## Convex Conjugate (Fenchel Dual Function)

$$f^*(x) = \max_y \langle x, y\rangle - f(y)$$

$f(x)$ is convex $\Rightarrow (f^*)^* = f$

Primal Problem (P)

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

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

$$\Rightarrow \begin{cases} f^*(y) \text{ is convex} \\ g^*(-y) \text{ is convex} \end{cases}$$

$\Leftrightarrow$ Dual Problem (D)

$$-\min_y f^*(y) + g^*(-y)$$

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

$$f^*(x) = \begin{cases} 0, & \|x\|_\infty \leq 1 \\ +\infty, & \|x\|_\infty > 1 \end{cases}$$

$\Leftrightarrow$ Primal Dual (PD)

$$\min_x \sup_y \langle x, y\rangle - f^*(y) + g(x)$$

Primal Dual Minizers Relation

$$\begin{cases} x^* \in \partial f^*(y^*) \\ y^* \in -\partial g(x^*) \end{cases}$$

---

In (PD), the cost function is

$$L(x,y) = \langle x, y\rangle - f^*(y) + g(x)$$

$$\frac{\partial L}{\partial x}(x,y) = y + \partial g(x)$$

$$\frac{\partial L}{\partial y}(x,y) = x - \partial f^*(y)$$

$$\boxed{\begin{array}{l}\text{Primal Dual (PD)} \\ \min_x \sup_y \langle x, y\rangle - f^*(y) + g(x)\end{array}}$$

For $\min_x \max_y L(x,y)$, a simple algorithm is to use gradient descent/ascent for approaching saddle point

$$\begin{cases} \dfrac{x_{k+1}-x_k}{\eta} = -\dfrac{\partial L}{\partial x}(x_{k+1}, y_k) \\[2ex] \dfrac{y_{k+1}-y_k}{\tau} = \dfrac{\partial L}{\partial y}(x_{k+1}, y_{k+1}) \end{cases}$$

$$\Leftrightarrow \begin{cases} x_{k+1} = x_k - \eta y_k - \eta \partial g(x_{k+1}) \\ y_{k+1} = y_k + \tau x_{k+1} - \tau \partial f^*(y_{k+1}) \end{cases}$$

$\Leftrightarrow$ ① Arrow-Hurwitz Algorithm ( 1958)

$$\begin{cases} x_{k+1} = (I+\eta\partial g)^{-1}[x_k - \eta y_k] \\ y_{k+1} = (I+\tau\partial f^*)^{-1}[y_k + \tau x_{k+1}] \end{cases} \quad \eta>0, \tau>0$$

② Primal-Dual-Hybrid-Gradient (PDHG, 2010)

$$\begin{cases} x_{k+1} = (I+\eta\partial g)^{-1}[x_k - \eta y_k] \\ y_{k+1} = (I+\tau\partial f^*)^{-1}[y_k + \tau(2x_{k+1} - x_k)] \end{cases} \quad \eta>0, \tau>0 \quad \eta\tau\le 1$$

$$y_{k+1} = \mathrm{Prox}_{f^*}^{1/\eta}\left[\tfrac{1}{\eta}(\eta y_k + 2x_{k+1} - x_k)\right]$$

③ PDHG with $\tau = \frac{1}{\eta}$

$\boxed{\text{Morean-Decomposition}}$ $f(x)$ is convex

$$\mathrm{Prox}_f^{\eta}(x) + \eta\, \mathrm{Prox}_{f^*}^{1/\eta}(x/\eta) = x$$

Let $V_k = x_k - \eta y_k \Leftrightarrow y_k = \dfrac{x_k - V_k}{\eta}$

$$\eta\,(I+\tfrac{1}{\eta}\partial f^*)^{-1}\left(\eta\left[y_k + \tfrac{1}{\eta}(2x_{k+1}-x_k)\right]/\eta\right)$$

