IPIANO: INERTIAL PROXIMAL ALGORITHM FOR NON-CONVEX OPTIMIZATION

DAVID STUTZ*

Composite function:

$$\min_{x\in\mathbb{R}^n} h(x) = \min_{x\in\mathbb{R}^n}\Big(\underbrace{f(x)}_{\substack{\in C^1\\ \nabla f \text{ Lipschitz-continuous}}} + \underbrace{g(x)}_{\substack{\text{proper closed convex}\\ \text{lower semicontinuous}}} \Big)$$

with

$$-\infty < h_{\min} := \inf_{x\in\mathbb{R}^n} h(x)$$

Iteration:

$$x^{(n+1)} = \text{prox}_{\alpha_n g}(x^{(n)} - \alpha_n \nabla f(x^{(n)}) + \beta_n(x^{(n)} - x^{(n-1)})).$$

Backtracking:

$$f(x^{(n+1)}) \leq f(x^{(n)}) + \nabla f(x^{(n)})^T(x^{(n+1)} - x^{(n)}) + \frac{L_n}{2}\|x^{(n+1)} - x^{(n)}\|_2^2$$

Remember:

$$\delta_n := \frac{1}{\alpha_n} - \frac{L_n}{2} - \frac{\beta_n}{2\alpha_n} \geq \gamma_n$$
$$\gamma_n := \frac{1}{\alpha_n} - \frac{L_n}{2} - \frac{\beta_n}{\alpha_n} \geq c_2$$

Proof Sketch (Monotonicity)
Using (**??**) and (**??**)

$$\gamma_n \geq c_2 \quad \Leftrightarrow \quad \alpha_n \leq \frac{1-\beta_n}{c_2 + \frac{L_n}{2}} < \frac{2(1-\beta_n)}{L_n}$$

To show monotonicity of δ_n:

$$\delta_{n-1} \geq \delta_n := \frac{1}{\alpha_n} - \frac{L_n}{2} - \frac{\beta_n}{2\alpha_n} \quad \Leftrightarrow \quad \alpha_n \geq \frac{1 - \frac{\beta_n}{2}}{\delta_{n-1} + \frac{L_n}{2}};$$

with (**??**):

$$\delta_{n-1} \geq \delta_n \;\Leftrightarrow\; \frac{1-\beta_n}{c_2 + \frac{L_n}{2}} \geq \alpha_n \geq \frac{1-\frac{\beta_n}{2}}{\delta_{n-1} + \frac{L_n}{2}} \;\Rightarrow\; \beta_n \leq \frac{b_n - 1}{b_n - \frac{1}{2}}$$

with $b_n := (\delta_{n-1} + \frac{L_n}{2})/(c_2 + \frac{L_n}{2})$.

Remember:

$$H_{\delta_{n+1}}(x^{(n+1)}, x^{(n)}) := h(x^{(n+1)}) + \delta_{n+1}\Delta_{n+1}^2$$

---

*Submitted as part of the seminar "Ausgewählte Themen der Bildverarbeitung", WS 2015/2016, and advised by Prof. Berkels, AICES, RWTH Aachen University.

and

$$(z^{(n+1)})_{n\in\mathbb{N}} = (x^{(n+1)}, x^{(n)})_{n\in\mathbb{N}}.$$

Proof Sketch (Condition (H1))
From (**??**):

$$w := \frac{x^{(n)} - x^{(n+1)}}{\alpha_n} - \nabla f(x^{(n)}) + \frac{\beta_n}{\alpha_n}(x^{(n)} - x^{(n-1)}) \in \partial g(x^{(n+1)})$$

With

$$g(x^{(n+1)}) \overset{convex}{\leq} g(x^{(n)}) - w^T(x^{(n)} - x^{(n-1)})$$

and

$$f(x^{(n+1)}) \leq f(x^{(n)}) - +\nabla f(x^{(n)})^T(x^{(n+1)} - x^{(n)}) + \frac{L_n}{2}\|x^{(n)} - x^{(n+1)}\|_2^2$$

the claim follows from

$$h(x^{(n+1)}) \leq h(x^{(n)}) - \delta_n\Delta_{n+1}^2 + \delta_n\Delta_n^2 - \gamma_n\Delta_n^2.$$

Proof Sketch (Condition (H2))
For $w^{(n+1)} \in \partial H_{\delta_{n+1}}(z^{(n+1)})$ it is $w^{(n+1)} = (w_1^{(n+1)}, w_2^{(n+1)})$ with

$$\begin{aligned} w_1^{(n+1)} &\in \partial g(x^{(n+1)}) + \nabla f(x^{(n+1)}) + 2\delta_n(x^{(n+1)} - x^{(n)}) \\ w_2^{(n+1)} &= -2\delta_n(x^{(n+1)} - x^{(n)}) \end{aligned}$$

and with Equation (**??**)

$$\|w^{(n+1)}\|_2 \leq ... \leq (\frac{1}{\alpha_n} + 4\delta_n + L_n)\Delta_{n+1} + \frac{\beta_n}{\alpha_n}\Delta_n \leq \frac{7}{c_1}(\Delta_{n+1} + \Delta_n)$$

Proof Sketch (Condition (H3))
Remember:

$$H_{\delta_{n+1}}(z^{(n+1)}) + \gamma_n\Delta_n^2 \leq H_{\delta_n}(z^{(n)}). \tag{H1}$$

Claim 1: $\lim_{n\to\infty}\Delta_n = 0$
Summing Condition (**??**) and deducing $\sum_{n=0}^{\infty}\Delta_n^2 < \infty$ it follows $\lim_{n\to\infty}\Delta_n = 0$.
Claim 2: it exists a subsequence $(x^{(n_j)})_{j\in\mathbb{N}}$ with $\tilde{x} = \lim_{j\to\infty} x^{(n_j)}$.
The convergence of $h(x^{(n)})$ follows from the squeeze theorem:

$$H_{-\delta_n}(x^{(n)}, x^{(n-1)}) \leq h(x^{(n)}) \leq H_{\delta_n}(x^{(n)}, x^{(n-1)})$$

as $H(x^{(n)}, x^{(n-1)})$ converges (monotonically decreasing and h bounded below). From the coercivity of h and the Bolzano-Weierstrass theorem the existence of the subsequence follows.
Then:

$$\lim_{j\to\infty} H_{\delta_{n_j+1}}(x^{(n_j+1)}, x^{(n_j)}) = \lim_{j\to\infty} h(x^{(n_j+1)}) + \underbrace{\delta_n\|x^{(n_j+1)} - x^{(n_j)}\|_2}_{\to 0} = H_\delta(\tilde{x}, \tilde{x}) = h(\tilde{x}).$$

Proof Sketch (Theorem)
Remember: it exists $w^{(n)} \in \partial H(z^{(n)})$ such that

$$\|w^{(n)}\|_2 \leq \frac{b}{2}(\Delta_n + \Delta_{n-1}). \tag{H2}$$

The proof is based on the following claim:

$$\sum_{i=1}^{n}\Delta_i \leq \frac{1}{2}(\Delta_0 - \Delta_n) + \frac{b}{a}\left[\phi(H(z^{(1)}) - H(\tilde{z})) - \phi(H(z^{(n+1)}) - H(\tilde{z}))\right]$$

which is shown by induction. Then, it follows $\sum_{n=0}^{\infty}\Delta_n < \infty$ and $x^{(n)} \to \tilde{x}$. Using the Kurdyka-Lojasiewicz property it can be shown that $H_{\delta_n}(z^{(n)}) \to H(\tilde{z})$. With Condition (**??**) it also follows that $\tilde{z}$ is a critical point of H.