

Plan

- Convergence of Peaceman-Rachford (PR)
- Convergence Rates of DR and PR

Lemma

Assume $f(x)$ and $g(x)$ are convex.

and x_* is a minimizer of $f(x) + g(x)$.

Let $y_* \in x_* + \eta \partial g(x_*)$, then

$$x_* = \text{Prox}_g^\eta(y_*) \quad \begin{array}{l} (I + \eta \partial g) \text{ is multi-valued} \\ (I + \eta \partial g)^{-1} \text{ is single-valued} \end{array}$$

$$0 \in \partial f(x_*) + \partial g(x_*) \Leftrightarrow R_f R_g(y_*) = y_*$$

Remark: y_* is a fixed point of

$$\left\{ \begin{array}{l} \text{Peaceman-Rachford} \quad y_{k+1} = R_f R_g(y_k) \quad \lambda = 2 \\ \text{Douglas-Rachford} \quad y_{k+1} = \frac{I + R_f R_g}{2}(y_k) \quad \lambda = 1 \\ \text{Generalized DR} \quad y_{k+1} = (1-\lambda)y_k + \lambda \frac{I + R_f R_g}{2}(y_k) \quad \lambda \in (0, 2) \end{array} \right.$$

Proof: $0 \in \partial f(x_*) + \partial g(x_*)$

$$\Leftrightarrow 0 \in -\eta \partial f(x_*) - \eta \partial g(x_*)$$

$$\Leftrightarrow 0 \in x_* - \eta \partial f(x_*) - \underbrace{[x_* + \eta \partial g(x_*)]}_{y_* \in}$$

$$\Leftrightarrow y_* \in x_* - \eta \partial f(x_*)$$

$$\Leftrightarrow -y_* \in -x_* + \eta \partial f(x_*)$$

$$\Leftrightarrow 2x_* - y_* \in (I + \eta \partial f)(x_*)$$

$$\Leftrightarrow (I + \eta \partial f)^{-1}(2x_* - y_*) = x_*$$

$$\Leftrightarrow (I + \eta \partial f)^{-1}(R_g(y_*)) = \text{Prox}_g^\eta(y_*)$$

$$\Leftrightarrow 2(I + \eta \partial f)^{-1}(R_g(y_*)) - [2\text{Prox}_g^\eta(y_*) - y_*] = y_*$$

$$\Leftrightarrow 2(I + \eta \partial f)^{-1}(R_g(y_*)) - R_g(y_*) = y_*$$

$$\Leftrightarrow R_f[R_g(y_*)] = y_*$$

So if $\begin{cases} \text{Peaceman-Rachford} & y_{k+1} = R_f R_g(y_k) \\ \text{Douglas-Rachford} & y_{k+1} = \frac{I + R_f R_g}{2}(y_k) \\ \text{Generalized DR} & y_{k+1} = (1-\lambda)y_k + \lambda \frac{I + R_f R_g}{2}(y_k) \end{cases}$
 $\lambda \in (0, 2)$
 converges, then $x_k = \text{Prox}_g^\eta(y_k)$
 converges to x_*

$y_{k+1} = R_f R_g(y_k)$ may NOT converge!

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

$$\|R_f(x) - R_f(y)\|^2 + 4\mu\eta \|\text{Prox}_f^\eta(x) - \text{Prox}_f^\eta(y)\|^2$$

$$\leq \|x - y\|^2$$

$$R_f(x) = 2\text{Prox}_f^\eta(x) - x$$

Remark: with only convexity ($\mu = 0$), we

get nonexpansiveness $\|R_f(x) - R_f(y)\| \leq \|x - y\|$.

Proof:

$$\begin{aligned}\|R_f(x) - R_f(y)\|^2 &= \|2\text{Prox}_f^\eta(x) - x - (2\text{Prox}_f^\eta(y) - y)\|^2 \\ &= 4\|\text{Prox}_f^\eta(x) - \text{Prox}_f^\eta(y)\|^2 + \|x - y\|^2 \\ &\quad - 4\langle \text{Prox}_f^\eta(x) - \text{Prox}_f^\eta(y), x - y \rangle\end{aligned}$$

Theorem $f(x)$ is strongly convex with $\mu > 0 \Rightarrow$
 $(1 + \mu\eta)\|\text{Prox}_f^\eta(x) - \text{Prox}_f^\eta(y)\|^2 \leq \langle \text{Prox}_f^\eta(x) - \text{Prox}_f^\eta(y), x - y \rangle$

\Downarrow

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

$$\leq -4\mu\eta\|\text{Prox}_f^\eta(x) - \text{Prox}_f^\eta(y)\|^2 + \|x - y\|^2$$

$$\begin{aligned}\Rightarrow \|R_f(x) - R_f(y)\|^2 + 4\mu\eta\|\text{Prox}_f^\eta(x) - \text{Prox}_f^\eta(y)\|^2 \\ \leq \|x - y\|^2\end{aligned}$$

Theorem (Peaceman-Rachford splitting with strong convexity)

Assume $f(x)$ is convex and $g(x)$ is strongly convex with $\mu > 0$.

