From perspective maps to epi-projections

Tim Hoheisel

McGill

joint work with

Michael P. Friedlander (UBC, Vancouver)

Ariel Goodwin (McGill, Montreal)

One World Optimization Seminar

September 27, 2021

Epigraphical projections via proximal map

Let $f \in \Gamma_0 := \left\{ f : \mathbb{E} \to \mathbb{R} \cup \{+\infty\} \;\middle|\; f \text{ closed, proper, convex} \right\}$, $\lambda > 0$:

$$P_\lambda f(x) := \underset{u \in \mathbb{E}}{\operatorname{argmin}} \left\{ f(u) + \frac{1}{2\lambda} \|x - u\|^2 \right\} \quad \text{(proximal map)}.$$

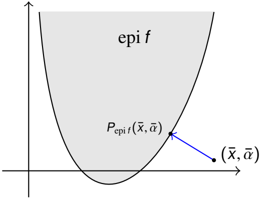
Let $f : \mathbb{E} \to \mathbb{R}$ convex, $(\bar{x}, \bar{\alpha}) \notin \operatorname{epi} f$. Then

$$P_{\operatorname{epi} f}(\bar{x}, \bar{\alpha}) = (P_{\lambda^*} f(\bar{x}), \bar{\alpha} + \lambda^*),$$

where $\lambda^* > 0$ is a positive root of

$$0 < \lambda \mapsto f(P_\lambda f(\bar{x})) - \lambda - \bar{\alpha}.$$

(Bauschke/Combettes, Theorem 29.35)



- Proximal map vs. projection onto the epigraph in the literature: E.g. Bauschke/Combettes ’10, Beck ’17, Chierca et al. ’15, Meng et al. ’05/’08.
- Goal: Understanding the convex-analytic properties of the proximal map (and Moreau envelope) as a function of the prox parameter λ (and the base point x simultaneously)!

A general framework via perspective maps

For $f \in \Gamma_0$, $\omega = \frac{1}{2}\|\cdot\|^2$:

$$P_\lambda f(x) = \operatorname*{argmin}_{u \in \mathbb{E}} \left\{ f(u) + \lambda\omega\left(\frac{x-u}{\lambda}\right)\right\} \quad \forall (x,\lambda) \in \mathbb{E} \times \mathbb{R}_{++}.$$

Definition 1 (Perspective map).

For $\omega \in \Gamma_0$, (the closure of) its *perspective map* is $\omega^\pi : \mathbb{E} \times \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$,

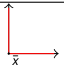
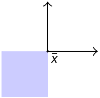
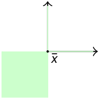
$$\omega^\pi(z,\lambda) = \begin{cases} \lambda\omega\left(\frac{z}{\lambda}\right) & \text{if} \quad \lambda > 0, \\ \omega^\infty(z) & \text{if} \quad \lambda = 0, \\ +\infty & \text{else.} \end{cases}$$

- $\omega^\pi(z,\lambda) = \sigma_{\mathrm{epi}\,\omega^*}(z,-\lambda)$, hence ω^π is lsc, proper and sublinear with $\mathrm{dom}\,\omega^\pi = \mathbb{R}_+(\mathrm{dom}\,\omega \times \{1\})$.
- Literature: E.g. Rockafellar '70, Aravkin et al '18, Combettes/Müller '18/'19.

For $f, \omega \in \Gamma_0$ we want to study the solution map

$$P_{\omega,f} : \mathbb{E} \times \mathbb{R} \rightrightarrows \mathbb{E}, \qquad P_{\omega,f}(x,\lambda) = \operatorname*{argmin}_{u \in \mathbb{E}} \{f(u) + \omega^\pi(x-u,\lambda)\}.$$

Interlude: Tools from variational analysis

Name	Definition	Properties	Example
tangent cone	$T_A(\bar{x}) := \operatorname{Lim\,sup}_{t\downarrow 0} \frac{A-\bar{x}}{t}$	closed	
regular normal cone	$\hat{N}_A(\bar{x}) := T_A(\bar{x})^\circ$	closed, convex	
limiting normal cone	$N_A(\bar{x}) := \operatorname{Lim\,sup}_{x\to\bar{x}} \hat{N}_A(x)$	closed	

$S : \mathbb{E}_1 \rightrightarrows \mathbb{E}_2$, $(\bar{x}, \bar{y}) \in \operatorname{gph} S$.

