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

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

Sufficient Decrease Lemma

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

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

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

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

Theorem [Prox-Grad Inequality]

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

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

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

Proof:

$$\phi(u) = g(y) + \langle \nabla g(y), u-y\rangle + f(u) + \underbrace{\frac{1}{2\eta}\|u-y\|^2}_{h(u)}$$

is strong convex w.r.t. u $\qquad \nabla^2 h(u) = \frac{1}{\eta} I$

$\Rightarrow$ $u^* = \arg\min_u \phi(u)$ satisfies

$$0 \in \nabla g(y) + \partial f(u^*) + \frac{1}{\eta}(u^* - y)$$

$$\Updownarrow$$

$$(I + \eta \partial f)(u^*) \ni y - \eta \nabla g(y)$$

$$\Updownarrow$$

$$u^* = (I + \eta \partial f)^{-1}(y - \eta \nabla g(y))$$

$\Rightarrow$ $\bar{y}$ minimizes $\phi(u)$

$$\Rightarrow \quad \phi(x) - \phi(\bar{y}) \geq \frac{1}{2\eta}\|x - \bar{y}\|^2$$

why?

$\phi(\cdot)$ is strongly convex

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

$$\forall g \in \partial\phi(y)$$

$$0 \in \partial\phi(\bar{y})$$

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

Descent Lemma

$$g(\bar{y}) \leq g(y) + \langle \nabla g(y), \bar{y} - y\rangle + \frac{L}{2}\|\bar{y} - y\|^2$$

$$\phi(\bar{y}) = \underline{g(y) + \langle \nabla g(y), \bar{y} - y\rangle + \frac{1}{2\eta}\|\bar{y} - y\|^2} + f(\bar{y})$$

$$\geq g(\bar{y}) + f(\bar{y}) = F(\bar{y})$$

$$\phi(x) - \phi(\bar{y}) \geq \frac{1}{2\eta}\|x - \bar{y}\|^2$$

$$-F(\bar{y}) \geq -\phi(\bar{y})$$

$$\Rightarrow \quad \phi(x) - F(\bar{y}) \geq \frac{1}{2\eta}\|x - \bar{y}\|^2$$

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

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

Remark: Let $x = y$, we get

$$F(y) - F(\bar{y}) \geq \frac{1}{2\eta}\|y - \bar{y}\|^2$$

Sufficient Decrease Lemma

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

For convenience, only consider $\eta = \frac{1}{L}$.

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

Proximal Gradient Method

$$\boxed{x_{k+1} = \mathrm{Prox}_f^{\eta}[x_k - \eta \nabla g(x_k)]}$$

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

$\eta \leq \frac{1}{L}$

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

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

Proof:

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

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

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

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

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

Assume
① $f(x)$ is convex
② $g(x)$ is convex
③ $\nabla g(x)$ is L-continuous with L
④ $F(x)$ has 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: $\quad 2\eta\left[F(x_*) - F(x_{k+1})\right] \geq \|x_* - x_{k+1}\|^2 - \|x_* - x_k\|^2$

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

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

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

$$\boxed{\begin{array}{l}\text{Sufficient Decrease Lemma} \\ F(x) - F(\bar{x}) \geq (\frac{1}{\eta} - \frac{L}{2})\|\bar{x} - x\|^2\end{array}} \Rightarrow F(x_{k+1}) \leq F(x_k)$$

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

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

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]

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

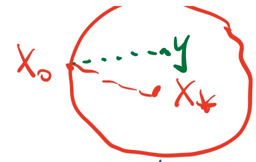
Proximal Gradient Method with $\eta = \frac{1}{L}$ satisfies:

$\{x_k\}$ converges to one minimizer of $F(x)$.

Proof:

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

$\Rightarrow \quad \|x_k - x_*\| \le \|x_0 - x_*\|, \ \forall k$



$\Rightarrow \quad \{x_k\}$ is in the ball centered at x_* with radius $\|x_0 - x_*\|$

Real Analysis

$\boxed{\text{Bounded Sequence in } \mathbb{R}^n \text{ has a convergent subsequence}}$

$\Rightarrow \quad x_{k_j} \to y, \ j \to \infty$

$F(x_{k_j}) \to F(x_*)$

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

$\Rightarrow \quad F(x_{k_j}) \to F(y)$

$\Rightarrow \quad F(y) = F(x_*)$

$\Rightarrow \quad y$ is a minimizer

$\Rightarrow \quad \|x_{k+1} - y\| \le \|x_k - y\|$

$$\Rightarrow \quad \left.\begin{array}{l} \|x_k - y\| \text{ is decreasing} \\ \|x_{k_j} - y\| \to 0 \end{array}\right\} \Rightarrow \|x_k - y\| \to 0$$

$\Rightarrow \quad \{x_k\}$ converges to y.

Strongly Convex Case

Theorem [Linear Rate] $\qquad F(x)$ is strongly convex

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

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

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

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

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

Proof:

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

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

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

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

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

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

$$\begin{aligned} F(x_{k+1}) - F(x_*) &\leq \frac{L-\mu}{2}\|x_k - x_*\|^2 - \frac{L}{2}\|x_{k+1} - x_*\|^2 \\ &\leq \frac{L-\mu}{2}\|x_k - x_*\|^2 \\ &= \frac{L}{2}(1-\frac{\mu}{L})\|x_k - x_*\|^2 \\ &\leq \frac{L}{2}(1-\frac{\mu}{L})^{k+1}\|x_0 - x_*\|^2 . \end{aligned}$$