More on proximal operator $f(x)$ is convex

Def Moreau-Yosida Regularization of $f(x)$ is

a.k.a. Moreau Envelope

$$f_\eta(x) = \min_u \left[f(u) + \frac{1}{2\eta}\|u-x\|^2 \right]$$

strong convexity $\Rightarrow$ existence of minimum

$$\mathrm{Prox}_f^\eta(x) = (I+\partial f)^{-1}(x) = \arg\min_u \left[f(u) + \frac{1}{2\eta}\|u-x\|^2 \right]$$

- It can be proven that $f_\eta(x)$ is convex & differentiable
- Theorem

① $$\nabla f_\eta(x) = \frac{x - \mathrm{Prox}_f^\eta(x)}{\eta}$$

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

② ∇f_η is L-Continuous with $L = \frac{1}{\eta}$

Proof: ① Let $u = \mathrm{Prox}_f^\eta(x)$

$v = \mathrm{Prox}_f^\eta(y)$

$$f_\eta(x) = \min_u \left[f(u) + \frac{\|u-x\|^2}{2\eta} \right]$$

$$f_\eta(y) = f(v) + \frac{\|v-y\|^2}{2\eta}$$

$$= f(v) + \frac{\|v-x\|^2}{2\eta} + \left\langle \frac{x-v}{\eta}, y-x \right\rangle + \frac{\|x-y\|^2}{2\eta}$$

$$\geq f(u) + \frac{\|u-x\|^2}{2\eta} + \frac{\|v-u\|^2}{2\eta} + \left\langle \frac{x-v}{\eta}, y-x \right\rangle + \frac{\|x-y\|^2}{2\eta}$$

u is minimizer of $g(w) = f(w) + \frac{1}{2\eta}\|w - x\|^2$ (Strongly convex)

$$\Rightarrow \quad g(v) \geq g(u) + \langle 0, v-u \rangle + \frac{1}{2\eta}\|v-u\|^2$$

$$= f(u) + \frac{\|u-x\|^2}{2\eta} + \langle \frac{x-u}{\eta}, y-x \rangle + \frac{\|x-y\|^2}{2\eta}$$

$$+ \langle \frac{u-v}{\eta}, y-x \rangle + \frac{\|u-v\|^2}{2\eta}$$

$$f_\eta(y) \geq = f_\eta(x) + \langle \frac{x-u}{\eta}, y-x \rangle + \frac{\eta}{2}\left\| \frac{x-y}{\eta} - \frac{u-v}{\eta} \right\|^2$$

$$\Rightarrow \qquad f_\eta(y) \geq f_\eta(x) + \langle \frac{x-u}{\eta}, y-x \rangle, \quad \forall x, y$$

$\Rightarrow$ $\frac{x-u}{\eta}$ is a subgradient of $f_\eta(x)$ at x.

$$\frac{x - \mathrm{Prox}_f^\eta(x)}{\eta} \in \partial f_\eta(x) \Rightarrow \nabla f_\eta(x) = \frac{x - \mathrm{Prox}_f^\eta(x)}{\eta}$$

② $f_\eta(y) \geq f_\eta(x) + \langle \frac{x-u}{\eta}, y-x \rangle + \frac{\eta}{2}\left\| \frac{x-u}{\eta} - \frac{y-v}{\eta} \right\|^2$

$$f_\eta(x) \geq f_\eta(y) + \langle \frac{y-v}{\eta}, x-y \rangle + \frac{\eta}{2}\left\| \frac{x-u}{\eta} - \frac{y-v}{\eta} \right\|^2$$

$$\Rightarrow \quad \langle \frac{x-u}{\eta} - \frac{y-v}{\eta}, x-y \rangle \geq \eta \left\| \frac{x-u}{\eta} - \frac{y-v}{\eta} \right\|^2$$

Let $G(x) = \dfrac{x - \mathrm{Prox}_f^{\eta}(x)}{\eta} = \nabla f_\eta(x)$

$\|G(x)-G(y)\| \cdot \|x-y\| \geq \langle G(x) - G(y), x-y \rangle \geq \eta \|G(x)-G(y)\|^2$

$\Rightarrow \|G(x) - G(y)\| \leq \frac{1}{\eta}\|x-y\|$

$\Rightarrow$ $G(x)$ is a Lip-continuous function with $L = \frac{1}{\eta}$

Proximal Point Method for $\min_x f(x)$

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

$\Updownarrow$

$$\boxed{\begin{array}{c} x_{k+1} = x_k - \eta \nabla f_\eta(x_k) \\ \text{Gradient Descent for } \min_x f_\eta(x) \end{array}}$$

$x_{k+1} = x_k - \sigma \nabla f_\eta(x_k)$

$\sigma < \frac{2}{L} = 2\eta$

$f_\eta(x) = \min_u \left[f(u) + \frac{1}{2\eta}\|u-x\|^2 \right]$

$\min_x f_\eta(x) = \min_{x,u} \left[\underbrace{f(u)}_{f(x_*)} + \underbrace{\frac{1}{2\eta}\|u-x\|^2}_{0} \right] = f_*$

$f_\eta(x)$ has the same minimizer as $f(x)$

Convergence Rate for nonsmooth $f(x)$

	Convexity	Strong Convexity
① Subgradient Method	$O(\frac{1}{\sqrt{k}})$	$O(\frac{1}{k})$
② Proximal Point Method	$O(\frac{1}{k})$	$O((\frac{1}{1+2\eta\mu})^k)$

