

## Prox of quadratic function

$$\arg \min_x \frac{1}{2} \langle x, Ax \rangle + \langle b, x \rangle + c + \frac{1}{2\tau} \|x - v\|^2 \quad (1)$$

$$\Leftrightarrow 0 = Ax + b + \frac{x - v}{\tau} \quad (2)$$

$$\Leftrightarrow 0 = \tau Ax + \tau b + x - v \quad (3)$$

$$\Leftrightarrow 0 = (I + \tau A)x + \tau b - v \quad (4)$$

$$\Leftrightarrow v - \tau b = (I + \tau A)x \quad (5)$$

$$\Leftrightarrow x = (I + \tau A)^{-1}(v - \tau b) \quad (6)$$

## Moreau decomposition

$$\begin{aligned} u &= \text{prox}_E(v) \\ \Leftrightarrow u &= \arg \min_w E(w) + \frac{1}{2} \|w - v\|^2 \\ \Leftrightarrow 0 &\in \partial E(u) + u - v \\ \Leftrightarrow v - u &\in \partial E(u) \\ \Leftrightarrow u &\in \partial E^*(v - u) \\ \Leftrightarrow 0 &\in -u + \partial E^*(v - u) \\ \Leftrightarrow v - u &= \arg \min_u E^*(u) + \frac{1}{2} \|u - v\|^2 \end{aligned}$$

## Extension of subspace decomposition

Let

$$f(x) = \begin{cases} 0, & \text{if } x \in V, \\ \infty, & \text{otherwise.} \end{cases}$$

Then the convex conjugate is given as

$$f^*(y) = \sup_{x \in V} \langle y, x \rangle = \sup_{x \in V} \langle y^1 + y^2, x \rangle = \sup_{x \in V} \langle y^1, x \rangle + \langle y^2, x \rangle = \begin{cases} 0, & \text{if } y \in V^\perp, \\ \infty, & \text{otherwise.} \end{cases}$$

In the above calculation we split up  $y = y^1 + y^2$  in  $y^1 \in V$  and  $y^2 \in V^\perp$ . The supremum evaluates to infinity since  $V$  is a subspace ( $x \in V \Rightarrow \alpha x \in V$ ).

Hence  $f(x) = \iota_V(x)$  and  $f^*(y) = \iota_{V^\perp}(y)$  and the Moreau decomposition states that

$$x = \text{prox}_f(x) + \text{prox}_{f^*}(x) = \Pi_V(x) + \Pi_{V^\perp}(x).$$

## Extended Moreau decomposition

We apply the standard Moreau decomposition to the function  $\tau E$ :

$$u = \text{prox}_{\tau E}(u) + \text{prox}_{(\tau E)^*}(u)$$

Due to the scaling rule from last lecture

$$(\tau E)^*(u) = \tau E^*\left(\frac{u}{\tau}\right),$$

we have with  $w = \frac{u}{\tau} \Leftrightarrow u = \tau w$

$$u^* = \arg \min_u \tau E^*\left(\frac{u}{\tau}\right) + \frac{1}{2} \|u - v\|^2 \xrightarrow{\text{substitution}} \quad (7)$$

$$u^* = \tau \arg \min_w \tau E^*(w) + \frac{1}{2} \|\tau w - v\|^2 = \quad (8)$$

$$u^* = \tau \arg \min_w \tau E^*(w) + \frac{\tau^2}{2} \|w - \frac{v}{\tau}\|^2 = \quad (9)$$

$$u^* = \tau \arg \min_w E^*(w) + \frac{\tau}{2} \|w - \frac{v}{\tau}\|^2. \quad (10)$$

$$\Rightarrow u^* = \tau \text{prox}_{\frac{1}{\tau} E^*}\left(\frac{v}{\tau}\right). \quad (11)$$

From that it follows that

$$u = \text{prox}_{\tau E}(u) + \tau \text{prox}_{\frac{1}{\tau} E^*}\left(\frac{v}{\tau}\right).$$

### Prox of $\ell_2$ -norm

We already know that for

$$E(u) = \|u\|,$$

we have that

$$E^*(u) = \begin{cases} 0 & \text{if } \|u\| \leq 1 \\ \infty & \text{otherwise.} \end{cases}$$