Then $f(x) + g(x)$ has a unique minimizer x_* . The iteration

$$\begin{cases} y_{k+1} = R_f R_g(y_k) \\ x_k = \text{Prox}_g^\eta(y_k) \end{cases}, \forall \eta > 0 \text{ converges and}$$

$$\min_{0 \leq k \leq N} \|x_k - x_*\|^2 \leq \frac{1}{N+1} \cdot \frac{1}{4\mu\eta} \|y_0 - y_*\|^2$$

Remark: ① $R_f R_g$ is nonexpansive $\Rightarrow \|y_k - y_*\| \searrow$

But $\|x_k - x_*\|$ is not \searrow

② This is the worst case rate.

Proof:

$$\begin{aligned} & \|R_f[R_g(y_k)] - R_f[R_g(y_*)]\|^2 + 4\mu\eta \|\text{Prox}_g^\eta(y_k) - \text{Prox}_g^\eta(y_*)\|^2 \\ & \leq \|R_g(y_k) - R_g(y_*)\|^2 + 4\mu\eta \|\text{Prox}_g^\eta(y_k) - \text{Prox}_g^\eta(y_*)\|^2 \\ & \leq \|y_k - y_*\|^2 \end{aligned}$$

$$\Rightarrow \|y_{k+1} - y_*\|^2 + 4\mu\eta \|x_k - x_*\|^2 \leq \|y_k - y_*\|^2$$

$$\begin{aligned} \Rightarrow 4\mu\eta \sum_{k=0}^N \|x_k - x_*\|^2 & \leq \|y_0 - y_*\|^2 - \|y_{N+1} - y_*\|^2 \\ & \leq \|y_0 - y_*\|^2 \end{aligned}$$

Theorem (Peaceman-Rachford splitting with strong convexity)

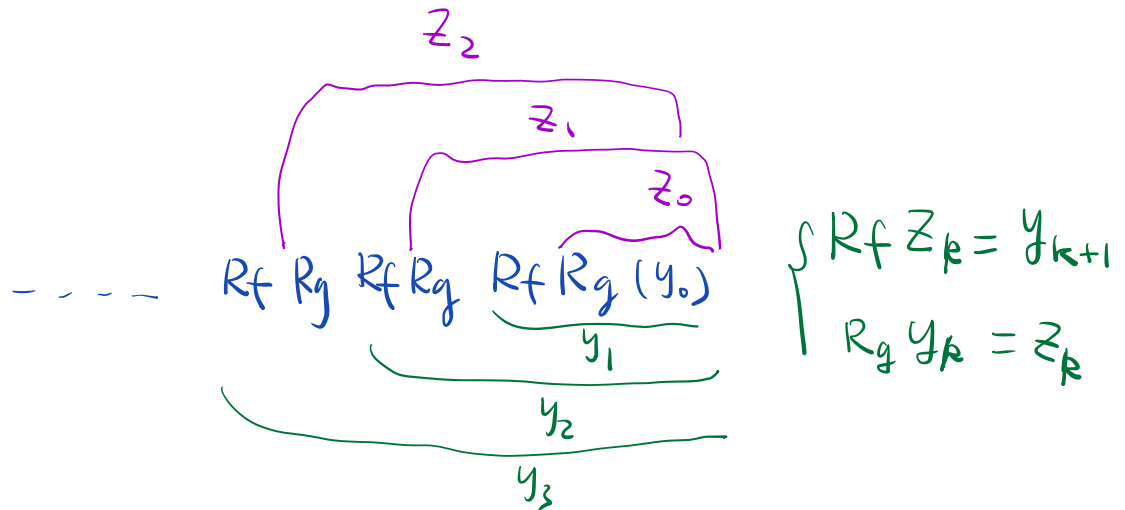
Assume $f(x)$ is convex and $g(x)$ is strongly convex with $\mu > 0$.

