A Newton-type Method for Non-smooth Multi-objective Optimization

Sorin-Mihai Grad, Titus Pinta

March 14, 2025
Spring School and Workshop on Variational Analysis and Optimization

Main Results

Given $F : \mathbb{R}^m \to \mathbb{R}^n$, $F(x) = [f_1(x), \dots, f_n(x)]$, with Newton differentiable Jacobian (e.g. piecewise $\mathcal{C}^2$),

Problem

Find

$$\underset{x \in \mathbb{R}^m}{\operatorname{argmin}}^* F(x) \tag{MOP}$$

Result

The Newton-type algorithm converges superlinearly

Outline

1. What is $\arg\min^*$?
2. Under-determined Systems of Equations
3. Non-smooth Equation
4. Constraints
5. Main Result

Efficient Points

No total order on $\mathbb{R}^m$

1. (local) strict Pareto minimizer:
   $\forall y (\in U): \quad F(y) \leq F(\bar{x}) \implies F(y) = F(\bar{x})$
2. (local) weak Pareto minimizer: $\nexists y (\in U): \quad F(y) < F(\bar{x})$
3. (local) weakly efficient $\exists \sigma \in \Delta^n$[1], $\bar{x}$ is a (local) minimizer of $\sigma^T F(x)$

Aim to replace (MOP) with an equation

Let $\sigma \in \Delta^n$, $G(x) = \sigma^T \nabla F(x)$ (first order scalarization)
NOT good: $F(x) = [x_1^2 - x_2, x_2]$ only one good guess for σ

[1] $\Delta^n = \{x \in [0,1]^n \mid \mathbf{1}^T x = 1\}$

First Order Equation

Given $F : \mathbb{R}^m \to \mathbb{R}^n$, $F(x) = [f_1(x), \ldots, f_n(x)]$, define

$$G : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^n, \quad G(x, \sigma) = \sigma^T \nabla F(x)$$

Problem

Find, $\bar{x} \in \mathbb{R}^m, \bar{\sigma} \in \Delta^n$

$$G(\bar{x}, \bar{\sigma}) = 0 \qquad \text{(FOE)}$$

Under-determined Equations

Given $G : \mathbb{R}^m \to \mathbb{R}^n$ in $\mathcal{C}^1$ with $m > n$ Find, $\bar{x} \in \mathbb{R}^m$

$$G(\bar{x}) = 0 \qquad \text{(UDE)}$$

Newton's method: linearize (UDE) and solve

$$\mathcal{A}(x) = \{y \in \mathbb{R}^m \mid G(x) + \nabla G(x)^T (y - x) = 0\} \qquad \text{(EUDE)}$$

$\mathcal{A}(x)$ is **not** a singelton

Under-determined Equations

We can project on $\mathcal{A}(x)$

$$x^{k+1} = \mathrm{P}_{\mathcal{A}(x^k)}\, x^k$$

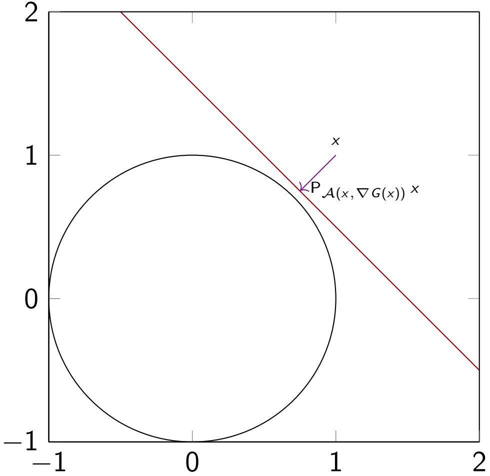
$$x^{k+1} = \underset{G(x)+\nabla G(x^k)^T(x-x^k)=0}{\operatorname{argmin}} \|x - x^k\|^2$$

First order optimality ($\nabla G(x^k)$ full rank)

$$x^{k+1} = x^k - \nabla G(x^k)^{+2} G(x^k)$$

[2] $H^+ = (HH^T)^{-1}H$

Under-determined Equations



Newton Differentiability

Definition

$G : \mathbb{R}^m \to \mathbb{R}^n$ is *pointwise Newton differentiable at $\bar{x}$* if there is $\mathcal{H}G : \mathbb{R}^m \rightrightarrows \mathbb{R}^{n \times m}$ with

$$\lim_{x \to \bar{x}} \sup_{H \in \mathcal{H}G(x)} \frac{\|G(x) - G(\bar{x}) - H(x - \bar{x})\|}{\|x - \bar{x}\|} = 0$$