- (Coderivative) $D^*S(\bar{x}|\bar{y}) : \mathbb{E}_2 \rightrightarrows \mathbb{E}_1$, $v \in D^*S(\bar{x}|\bar{y})(y) \;:\Longleftrightarrow\; (v, -y) \in N_{\operatorname{gph} S}(\bar{x}, \bar{y})$.
- (Graphical derivative) $DS(\bar{x}|\bar{y}) : \mathbb{E}_1 \rightrightarrows \mathbb{E}_2$, $v \in DS(\bar{x}|\bar{y})(u) \;:\Longleftrightarrow\; (u, v) \in T_{\operatorname{gph} S}(\bar{x}, \bar{y})$.

Variational Analysis of $P_{\omega,f}$: Part I

Recall: $P_{\omega,f} : \mathbb{E}\times\mathbb{R} \rightrightarrows \mathbb{E}$, $P_{\omega,f}(x,\lambda) = \operatorname{argmin}_{u\in\mathbb{E}} \{f(u) + \omega^\pi(x-u,\lambda)\}$.

Theorem 2 (Friedlander, Goodwin, H.).

Let $\omega : \mathbb{E}\to\mathbb{R}$ be strictly convex, level-bounded and C^2. Then:

a) $\operatorname{dom} P_{\omega,f} \subset \mathbb{E}\times\mathbb{R}_+$ *and $P_{\omega,f}$ is single-valued for all $\lambda > 0$.*

b) *For $\bar\lambda > 0$, if $\nabla^2\omega\left(\frac{\bar x - P_{\omega,f}(\bar x,\bar\lambda)}{\bar\lambda}\right) \succ 0$, then $P_{\omega,f}$ is locally Lipschitz at $(\bar x,\bar\lambda)$.*

c) *Under the conditions in b), assume, in addition, that ∂f is proto-differentiable at $\left(P_{\omega,f}(\bar x,\bar\lambda), \nabla\omega\left(\frac{\bar x - P_{\omega,f}(\bar x,\bar\lambda)}{\bar\lambda}\right)\right)$. Then $P_{\omega,f}$ is directionally differentiable at $(\bar x,\bar\lambda)$ with*

$$P'_{\omega,f}((\bar x,\bar\lambda);(d,\Delta)) = \left[\bar\lambda D(\partial f)\left(P_{\omega,f}(\bar x,\bar\lambda)\middle|\nabla\omega\left(\frac{\bar x - P_{\omega,f}(\bar x,\bar\lambda)}{\bar\lambda}\right)\right) + \bar V\right]^{-1}\left(\bar V d - \frac{\Delta}{\bar\lambda}\bar V(\bar x - P_{\omega,f}(\bar x,\bar\lambda))\right)$$

with $\bar V = \nabla^2\omega\left(\frac{\bar x - P_{\omega,f}(\bar x,\bar\lambda)}{\bar\lambda}\right)$, for any $(d,\Delta)\in\mathbb{E}\times\mathbb{R}$.

- ω strongly convex $\Rightarrow$ level-bounded (in fact, supercoercive) and strictly convex.
- ∂f is proto-differentiable e.g. if $f = g\circ F$, with g PLQ and $F\in C^2$ such that BCQ holds.

Directional normal cone and semismoothness* (Gfrerer et al.)

<u>*Directional normal cone*[1]</u> of A at $\bar{x}$ in direction $\bar{u}$:

$$N_A(\bar{x};\bar{u}) := \operatorname*{Lim\,sup}_{u\to\bar{u},\ t\downarrow 0} \hat{N}_A(\bar{x}+tu).$$

- $N(\bar{x};\bar{u}) = \emptyset$ if $\bar{u} \notin T_A(\bar{x})$;
- $N(\bar{x};\bar{u}) \subset N_A(\bar{x})$ for all $u \in \mathbb{E}$.

