[f(x^{t+1})] \leq E\left[f(x^t) - \frac{1}{2L}\|\nabla_{j_t} f(x^t)\|^2\right]$$

wrt to j_t doesn't depend on j_t

$$\Rightarrow E\left[f(x^{t+1})\right] \leq f(x^t) - \frac{1}{2dL}\sum_{j=1}^{d} \|\nabla_{j_t} f(x^t)\|^2$$

$$\leq f(x^t) - \frac{1}{2dL}\|f(x^t)\|^2$$

$$\Rightarrow E\left[f(x^{t+1})\right] - f(x^*) \leq \left(1 - \frac{\mu}{dL}\right)\left[f(x^t) - f(x^*)\right]$$

using strong - convexity

GD using PL-inequality → weaker condition compared to strong-convexity

$\forall x$: $\frac{1}{2}\|\nabla f(x)\|^2 \geq \mu\left[f(x) - f^*\right]$ (1) $\quad \left[f^* = f(x^*)\right]$

(2) $f(x^{t+1}) - f(x^t) \leq \frac{-1}{2L}\|\nabla f(x^t)\|^2$ [using Lipschitz continuity as before]

From (1)

$$\Rightarrow f(x^{t+1}) - f(x^t) \leq \frac{-1}{2L}\left[2\mu\left(f(x^t) - f^*\right)\right]$$

$$\Rightarrow f(x^{t+1}) - f(x^t) \leq \frac{-\mu}{2}\left[f(x^t) - f^*\right]$$

$$\Rightarrow f(x^{t+1}) - f^* \leq \left(1 - \frac{\mu}{2}\right)\left[f(x^t) - f^*\right]$$