Then the proximal operator of  $E^*$  is simply the projection onto the unit ball

$$\text{prox}_{E^*}(v) = \begin{cases} v/\|v\| & \text{if } \|v\| \geq 1 \\ v & \text{otherwise.} \end{cases}$$

Then using the Moreau decomposition we have

$$\text{prox}_{\tau E}(v) = v - \tau \text{prox}_{\frac{1}{\tau} E^*}\left(\frac{v}{\tau}\right) \quad (12)$$

$$= v - \tau \begin{cases} v/\|v\| & \text{if } \|v\| \geq \tau \\ v/\tau & \text{otherwise.} \end{cases} \quad (13)$$

$$= \begin{cases} v - \tau v/\|v\| & \text{if } \|v\| \geq \tau \\ 0 & \text{otherwise.} \end{cases} \quad (14)$$

$$= \begin{cases} (1 - \tau/\|v\|)v & \text{if } \|v\| \geq \tau \\ 0 & \text{otherwise.} \end{cases} \quad (15)$$

### MM interpretation

$$\text{prox}_{(1/L)G}(u^k - (1/L)\nabla F(u^k)) \quad (16)$$

$$= \arg \min_u G(u) + \frac{L}{2} \|u - u^k + (1/L)\nabla F(u^k)\|^2 \quad (17)$$

$$= \arg \min_u G(u) + \frac{L}{2} (\|u - u^k\|^2 + 2\langle u - u^k, (1/L)\nabla F(u^k) \rangle + \|(1/L)\nabla F(u^k)\|^2) \quad (18)$$

$$= \arg \min_u G(u) + \langle \nabla F(u^k), u - u^k \rangle + \frac{L}{2} \|u - u^k\|^2 \quad (19)$$

## Nonexpansiveness of prox

Let  $x = \text{prox}_E(u)$  and  $y = \text{prox}_E(v)$ . Then we want to show

$$\langle u - v, x - y \rangle \geq \|x - y\|^2. \quad (20)$$

Since  $x = \arg \min_z E(z) + \frac{1}{2}\|z - u\|^2$  and  $y = \arg \min_z E(z) + \frac{1}{2}\|z - v\|^2$  we have that

$$u - x \in \partial E(x) \quad (21)$$

$$v - y \in \partial E(y) \quad (22)$$

From that it follows

$$E(z) - E(x) \geq \langle u - x, z - x \rangle, \forall z \quad (23)$$

$$E(z) - E(y) \geq \langle v - y, z - y \rangle, \forall z \quad (24)$$

With special choice of  $z$  it follows

$$E(y) - E(x) \geq \langle u - x, y - x \rangle, \quad (25)$$

$$E(x) - E(y) \geq \langle v - y, x - y \rangle, \quad (26)$$

and adding these two inequalites gives

$$0 \geq \langle u - x, y - x \rangle + \langle v - y, x - y \rangle = \langle v - y + x - u, x - y \rangle, \quad (27)$$

And hence

$$\langle y - v + u - x, x - y \rangle \geq 0 \quad (28)$$

$$\Leftrightarrow \langle u - v, x - y \rangle \geq \|x - y\|^2 \quad (29)$$

$$\Leftrightarrow \langle u - v, \text{prox}_E(u) - \text{prox}_E(v) \rangle \geq \|\text{prox}_E(u) - \text{prox}_E(v)\|^2 \quad (30)$$

With Cauchy-Schwarz it follows

$$\|u - v\| \|\text{prox}_E(u) - \text{prox}_E(v)\| \geq \langle u - v, \text{prox}_E(u) - \text{prox}_E(v) \rangle \geq \|\text{prox}_E(u) - \text{prox}_E(v)\|^2 \quad (31)$$

$$\|u - v\| \geq \|\text{prox}_E(u) - \text{prox}_E(v)\| \quad (32)$$

## Convergence of Proximal gradient method

### Gradient map in subdifferential

Since  $x = \text{prox}_E(u) \Rightarrow u - x \in \partial E(u)$  we have for  $u - \tau \varphi_\tau(u) = \text{prox}_{\tau G}(u - \tau \nabla F(u))$  that:

$$u - \tau \nabla F(u) - (u - \tau \varphi_\tau(u)) \in \tau \partial G(u - \tau \nabla F(u)) \quad (33)$$

$$\Leftrightarrow \varphi_\tau(u) \in \nabla F(u) + \partial G(u - \tau \nabla F(u)) \quad (34)$$