Definition 3 (Semismothness*).

i) $A \subset \mathbb{E}$ *semismooth** at $\bar{x} \in A$ $\quad :\Longleftrightarrow \quad \langle x^*, u\rangle = 0 \quad \forall u \in \mathbb{E},\ x^* \in N_A(\bar{x};u)$.

ii) $S : \mathbb{E}_1 \rightrightarrows \mathbb{E}_2$ *semismooth** at $(\bar{x},\bar{y}) \in \operatorname{gph} S$ $\quad :\Longleftrightarrow \quad \operatorname{gph} S$ semismooth* at $(\bar{x},\bar{y})$.

- A convex $\Longrightarrow$ A semismooth*;
- A_i semismooth* at $\bar{x}$ (i=1,. . . ,n) $\Longrightarrow$ $\bigcup_{i=1}^n A_i$ semismooth* at $\bar{x}$.

[1] Mordukhovich and Ginchev

Semismoothness(*) of $P_{\omega,f}$

Proposition 4 (Gfrerer and Outrata '19).

Let $F : D \subset \mathbb{E}_1 \to \mathbb{E}_2$ be locally Lipschitz at $\bar{x} \in \operatorname{int} D$. Then the following are equivalent:

- *F semismooth (in the sense of Qi and Sun) at $\bar{x}$, i.e.* $\lim\limits_{\substack{V \in \partial_C F(\bar{x}+td)^2 \\ t \downarrow 0, d \to \bar{d}}} Vd$ *exists for all $\bar{d}$.*
- *F semismooth* and directionally differentiable at $\bar{x}$.*

Proposition 5 (Friedlander, Goodwin, H.).

Let $f \in \Gamma_0$, let $(\bar{x}, \bar{\lambda}) \in \mathbb{E} \times \mathbb{R}_{++}$ such that ∂f is semismooth at $\left(P_{\omega,f}(\bar{x}, \bar{\lambda}), \nabla\omega\left(\frac{\bar{x} - P_{\omega,f}(\bar{x},\bar{\lambda})}{\bar{\lambda}}\right)\right)$ and let ω be strongly convex and C^2. Then:*

- *$P_{\omega,f}$ f is semismooth* at $((\bar{x}, \bar{\lambda}), P_{\omega,f}(\bar{x}, \bar{\lambda}))$.*
- *If, also, ∂f is proto-differentiable at $\left(P_{\omega,f}(\bar{x}, \bar{\lambda}), \nabla\omega\left(\frac{\bar{x} - P_{\omega,f}(\bar{x},\bar{\lambda})}{\bar{\lambda}}\right)\right)$, then $P_{\omega,f}$ is semismooth at $(\bar{x}, \bar{\lambda})$.*

- $f \in C^2$ or f PLQ $\Longrightarrow \partial f$ proto-differentiable and semismooth*.
- $\omega = \langle H(\cdot), (\cdot)\rangle$ with $H \succ 0 \Longrightarrow \omega$ strongly convex + C^2.

[2] $\partial_C F(x) := \operatorname{conv}\left\{V \mid \exists \{x_k \in D_F\} \to x : F'(x_k) \to V\right\}$ (Clarke Jacobian).

The proximal map revisited I - continuity properties

Proposition 6 (Attouch, Rockafellar/Wets, Friedlander/Goodwin, H.).

Let $f \in \Gamma_0$ and $\bar{x} \in \mathbb{E}$. Then

$$\lim_{\substack{\lambda \downarrow 0, \\ x \to \bar{x}}} P_\lambda f(x) = P_{\mathrm{cl}\,(\mathrm{dom}\, f)}(\bar{x}).$$

Natural extension of the proximal operator of $f \in \Gamma_0$:

$$P_f : \mathbb{E} \times \mathbb{R} \rightrightarrows \mathbb{E}, \quad P_f(x, \lambda) := \begin{cases} P_\lambda f(x), & \lambda > 0, \\ P_{\mathrm{cl}\,(\mathrm{dom}\, f)}(x), & \lambda = 0, \\ \emptyset, & \lambda < 0. \end{cases}$$

Corollary 7 (Friedlander/Goodwin/H., Milzarek, Strömberg).

