Otimização Contínua: Aspectos teóricos e computacionais

Ademir Alves Ribeiro
Elizabeth Wegner Karas

Capítulo 8 -Métodos para Otimização com Restrições

- Otimização não linear com restrições.
- Princípio: A solução de um problema “difícil” é aproximada por uma sequência de pontos obtidos como solução de um problema “fácil”, que muda a cada iteração de acordo com as informações disponíveis no ponto corrente.
- A cada iteração, substituir a função objetivo por um modelo quadrático do Lagrangiano e as restrições por aproximações de Taylor de primeira ordem em torno do ponto corrente.

Método de Programação Quadrática Sequencial (PQS)

- Otimização não linear com restrições.
- Princípio: A solução de um problema “difícil” é aproximada por uma sequência de pontos obtidos como solução de um problema “fácil”, que muda a cada iteração de acordo com as informações disponíveis no ponto corrente.
- A cada iteração, substituir a função objetivo por um modelo quadrático do Lagrangiano e as restrições por aproximações de Taylor de primeira ordem em torno do ponto corrente.

Método de Programação Quadrática Sequencial (PQS)

- Otimização não linear com restrições.
- Princípio: A solução de um problema "difícil" é aproximada por uma sequência de pontos obtidos como solução de um problema "fácil", que muda a cada iteração de acordo com as informações disponíveis no ponto corrente.
- A cada iteração, substituir a função objetivo por um modelo quadrático do Lagrangiano e as restrições por aproximações de Taylor de primeira ordem em torno do ponto corrente.

O problema

$$\begin{array}{ll} \text{minimizar} & f(x) \\ \text{sujeito a} & c(x) = 0 \end{array}$$

onde $f : \mathbb{R}^n \to \mathbb{R}$ e $c : \mathbb{R}^n \to \mathbb{R}^m$ são funções de classe $\mathcal{C}^2$.

Lagrangiano associado ao problema

$$x \in \mathbb{R}^n, \lambda \in \mathbb{R}^m \mapsto \ell(x,\lambda) = f(x) + \lambda^T c(x)$$

onde λ é o multiplicador de Lagrange.

Ideia do algoritmo

Dados x^k e λ^k, considere d^k solução do

Subproblema quadrático

$$\begin{array}{ll} \text{minimizar} & \ell(x^k, \lambda^k) + \nabla_x \ell(x^k, \lambda^k)^T d + \frac{1}{2} d^T B(x^k, \lambda^k) d \\ \text{sujeito a} & A(x^k) d + c(x^k) = 0, \end{array}$$

onde

- $A(x)$: matriz Jacobiana de c no ponto x;
- $B(x, \lambda)$: Hessiana parcial do lagrangiano, $\nabla^2_{xx} \ell(x, \lambda)$.

Definimos então $x^{k+1} = x^k + d^k$, calculamos o multiplicador de Lagrange λ^{k+1} e repetimos o processo com o novo problema quadrático.

Ideia do algoritmo

Dados x^k e λ^k, considere d^k solução do

Subproblema quadrático

$$\begin{array}{ll} \text{minimizar} & \ell(x^k,\lambda^k) + \nabla_x \ell(x^k,\lambda^k)^T d + \frac{1}{2} d^T B(x^k,\lambda^k) d \\ \text{sujeito a} & A(x^k)d + c(x^k) = 0, \end{array}$$

onde

- $A(x)$: matriz Jacobiana de c no ponto x;
- $B(x,\lambda)$: Hessiana parcial do lagrangiano, $\nabla^2_{xx}\ell(x,\lambda)$.

Definimos então $x^{k+1} = x^k + d^k$, calculamos o multiplicador de Lagrange λ^{k+1} e repetimos o processo com o novo problema quadrático.