Then $f(x) + g(x)$ has a unique minimizer x_* . The iteration

$$\begin{cases} z_{k+1} = R_g R_f(z_k) \\ w_{k+1} = \text{Prox}_g^\eta(R_f(z_k)) \end{cases} \quad \forall \eta > 0 \text{ converges and}$$

$$\min_{0 \leq k \in \mathbb{N}} \|w_k - x_*\|^2 \leq \frac{1}{N} \cdot \frac{1}{4\mu\eta} \|R_f(z_0) - y_*\|^2$$

Proof:



Theorem [rate of Generalized Douglas-Rachford]

Assume ① $f(x)$ and $g(x)$ are convex,

$$0 < \theta < 1$$

② $f(x) + g(x)$ has a minimizer $(1-\theta)I + \theta R_f R_g$

then the iteration

$$\begin{cases} y_{k+1} = (1-\lambda)y_k + \lambda \frac{I + R_f R_g}{2}(y_k) \\ x_{k+1} = \text{Prox}_g^\eta(y_{k+1}) \end{cases}$$

$$\theta = \frac{1}{2}$$

converges >

$$\forall \eta > 0 \\ \lambda \in (0, 2)$$

and $\{x_k\}$ converges to one minimizer of $f(x) + g(x)$ with

$$1) \quad \underline{\|y_{k+1} - y_k\|^2} \leq \frac{1}{(k+1)} \left(\frac{1}{1-\lambda/2} - 1 \right) \|y_0 - y^*\|^2$$

$$2) \quad \underline{\|x_{k+1} - x_k\|^2} \leq \frac{1}{(k+1)} \left(\frac{1}{1-\lambda/2} - 1 \right) \|y_0 - y^*\|^2$$

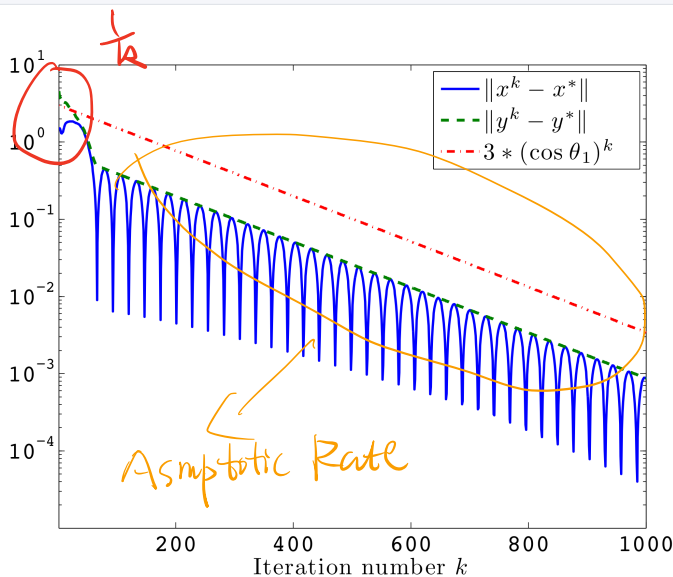
Proof: It is trivially implied by
the fixed point iteration:

Theorem Assume $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ is nonexpansive,
then T has at least one fixed point.

$$x_{k+1} = \theta x_k + (1-\theta) T(x_k), \quad \theta \in (0,1) \text{ satisfies}$$

1) $\{x_k\}$ converges to one fixed point y of T

$$2) \quad \|x_{k+1} - x_k\|^2 \leq \frac{1}{(k+1)} \left(\frac{1}{\theta} - 1 \right) \|x_0 - y\|^2$$



$$\min \|x\|_1 + \tau \sum_{Ax=b}$$

$$\begin{aligned} & \|x_k - x^*\| \\ &= \| \text{Prox}_g(y_k) - \text{Prox}_g(y^*) \| \\ &\leq \| y_k - y^* \| \end{aligned}$$

Provable linear rate if $\begin{cases} f(x) \text{ is strongly convex with } \mu > 0 \\ \nabla f(x) \text{ is } L\text{-cont with } L \end{cases}$

$$\|\nabla f(x) - \nabla f(y)\|^2 \leq L^2 \|x - y\|^2$$

$$\Rightarrow \|(I + \eta \nabla f)(x) - (I + \eta \nabla f)(y)\|^2$$

$$= \|x - y\|^2 + \eta^2 \|\nabla f(x) - \nabla f(y)\|^2 + 2\eta \langle x - y, \nabla f(x) - \nabla f(y) \rangle$$

$$\leq \|x - y\|^2 + \eta^2 L^2 \|x - y\|^2 + 2\eta \|x - y\| \cdot \|\nabla f(x) - \nabla f(y)\|$$

$$\leq [1 + \eta^2 L^2 + 2\eta L] \|x - y\|^2 \quad \begin{array}{l} (I + \eta \nabla f)(x) = u \\ (I + \eta \nabla f)(y) = v \end{array}$$

$$\Rightarrow \|\text{Prox}_f^\eta(u) - \text{Prox}_f^\eta(v)\|^2 \geq \frac{1}{[1 + \eta L]^2} \|x - y\|^2$$

