

$$\min_x F(x) := f(x) + g(x)$$

$$\begin{aligned} x_{k+1} &= (I + \eta \partial f)^{-1} (I - \eta \nabla g)(x_k) \\ &= \text{Prox}_f^\eta [x_k - \eta \nabla g(x_k)] \end{aligned}$$

Sufficient Decrease Lemma

$$\text{Prox}_f^\eta(x) = \arg \min_u f(u) + \frac{1}{2\eta} \|x - u\|^2$$

Assume $\begin{cases} \textcircled{1} f(x) \text{ is } \gamma\text{-convex} \\ \textcircled{2} g(x) \text{ is convex} \\ \textcircled{3} \nabla g(x) \text{ is } L\text{-continuous with } L \end{cases}$

$$\bar{x} = \text{Prox}_f^\eta(x - \eta \nabla g(x))$$

Stable step size
 $\eta \leq \frac{2}{L}$

$$F(x) - F(\bar{x}) \geq \left(\frac{1}{\eta} - \frac{L}{2}\right) \|\bar{x} - x\|^2$$

$$g(x) \equiv 0 \Rightarrow f(x) - f(\text{Prox}_f^\eta(x)) \geq \frac{1}{\eta} \|\text{Prox}_f^\eta(x) - x\|^2, \forall \eta > 0$$

Theorem [Prox-Grad Inequality]

Assume $\begin{cases} \textcircled{1} f(x) \text{ is convex} \\ \textcircled{2} \nabla g(x) \text{ is } L\text{-continuous with } L \end{cases}$

$$\text{Let } \bar{y} = \text{Prox}_f^\eta(y - \eta \nabla g(y)), \text{ and } \eta \leq \frac{1}{L}$$

$$\begin{aligned} F(x) - F(\bar{y}) &\geq \frac{L}{2} \|x - \bar{y}\|^2 - \frac{L}{2} \|x - y\|^2 \\ &\quad + g(x) - g(y) - \langle \nabla g(y), x - y \rangle \end{aligned}, \forall x$$

$$F(x) = f(x) + g(x)$$

Theorem Assume {

- ① $f(x)$ is convex
- ② $g(x)$ is convex
- ③ $\nabla g(x)$ is L -continuous with L

$F(x) = f(x) + g(x)$ is convex

$\{x_k\}$ generated by Proximal Gradient Method satisfies:

$$\|x_{k+1} - x_*\| \leq \|x_k - x_*\| \quad \text{for any minimizer } x_*$$

Proof:

$$F(x) - F(\bar{y}) \geq \frac{1}{2\eta} \|x - \bar{y}\|^2 - \frac{1}{2\eta} \|x - y\|^2 + g(x) - g(y) - \langle \nabla g(y), x - y \rangle$$

$$0 \geq 2\eta [F(x_*) - F(x_{k+1})] \geq \|x_* - x_{k+1}\|^2 - \|x_* - x_k\|^2$$

$$y = x_k \Rightarrow \bar{y} = x_{k+1} \quad x = x_*$$

because

$$g(x) - g(y) - \langle \nabla g(y), x - y \rangle \geq 0$$

Theorem [$O(\frac{1}{k})$ rate]

Assume {

- ① $f(x)$ is convex
- ② $g(x)$ is convex
- ③ $\nabla g(x)$ is L -continuous with L
- ④ $F(x)$ has at least one minimizer

Proximal Gradient Method with $\eta = \frac{1}{L}$ satisfies:

$$F(x_k) - F(x_*) \leq \frac{L}{2} \|x_0 - x_*\|^2 \cdot \frac{1}{k}$$

Proof: $2\eta [F(x_*) - F(x_{k+1})] \geq \|x_* - x_{k+1}\|^2 - \|x_* - x_k\|^2$

$$\Rightarrow \frac{2}{L} \sum_{k=0}^{n-1} [F(x_*) - F(x_{k+1})] \geq \|x_* - x_n\|^2 - \|x_* - x_0\|^2$$

$$\geq -\|x_* - x_0\|^2$$

$$\Rightarrow \frac{2}{L} \sum_{k=0}^{n-1} [F(x_{k+1}) - F(x_*)] \leq \|x_0 - x_*\|^2$$

Sufficient Decrease Lemma

$$F(x) - F(\bar{x}) \geq \left(\frac{1}{\eta} - \frac{L}{2}\right) \|\bar{x} - x\|^2$$

$$\Rightarrow F(x_{k+1}) \leq F(x_k)$$

$$\Rightarrow \text{LHS} \geq \frac{2}{L} \cdot n \cdot [F(x_n) - F(x_*)]$$

$$\Rightarrow \frac{2}{L} \cdot n \cdot [F(x_n) - F(x_*)] \leq \|x_0 - x_*\|^2$$

Strongly Convex Case

Theorem [Linear Rate]

$F(x)$ is strongly convex

$$\text{Assume} \begin{cases} \textcircled{1} f(x) \text{ is convex} \\ \textcircled{2} g(x) \text{ is strongly convex with } \mu \\ \textcircled{3} \nabla g(x) \text{ is } L\text{-continuous with } L \end{cases}$$

