Global convergence of filter methods for nonlinear programming

Ademir Alves Ribeiro

Optimization 2007 - Porto - Portugal
1. A general filter algorithm
   - The problem
   - Filter criteria
   - The general algorithm
   - Convergence results

2. Internal algorithms
   - Sequential quadratic programming
   - Inexact restoration

3. Conclusions
The problem

\[
\begin{align*}
\min & \quad f(x) \\
\text{s. t.} & \quad c_E(x) = 0 \\
& \quad c_I(x) \leq 0
\end{align*}
\]

- \( f, c_i : \mathbb{R}^n \to \mathbb{R} \in C^1, i \in E \cup I \)
- \( m = \text{card}(E \cup I) \)
Consider the function $c^+ : \mathbb{R}^n \to \mathbb{R}^m$ given by

$$c_i^+(x) = \begin{cases} 
  c_i(x) & \text{if } i \in E \\
  \max\{0, c_i(x)\} & \text{if } i \in I
\end{cases}$$

and define the infeasibility measure $h : \mathbb{R}^n \to \mathbb{R}$ by

$$h(x) = \|c^+(x)\|$$

**Exact penalty**

$$h(x) = 0 \iff x \text{ is feasible}$$
Consider the function $c^+ : \mathbb{R}^n \rightarrow \mathbb{R}^m$ given by

$$c_i^+(x) = \begin{cases} 
  c_i(x) & \text{if } i \in \mathcal{E} \\
  \max\{0, c_i(x)\} & \text{if } i \in \mathcal{I}
\end{cases}$$

and define the infeasibility measure $h : \mathbb{R}^n \rightarrow \mathbb{R}$ by

$$h(x) = \|c^+(x)\|$$

**Exact penalty**

$$h(x) = 0 \iff x \text{ is feasible}$$
Two goals to minimize:
- objective function $f$
- infeasibility measure $h$

Iterative algorithms: generate a sequence of points

How to decide if $x^+$ is better than $x$?
Two goals to minimize:
- objective function $f$
- infeasibility measure $h$

Iterative algorithms: generate a sequence of points

How to decide if $x^+$ is better than $x$?
Step acceptance criteria

- Two goals to minimize:
  - objective function $f$
  - infeasibility measure $h$

- Iterative algorithms: generate a sequence of points

```
current point $x$
↓
trial point $x^+$
↓
How to decide if $x^+$ is better than $x$?
```
Step acceptance criteria

- Two goals to minimize:
  - objective function $f$
  - infeasibility measure $h$

- Iterative algorithms: generate a sequence of points

```
current point $x$

↓

trial point $x^+$

↓

How to decide if $(f(x^+), h(x^+))$ is better than $(f(x), h(x))$?
```
Step acceptance criteria

Merit function: $\phi_\sigma(x) = f(x) + \sigma h(x)$

$(f(x^+), h(x^+))$ is dominated by $(f(x), h(x))$ if and only if

$\phi_\sigma(x^+) \geq \phi_\sigma(x)$. 

$$
\phi_\sigma(x) = f(x) + \sigma h(x)
$$
Step acceptance criteria

Merit function: \( \phi_\sigma(x) = f(x) + \sigma h(x) \)

\((f(x^+), h(x^+))\) is dominated by \((f(x), h(x))\)

\[\iff \phi_\sigma(x^+) \geq \phi_\sigma(x).\]

Filter - Pareto Domination

\((f(x^+), h(x^+))\) is dominated by \((f(x), h(x))\)

\[\iff f(x^+) \geq f(x) \text{ and } h(x^+) \geq h(x).\]
Filter algorithms

- Fletcher and Leyffer (1997)

- Set of pairs
  \[ F = \{(f^j, h^j), \quad j = 1, \ldots, n_F\} \]

- Forbidden region in \( \mathbb{R}^2 \)
Filter algorithms

- Fletcher and Leyffer (1997)

- Set of pairs
  \[ F = \{(f^j, h^j), \ j = 1, \ldots, n_F\} \]

- Forbidden region in \( \mathbb{R}^2 \)
Filter algorithms

- Fletcher and Leyffer (1997)

- Set of pairs
  \[ F = \{(f^j, h^j), \ j = 1, \ldots, n_F\} \]

- Forbidden region in \( \mathbb{R}^2 \)

**Definition**

\((f, h)\) is forbidden if it is dominated by some pair of \( F \).
Filter criteria

\[ f(x^+) \geq f(x) \]
\[ h(x^+) \geq h(x) \]
Filter criteria

Domination - original filter

\[ f(x^+) \geq f(x) - \alpha h(x) \]
\[ h(x^+) \geq (1 - \alpha)h(x) \]
Filter criteria

**Domination - original filter**

\[
\begin{align*}
    f(x^+) & \geq f(x) - \alpha h(x) \\
    h(x^+) & \geq (1 - \alpha) h(x)
\end{align*}
\]

**Domination - slanting filter**

\[
\begin{align*}
    f(x^+) & \geq f(x) - \alpha h(x^+) \\
    h(x^+) & \geq (1 - \alpha) h(x)
\end{align*}
\]
Filter criteria

**Domination - original filter**
\[
f(x^+) \geq f(x) - \alpha h(x) \\
h(x^+) \geq (1 - \alpha)h(x)
\]

**Domination - slanting filter**
\[
f(x^+) \geq f(x) - \alpha h(x^+) \\
h(x^+) \geq (1 - \alpha)h(x)
\]
The general filter algorithm

Data: $x^0 \in \mathbb{R}^n$, $F_0 = \emptyset$ (initial filter)

$k = 0$

REPEAT (while $x^k$ is not stationary)
   define the temporary filter
   compute $x^{k+1}$ not forbidden by the filter
   update the filter
   $k = k + 1$. 