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

$$\|R_f(x) - R_f(y)\|^2 + 4\mu\eta \|\text{Prox}_f^\eta(x) - \text{Prox}_f^\eta(y)\|^2 \leq \|x - y\|^2$$

$$R_f(x) = 2\text{Prox}_f^\eta(x) - x$$



$$\|R_f(x) - R_f(y)\|^2 + 4\mu\eta \|\text{Prox}_f^\eta(x) - \text{Prox}_f^\eta(y)\|^2 - \|x - y\|^2 \leq 0$$

$$\Rightarrow \|R_f(x) - R_f(y)\|^2 + \left[\frac{4\mu\eta}{[1 + \eta L]^2} - 1 \right] \|x - y\|^2 \leq 0$$

$$\Rightarrow \|R_f(x) - R_f(y)\|^2 \leq \left[1 - \frac{4\mu\eta}{[1 + \eta L]^2} \right] \|x - y\|^2$$

Recall $\mu \leq L$, so $1 - \frac{4\mu\eta}{[1+\eta L]^2} \geq 1 - \frac{4\eta L}{[1+\eta L]^2} = \frac{[1-\eta L]^2}{[1+\eta L]^2}$

Lemma Assume $\begin{cases} f(x) \text{ is strongly convex with } \mu > 0 \\ \nabla f(x) \text{ is } L\text{-cont with } L \end{cases}$

$$\|R_f(x) - R_f(y)\|^2 \leq \left[1 - \frac{4\mu\eta}{[1+\eta L]^2} \right] \|x - y\|^2$$

where $1 - \frac{4\mu\eta}{[1+\eta L]^2} \geq \frac{[1-\eta L]^2}{[1+\eta L]^2}$

So R_f is a contraction

$$\Rightarrow \begin{cases} R_f R_g \text{ is a contraction} \\ \frac{I + R_f R_g}{2} \text{ is a contraction} \\ (1-\lambda)I + \lambda \frac{I + R_f R_g}{2} \text{ is a contraction, } \lambda \in (0, 2) \end{cases}$$

Theorem (Peaceman-Rackford splitting Lions & Mercier 1979)

Assume ① $f(x)$ and $g(x)$ are convex

② $g(x)$ is strongly convex with $\mu > 0$

③ $\nabla g(x)$ is Lipschitz continuous with $L > 0$.

then the iteration

$$\begin{cases} y_{k+1} = R_f R_g (y_k) \\ x_{k+1} = \text{Prox}_g^\eta (y_{k+1}) \end{cases} \text{ converges, } \forall \eta > 0$$

and $\{x_k\}$ converges to the minimizer of $f(x) + g(x)$.

$$\Rightarrow \|y_{k+1} - y_*\| \leq \left(\sqrt{1 - \frac{4\mu\eta}{(1+\eta L)^2}} \right) \|y_k - y_*\|$$

$$\Rightarrow \|x_k - x_*\|^2 \leq \left(1 - \frac{4\mu\eta}{(1+\eta L)^2} \right)^k \|y_0 - y_*\|^2$$

Remark: $\{\|y_k - y_*\|\}$ is \downarrow

$\{\|x_k - x_*\|\}$ is not monotone!

Proof:

$$c = \sqrt{1 - \frac{4\mu\eta}{(1+\eta L)^2}}$$

$$\textcircled{1} \|R_f(x) - R_f(y)\| \leq c \|x - y\|$$

$$\textcircled{2} R_g \text{ is a contraction } \Rightarrow \text{ so is } R_f R_g$$

$$\Rightarrow \|R_f R_g (y_k) - R_f R_g (y_*)\| \leq c \|y_k - y_*\|$$

$$\Rightarrow \|y_{k+1} - y_*\| \leq c \|y_k - y_*\|$$

$$\Rightarrow \|y_k - y_*\| \leq c^k \|y_0 - y_*\|$$

$$\textcircled{3} \|x_k - x_*\| = \|\text{Prox}_g^\eta (y_k) - \text{Prox}_g^\eta (y_*)\|$$

$$\leq \|y_k - y_*\| \leq c^k \|y_0 - y_*\|$$

Theorem (Linear Rate of Generalized Douglas-Rachford)

