3.5 | Sea $f: \mathbb{R}^n \to \mathbb{R}$ m-fuertemente convexa y L-suave (1)
con (único) minimizador x^*. Pruebe que la k-esima iteración del metodo del máximo descenso aplicado a f con paso $\frac{2}{m+L}$ cumple

$$\| x^n - x^* \| \leq \left(\frac{\kappa - 1}{\kappa + 1} \right)^n \| x^0 - x^* \|$$

En que $\kappa = L/m$.

Dem: Por pregunta anterior, conocemos la sgte propiedad: $\forall x, y \in \mathbb{R}^n$:

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

Haciendo $y = x^*$, usando que $\nabla f(x^*) = 0$, obtenemos

$$(\nabla f(x))^T (x - x^*) \geq \frac{mL}{m+L} \| x - x^* \|^2 + \frac{1}{m+L} \| \nabla f(x) \|^2 . \quad (*)$$

Ahora, probamos la propiedad por inducción:

$n = 0$ | ✓

Asumimos entonces

$$\| x^n - x^* \|^2 \leq \left(\frac{\kappa - 1}{\kappa + 1} \right)^{2n} \| x^0 - x^* \|^2$$

y evaluando el metodo del máximo descenso, obtenemos

$$\| x^{n+1} - x^* \|^2 = \| x^n - \frac{2}{m+L} \nabla f(x) - x^* \|^2$$

$$= \| (x^n - x^*) - \frac{2}{m+L} \nabla f(x) \|^2$$

$$= \| x^n - x^* \|^2 - \frac{4}{m+L} \nabla f(x)^T (x^n - x^*) + \frac{4}{(m+L)^2} \| \nabla f(x) \|^2$$

$$\leq \| x^n - x^* \|^2 - \frac{4}{m+L} \left[\frac{mL}{m+L} \| x^n - x^* \|^2 + \cancel{\frac{1}{m+L} \| \nabla f(x^n) \|^2} \right] + \cancel{\frac{4}{(m+L)^2} \| \nabla f(x) \|^2}$$

$$= \left(1 - \frac{4mL}{(m+L)^2} \right) \| x^n - x^* \|^2 = \left(\frac{m-L}{m+L} \right)^2 \| x^n - x^* \|^2$$

Obtenemos que

$$\|x^{n+1} - x^*\|^2 \leq \left(\frac{k-1}{k+1}\right)^2 \|x^n - x^*\|^2$$

$$(H.I) \Leftarrow \leq \left(\frac{k-1}{k+1}\right)^2 \left(\frac{k-1}{k+1}\right)^{2n} \|x^0 - x^*\|^2$$

$$= \left(\frac{k-1}{k+1}\right)^{2(n+1)} \|x^0 - x^*\|^2 \qquad \square.$$

3.6| Sea f convexa con gradiente L-Lipschitz. Asuma que sabemos que $x^* \in B(0,R)$. Considere el método del máximo descenso sobre la función fuertemente convexa

$$f_\varepsilon(x) = f(x) + \frac{\varepsilon}{2R^2}\|x\|^2.$$

En que $0 < \varepsilon << L$. Inicializando en $x^0 \in B(0,R)$, sea x_ε^* el minimizador de f_ε.

(a) pruebe que

$$\forall z \in B(0,R): f(z) - f(x^*) \leq f_\varepsilon(z) - f_\varepsilon(x_\varepsilon^*) + \frac{\varepsilon}{2}.$$

Dem: ya que

$$f_\varepsilon(x) = f(x) + \frac{\varepsilon}{2R^2}\|x\|^2$$

$$\Leftrightarrow f(x) = f_\varepsilon(x) - \frac{\varepsilon}{2R^2}\|x\|^2$$

tenemos que

$$f(z) - f(x^*) = \left(f_\varepsilon(z) - \frac{\varepsilon}{2R^2}\|z\|^2\right) - f(x^*)$$

$$= f_\varepsilon(z) - \left(f(x^*) + \frac{\varepsilon}{2R^2} \| x^* - (x^* - z) \|^2 \right)$$

$$= f_\varepsilon(z) - \left(f(x^*) + \frac{\varepsilon}{2R^2} \left(\|x^*\|^2 - 2 x^{*T}(x^* - z) + \|x^* - z\|^2 \right) \right)$$

$$= f_\varepsilon(z) - \left(f(x^*) + \frac{\varepsilon}{2R^2} \|x^*\|^2 \right) + \frac{\varepsilon}{R^2} x^{*T}(x^* - z) - \frac{\varepsilon}{2R^2} \|x^* - z\|^2$$

$$= f_\varepsilon(z) - f_\varepsilon(x^*) + \frac{\varepsilon}{R^2} x^{*T}(x^* - z) - \frac{\varepsilon}{2R^2} \|x^* - z\|^2$$

$$\leq f_\varepsilon(z) - f_\varepsilon(x^*) + \frac{\varepsilon}{2R^2} (x^* - z)^T (2x^* - (x^* - z))$$

$$= f_\varepsilon(z) - f_\varepsilon(x^*) + \frac{\varepsilon}{2R^2} (x^* - z)^T (x^* + z)$$

$$= f_\varepsilon(z) - f_\varepsilon(x^*) + \frac{\varepsilon}{2R^2} \left(\underbrace{\|x^*\|^2}_{\leq R^2} - \underbrace{\|z\|^2}_{\geq 0} \right)$$