Convergence Rate for smoother $f_\eta(x)$

$\nabla f_\eta(x)$ is Lip-cont. with $L = \frac{1}{\eta}$

	Convexity	Strong Convexity
Gradient Descent	$O(\frac{1}{k})$	$O\left[\left(1 - \frac{2\eta\mu L}{L+\mu}\right)^k\right]$

Implementing Proximal Point Method without formula for Prox:

for $k = 1, 2, \dots$

$$\boxed{\begin{array}{l}\text{Approximately solve } x_{k+1} = \arg\min_u \left[f(u) + \frac{1}{2\eta}\|u - x_k\|^2\right] \\ \text{Use Subgradient (or Accelerated Gradient) for} \\ \quad \min_u \left[f(u) + \frac{1}{2\eta}\|u - x_k\|^2\right]\end{array}}$$

for $j = 1, 2, \dots N$

$$u_{j+1} = u_j - \tau\left[\partial f(u_j) + \frac{1}{\eta}(u_j - x_k)\right]$$

end

$x_{k+1} \approx u_N$

end

Is this better than subgradient method for $\min_x f(x)$?

<u>Theorem</u> [Prox is firmly nonexpansive] $f(x)$ is convex

$$\|\mathrm{Prox}_f^\eta(x) - \mathrm{Prox}_f^\eta(y)\|^2 \le \langle \mathrm{Prox}_f^\eta(x) - \mathrm{Prox}_f^\eta(y), x - y \rangle$$

It implies $\|\mathrm{Prox}_f^{\eta}(x) - \mathrm{Prox}_f^{\eta}(y)\| \leq \|x-y\|$ (nonexpansive)

$\hookrightarrow$ L-cont. with $L=1$

Proof:

$$u = \mathrm{Prox}_f^{\eta}(x) \Leftrightarrow u = (I+\eta\partial f)^{-1}(x) \Leftrightarrow \frac{x-u}{\eta} \in \partial f(u)$$

$$v = \mathrm{Prox}_f^{\eta}(y) \Leftrightarrow v = (I+\eta\partial f)^{-1}(y) \Leftrightarrow \frac{y-v}{\eta} \in \partial f(v)$$

$$\left.\begin{array}{l} f(u) \geq f(v) + \langle \partial f(v), u-v\rangle \\ f(v) \geq f(u) + \langle \partial f(u), v-u\rangle \end{array}\right\} \Rightarrow \langle \partial f(u) - \partial f(v), u-v\rangle \geq 0$$

$$\Rightarrow \left\langle \frac{x-u}{\eta} - \frac{y-v}{\eta}, u-v \right\rangle \geq 0$$

$$\Rightarrow \langle x-y, u-v\rangle \geq \|u-v\|^2$$

1) $\min_x \|x\|_1 + \|Ax-b\|^2$

2) $\min_x \|x\|_1 + i_S$

3) $\min_x \|Ax\|_1 + i_S$

Def An operator T is

1) Firmly nonexpansive if $\|Tx - Ty\|^2 \leq \langle Tx - Ty, x-y\rangle$

2) Nonexpansive if $\|Tx - Ty\| \leq \|x-y\|$

3) Contraction if $\|Tx - Ty\| < \|x-y\|$

Example: $T(x)$ is L-cont. with $L<1$ is a contraction.

Theorem [Prox is a contraction for strongly convex function]

If $f(x)$ is strongly convex with $\mu > 0$

$$(1+\eta\mu)\|\mathrm{Prox}_f^{\eta}(x) - \mathrm{Prox}_f^{\eta}(y)\|^2 \leq \langle \mathrm{Prox}_f^{\eta}(x) - \mathrm{Prox}_f^{\eta}(y), x-y\rangle$$

It implies $\|\mathrm{Prox}_f^{\eta}(x) - \mathrm{Prox}_f^{\eta}(y)\| \leq \frac{1}{1+\eta\mu}\|x-y\|$

$$\Rightarrow \|x_{k+1} - x_*\| \le \frac{1}{1+\eta\mu} \|x_k - x_*\|$$

$$\left(\frac{1}{1+\eta\mu}\right)^2 < \frac{1}{1+2\eta\mu} \Leftrightarrow 1+2\eta\mu < 1+2\eta\mu + \eta^2\mu^2$$

Proof: $u = \text{Prox}^{\eta}_{f}(x) \Leftrightarrow u = (I+\eta\partial f)^{-1}(x) \Leftrightarrow \frac{x-u}{\eta} \in \partial f(u)$

$v = \text{Prox}^{\eta}_{f}(y) \Leftrightarrow v = (I+\eta\partial f)^{-1}(y) \Leftrightarrow \frac{y-v}{\eta} \in \partial f(v)$

$$f(u) \ge f(v) + \langle \partial f(v), u-v\rangle + \frac{\mu}{2}\|u-v\|^2$$

$$f(v) \ge f(u) + \langle \partial f(u), v-u\rangle + \frac{\mu}{2}\|u-v\|^2$$

$$\Rightarrow \langle \partial f(u) - \partial f(v), u-v\rangle \ge \mu\|u-v\|^2$$

$$\Rightarrow \left\langle \frac{x-u}{\eta} - \frac{y-v}{\eta}, u-v\right\rangle \ge \mu\|u-v\|^2$$

$$\Rightarrow \langle x-y, u-v\rangle \ge [1+\eta\mu]\|u-v\|^2$$

Convergence Rate for nonsmooth problems

	Convexity	Strong Convexity
① Subgradient Method	$O(\frac{1}{\sqrt{k}})$	$O(\frac{1}{k})$
② Proximal Point Method	$O(\frac{1}{k})$	$O\left(\frac{1}{[1+\eta\mu]^2}\right)^k$

$$\underbrace{\qquad\qquad}_{f(x_k) - f(x_*)}$$