$$= \eta\left[y_k + \frac{1}{\eta}(2x_{k+1} - x_k)\right] - \mathrm{Prox}_f^{\eta}(2x_{k+1} - x_k + \eta y_k)$$

$$= 2x_{k+1} - (x_k - \eta y_k) - \mathrm{Prox}_f^{\eta}(2x_{k+1} - (x_k - \eta y_k))$$

$$\begin{cases} x_{k+1} = (I + \eta\partial g)^{-1}[x_k - \eta y_k] \\ y_{k+1} = (I + \frac{1}{\eta}\partial f^*)^{-1}\left[y_k + \frac{1}{\eta}(2x_{k+1} - x_k)\right] \end{cases}$$

$$\Rightarrow \begin{cases} x_{k+1} = \mathrm{Prox}_g^{\eta}(v_k) \\ y_{k+1} = \dfrac{2x_{k+1} - v_k - \mathrm{Prox}_f^{\eta}(2x_{k+1} - v_k)}{\eta} \end{cases}$$

$$\Rightarrow \begin{cases} x_{k+1} = \mathrm{Prox}_g^{\eta}(v_k) \\ \mathrm{Prox}_f^{\eta}(2x_{k+1} - v_k) = x_{k+1} - v_k + (x_{k+1} - \eta y_{k+1}) \end{cases}$$

$$\Rightarrow \begin{cases} x_{k+1} = \mathrm{Prox}_g^{\eta}(v_k) \quad \boxed{v_k = x_k - \eta y_k} \\ v_{k+1} = v_k - x_{k+1} + \mathrm{Prox}_f^{\eta}(2x_{k+1} - v_k) \end{cases}$$

$$\frac{I + R_f R_g}{2}[v_k] = \frac{v_k + R_f[R_g(v_k)]}{2} = \frac{1}{2}v_k + \frac{1}{2}R_f[2x_{k+1} - v_k]$$

$$= \frac{1}{2}v_k + \frac{1}{2}\left[2\mathrm{Prox}_f^{\eta}(2x_{k+1} - v_k) - (2x_{k+1} - v_k)\right]$$

So PDHG ($\eta\tau \le 1$) with $\tau = \frac{1}{\eta}$

is Douglas-Rachford on (P)

thus converges with $\tau = \frac{1}{\eta}$, $\forall \eta > 0$.

④ ADMM (Alternating Direction Method of Multipliers)

since 1975

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

$$\Longleftrightarrow \min f(w) + g(z) \quad \text{s.t. } w = z$$

The Lagrangian for constrained minimization:

$$L(w, z, y) = f(w) + g(z) - \langle \underbrace{y}_{\text{Lagrangian multiplier}}, w - z \rangle$$

| |
|---|
| If $(w_*, z_*, y_*)$ is a solution (saddle point) to $\min_{w,z} \max_y L(w,z,y)$ then $w_* = z_*$ is a solution to $\min_{w,z} f(w) + g(z)$ s.t. $w = z$ |

Augmented Lagrangian is

$$\mathcal{L}(w, z, y) = f(w) + g(z) - \langle y, w - z \rangle + \frac{\tau}{2} \|w - z\|^2$$

The saddle point of Augmented Lagrangian still gives minimizer to (P)

ADMM :

$$\min_{w,z} \max_{y} \mathcal{L}(w,z,y)$$

$$\begin{cases} z_{k+1} = \arg\min_{z} \mathcal{L}(w_k, z, y_k) \\ w_{k+1} = \arg\min_{w} \mathcal{L}(w, z_{k+1}, y_k) \\ y_{k+1} = y_k + \sigma \frac{\partial}{\partial y}\mathcal{L}(w_{k+1}, z_{k+1}, y_k) \end{cases}$$

$$\mathcal{L}(w,z,y) = f(w) + g(z) - \langle y, w-z\rangle + \frac{\tau}{2}\|w-z\|^2$$