Ribeiro, A. A.  Filter Methods for NLP
The temporary filter

Original filter

Slanting filter
Filter update

The permanent filter does not change

The temporary filter becomes permanent

Ribeiro, A. A.
Filter Methods for NLP
Filter update

\[ f \text{-iteration} \]

The permanent filter does not change

\[ h \text{-iteration} \]
The problem
Filter criteria
The general algorithm
Conclusions

A general filter algorithm
Internal algorithms
Conclusions

Filter update

\[ f \text{-iteration} \]
The permanent filter
does not change

\[ h \text{-iteration} \]
Filter update

- **f-iteration**
  - The permanent filter does not change

- **h-iteration**
  - The temporary filter becomes permanent
Classical hypotheses

(1) All functions $f, c_i(\cdot), i = 1, \ldots, m$, are Lipschitz continuously differentiable.

(2) The sequence $(x^k)_{k \in \mathbb{N}}$ remains in a convex compact domain $X \subset \mathbb{R}^n$.

(3) Every feasible accumulation point $\bar{x} \in X$ of $(x^k)_{k \in \mathbb{N}}$ satisfies the Mangasarian-Fromovitz constraint qualification.
Feasibility results

Original filter

If the set of $h$-iterations $\mathcal{K}_a$ is infinite, then $h(x^k) \xrightarrow{\mathcal{K}_a} 0$. 
Feasibility results

Original filter

If the set of $h$-iterations $\mathcal{K}_x$ is infinite, then $h(x^k) \xrightarrow{\mathcal{K}_x} 0$.

Slanting filter

The whole sequence $h(x^k)$ tends to zero.
(H) Given a feasible nonstationary point $\bar{x} \in X$, there exist $M > 0$ and a neighborhood $V$ of $\bar{x}$ such that if $x^k \in V$, then

$$f(x^k) - f(x^{k+1}) \geq Mv_k.$$ 

$$v_k = \min \left\{ 1, \min \left\{ \tilde{h}^j \mid (\tilde{f}^j, \tilde{h}^j) \in F_k \right\} \right\}$$
(H) Given a feasible nonstationary point \( \bar{x} \in X \), there exist \( M > 0 \) and a neighborhood \( V \) of \( \bar{x} \) such that if \( x^k \in V \), then

\[
f(x^k) - f(x^{k+1}) \geq Mv_k.
\]

\(v_k = \min \left\{ 1, \min \{ \tilde{h}^j \mid (\tilde{f}^j, \tilde{h}^j) \in F_k \} \right\}\)
(H) Given a feasible nonstationary point $\bar{x} \in X$, there exist $M > 0$ and a neighborhood $V$ of $\bar{x}$ such that if $x^k \in V$, then

$$f(x^k) - f(x^{k+1}) \geq M v_k.$$
Theorem

The sequence generated by the algorithm has a stationary accumulation point.
Internal algorithms

- Sequential quadratic programming
- Inexact restoration
Sequential quadratic programming

- Fletcher, Gould, Leyffer, Toint and Wächtter (2002)
Sequential quadratic programming

Result

If the step is obtained by sequential quadratic programming, then the Hypothesis (H) is satisfied.

- Slanting filter: Karas and Ribeiro (2007)
Inexact restoration

Martínez and Pilotta (1999)
Result

If the step is obtained by inexact restoration, then the Hypothesis (H) is satisfied.

- Slanting filter: Karas, Oening and Ribeiro (2007)
Inexact restoration

Global convergence - stronger result

*If the step is obtained by inexact restoration and the criterion is the slanting filter, then any accumulation point is stationary.*

- Slanting filter: Karas, Oening and Ribeiro (2007)
A general filter algorithm for NLP

- Features: great deal of freedom in the step computation
- Step acceptance: original or slanting filter
- Main hypothesis: near a feasible nonstationary point, the reduction of the objective function $f$ is large
- Result: Global convergence of the algorithm

Step computation

- Sequential quadratic programming
- Inexact Restoration
Conclusions

- A general filter algorithm for NLP
  - Features: great deal of freedom in the step computation
  - Step acceptance: original or slanting filter
  - Main hypothesis: near a feasible nonstationary point, the reduction of the objective function $f$ is large
  - Result: Global convergence of the algorithm

- Step computation
  - Sequential quadratic programming
  - Inexact Restoration
References on filter methods

R. Fletcher and S. Leyffer.
Nonlinear programming without a penalty function.

Global convergence of trust-region and SQP-filter algorithms for general nonlinear programming.
References on slanting filter

C. M. Chin and R. Fletcher.
On the Global Convergence of an SLP-Filter Algorithm that takes EQP steps.

R. Fletcher, S. Leyffer, and P. L. Toint.
On the Global Convergence of a Filter-SQP Algorithm.
J. M. Martínez and E. A. Pilotta.

C. C. Gonzaga, E. W. Karas, and M. Vanti.
References of our research

Global convergence of filter methods for nonlinear programming.

E. W. Karas, A. P. Oening and A. A. Ribeiro.
Global convergence of slanting filter methods for nonlinear programming.
Contact

Ademir Alves Ribeiro

Federal University of Paraná
ademir@mat.ufpr.br
www.mat.ufpr.br/∼ademir