Definition

$G : \mathbb{R}^m \to \mathbb{R}^n$ is *uniformly Newton differentiable on V* if there is $\mathcal{H}G : \mathbb{R}^m \rightrightarrows \mathbb{R}^{n \times m}$ such that $\forall \varepsilon > 0\ \exists \delta$, $\forall x \in \mathbb{R}^m$, $\forall y \in V$ with $\|x - y\| \leq \delta$,

$$\sup_{H \in \mathcal{H}G(x)} \frac{\|G(x) - G(y) - H(x - y)\|}{\|x - y\|} < \varepsilon.$$

Examples

Proposition

$G : \mathbb{R}^m \to \mathbb{R}^n$, K compact, $G \in \mathcal{C}^1(K)$, then G is uniformly Newton differentiable on K with a Newton differential $\mathcal{H}(x) := \{\nabla G(x)^T\}$

Definition

$G : \mathbb{R}^m \to \mathbb{R}^n$ is Lipschitz *uniformly semi-smooth* on V* if $\forall \varepsilon > 0$ $\exists \delta$ such that $\forall x$, $\forall y \in V$ with $\|x - y\| \leq \delta$,

$$\frac{\|F'(x; x-y) - F'(y; x-y)\|}{\|x-y\|} \leq \varepsilon$$

Proposition

$G : \mathbb{R}^m \to \mathbb{R}^n$, G uniformly semi-smooth* on V, then G is uniformly Newton differentiable on V with a Newton differential $\mathcal{H}(x) := \overline{\text{conv}} \left\{ H \in \mathbb{R}^{n \times m} \mid \exists \{x^k\}_{k \in \mathbb{N}}, \lim_{k \to \mathbb{N}} \nabla G(x^k)^T = H \right\}$

Non-smooth Newton

Given $G : \mathbb{R}^m \to \mathbb{R}^n$ Newton differentiable with $\mathcal{H}G$ Find, $\bar{x} \in \mathbb{R}^m$

$$G(\bar{x}) = 0 \tag{UDE}$$

Denote $\mathcal{Z} = \{x \mid G(x) = 0\}$

Algorithm: $\mathcal{N} : \mathbb{R}^m \rightrightarrows \mathbb{R}^m$,

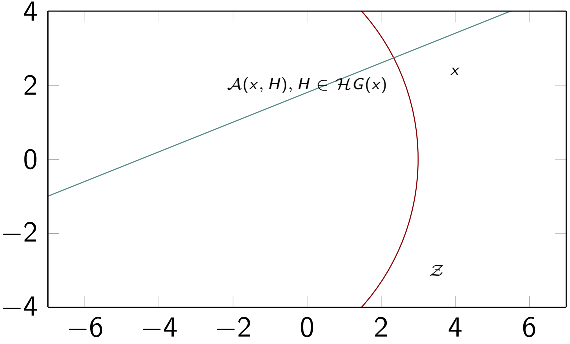
$$\mathcal{N}(x) = \{x - H^+ G(x) \mid H \in \mathcal{H}G(x)\}$$

Geometrically Compatible

Definition

$\mathcal{H}G$ is *geometrically compatible* if

$$\| \mathrm{P}_{\mathcal{A}(x,H)} x - \mathrm{P}_{\mathcal{Z}} \mathrm{P}_{\mathcal{A}(x,H)} x \| \leq P \| \mathrm{P}_{\mathcal{Z}} \mathrm{P}_{\mathcal{A}(x,H)} x - \mathrm{P}_{\mathcal{A}(x,H)} \mathrm{P}_{\mathcal{Z}} \mathrm{P}_{\mathcal{A}(x,H)} x \|$$



If $\mathcal{H}G$ is single-valued and uniformly continuous then it is geometrically compatible

Convergence

Theorem

- *$G : \mathbb{R}^n \to \mathbb{R}^m$ is* uniformly Newton differentiable on $\mathcal{Z}$ *with $\mathcal{H}G$ and Lipschitz*
- *$\mathcal{Z} := \{x \mid G(x) = 0\} \neq \emptyset$, $\mathrm{P}_{\mathcal{Z}}$ is Lipschitz*
- $\sup_{x \in U \in \mathcal{V}(\bar{x})} \sup_{H \in \mathcal{H}G(x)} \|H^+\| \leq \Omega$ *and H full rank*
- *$\mathcal{H}G$ is* geometrically compatible

Then the Newton-type method approaches $\mathcal{Z}$ superlinearly **for any x^0 close enough to $\mathcal{Z}$**

$$\lim_{k \to \infty} \frac{\mathrm{dist}(x^{k+1}, \mathcal{Z})}{\mathrm{dist}(x^k, \mathcal{Z})} = 0$$