$$\Leftrightarrow \begin{cases} z_{k+1} = \arg\min_{z} \; g(z) - \langle y_k, w_k - z\rangle + \frac{\tau}{2}\|w_k - z\|^2 \\ w_{k+1} = \arg\min_{w} \; f(w) - \langle y_k, w - z_{k+1}\rangle + \frac{\tau}{2}\|w - z_{k+1}\|^2 \\ y_{k+1} = y_k - \sigma(w_{k+1} - z_{k+1}) \end{cases}$$

$$\Leftrightarrow \begin{cases} z_{k+1} = \arg\min_{z} \; g(z) + \langle y_k, z\rangle + \frac{\tau}{2}\|z\|^2 - \tau\langle w_k, z\rangle \\ w_{k+1} = \arg\min_{w} \; f(w) - \langle y_k, w\rangle + \frac{\tau}{2}\|w\|^2 - \tau\langle w, z_{k+1}\rangle \\ y_{k+1} = y_k - \sigma(w_{k+1} - z_{k+1}) \end{cases}$$

Let $\sigma = \tau$ $\qquad \frac{v_{k+1} - y_k}{\eta} = z_{k+1}$ & $\frac{v_k - y_k}{\eta} = w_k$

$$\Leftrightarrow \begin{cases} z_{k+1} = \arg\min_{z} \; g(z) + \frac{\tau}{2}\|z - w_k + \frac{1}{\tau}y_k\|^2 \\ w_{k+1} = \arg\min_{w} \; f(w) + \frac{\tau}{2}\|w - z_{k+1} - \frac{1}{\tau}y_k\|^2 \\ y_{k+1} = y_k - \tau(w_{k+1} - z_{k+1}) \quad \text{(ADMM)} \end{cases}$$

Claim : (ADMM) $\Leftrightarrow$ Douglas-Rachford on (D)

$$\text{(ADMM)} : \begin{cases} z_{k+1} = \arg\min_{z} \; g(z) + \frac{\tau}{2}\|z - (w_k - \frac{1}{\tau}y_k)\|^2 \\ w_{k+1} = \arg\min_{w} \; f(w) + \frac{\tau}{2}\|w - (z_{k+1} + \frac{1}{\tau}y_k)\|^2 \\ y_{k+1} = y_k - \tau(w_{k+1} - z_{k+1}) \end{cases}$$

$$\iff \begin{cases} z_{k+1} = \mathrm{Prox}_g^{1/\tau}\left(w_k - \frac{1}{\tau} y_k\right) \\ w_{k+1} = \mathrm{Prox}_f^{1/\tau}\left(z_{k+1} + \frac{1}{\tau} y_k\right) \\ y_{k+1} = y_k - \tau(w_{k+1} - z_{k+1}) \end{cases}$$

Consider $\min_y \underbrace{f^*(y)}_{G(y)} + \underbrace{g^*(-y)}_{F(y)}$ $\quad \frac{I + R_F R_G}{2}$

**Morean Decomposition**

$$\boxed{\mathrm{Prox}_{f^*}^{\eta}(x) + \eta\, \mathrm{Prox}_f^{1/\eta}\left(\frac{x}{\eta}\right) = x}$$

$$\mathrm{Prox}_G^{\eta}(x) = \mathrm{Prox}_{f^*}^{\eta}(x) = x - \eta\, \mathrm{Prox}_f^{1/\eta}\left(\frac{x}{\eta}\right)$$

$F(x) = g^*(-x)$

$$\mathrm{Prox}_F^{\eta}(x) = \arg\min_u \left[ g^*(-u) + \frac{1}{2\eta}\|u - x\|^2 \right]$$