a) *P_f is continuous on $\mathrm{dom}\, P_f = \mathbb{E} \times \mathbb{R}_+$, and locally Lipschitz on $\mathrm{int}\,(\mathrm{dom}\, P)$.*

b) *For $\bar{x} \in \mathrm{dom}\, \partial f$ (hence $P_f(\bar{x}, 0) = \bar{x} \in \mathrm{dom}\, f$), P is calm at $(\bar{x}, 0)$, i.e.*

$$\exists \kappa, \varepsilon > 0 : \ \|P_f(\bar{x}, 0) - P_f(x, \lambda)\| \le \kappa \|(\bar{x} - x, \lambda)\| \quad ((x, \lambda) \in B_\varepsilon(\bar{x}, 0) \cap \mathrm{dom}\, P_f).$$

c) *Under the assumptions of b) the map $P_f(\bar{x}, \cdot) : \mathbb{R}_+ \to \mathbb{E}$ is locally Lipschitz at 0, i.e.*

$$\exists \kappa, \varepsilon > 0 : \ \|P_f(\bar{x}, \lambda) - P_f(\bar{x}, \mu)\| \le \kappa |\mu - \lambda| \quad \forall \lambda, \mu \in [0, \varepsilon].$$

The proximal map revisited II - directional differentiability

Corollary 8 (Friedlander, Goodwin, H.).

Let $f \in \Gamma_0$ and let $(\bar{x}, \bar{\lambda}) \in \mathbb{E} \times \mathbb{R}_{++}$. If ∂f is proto-differentiable at $\left(P_f(\bar{x}, \bar{\lambda}), \frac{\bar{x} - P_f(\bar{x}, \bar{\lambda})}{\bar{\lambda}}\right)$, then P_f is directionally differentiable at $(\bar{x}, \bar{\lambda})$ with

$$P_f'((\bar{x}, \bar{\lambda}); (d, \Delta)) = \left[\bar{\lambda} D(\partial f)\left(P_f(\bar{x}, \bar{\lambda}) \middle| \frac{\bar{x} - P_f(\bar{x}, \bar{\lambda})}{\bar{\lambda}}\right) + I\right]^{-1} \left(d - \Delta \frac{\bar{x} - P_f(\bar{x}, \bar{\lambda})}{\bar{\lambda}}\right) \quad \forall (d, \Delta) \in \mathbb{E} \times \mathbb{R}.$$

In particular, for any $\bar{\lambda} > 0$, we have

$$(P_{\bar{\lambda}} f)'(\bar{x}; d) = \left[\bar{\lambda} D(\partial f)\left(P_f(\bar{x}, \bar{\lambda}) \middle| \frac{\bar{x} - P_f(\bar{x}, \bar{\lambda})}{\bar{\lambda}}\right) + I\right]^{-1} (d) \quad \forall d \in \mathbb{E}.$$

Observe: Projection onto a closed, convex set is not necessarily directionally differentiable. For

$$S = \operatorname{conv}\left[\left\{\left(\cos\frac{\pi}{2^k}, \sin\frac{\pi}{2^k}\right) \mid k \in \mathbb{N}\right\} \cup \{(0,0), (1,0)\}\right],$$

the projection P_S is not directionally differentiable at $(2, 0)$ in direction $(0, 1)$.

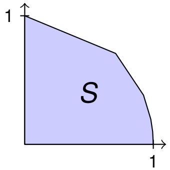


Figure: Shapiro's example (Shapiro '94)

The proximal map revisited III - semismoothness(*) properties

Proposition 9 (Friedlander, Goodwin, H.).

Let $f \in \Gamma_0$ and let $(\bar{x}, \bar{\lambda}) \in \mathbb{E} \times \mathbb{R}_{++}$. Then:

a) *If ∂f semismooth* at $\left(P_f(\bar{x}, \bar{\lambda}), \frac{\bar{x} - P_f(\bar{x}, \bar{\lambda})}{\bar{\lambda}}\right)$ then P_f is semismooth* at $(\bar{x}, \bar{\lambda})$.*

b) *The map $P_{\bar{\lambda}} f (= P_f(\cdot, \bar{\lambda}))$ is semismooth* at $\bar{x}$ if and only if ∂f is semismooth* at $\left(P_{\bar{\lambda}} f(\bar{x}), \frac{\bar{x} - P_{\bar{\lambda}} f(\bar{x})}{\bar{\lambda}}\right)$.*

c) *If ∂f is proto-differentiable and semismooth* at $\left(P_f(\bar{x}, \bar{\lambda}), \frac{\bar{x} - P_f(\bar{x}, \bar{\lambda})}{\bar{\lambda}}\right)$, then P_f is semismooth at $(\bar{x}, \bar{\lambda})$.*

- The assumption in a) are satisfied e.g. if $f = g \circ H$ with g PLQ and $H \in C^2$ such that BCQ holds.
- The assumptions in c) are satisfied if f is PLQ or twice continuously differentiable at $P_f(\bar{x}, \bar{\lambda})$ (in which case P_f is continuously differentiable at $(\bar{x}, \bar{\lambda})$).

