A survey on Descent Method in Multiobjective optimization

Benar F. Svaiter
IMPA

X Brazilian Workshop on Continuous Optimization
Celebrating Clovis Gonzaga's 70th birthday

Florianópolis, Santa Catarina, March 2014

The problem

- Given $F : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$, $F(x) = (F_1(x), \ldots, F_m(x))$ ($\mathcal{C}^1$ or $\mathcal{C}^2$ or componentwise convex, pseudoconvex etc.)

The problem

- Given $F : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$, $F(x) = (F_1(x), \dots, F_m(x))$ ($\mathcal{C}^1$ or $\mathcal{C}^2$ or componentwise convex, pseudoconvex etc.)
- minimize $F_1(x)$ *and* $F_2(x)$ *and*... *and* $F_m(x)$ for $x \in \Omega$

The problem

- Given $F : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$, $F(x) = (F_1(x), \ldots, F_m(x))$ ($\mathcal{C}^1$ or $\mathcal{C}^2$ or componentwise convex, pseudoconvex etc.)
- minimize $F_1(x)$ *and* $F_2(x)$ *and*... *and* $F_m(x)$ for $x \in \Omega$
- what is an optimum?

The problem

- Given $F : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$, $F(x) = (F_1(x), \dots, F_m(x))$ ($\mathcal{C}^1$ or $\mathcal{C}^2$ or componentwise convex, pseudoconvex etc.)
- minimize $F_1(x)$ *and* $F_2(x)$ *and*... *and* $F_m(x)$ for $x \in \Omega$
- what is an optimum?
- $x^* \in \Omega$ is *Pareto* optimum if:
  $y \in \Omega,\ F(y) \leq F(x^*)$ (componentwise) $\Rightarrow F(y) = F(x^*)$

F is C^1 and $\Omega = \mathbb{R}^n$

- d is a descent direction at x if

F is C^1 and $\Omega = \mathbb{R}^n$

- d is a descent direction at x if $\langle \nabla F_i(x), d \rangle < 0 \;\; i = 1, \ldots, m$

F is $\mathcal{C}^1$ and $\Omega = \mathbb{R}^n$

- d is a descent direction at x if $\langle \nabla F_i(x), d \rangle < 0$ $i = 1, \ldots, m$
- the steepest descent direction?

F is $\mathcal{C}^1$ and $\Omega = \mathbb{R}^n$

- d is a descent direction at x if $\langle \nabla F_i(x), d \rangle < 0 \; i = 1, \dots, m$
- the steepest descent direction?
  d_s solution of

$$\min \max_{i=1,\dots,m} \langle \nabla F_i(x), d \rangle + \|d\|^2/2 \qquad d \in \mathbb{R}^m$$

F is $\mathcal{C}^1$ and $\Omega = \mathbb{R}^n$

- d is a descent direction at x if $\langle \nabla F_i(x), d\rangle < 0$ $i = 1, \ldots, m$
- the steepest descent direction?
  d_s solution of

$$\min \max_{i=1,\ldots,m} \langle \nabla F_i(x), d\rangle + \|d\|^2/2 \qquad d \in \mathbb{R}^m$$

- $m = 1,$

F is $\mathcal{C}^1$ and $\Omega = \mathbb{R}^n$

- d is a descent direction at x if $\langle \nabla F_i(x), d \rangle < 0 \;\; i = 1, \ldots, m$
- the steepest descent direction?
  d_s solution of

$$\min \max_{i=1,\ldots,m} \langle \nabla F_i(x), d \rangle + \|d\|^2/2 \qquad d \in \mathbb{R}^m$$

- $m = 1$, scalar minimization, $F : \mathbb{R}^n \to \mathbb{R}$

F is $\mathcal{C}^1$ and $\Omega = \mathbb{R}^n$

- d is a descent direction at x if $\langle \nabla F_i(x), d \rangle < 0 \;\; i = 1, \dots, m$
- the steepest descent direction?
  d_s solution of

$$\min \max_{i=1,\dots,m} \langle \nabla F_i(x), d \rangle + \|d\|^2/2 \qquad d \in \mathbb{R}^m$$

- $m = 1$, scalar minimization, $F : \mathbb{R}^n \to \mathbb{R}$
  we retrieve Cauchy direction, $d_s = -\nabla F(x)$

Critical points

Critical points

x is critical if $\nexists d$ desc. dir.