Assume ① $f(x)$ and $g(x)$ are convex

② $g(x)$ is strongly convex with $\mu > 0$

③ $\nabla g(x)$ is Lipschitz continuous with $L > 0$.

then the iteration

$$\begin{cases} y_{k+1} = (1-\lambda)y_k + \lambda \frac{I + R_f R_g}{2}(y_k) \\ x_{k+1} = \text{Prox}_g^\eta(y_{k+1}) \end{cases} \text{ converges,}$$

$\forall \eta > 0$
 $\lambda \in (0, 2]$

and $\{x_k\}$ converges to one minimizer of $f(x) + g(x)$:

$$1) \|y_{k+1} - y_*\| \leq \left[1 - \frac{\lambda}{2} + \frac{\lambda}{2}c\right] \|y_k - y_*\| \quad c = \sqrt{1 - \frac{4\mu\eta}{(1+\eta L)^2}}$$

$$2) \|x_k - x_*\|^2 \leq \left(\left[1 - \frac{\lambda}{2} + \frac{\lambda}{2}c\right]^2\right)^k \|y_0 - y_*\|^2$$

Remark: Be aware that actual convergence rate can be faster than this provable rate.

In other words, $\lambda = 2$ is not necessarily the best in practice.

Proof: $T = R_f R_g$ is a contraction with

$$\|T(x) - T(y)\| \leq c \|x - y\|,$$

$$S = (1-\lambda)I + \lambda \frac{I+T}{2}, \quad \lambda \in (0, 2]$$

$$= \left(1 - \frac{\lambda}{2}\right)I + \frac{\lambda}{2}T$$

$$\begin{aligned} \|S(x) - S(y)\| &\leq \left(1 - \frac{\lambda}{2}\right) \|x - y\| + \frac{\lambda}{2} \|T(x) - T(y)\| \\ &\leq \left(1 - \frac{\lambda}{2} + \frac{\lambda}{2}c\right) \|x - y\| \end{aligned}$$

Convergence rate of either $F(x_k) - F(x_*)$
or $\|x_k - x_*\|^2$

$\min_x f(x)$ ∇f

Gradient Descent

Accelerated GD

Convexity

$$O\left(\frac{1}{k}\right)$$

$$O\left(\frac{1}{k^2}\right)$$

Strong Convexity

$$O\left(\left[\frac{L-\mu}{L+\mu}\right]^2\right)^k \quad \eta = \frac{2}{L+\mu}$$

$$O\left((1 - \sqrt{\frac{\mu}{L}})^k\right) \quad \eta = \frac{1}{L}$$

$\min_x f(x)$ ∂f

Subgradient Method

Proximal Point Method

$$O\left(\frac{1}{\sqrt{k}}\right)$$

$$O\left(\frac{1}{k}\right)$$

$$O\left(\frac{1}{k}\right)$$

$$\eta \leq \dots$$

$$O\left(\left[\frac{1}{1+\eta\mu}\right]^2\right)^k \quad \forall \eta > 0$$

$\min_x f(x) + g(x)$ $\partial f \quad \nabla g$

Proximal Gradient

Accelerated Prox Grad

$$O\left(\frac{1}{k}\right)$$

$$O\left(\frac{1}{k^2}\right)$$

$$O\left((1 - \frac{\mu}{L})^k\right) \quad \eta = \frac{1}{L}$$

$$O\left((1 - \sqrt{\frac{\mu}{L}})^k\right) \quad \eta = \frac{1}{L}$$

$\min_x f(x) + g(x)$ $\partial f \quad \partial g$

Generalized Douglas-Rachford $\lambda \in (0, 2)$

Peaceman-Rachford $\lambda = 2$ divergent

$$O\left(\frac{1}{k}\right)$$

$$O\left(\frac{1}{k}\right)$$

$$\forall \eta > 0$$

$$O\left(\frac{1}{k}\right)$$

$$\forall \eta > 0$$

$\min_x f(x) + g(x)$ $\partial f \quad \nabla g$

Generalized Douglas-Rachford $\lambda \in (0, 2]$

$$O\left(\frac{1}{k}\right)$$

$$O\left(\left[1 - \frac{\lambda}{2} + \frac{\lambda}{2}c\right]^{2k}\right)$$

$$c = \sqrt{1 - \frac{4\mu\eta}{(1+\eta L)^2}} \quad \forall \eta > 0$$