Constraints

Given $G : \mathbb{R}^m \to \mathbb{R}^n$ Newton differentiable with $\mathcal{H}G$ Find, $\bar{x} \in K$, K close and convex

$$G(\bar{x}) = 0 \tag{CUDE}$$

Denote $\mathcal{Z} = \{x \mid G(x) = 0\}, \mathcal{K} = K \cap \mathcal{Z}$

Algorithm: $\mathcal{M} : \mathbb{R}^m \rightrightarrows \mathbb{R}^m$,

$$\mathcal{M} = \mathrm{P}_K \circ \mathcal{M}$$

Using non-expansiveness of P_K, $y \in \mathcal{N}(x)$

$$\mathrm{dist}(\mathrm{P}_K\, y, \mathcal{K}) \leq \| \mathrm{P}_K\, y - \mathrm{P}_{\mathcal{K}}\, x\| \leq \|y - \mathrm{P}_{\mathcal{K}}\, x\| \leq c\|x - \mathrm{P}_{\mathcal{K}}\, x\|$$

Back to Multi-objective

Problem
Find, $\bar{x} \in \mathbb{R}^m, \bar{\sigma} \in \Delta^n$

$$\bar{\sigma}^T \nabla F(\bar{x}) = 0 \tag{FOE}$$

Denote $K = \mathbb{R}^m \times \Delta^n, \mathcal{K} = \{x, \sigma \mid \sigma \in \Delta^n, \sigma^T \nabla F(x) = 0\}$

Assume: ∇F is uniformly Newton differentiable on $\mathcal{Z}$ with $\mathcal{H}F$

Algorithm: $H^k \in \mathcal{H}F(x^k)$

$$\begin{bmatrix} x^{k+1} \\ \sigma^{k+1} \end{bmatrix} = \mathrm{P}_K \left(\begin{bmatrix} x^k \\ \sigma^k \end{bmatrix} - \begin{bmatrix} {\sigma^k}^T H^k & \nabla F(x^k) \end{bmatrix}^+ {\sigma^k}^T \nabla F(x^k) \right)$$

Main Result

Theorem

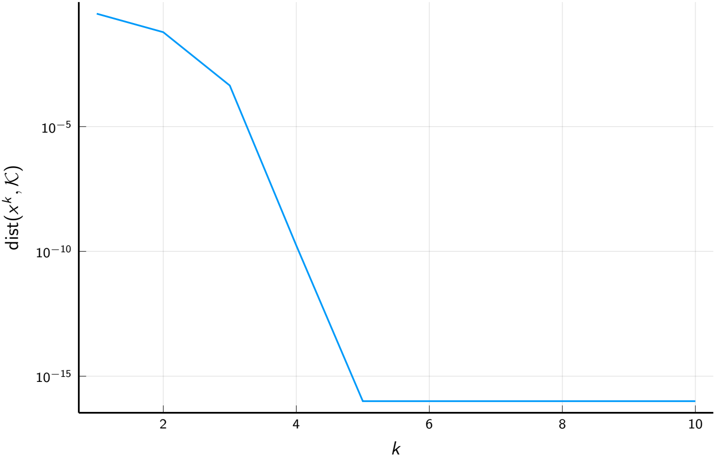
- *$F : \mathbb{R}^n \to \mathbb{R}^m$ is* uniformly Newton differentiable on $\mathcal{Z}$ *with $\mathcal{H}F$ and Lipschitz*
- $\mathcal{Z} := \{x, \sigma \mid \sigma^T \nabla F(x) = 0\} \neq \emptyset$, $\mathrm{P}_{\mathcal{Z}}$ *is Lipschitz*
- $\sup_{x \in U \in \mathcal{V}(\bar{x})} \sup_{H \in \mathcal{H}G(x)} \|H^+\| \leq \Omega$ *and H full rank*
- *$\mathcal{H}F$ is* geometrically compatible

Then the Newton-type method approaches $\mathcal{Z} \cap \mathbb{R}^m \times \Delta^n$ superlinearly **for any x^0 close enough to $\mathcal{Z}$**

$$\lim_{k \to \infty} \frac{\mathrm{dist}(x^{k+1}, \mathcal{Z} \cap \mathbb{R}^m \times \Delta^n)}{\mathrm{dist}(x^k, \mathcal{Z} \cap \mathbb{R}^m \times \Delta^n)} = 0$$

Numerical Example

$$F(x) = [x_1^2 - x_2, x_2], \quad \mathcal{K} = \{0\} \times \mathbb{R} \times \{.5\} \times \{.5\}$$



Thank you!