$u = -v$

$u^* = -v^* \Leftarrow$

$$= -\arg\min_v \left[ g^*(v) + \frac{1}{2\eta}\|-v - x\|^2 \right]$$

$$= -\arg\min_v \left[ g^*(v) + \frac{1}{2\eta}\|v - (-x)\|^2 \right]$$

$$= -\mathrm{Prox}_{g^*}^{\eta}(-x)$$

$$= \eta\, \mathrm{Prox}_g^{1/\eta}\left(\frac{-x}{\eta}\right) - (-x)$$

$$= \eta\, \mathrm{Prox}_g^{1/\eta}\left(-\frac{x}{\eta}\right) + x$$

DR:

$$\begin{cases} v_{k+1} = \frac{I + R_F R_G}{2}(v_k) \\ y_k = \mathrm{Prox}_G(v_k) \end{cases}$$

$$\begin{cases} y = \mathrm{Prox}^{\eta}_{G}(v) = v - \eta\, \mathrm{Prox}^{1/\eta}_{f}(v/\eta) \\ \dfrac{I + R_F R_G}{2}(v) = v - y + \mathrm{Prox}_F[2y - v] \end{cases}$$

$$= v - y + (2y - v) + \eta\, \mathrm{Prox}^{1/\eta}_{g}\left(-\frac{2y}{\eta} + \frac{v}{\eta}\right)$$

$$= y + \eta\, \mathrm{Prox}^{1/\eta}_{g}\left(-\frac{2y}{\eta} + \frac{v}{\eta}\right)$$

$$\begin{cases} v_{k+1} = y_k + \eta\, \mathrm{Prox}^{1/\eta}_{g}\left(-\dfrac{2y_k}{\eta} + \dfrac{v_k}{\eta}\right) \\ y_{k+1} = v_{k+1} - \eta\, \mathrm{Prox}^{1/\eta}_{f}(v_{k+1}/\eta) \end{cases}$$

$y_k \to y_*$

$\min_y f^*(y) + g^*(-y)$

Let $\dfrac{v_{k+1} - y_k}{\eta} = z_{k+1}$ & $\dfrac{v_k - y_k}{\eta} = w_k$

Then
$$\begin{cases} z_{k+1} = \mathrm{Prox}^{1/\eta}_{g}\left(w_k - \frac{1}{\eta} y_k\right) \\ w_{k+1} = \mathrm{Prox}^{1/\eta}_{f}\left(z_{k+1} + \frac{1}{\eta} y_k\right) \\ y_{k+1} = y_k - \tau(z_{k+1} - w_{k+1}) \end{cases}$$

DR

$\min_y \underbrace{f^*(y)}_{G(y)} + \underbrace{g^*(-y)}_{F(y)}$

$(I + R_F R_G)/2$

$\Updownarrow$

(ADMM)

$\Updownarrow$

$$\begin{cases} z_{k+1} = \mathrm{Prox}^{1/\tau}_{g}\left(w_k - \frac{1}{\tau} y_k\right) \\ w_{k+1} = \mathrm{Prox}^{1/\tau}_{f}\left(z_{k+1} + \frac{1}{\tau} y_k\right) \\ y_{k+1} = y_k - \tau(w_{k+1} - z_{k+1}) \end{cases}$$

$$\begin{cases} w_{k+1} = \mathrm{Prox}_f^{1/\tau}\left(z_k + \frac{1}{\tau} y_k\right) \\ z_{k+1} = \mathrm{Prox}_g^{1/\tau}\left(w_{k+1} - \frac{1}{\tau} y_k\right) \\ y_{k+1} = y_k + \tau (w_{k+1} - z_{k+1}) \end{cases}$$

(ADMM 2)

$\Updownarrow$

$$\min_y \underbrace{f^*(y)}_{F(y)} + \underbrace{g^*(-y)}_{G(y)}$$

$(I + R_F R_G)/2$

$\Updownarrow$

Conclusion: Convergence of Douglas-Rachford

$\Rightarrow$ Convergence of $\begin{cases} \text{1) ADMM} \\ \text{2) PDHG with } \tau = \frac{1}{\eta} \end{cases}$