The proximal value map $\lambda \mapsto f(P_\lambda f(\bar{x}))$

For $f \in \Gamma_0$ and $\bar{x} \in \mathbb{E}$, consider $0 < \lambda \mapsto f(P_\lambda f(\bar{x}))$.

Example: Let $f := |\cdot| + \delta_{[-1,1]} \in \Gamma_0(\mathbb{R})$. Then

$P_\lambda f(x) = \min\{\max\{|x| - \lambda, 0\}, 1\} \cdot \mathrm{sgn}(x) \quad (x \in \mathbb{R}, \lambda > 0)$.

For $\bar{x} = 2$:

$$f(P_\lambda f(\bar{x})) = \begin{cases} 1, & \lambda \in (0,1], \\ 2 - \lambda, & \lambda \in (1,2], \\ 0, & \lambda > 2. \end{cases}$$

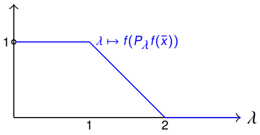


Figure: Nonconvex example

Proposition 10 (Friedlander/Goodwin/H., Attouch '84).

Let $f \in \Gamma_0$ and $\bar{x} \in \mathbb{E}$. Then:

a) *The function $0 < \lambda \mapsto f(P_\lambda f(\bar{x}))$ is decreasing (i.e. increasing as $\lambda \downarrow 0$).*

b) $\lim_{\lambda \to 0} f(P_\lambda f(\bar{x})) = f(P_{\mathrm{cl}\,(\mathrm{dom}\, f)}(\bar{x}))$.

The map $\bar{\phi}^f_{\bar{x}}$

Proposition 11 (Friedlander, Goodwin, H.).

Let $f \in \Gamma_0$, $\bar{x} \in \mathbb{E}$ and define $\bar{\phi}^f_{\bar{x}} : \mathbb{R} \to \overline{\mathbb{R}}$ by

$$\bar{\phi}^f_{\bar{x}}(\lambda) := \begin{cases} -\lambda e_\lambda f(\bar{x}), & \lambda > 0,^3 \\ -\frac{1}{2} d^2_{\mathrm{cl}\,(\mathrm{dom}\, f)}(\bar{x}), & \lambda = 0, \\ -f(P_{\mathrm{cl}\,(\mathrm{dom}\, f)}(\bar{x}))\lambda - \frac{1}{2} d^2_{\mathrm{cl}\,(\mathrm{dom}\, f)}(\bar{x}), & \lambda < 0. \end{cases}$$

Then the following hold:

a) $\bar{\phi}^f_{\bar{x}}$ *is proper, convex and continuous (possibly in an extended real-valued sense), and continuously differentiable on* $\mathbb{R}_{++}$ *with* $\frac{d}{d\lambda}\bar{\phi}^f_{\bar{x}}(\lambda) = -f(P_\lambda f(\bar{x}))$ *locally Lipschitz for all* $\lambda > 0$.

b) *If* $\bar{x} \in \mathrm{dom}\, f$, *then* $\bar{\phi}^f_{\bar{x}}$ *is continuously differentiable (on* $\mathbb{R}$*) with derivative given by*

$$\frac{d}{d\lambda}\bar{\phi}^f_{\bar{x}}(\lambda) = \begin{cases} -f(P_\lambda f(\bar{x})), & \lambda > 0, \\ -f(P_{\mathrm{cl}\,(\mathrm{dom}\, f)}(\bar{x})), & \lambda \leq 0. \end{cases}$$

If, more strictly, $\bar{x} \in \mathrm{dom}\, \partial f$, *then* $\frac{d}{d\lambda}\bar{\phi}^f_{\bar{x}}(\lambda)$ *is locally Lipschitz on all of* $\mathbb{R}$.