Critical points

x is critical if $\nexists d$ desc. dir.

equivalent conditions

Critical points

x is critical if $\nexists d$ desc. dir.

equivalent conditions

- $\text{range}(DF(x)) \bigcap \mathbb{R}^m_{++} = \emptyset$

Critical points

x is critical if $\nexists d$ desc. dir.

equivalent conditions

- range$(DF(x)) \bigcap \mathbb{R}^m_{++} = \emptyset$
- $d_s = 0$

Critical points

x is critical if $\nexists d$ desc. dir.

equivalent conditions

- range$(DF(x)) \bigcap \mathbb{R}^m_{++} = \emptyset$
- $d_s = 0$
- for $m = 1$, $\nabla F(x) = 0$

Critical points

x is critical if $\nexists d$ desc. dir.

equivalent conditions

- $\text{range}(DF(x)) \bigcap \mathbb{R}^m_{++} = \emptyset$
- $d_s = 0$
- for $m = 1$, $\nabla F(x) = 0$

is a necessary condition for (Pareto) optimality

Armijo rule (for a descent direction)

$0 < \beta < 1$,

Armijo rule (for a descent direction)

$0 < \beta < 1$ $(\beta < 1/2)$,

Armijo rule (for a descent direction)

$0 < \beta < 1$ $(\beta < 1/2)$,

d desc. dir. at x

Armijo rule (for a descent direction)

$0 < \beta < 1$ ($\beta < 1/2$),

d desc. dir. at x

$F_j(x + td) \leq F_j(x) + \beta t \max_{i=1,\dots,m} \langle \nabla F_i(x), d \rangle$

Steepest Descent Method

steepest desc. dir.

Steepest Descent Method

steepest desc. dir. + backtracking w. Armijo rule

Steepest Descent Method

steepest desc. dir. + backtracking w. Armijo rule
=Steepest desc. method for multiobjective optimization

Steepest Descent Method

steepest desc. dir. + backtracking w. Armijo rule
=Steepest desc. method for multiobjective optimization

generates $\{x_k\}$

Steepest Descent Method

steepest desc. dir. + backtracking w. Armijo rule
=Steepest desc. method for multiobjective optimization

generates $\{x_k\} \Rightarrow$ cluster points are critical

Steepest Descent Method

steepest desc. dir. + backtracking w. Armijo rule
=Steepest desc. method for multiobjective optimization

generates $\{x_k\} \Rightarrow$ cluster points are critical

(additionally) F componentwise convex

Steepest Descent Method

steepest desc. dir. + backtracking w. Armijo rule
=Steepest desc. method for multiobjective optimization

generates $\{x_k\}\Rightarrow$ cluster points are critical

(additionally) F componentwise convex $\Rightarrow$ convergence to a critical point

Projected Gradient

- Ω is a closed convex set

Projected Gradient

- Ω is a closed convex set
- search direction sol. of

Projected Gradient

- Ω is a closed convex set
- search direction sol. of

$$\min \max_{i=1,\ldots,m} \langle \nabla F_i(x), d \rangle + \|d\|^2/2 \qquad x + d \in \Omega$$

- similar conv. results

Newton Method

Newton Method

- F is $\mathcal{C}^2$

Newton Method

- F is $\mathcal{C}^2$
- quadratic model of F's variation

Newton Method

- F is $\mathcal{C}^2$
- quadratic model of F's variation

$$q_{x,i}(d) = \langle \nabla F_i(x), d \rangle + \frac{1}{2} \langle \nabla^2 F_i(x) d, d \rangle$$

Newton Method

- F is $\mathcal{C}^2$
- quadratic model of F's variation

$$q_{x,i}(d) = \langle \nabla F_i(x), d\rangle + \frac{1}{2}\langle \nabla^2 F_i(x) d, d\rangle$$

- Newton direction,

Newton Method

- F is $\mathcal{C}^2$
- quadratic model of F's variation

$$q_{x,i}(d) = \langle \nabla F_i(x), d \rangle + \frac{1}{2} \langle \nabla^2 F_i(x) d, d \rangle$$

- Newton direction, ($\nabla^2 F_i$ positive def. for $i = 1, \ldots, m$)

Newton Method

- F is $\mathcal{C}^2$
- quadratic model of F's variation

$$q_{x,i}(d) = \langle \nabla F_i(x), d \rangle + \frac{1}{2} \langle \nabla^2 F_i(x) d, d \rangle$$

- Newton direction, ($\nabla^2 F_i$ positive def. for $i = 1, \ldots, m$)
  d_N sol. of

$$\min \max_{i=1,\ldots,m} q_{x,i}(d) \qquad d \in \mathbb{R}^n$$

Newton Method

- F is $\mathcal{C}^2$
- quadratic model of F's variation

$$q_{x,i}(d) = \langle \nabla F_i(x), d \rangle + \frac{1}{2} \langle \nabla^2 F_i(x) d, d \rangle$$

- Newton direction, ($\nabla^2 F_i$ positive def. for $i = 1, \dots, m$)
  d_N sol. of

$$\min \max_{i=1,\dots,m} q_{x,i}(d) \qquad d \in \mathbb{R}^n$$

- this problem is unbounded?