## Split Bregman (2009)

$$\min f(\mathbf{w}) + g(\mathbf{z}), \quad \mathbf{w} = \mathbf{z}.$$

Consider an unconstrained problem

$$\min_{\mathbf{w},\mathbf{z}} f(\mathbf{w}) + g(\mathbf{z}) + \frac{\tau}{2}\|\mathbf{w} - \mathbf{z}\|^2,$$

for which the Bregman iteration is given as

$$\text{(Split Bregman)}: \quad \begin{aligned} \mathbf{z}_{k+1} &= \operatorname*{argmin}_{\mathbf{z}} g(\mathbf{z}) + \frac{\tau}{2}\|\mathbf{w}_k - \mathbf{z} - \mathbf{u}_k\|^2 \\ \mathbf{w}_{k+1} &= \operatorname*{argmin}_{\mathbf{w}} f(\mathbf{w}) + \frac{\tau}{2}\|\mathbf{w} - \mathbf{z}_{k+1} - \mathbf{u}_k\|^2 \\ \mathbf{u}_{k+1} &= \mathbf{u}_k + (\mathbf{w}_{k+1} - \mathbf{z}_{k+1}) \end{aligned}$$

Plug in $\mathbf{u} = \frac{1}{\tau}\mathbf{y}$, then it can be written as

$$\text{(Split Bregman)}: \quad \begin{aligned} \mathbf{z}_{k+1} &= \operatorname*{argmin}_{\mathbf{z}} g(\mathbf{z}) - \langle \mathbf{y}_k, \mathbf{w}_k - \mathbf{z}\rangle + \frac{\tau}{2}\|\mathbf{w}_k - \mathbf{z}\|^2 \\ \mathbf{w}_{k+1} &= \operatorname*{argmin}_{\mathbf{w}} f(\mathbf{w}) - \langle \mathbf{y}_k, \mathbf{w} - \mathbf{z}_{k+1}\rangle + \frac{\tau}{2}\|\mathbf{w} - \mathbf{z}_{k+1}\|^2 \\ \mathbf{y}_{k+1} &= \mathbf{y}_k + \tau(\mathbf{w}_{k+1} - \mathbf{z}_{k+1}) \end{aligned}$$

So this version of split Bregman method is equivalent to ADMM with $\sigma = \tau$, thus also equivalent to the DR spliting on the dual.

$$\begin{cases} \text{Primal Problem (P)} \\ \min_{x} f(x) + g(x) \end{cases}$$

$$\Leftrightarrow \begin{cases} \text{Dual Problem (D)} \\ -\min_{y} f^*(y) + g^*(-y) \end{cases}$$

$$\boxed{\begin{array}{c}\text{PDHG}\\ \text{with } \tau = \frac{1}{\eta}\end{array}} \Longleftrightarrow \boxed{\begin{array}{c}\text{DR on (P)}\\ \text{with } \eta\end{array}} \Longleftrightarrow \boxed{\begin{array}{c}\text{ADMM}\\ \text{on (D)}\\ \text{with } \tau=\sigma=\eta\end{array}} \Longleftrightarrow \boxed{\begin{array}{c}\text{Split Bregman}\\ \text{on (D)}\\ \tau=\eta\end{array}}$$

$$\Big\Updownarrow$$

$$\boxed{\begin{array}{c}\text{DR on (D)}\\ \text{with } \frac{1}{\eta}>0\end{array}} \Longleftrightarrow \boxed{\begin{array}{c}\text{ADMM}\\ \text{on (P)}\\ \text{with } \tau=\sigma=\frac{1}{\eta}\end{array}} \Longleftrightarrow \boxed{\begin{array}{c}\text{Split Bregman}\\ \text{on (P)}\\ \tau=\frac{1}{\eta}\end{array}}$$

Excercise:

For $\min_x \underbrace{\|x\|_1}_{f(x)} + \underbrace{\iota_{\{x:Ax=b\}}(x) + \frac{1}{2\alpha}\|x\|_2^2}_{g(x)}$

how many different algorithms we have learned?

Which one is the best?