c) *If* $P_{\mathrm{cl}\,(\mathrm{dom}\, f)}(\bar{x}) \notin \mathrm{dom}\, f$, *then* $\mathrm{dom}\, \bar{\phi}^f_{\bar{x}} = \mathbb{R}_+$ *and* $\partial\bar{\phi}^f_{\bar{x}}(\lambda) = \begin{cases} -f(P_\lambda f(\bar{x})), & \lambda > 0, \\ \emptyset, & \lambda \leq 0. \end{cases}$

[3] For $\varphi \in \Gamma_0$, $\lambda > 0$, the function $e_\lambda\varphi(\bar{x}) = \varphi(P_\lambda\varphi(\bar{x})) + \frac{1}{2\lambda}\|\bar{x} - P_\lambda\varphi(\bar{x})\|^2$ is the Moreau envelope of φ at $\bar{x}$.

Proximal operators for post-compositions

Recall that

$$\phi_{\bar{x}}^{\psi}(\lambda) = \begin{cases} -\lambda e_{\lambda}\psi(\bar{x}), & \lambda > 0, \\ -\frac{1}{2} d^2_{\mathrm{cl}\,(\mathrm{dom}\,\psi)}(\bar{x}), & \lambda = 0. \end{cases}$$

Proposition 12 (Friedlander, Goodwin, H.).

Let $g \in \Gamma_0(\mathbb{R})$ be increasing and let $\psi \in \Gamma_0(\mathbb{E})$ such that

$$\mathrm{ri}\,(\mathrm{dom}\,g) \cap \psi(\mathrm{ri}\,(\mathrm{dom}\,\psi)) \neq \emptyset. \tag{1}$$

Then[4] $g \circ \psi \in \Gamma_0$ and:

a) $e_1(g \circ \psi)(\bar{x}) = -\min_{\lambda \geq 0} g^*(\lambda) + \bar{\phi}_{\bar{x}}^{\psi}(\lambda)$;

b) $P_1(g \circ \psi)(\bar{x}) = P_1(\bar{\lambda} \cdot \psi)(\bar{x})$ *for every* $\bar{\lambda} \in \mathrm{argmin}_{\lambda \geq 0} \left\{ g^*(\lambda) + \bar{\phi}_{\bar{x}}^{\psi}(\lambda) \right\} \neq \emptyset$;

c) *If* $\psi(P_{\mathrm{cl}\,(\mathrm{dom}\,\psi)}(\bar{x})) \notin \partial g^*(0)$, *then* $\mathrm{argmin}_{\lambda \geq 0} \left\{ g^*(\lambda) + \bar{\phi}_{\bar{x}}^{\psi}(\lambda) \right\} \subset \mathbb{R}_{++}$. *This is, in particular, the case if* $P_{\mathrm{cl}\,(\mathrm{dom}\,\psi)}(\bar{x}) \notin \mathrm{dom}\,\psi$.

[4] Set $g(+\infty) := +\infty$.

Level-set and epigraphical projections

Setting $g := \delta_{\mathbb{R}_-}$ and $\psi := f - \bar{\alpha}$ in in Theorem 12 gives:

Corollary 13 (Level set case).

Let $f \in \Gamma_0$, $\bar{\alpha} \in \mathbb{R}, \bar{x} \in \mathbb{E}$, and assume that there exists $\hat{x} \in \mathbb{E}$ such that $f(\hat{x}) < \bar{\alpha}$. Then

$$P_{\mathrm{lev}_{\bar{\alpha}} f}(\bar{x}) = \begin{cases} P_{\mathrm{cl}\,(\mathrm{dom}\, f)}(\bar{x}), & f(P_{\mathrm{cl}\,(\mathrm{dom}\, f)}(\bar{x})) \leq \bar{\alpha}, \\ P_{\bar{\lambda}} f(\bar{x}), & \textit{else} \end{cases}$$

for any $0 < \bar{\lambda} \in \mathrm{argmin}_{\lambda \geq 0} \{\bar{\phi}^f_{\bar{x}}(\lambda) + \bar{\alpha}\lambda\} = \left\{\lambda \geq 0 \;\middle|\; f(P_\lambda f(\bar{x})) = \bar{\alpha}\right\} \neq \emptyset$.