Newton Method

- F is $\mathcal{C}^2$
- quadratic model of F's variation

$$q_{x,i}(d) = \langle \nabla F_i(x), d \rangle + \frac{1}{2} \langle \nabla^2 F_i(x) d, d \rangle$$

- Newton direction, ($\nabla^2 F_i$ positive def. for $i = 1, \ldots, m$)
  d_N sol. of

$$\min \max_{i=1,\ldots,m} q_{x,i}(d) \qquad d \in \mathbb{R}^n$$

- this problem is unbounded?
  use

$$\min \max_{i=1,\ldots,m} q_{x,i}(d) + \mu \|d\|^2/2 \qquad d \in \mathbb{R}^n$$

Newton Method

- F is $\mathcal{C}^2$
- quadratic model of F's variation

$$q_{x,i}(d) = \langle \nabla F_i(x), d \rangle + \frac{1}{2} \langle \nabla^2 F_i(x) d, d \rangle$$

- Newton direction, ($\nabla^2 F_i$ positive def. for $i = 1, \ldots, m$)
  d_N sol. of

$$\min \max_{i=1,\ldots,m} q_{x,i}(d) \qquad d \in \mathbb{R}^n$$

- this problem is unbounded?
  use

$$\min \max_{i=1,\ldots,m} q_{x,i}(d) + \mu \|d\|^2/2 \qquad d \in \mathbb{R}^n$$

for a "large" $\mu \geq 0$,

Newton Method

- F is $\mathcal{C}^2$
- quadratic model of F's variation

$$q_{x,i}(d) = \langle \nabla F_i(x), d \rangle + \frac{1}{2} \langle \nabla^2 F_i(x) d, d \rangle$$

- Newton direction, ($\nabla^2 F_i$ positive def. for $i = 1, \ldots, m$)
  d_N sol. of

$$\min \max_{i=1,\ldots,m} q_{x,i}(d) \qquad d \in \mathbb{R}^n$$

- this problem is unbounded?
  use

$$\min \max_{i=1,\ldots,m} q_{x,i}(d) + \mu \|d\|^2/2 \qquad d \in \mathbb{R}^n$$

for a "large" $\mu \geq 0$, not "too large"

Armijo rule for quad. models (in multiob. opt.)

Armijo rule for quad. models (in multiob. opt.)

d desc. dir at x

Armijo rule for quad. models (in multiob. opt.)

d desc. dir at x

- $\theta = \max_{i=1,\ldots,m} q_{x,i}(d)$

Armijo rule for quad. models (in multiob. opt.)

d desc. dir at x

- $\theta = \max_{i=1,\ldots,m} q_{x,i}(d)$
- $F_i(x + td) \leq F_i(x) + \beta\, t\theta$ for $i = 1, \ldots, m$

prox. for multiob. optim.

$$x_{k+1} \in \arg\min F(x) + w\|x - x_k\|^2/2, \qquad w \in \mathbb{R}^m, w > 0$$

Extension to vector optim.

Extension to vector optim.

$$F(y) \leq F(x) \iff F(x) - F(y) \in \mathbb{R}^m_+$$

Extension to vector optim.

$$F(y) \leq F(x) \iff F(x) - F(y) \in \mathbb{R}^m_+$$

- K closed convex cone

Extension to vector optim.

$$F(y) \leq F(x) \iff F(x) - F(y) \in \mathbb{R}^m_+$$

- K closed convex cone
- $F(y) \preceq_K F(x) \iff F(x) - F(y) \in K$

steepest descent OK

Extension to vector optim.

$$F(y) \leq F(x) \iff F(x) - F(y) \in \mathbb{R}^m_+$$

- K closed convex cone
- $F(y) \preceq_K F(x) \iff F(x) - F(y) \in K$
- x^* is K-optimum if $F(y) \preceq_K F(x^*) \Rightarrow F(y) = F(x^*)$

steepest descent OK
Newton OK

Some open problems

Some open problems

- Duality in multiobjective optim.

Some open problems

- Duality in multiobjective optim.
- Augmented Lagrangian methods

Some open problems

- Duality in multiobjective optim.
- Augmented Lagrangian methods
- Decomposition methods

Some open problems

- Duality in multiobjective optim.
- Augmented Lagrangian methods
- Decomposition methods
- secant methods

- Steepest Descent Methods for Multicriteria Optimization
Jorg Fliege and B. F. Svaiter. *Mathematical Methods of Operations Research*, **51** (2000), 479–494.

- A Projected Gradient Method for Vector Optimization Problems
L. M. Graña Drummond and A. N. Iusem, *Computational Optimization and Applications*, **28** (2004), 5–30.

- A Steepest Descent Method for Vector Optimization
L. M. Grana Drummond e B. F. Svaiter. *Journal of Computational and Aplied Mathematics*, **175** (2005), 395–414.

- Proximal methods in vector optimization
Henri Bonnel, A. N. Iusem e B. F. Svaiter. *SIAM Journal on Optimization*, **15** (2005), 953–970.

- Newton's Method for Multiobjective Optimization
Joerg Fliege, L. M. Graña Drummond and B. F. Svaiter. *SIAM Journal on Optimization*, **20** (2009), 602–626.

- On the Convergence of the Projected Gradient Method for Vector Optimization
E. H. Fukuda and M. L. Graña Drummond. *Optimization*, **60** (2011), 1009–1021.

- A quadratically convergent Newton method for vector optimization
L. M. Grana Drummond, F. M. P. Raupp e B. F. Svaiter. *Optimization*.