Exercise 12 - Coordinate Descent

① Prove linear convergence for gradient descent under the PL condition.

Function f satisfies the PL condition if the following holds for $\mu > 0$:

$$\boxed{\frac{1}{2}\|\nabla f(x)\|^2 \geq \mu (f(x) - f^*)}$$

Suppose f is smooth with a constant L, then the stepsize for GD is $1/L$:

$$x_{k+1} = x_k - \frac{1}{L}\nabla f(x_k) \rightarrow x_{k+1} - x_k = -\frac{1}{L}\nabla f(x_k)$$

Since f is smooth, then for x_k and x_{k+1} it holds:

$$f(x_{k+1}) \leq f(x_k) + \langle \nabla f(x_k), x_{k+1} - x_k \rangle + \frac{L}{2}\|x_{k+1} - x_k\|^2$$

$$f(x_{k+1}) \leq f(x_k) + \langle \nabla f(x_k), -\frac{1}{L}\nabla f(x_k) \rangle + \frac{L}{2}\|-\frac{1}{L}\nabla f(x_k)\|^2 \quad -\frac{1}{L} \text{ is a constant}$$

$$f(x_{k+1}) \leq f(x_k) - \frac{1}{L}\langle \nabla f(x_k), \nabla f(x_k) \rangle + \frac{\cancel{L}}{2}\cdot\frac{1}{L^{\cancel{2}}}\|\nabla f(x_k)\|^2$$

$$f(x_{k+1}) \leq f(x_k) - \frac{1}{L}\|\nabla f(x_k)\|^2 + \frac{1}{2L}\|\nabla f(x_k)\|^2$$

$$f(x_{k+1}) \leq f(x_k) - \frac{1}{2L}\|\nabla f(x_k)\|^2$$

We plug in the PL condition:

$$f(x_{k+1}) \leq f(x_k) - \frac{1}{L}\mu(f(x_k) - f^*)$$

Reformulating, to get the linear rate statement: $(-f^*)$

$$f(x_{k+1}) - f^* \leq f(x_k) - \frac{\mu}{L}f(x_k) + \frac{\mu}{L}f^* - f^*$$

$$f(x_{k+1}) - f^* \leq f(x_k)\left(1 - \frac{\mu}{L}\right) - f^*\left(1 - \frac{\mu}{L}\right)$$

$$\boxed{f(x_{k+1}) - f^* \leq \left(1 - \frac{\mu}{L}\right)(f(x_k) - f^*)}$$

$\Rightarrow$ In recursive form:

$$\boxed{f(x_k) - f^* \leq \left(1 - \frac{\mu}{L}\right)^k (f(x_0) - f^*)}$$