problem 1.5

In the convergence proof of GD with constant step size and strongly convex objective function proof the coercivity of the gradient:

$$(\nabla f(x) - \nabla f(y))^T (x-y) \geq \frac{\mu L}{\mu + L} \|x-y\|_2^2 + \frac{1}{\mu + L} \|\nabla f(x) - \nabla f(y)\|_2^2$$

Solution:

Let define $g(x) = f(x) - \frac{\mu}{2} \|x\|_2^2$, we get $g(x)$ is convex from strong convexity of $f(x)$, we get $g(x)$ is convex

based on convergence of GD with constant step size and μ-strongly convex and L-smooth, we can say that the function $f(x)$, is differentiable and it's gradient is L-Lipschitz continuous,

So we get $g(x)$ is also L-Lipschitz continuous and smooth with parameter $(L-\mu)$

Now we can apply inequality in problem 1.2c to $g(x)$: co-coercivity

$$(\nabla g(x) - \nabla g(y))^T (x-y) \geq \frac{1}{L-\mu} \|\nabla g(x) - \nabla g(y)\|_2^2$$

$$(\nabla f(x) - \mu x - \nabla f(y) + \mu y)^T (x-y) \geq \frac{1}{L-\mu} \|\nabla f(x) - \nabla f(y) - \mu(x-y)\|_2^2$$

$$(\nabla f(x) - \nabla f(y) - \mu(x-y))^T (x-y) \geq \frac{1}{L-\mu} \|\nabla f(x) - \nabla f(y) - \mu(x-y)\|$$

$$\Rightarrow \left(\nabla f(x) - \nabla f(y)\right)^T (x-y) - \mu \|x-y\|_2^2 \geq \frac{1}{L-\mu}\Big\{ \|\nabla f(x) - \nabla f(y)\|^2 + \mu^2 \|x-y\|_2^2 - 2\mu \left(\nabla f(x) - \nabla f(y)\right)^T (x-y) \Big\}$$

$$\Rightarrow \left(\nabla f(x) - \nabla f(y)\right)^T (x-y) + \frac{2\mu}{L-\mu}\left(\nabla f(x) - \nabla f(y)\right)^T (x-y) \geq \frac{1}{L-\mu}\|\nabla f(x) - \nabla f(y)\|_2^2 + \frac{\mu^2}{L-\mu}\|x-y\|_2^2 + \mu\|x-y\|_2^2$$

$$\left(\frac{L+\mu}{L-\mu}\right)\left(\nabla f(x) - \nabla f(y)\right)^T (x-y) \geq \frac{1}{L-\mu}\|\nabla f(x) - \nabla f(y)\|_2^2 + \frac{\mu^2}{L-\mu}\|x-y\|_2^2 + \mu\|x-y\|_2^2$$

$$\Rightarrow \left(\frac{L+\mu}{L-\mu}\right)\left(\nabla f(x) - \nabla f(y)\right)^T (x-y) \geq \frac{1}{L-\mu}\|\nabla f(x) - \nabla f(y)\|_2^2 + \frac{\mu L}{L-\mu}\|x-y\|_2^2$$

$$\Rightarrow \left(\nabla f(x) - \nabla f(y)\right)^T (x-y) \geq \frac{1}{L+\mu}\|\nabla f(x) - \nabla f(y)\|_2^2 + \frac{\mu L}{L+\mu}\|x-y\|_2^2$$