Letting $g := \delta_{\mathbb{R}_-}$ and $\psi : (x, \alpha) \in \mathbb{E} \times \mathbb{R} \mapsto f(x) - \alpha$ in Theorem 12 yields:

Corollary 14 (Projection onto epigraph).

Let $f \in \Gamma_0$, $\bar{\alpha} \in \mathbb{R}$ and $\bar{x} \in \mathbb{E}$. Then

$$P_{\mathrm{epi}\, f}(\bar{x}, \bar{\alpha}) = \begin{cases} (P_{\mathrm{cl}\,(\mathrm{dom}\, f)}(\bar{x}), \bar{\alpha}), & f(P_{\mathrm{cl}\,(\mathrm{dom}\, f)}(\bar{x})) \leq \bar{\alpha}, \\ (P_{\bar{\lambda}} f(\bar{x}), \bar{\alpha} + \bar{\lambda}), & \textit{else}, \end{cases}$$

where $\bar{\lambda} > 0$ is the unique minimizer of the (strongly convex) optimization problem

$$\min_{\lambda \geq 0} \frac{1}{2}\lambda^2 + \bar{\alpha}\lambda + \bar{\phi}^f_{\bar{x}}(\lambda).$$

SC¹ optimization framework

For computing level-set and epigraphical projections, respectively, we have to solve the scalar problems

$$\min_{\lambda \geq 0} \theta_\xi(\lambda) = \begin{cases} \bar{\alpha} \cdot \lambda + \bar{\phi}^f_{\bar{x}}(\lambda), & \xi = \mathrm{lev}, \\ \frac{1}{2}\lambda^2 + \bar{\alpha} \cdot \lambda + \bar{\phi}^f_{\bar{x}}(\lambda), & \xi = \mathrm{epi}\,. \end{cases}$$

The variational properties of $\phi^f_{\bar{x}}$ suggest the SC¹ optimization framework à la Pang and Qi (1995).

Algorithm 1 SC¹ Newton method for minimizing θ_ξ

(S.0) Choose $\lambda_0, \delta > 0$, $\{\varepsilon_k\} \downarrow 0$, and let $\beta, \sigma \in (0,1)$. Set $k := 0$.

(S.1) If $|\theta'(\lambda_k)| \leq \delta$: STOP.

(S.2) Choose $g_k \in \partial_C(\theta'_\xi)(\lambda_k)$ and set

$$\Delta_k := P_{[-\lambda_k,\infty)}\left(-\frac{\theta'_\xi(\lambda_k)}{g_k + \varepsilon_k}\right).$$

(S.3) Set

$$t_k := \max_{l \in \mathbb{N}_0}\left\{\beta^l \;\middle|\; \theta_\xi(\lambda_k + \beta^l \Delta_k) \leq \theta_\xi(\lambda_k) + \beta^l \sigma \theta'_\xi(\lambda)\Delta_k\right\}.$$

(S.4) Set $\lambda_{k+1} := \lambda_k + t_k\Delta_k$, $k \leftarrow k+1$, and go to (S.1).

- For epigraphical projections ($\xi = \mathrm{epi}$) use $\varepsilon_k := 0$ to improve accuracy.
- Computational bottleneck: Backtracking in (S.3).

A full Newton step algorithm

Algorithm 2 Full step SC^1 Newton method

(S.0) Choose $\lambda_0 > 0, \delta > 0$, and $\{\varepsilon_k\} \downarrow 0$. Set $k := 0$.
(S.1) If $|\theta'_\xi(\lambda_k)| \leq \delta$: STOP.
(S.2) Choose $g_k \in \partial_C(\theta'_\xi)(\lambda_k)$ and set

$$\Delta_k := \max\left\{\frac{-\lambda_k}{2}, -\frac{\theta'_\xi(\lambda_k)}{g_k + \varepsilon_k}\right\}.$$

(S.3) Set $\lambda_{k+1} := \lambda_k + \Delta_k$, $k \leftarrow k+1$, and go to (S.1).

Slogan: *Take full Newton step in every iteration while respecting positivity of the iterates!*

Set $[\lambda_l, \lambda_u] := \left\{\lambda > 0 \,\middle|\, \theta'_\xi(\lambda) = 0\right\}$.

- Assume θ'_ξ concave on $(0, \lambda_l)$.

