Previously on nonsmooth case:

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

$g(x)$ is differentiable

$$\min_x f(x)$$

Example: $f(x) = \|x\|_1$

① Subgradient Method

$$x_{k+1} = x_k - \eta \partial f(x_k)$$

② Proximal Point Method

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

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

Example: $\|x\|_1 + \frac{\lambda}{2}\|Ax-b\|^2$

③ Proximal Gradient Method

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

④ Fast Proximal Gradient

Plan

$$\min_x f(Ax) + g(x)$$

Example: ROF (Rudin, Osher, Fatemi 1992) model

1) 1D signal

$$\min_{u \in \mathbb{R}^n} \|u\|_{TV} + \frac{\lambda}{2}\|u-d\|^2$$

$$\|Du\|_1 + \frac{\lambda}{2}\|u-d\|^2$$

$$\boxed{\begin{aligned} \|u\|_{TV} &= \sum_i |u_{i+1} - u_i| \\ &= \|Du\|_1 \end{aligned} \qquad Du = \underbrace{\begin{bmatrix} -1 & 1 & & & \\ & -1 & 1 & & \\ & & \ddots & \ddots & \\ & & & -1 & 1 \end{bmatrix}}_{(n-1)\times n} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}}$$

2) 2D image $\min_{u \in \mathbb{R}^{n\times n}} \|u\|_{TV} + \frac{\lambda}{2}\|u-d\|^2$

$$\|u\|_{TV} = \sum_{i,j} \sqrt{|u_{i+1,j} - u_{i,j}|^2 + |u_{i,j+1} - u_{i,j}|^2}$$

$$Au = (Du, uD^T)$$

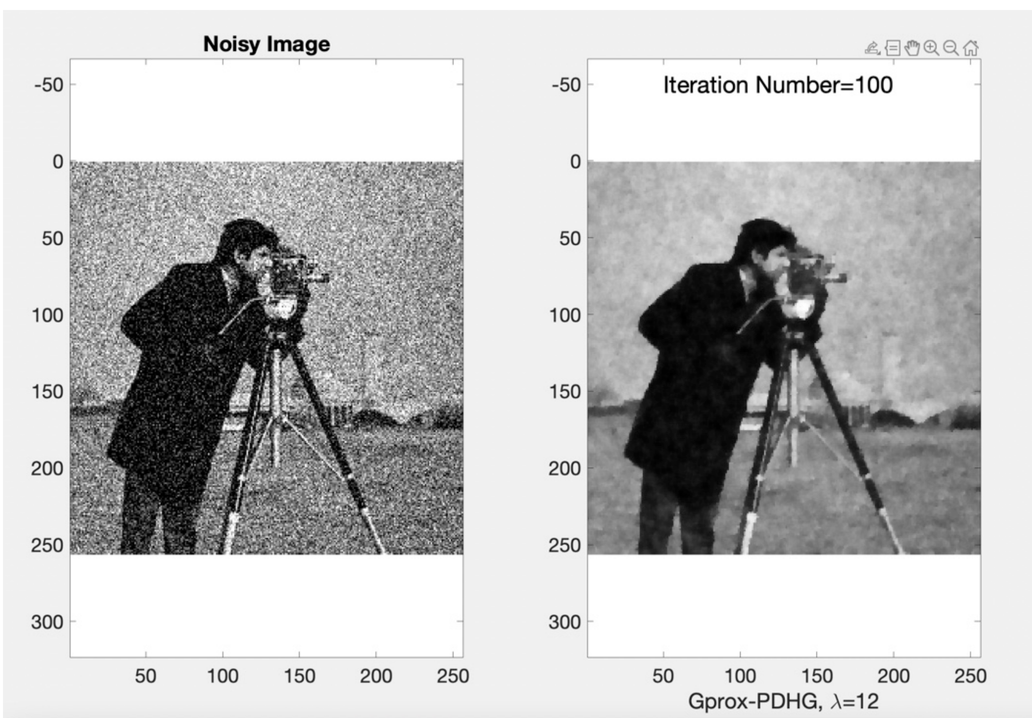
$$\vec{w} = (U, V), \quad U \in \mathbb{R}^{n\times n}, \; V \in \mathbb{R}^{n\times n}$$

$$\|\vec{w}\|_1 = \sum_{i,j} \sqrt{U_{ij}^2 + V_{ij}^2}$$

Algorithms for $\min_u \|Du\|_1 + \frac{\lambda}{2}\|u-d\|^2$

① Primal Dual Hybrid Gradient (PDHG)

