Prox of quadratic function

$$\begin{aligned}
&\arg\min_x \ \frac{1}{2}\langle x, Ax\rangle + \langle b, x\rangle + c + \frac{1}{2\tau}\|x - v\|^2 && (1)\\
\Leftrightarrow \quad & 0 = Ax + b + \frac{x - v}{\tau} && (2)\\
\Leftrightarrow \quad & 0 = \tau Ax + \tau b + x - v && (3)\\
\Leftrightarrow \quad & 0 = (I + \tau A)x + \tau b - v && (4)\\
\Leftrightarrow \quad & v - \tau b = (I + \tau A)x && (5)\\
\Leftrightarrow \quad & x = (I + \tau A)^{-1}(v - \tau b) && (6)
\end{aligned}$$

Moreau decomposition

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

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

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

Extended Moreau decomposition

We apply the standard Moreau decomposition to the function τE:

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

Due to the scaling rule from last lecture

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

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

$$u^* = \arg\min_u \; \tau E^*(\frac{u}{\tau}) + \frac{1}{2}\|u - v\|^2 \overset{\text{substitution}}{\rightarrow} \tag{7}$$

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

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

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

$$\Rightarrow u^* = \tau \mathsf{prox}_{\frac{1}{\tau}E^*}(\frac{v}{\tau}). \tag{11}$$

From that it follows that

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

Prox of ℓ_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

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

Then using the Moreau decomposition we have

$$\mathsf{prox}_{\tau E}(v) = v - \tau \mathsf{prox}_{\frac{1}{\tau}E^*}(\frac{v}{\tau}) \tag{12}$$

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

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

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

MM interpretation

$$\mathsf{prox}_{(1/L)G}\left(u^k - (1/L)\nabla F(u^k)\right) \tag{16}$$

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

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

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

Nonexpansiveness of prox

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

$$\langle u - v, x - y \rangle \geq \|x - y\|^2. \tag{20}$$

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

$$u - x \in \partial E(x) \tag{21}$$

$$v - y \in \partial E(y) \tag{22}$$

From that it follows

$$E(z) - E(x) \geq \langle u - x, z - x \rangle, \forall z \tag{23}$$

$$E(z) - E(y) \geq \langle v - y, z - y \rangle, \forall z \tag{24}$$

With special choice of z it follows

$$E(y) - E(x) \geq \langle u - x, y - x \rangle, \tag{25}$$

$$E(x) - E(y) \geq \langle v - y, x - y \rangle, \tag{26}$$

and adding these two inequalites gives

$$0 \geq \langle u - x, y - x \rangle + \langle v - y, x - y \rangle = \langle v - y + x - u, x - y \rangle, \tag{27}$$

And hence

$$\langle y - v + u - x, x - y \rangle \geq 0 \tag{28}$$

$$\Leftrightarrow \quad \langle u - v, x - y \rangle \geq \|x - y\|^2 \tag{29}$$

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

With Cauchy-Schwarz it follows

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

$$\|u - v\| \geq \|\mathsf{prox}_E(u) - \mathsf{prox}_E(v)\| \tag{32}$$

Convergence of gradient projection

Gradient map in subdifferential

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

$$u - \tau \nabla F(u) - (u - \tau \varphi_\tau(u)) \in \tau \partial G(u - \tau \nabla F(u)) \tag{33}$$

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