### Global estimate

Define  $p = \varphi_\tau(u) - \nabla F(u) \in \partial G(u - \tau\varphi_\tau(u))$ :

$$E(u - \tau\varphi_\tau(u)) \leq F(u) - \tau\langle \nabla F(u), \varphi_\tau(u) \rangle + \frac{\tau}{2}\|\varphi_\tau(u)\|^2 + G(u - \tau\varphi_\tau(u)) \quad (35)$$

$$\stackrel{\text{convexity } (*), (**)}{\leq} F(w) + \langle \nabla F(u), u - w \rangle - \tau\langle \nabla F(u), \varphi_\tau(u) \rangle + \frac{\tau}{2}\|\varphi_\tau(u)\|^2 \quad (36)$$

$$+ G(w) + \langle p, u - w - \tau\varphi_\tau(u) \rangle \quad (37)$$

$$= F(w) + \langle \nabla F(u), u - w \rangle - \tau\langle \nabla F(u), \varphi_\tau(u) \rangle + \frac{\tau}{2}\|\varphi_\tau(u)\|^2 \quad (38)$$

$$+ G(w) + \langle \varphi_\tau(u) - \nabla F(u), u - w - \tau\varphi_\tau(u) \rangle \quad (39)$$

$$= F(w) + G(w) + \langle \varphi_\tau(u), u - w \rangle - \frac{\tau}{2}\|\varphi_\tau(u)\|^2 \quad (40)$$

$$= E(w) + \langle \varphi_\tau(u), u - w \rangle - \frac{\tau}{2}\|\varphi_\tau(u)\|^2 \quad (41)$$

$$(42)$$

$$(*) \quad F(w) \geq F(u) - \langle \nabla F(u), u - w \rangle, \forall w$$

$$(**) \quad G(w) \geq G(u - \tau\varphi_\tau(u)) - \langle p, u - \tau\varphi_\tau(u) - w \rangle, p \in \partial G(u - \tau\varphi_\tau(u))$$

**Choice**  $w = u^*$

$$E(u^+) - E(u^*) \leq \langle \varphi_\tau(u), u - u^* \rangle - \frac{\tau}{2}\|\varphi_\tau(u)\|^2 \quad (43)$$

$$= \frac{1}{2\tau} (2\tau\langle \varphi_\tau(u), u - u^* \rangle - \tau^2\|\varphi_\tau(u)\|^2) \quad (44)$$

$$= \frac{1}{2\tau} (\|u - u^*\|^2 - \|u - u^*\|^2 + 2\tau\langle \varphi_\tau(u), u - u^* \rangle - \|\tau\varphi_\tau(u)\|^2) \quad (45)$$

$$= \frac{1}{2\tau} (\|u - u^*\|^2 - (\|u - u^*\|^2 - 2\tau\langle \varphi_\tau(u), u - u^* \rangle + \|\tau\varphi_\tau(u)\|^2)) \quad (46)$$

$$= \frac{1}{2\tau} (\|u - u^*\|^2 - \|u - u^* - \tau\varphi_\tau(u)\|^2) \quad (47)$$

$$= \frac{1}{2\tau} (\|u - u^*\|^2 - \|u^+ - u^*\|^2) \quad (48)$$

### Final convergence estimate

Summing the above inequalities with  $u^+ = u^i$ ,  $u = u^{i-1}$  yields

$$\sum_{i=1}^k E(u^i) - E(u^*) \leq \frac{1}{2\tau} \sum_{i=1}^k (\|u - u^*\|^2 - \|u^+ - u^*\|^2) \quad (49)$$

$$= \frac{1}{2\tau} (\|u^0 - u^*\|^2 - \|u^k - u^*\|^2) \quad (50)$$

$$\leq \frac{1}{2\tau} \|u^0 - u^*\|^2 \quad (51)$$

Since the energy is non-increasing

$$E(u^k) - E(u^*) \leq \frac{1}{k} \sum_{i=1}^k E(u^i) - E(u^*) \leq \frac{1}{2\tau k} \|u^0 - u^*\|^2.$$