② Fast PDHG (Chambolle & Pock 2011)



$$\min_x f(x) + g(x) \qquad f \,\&\, g \text{ are both nonsmooth}$$

Example: $\min_x \|x\|_1 + \iota_{\{x : Ax = b\}}$

③ ADMM (1975)

④ Douglas-Rachford splitting (1979)

<u>Theorem</u> If $f(x)$ & $g(x)$ are convex, the Douglas-Rachford splitting iteration converges.

$$\begin{cases} y_{k+1} = \dfrac{I + R_f R_g}{2}(y_k) \\ R_f = 2 \cdot \mathrm{Prox}_f^{\eta} - I \\ R_g = 2 \cdot \mathrm{Prox}_g^{\eta} - I \end{cases}$$

Proof:

Theorem

The following are equivalent:

① T is firmly nonexpansive

② $I - T$ is firmly nonexpansive

③ $2T - I$ is nonexpansive

$f(x)$ is convex $\Rightarrow$ Prox_f^{η} is firmly nonexpansive

$\Leftrightarrow R_f = 2\,\mathrm{Prox}_f^{\eta} - I$ is nonexpansive

$\Rightarrow T = R_f R_g$ is nonexpansive

$\Rightarrow Y_{k+1} = \left[(1-\theta)I + \theta T\right](Y_k)$ converges

$\theta \in (0,1)$

$\theta = \frac{1}{2}$ gives DR Splitting

$S = \frac{1}{2}I + \frac{1}{2}R_f R_g$ is firmly nonexpansive

$\Updownarrow$

$2S - I = R_f R_g$ is nonexpansive

Developments of Douglas-Rachford Splitting

$u_t = u_{xx} + u_{yy}$

1. Peaceman and Rachford, 1955; Douglas and Rachford, 1956: implicit finite difference in solving heat equations. $\longrightarrow$ ADI
2. Lions and Mercier, 1979: extension to maximal monotone operators; DR is firmly nonexpansive thus convergent.
3. Glowinski and Marroco, 1975; Gabay, 1983: the alternating direction method of multipliers (ADMM) is equivalent to DR. ADMM has been widely used in nonlinear mechanics and convex optimization.
4. Goldstein and Osher, 2009: split Bregman method widely used in image processing problems, which is the same as ADMM.
5. Bauschke, Combettes, and Luke, 2002: the widely used Fienup's Hybrid Input-Output (1982) algorithm for phase retrieval problem can be viewed as Douglas-Rachford splitting.

The following three algorithms are exactly the same:

- Douglas-Rachford Splitting on $\min_x f(x) + g(x)$.
- ADMM on Fenchel Dual $\min_y f^*(y) + g^*(-y)$.
- Split Bregman on Fenchel Dual $\min_y f^*(y) + g^*(-y)$.

③ Generalized DR (1992)

$$y_{k+1} = \left[(1-\lambda)I + \lambda \frac{I + R_f R_g}{2}\right](y_k)$$

$$\lambda \in (0,2)$$

Theorem Generalized DR converges for any convex $f(x)$ & $g(x)$.

Proof: $$T = \frac{I + R_f R_g}{2}$$

T is firmly nonexpansive

$\Updownarrow$

$2T - I$ is nonexpansive

$$S = (1-\lambda)I + \lambda T = (1-2\theta)I + 2\theta T, \quad \theta \in (0,1)$$

$$= (1-\theta)\cdot I + \theta \cdot (2T - I)$$