$$\leq f_\varepsilon(z) - f_\varepsilon(x^*) + \frac{\varepsilon}{2} \quad \square$$

b) pruebe que, para un paso a especificar, el metodo del máximo descenso aplicado a f_ε encuentra una solucion z tal que

$$f_\varepsilon(z) - f_\varepsilon(x_\varepsilon^*) \leq \frac{\varepsilon}{2}$$

en a lo más

$$\left(1 + \frac{R^2 L}{\varepsilon}\right) 2 \log\left(\frac{2R^2 L}{\varepsilon}\right)$$

$$\cancel{\frac{R^2 L}{\varepsilon} \log\left(\frac{8R^2 L}{\varepsilon}\right)}$$ iteraciones

sol: obtengamos el modulo de convexidad de la funcion

$$f_\varepsilon(x) = f(x) + \frac{\varepsilon}{2R^2} \|x\|^2$$

para esto, sean $x, y \in \mathbb{R}^n$, $\alpha \in (0,1)$:

$$f_\varepsilon((1-\alpha)x+\alpha y) = f((1-\alpha)x+\alpha y) + \frac{\varepsilon}{2R^2}\|(1-\alpha)x+\alpha y\|^2$$

$$\leq (1-\alpha)f(x) + \alpha f(y) + \frac{\varepsilon}{R^2}\left(\|1-\alpha\|^2\|x\|^2 + 2\alpha(1-\alpha)x^T y + \alpha^2\|y\|^2\right)$$

$$= (1-\alpha)\left(f(x) + \frac{\varepsilon}{2R^2}\|x\|^2\right) + \alpha\left(f(y) + \frac{\varepsilon}{2R^2}\|y\|^2\right)$$

$$+ \frac{\varepsilon}{R^2}\left((1-\alpha)^2\|x\|^2 - (1-\alpha)\|x\|^2 + \alpha^2\|y\|^2 - \alpha\|y\|^2 + 2\alpha(1-\alpha)x^T y\right)$$

$$= (1-\alpha)f_\varepsilon(x) + \alpha f_\varepsilon(y) + \frac{\varepsilon}{2R^2}\left(-\|x\|^2(1-\alpha)\alpha - \|y\|^2\alpha(1-\alpha) + 2\alpha(1-\alpha)x^T y\right)$$

$$= (1-\alpha)f_\varepsilon(x) + \alpha f_\varepsilon(y) \cancel{+\frac{\varepsilon}{2R^2}} - \alpha(1-\alpha)\frac{\varepsilon}{2R^2}\left(\|x\|^2 - 2x^T y + \|y\|^2\right)$$

$$= (1-\alpha)f_\varepsilon(x) + \alpha f_\varepsilon(y) - \frac{1}{2}\underbrace{\left(\frac{\varepsilon}{R^2}\right)}_{m}\alpha(1-\alpha)\|x-y\|^2$$

De modo que f_ε es $\left(\frac{\varepsilon}{R^2}\right)$-fuertemente convexa.
Ahora, obtenemos el módulo de suavidad como sigue:

$$\|\nabla f_\varepsilon(x) - \nabla f_\varepsilon(y)\| = \|\nabla f(x) - \nabla f(y) + \frac{\varepsilon}{R^2}(x-y)\|$$

$$\leq \left(L + \frac{\varepsilon}{R^2}\right)\|x-y\|$$

Así que f_ε es $\left(L + \frac{\varepsilon}{R^2}\right)$-suave.

Como vimos en clases, teenos qe para una función g L-suave y m fuertemente convexa, se tiee qe la iteración de máximo descenso con paso constante $\alpha_k \equiv 1/L$ genera x^k tal qe (5)

$$g(x^k) - g^* \leq \varepsilon$$

siempre qe

$$k \geq \frac{L}{m} \log\left(\frac{g(x^0) - g^*}{\varepsilon}\right).$$

Aplicamos esto en nuestro caso para $\varepsilon \to \frac{\varepsilon}{2}$ obteniendo qe

$$k \geq \frac{L + \frac{\varepsilon}{R^2}}{\frac{\varepsilon}{R^2}} \log\left(\frac{f_\varepsilon(x^0) - f_\varepsilon^*}{\varepsilon/2}\right)$$

$$= \left(1 + \frac{R^2 L}{\varepsilon}\right) \log\left(\frac{2\,(f_\varepsilon(x^0) - f_\varepsilon^*)}{\varepsilon}\right)$$

ya qe f_ε es $\left(\frac{\varepsilon}{R^2}\right)$-fuertemente convexa, teenos la estimación $\leq 4R^2$

$$f_\varepsilon(x^0) - f_\varepsilon^* \leq \frac{\|\nabla f(x^0) - \nabla f(x^*)\|^2}{2\left(\frac{\varepsilon}{R^2}\right)} \leq \frac{R^2 L^2 \|x^0 - x^*\|^2}{2\varepsilon}$$

$$\leq \frac{2R^4 L^2}{\varepsilon}$$

Así obtenemos qe la cota se cumple para

$$k \geq \left(1 + \frac{R^2 L}{\varepsilon}\right) \log\left(4\frac{R^4 L^2}{\varepsilon^2}\right)$$

$$= \left(1 + \frac{R^2 L}{\varepsilon}\right) \cdot 2 \log\left(\frac{2R^2 L}{\varepsilon}\right) \quad \square$$