Proximal Gradient Method with $\eta = \frac{1}{L}$ satisfies:

$$\textcircled{1} \|x_{k+1} - x_*\|^2 \leq \left(1 - \frac{\mu}{L}\right) \|x_k - x_*\|^2$$

$$\textcircled{2} \|x_k - x_*\|^2 \leq \left(1 - \frac{\mu}{L}\right)^k \|x_0 - x_*\|^2$$

$$\textcircled{3} F(x_k) - F(x_*) \leq \frac{L}{2} \left(1 - \frac{\mu}{L}\right)^k \|x_0 - x_*\|^2$$

$$F(x_k) - F(x_*) \geq \langle \nabla F(x_*), x_k - x_* \rangle + \frac{\mu}{2} \|x_k - x_*\|^2$$

Proof:

$$\begin{aligned} F(x) - F(y) &\geq \frac{1}{2\eta} \|x - y\|^2 - \frac{1}{2\eta} \|x - y\|^2 \\ &\quad + g(x) - g(y) - \langle \nabla g(y), x - y \rangle \end{aligned}$$

$$\begin{aligned} \Rightarrow F(x_*) - F(x_{k+1}) &\geq \frac{L}{2} \|x_* - x_{k+1}\|^2 - \frac{L}{2} \|x_* - x_k\|^2 \\ &\quad + \underbrace{g(x_*) - g(x_k) - \langle \nabla g(x_k), x_* - x_k \rangle}_{\geq \frac{\mu}{2} \|x_* - x_k\|^2} \end{aligned}$$

$$\Rightarrow F(x_*) - F(x_{k+1}) \geq \frac{L}{2} \|x_* - x_{k+1}\|^2 - \frac{L - \mu}{2} \|x_* - x_k\|^2$$

$$\Rightarrow \|x_{k+1} - x_*\|^2 \leq \left(1 - \frac{\mu}{L}\right) \|x_k - x_*\|^2$$

$$\Rightarrow \|X_k - X_*\|^2 \leq \left(1 - \frac{\mu}{L}\right)^k \|X_0 - X_*\|^2$$

$$\begin{aligned} F(X_{k+1}) - F(X_*) &\leq \frac{L-\mu}{2} \|X_k - X_*\|^2 - \frac{L}{2} \|X_{k+1} - X_*\|^2 \\ &\leq \frac{L-\mu}{2} \|X_k - X_*\|^2 \\ &= \frac{L}{2} \left(1 - \frac{\mu}{L}\right) \|X_k - X_*\|^2 \\ &\leq \frac{L}{2} \left(1 - \frac{\mu}{L}\right)^{k+1} \|X_0 - X_*\|^2. \end{aligned}$$

Def $\{X_k\}$ is Fejer Monotone if there is y s.t.

$$\|X_{k+1} - y\| \leq \|X_k - y\|, \quad \forall k$$

Theorem [Convergence]

$$\text{Assume } \left\{ \begin{array}{l} \textcircled{1} f(x) \text{ is convex} \\ \textcircled{2} g(x) \text{ is convex} \\ \textcircled{3} \nabla g(x) \text{ is } L\text{-continuous with } L \\ \textcircled{4} F(x) \text{ has a minimizer} \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} \textcircled{1} \|X_{k+1} - X_*\| \leq \|X_k - X_*\| \\ \textcircled{2} F(X_k) - F(X_*) = O\left(\frac{1}{k}\right) \end{array} \right.$$

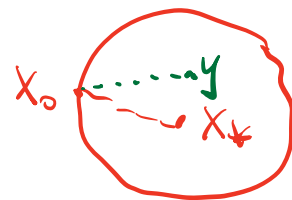
Proximal Gradient Method with $\eta = \frac{1}{L}$ satisfies:

$\{X_k\}$ converges to one of the minimizers of $F(x)$.

Proof:

$$\|X_{k+1} - X_*\| \leq \|X_k - X_*\| \quad \text{for any minimizer } X_*$$

$$\Rightarrow \|X_k - X_*\| \leq \|X_0 - X_*\|, \forall k$$



$\Rightarrow \{X_k\}$ is in the ball centered at X_* with radius $\|X_0 - X_*\|$

Real Analysis Bounded Sequence in $\mathbb{R}^n$ has a convergent subsequence

$$\Rightarrow X_{k_j} \rightarrow y, j \rightarrow \infty$$

Want to show y is a minimizer

$F(x)$ on $\mathbb{R}^n$ is convex $\Rightarrow F(x)$ is continuous

$$\Rightarrow F(X_{k_j}) \rightarrow F(y)$$

$$F(X_{k_j}) \rightarrow F(X_*) \Rightarrow F(y) = F(X_*)$$

$$\begin{matrix} \uparrow \\ F(X_{k_j}) \rightarrow F(X_*) \end{matrix} \Rightarrow y \text{ is a minimizer}$$

$$\Rightarrow \|X_{k+1} - y\| \leq \|X_k - y\|$$

$$\Rightarrow \left. \begin{array}{l} \|X_k - y\| \text{ is decreasing} \\ \|X_{k_j} - y\| \rightarrow 0 \end{array} \right\} \Rightarrow \|X_k - y\| \rightarrow 0$$

$\Rightarrow \{X_k\}$ converges to y .

$$X_1, X_2, X_3, X_4, \dots, X_{2k}, X_{2k+1}$$

$$\|X_{2k} - y\| \leq \|X_{2k-1} - y\|$$