Then Algorithm 2 converges to a minimizer of θ_ξ.

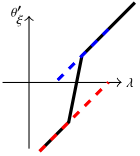


Figure: Cycling without concavity assumption

Projection onto the l_1-unit ball

For $\mathrm{lev}_1 \|\cdot\|_1 = \{x \in \mathbb{R}^n \mid \|x\|_1 \leq 1\}$ we have

$$\theta'_{\mathrm{lev}}(\lambda) = \begin{cases} 1 - \sum_{i=1}^{n} \max\{|x_i| - \lambda, 0\}, & \lambda \geq 0, \\ 1 - \|x\|_1, & \lambda < 0, \end{cases}$$

which has θ'_{lev} concave on $(0, \lambda_l)$.

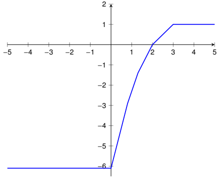


Figure: The function θ'_{ξ} for projecting onto the l_1-unit ball.

Table: Time (seconds) for projecting onto the l_1-unit ball in dimension N with coordinates chosen using a (component-wise) Gaussian distribution with $\sigma = 0.1$.

N	Warm Newton	Condat	IBIS
20	1.44×10^{-6}	1.53×10^{-6}	1.83×10^{-6}
10^3	1.83×10^{-5}	2.11×10^{-5}	3.65×10^{-5}
10^6	1.38×10^{-2}	1.44×10^{-2}	2.89×10^{-2}
10^7	1.51×10^{-1}	1.43×10^{-1}	2.85×10^{-1}

Takeaway: Our algorithm performs better than alternative methods when the starting point is good!

Future directions

- Understand semismoothness* of ∂f better:
  - Establish more powerful checkable conditions;
  - extend to non-convex functions f;
  - extend to (maximally) monotone operators.
- Generalize the result for post-composition setting to the convex convex-composite setting (e.g. Burke, H., Nguyen, MOR 2020): $f : \mathbb{E}_1 \to \mathbb{R} \cup \{+\infty\}$ with $f = g \circ H$
  - $K \subset \mathbb{E}_2$ (closed) convex cone;
  - $H : \mathbb{E}_1 \to \mathbb{E}_2^\bullet$ K-convex, i.e. $\left\{(X, Y) \mid Y - H(x) \in K\right\}$ (K-epigraph) is convex;
  - $g \in \Gamma_0(\mathbb{E}_2)$ K-increasing.

Goal: Establish theory and algorithms for computing proximal operators of $g \circ F$ and K-epigraphical projections.

References

- H. Attouch: *Variational Convergence for Functions and Operators.* Applied Mathematics Series, Pittman, Boston, 1984.
- H.H. Bauschke and P.L. Combettes: *Convex analysis and Monotone Operator Theory in Hilbert Spaces.* CMS Books in Mathematics, Springer-Verlag, New York, 2nd Edition, 2017.
- P.L. Combettes: *Perspective functions: properties, constructions, and examples.* Set-Valued and Variational Analysis 26, 2019, pp. 247–264.
- L. Condat: *Fast projection onto the simplex and l1 ball.* Mathematical Programming, Series A, Springer, 2016, 158 (1), pp. 575–585.
- H. Gfrerer: *On directional metric regularity, subregularity and optimality conditions for nonsmooth mathematical programs.* Set-Valued and Variational Analysis 21, 2013, pp. 151–176.
- H. Gfrerer and J.V. Outrata: *On a semismooth* Newton method for solving generalized equations.* arXiv, 2019.
- J. Liu and J. Ye: *Efficient Euclidean projections in linear time.* Proceedings of the 26th Annual International Conference on Machine Learning, 2009, pp. 657–664.
- J.S. Pang and L. Qi: *A Globally convergent Newton method for convex SC^1 minimization problems.* Journal of Optimization Theory and Applications 85(3), 1995, pp. 633–648.
- L. Qi and J. Sun: *A nonsmooth version of Newton's method.* Mathematical Programming 58, 1993, pp. 353–367.
- R.T. Rockafellar and R.J.-B. Wets: *Variational Analysis.* Grundlehren der Mathematischen Wissenschaften, Vol. 317, Springer-Verlag, Berlin, 19