O algoritmo básico de Programação Quadrática Sequencial

Dados: $(x^0, \lambda^0) \in \mathbb{R}^n \times \mathbb{R}^m$
$k = 0$
ENQUANTO $\nabla \ell(x^k, \lambda^k) \neq 0$

  Resolva o subproblema quadrático,

    obtendo uma solução primal-dual (d^k, ξ^k)

  Faça $x^{k+1} = x^k + d^k$

  Defina $\lambda^{k+1} = \lambda^k + \xi^k$

  $k = k + 1$

Solução primal-dual (d^k, ξ^k) do subproblema quadrático

Subproblema quadrático

$$\begin{array}{ll} \text{minimizar} & \ell(x^k,\lambda^k) + \nabla_x \ell(x^k,\lambda^k)^T d + \dfrac{1}{2} d^T B(x^k,\lambda^k) d \\ \text{sujeito a} & A(x^k)d + c(x^k) = 0, \end{array}$$

Condições de otimalidade (KKT) do subproblema

$$\begin{array}{lcl} B(x^k,\lambda^k)d + A(x^k)^T \xi & = & -\nabla_x \ell(x^k,\lambda^k) \\ A(x^k)d & = & -c(x^k). \end{array}$$

Seja (x^*, λ^*) uma solução primal-dual do problema original.

Hipóteses

- As funções $\nabla^2 f$ e $\nabla^2 c_i$, $i = 1, \ldots, m$, são lipschitzianas em uma vizinhança de x^*.
- A Jacobiana das restrições, $A(x^*)$, tem posto linha completo e a Hessiana parcial $B(x^*, \lambda^*)$ é definida positiva no espaço nulo de $A(x^*)$, isto é, $d^T B(x^*, \lambda^*) d > 0$ para todo $d \neq 0$, $d \in \mathcal{N}(A(x^*))$.

Lema

Existe uma vizinhança V_1 de (x^*, λ^*), tal que se $(x^k, \lambda^k) \in V_1$, o sistema KKT do subproblema quadrático tem uma única solução (d^k, ξ^k).

Teorema

Existe uma vizinhança V de (x^*, λ^*), tal que se $(x^0, \lambda^0) \in V$, o Algoritmo de PQS está bem definido e, se o critério de parada não for satisfeito, gera uma sequência $(x^k, \lambda^k)_{k \in \mathbb{N}}$ que converge quadraticamente para esta solução.

Lema

Existe uma vizinhança V_1 de (x^*, λ^*), tal que se $(x^k, \lambda^k) \in V_1$, o sistema KKT do subproblema quadrático tem uma única solução (d^k, ξ^k).

Teorema

Existe uma vizinhança V de (x^*, λ^*), tal que se $(x^0, \lambda^0) \in V$, o Algoritmo de PQS está bem definido e, se o critério de parada não for satisfeito, gera uma sequência $(x^k, \lambda^k)_{k \in \mathbb{N}}$ que converge quadraticamente para esta solução.

Condições de otimalidade (KKT) do subproblema

$$\begin{array}{rcl} B(x^k,\lambda^k)d + A(x^k)^T\xi & = & -\nabla_x \ell(x^k,\lambda^k) \\ A(x^k)d & = & -c(x^k). \end{array}$$

Faça $\mu = \xi + \lambda^k$.

$$\left\{ \begin{array}{rcl} B(x^k,\lambda^k)d + \nabla f(x^k) + A(x^k)^T\mu & = & 0 \\ A(x^k)d + c(x^k) & = & 0, \end{array} \right.$$

que representa as condições de otimalidade do problema quadrático

$$\begin{array}{ll} \text{minimizar} & f(x^k) + \nabla f(x^k)^T d + \frac{1}{2} d^T B(x^k,\lambda^k) d \\ \text{sujeito a} & A(x^k)d + c(x^k) = 0. \end{array}$$

Interpretação: Minimizamos a cada iteração um modelo quadrático da função objetivo, sujeito à linearização das restrições. Mas na Hessiana do modelo incorporamos informações sobre a curvatura das restrições.

Outra interpretação

Condições de otimalidade (KKT) do subproblema

$$\begin{array}{rcl} B(x^k,\lambda^k)d + A(x^k)^T\xi & = & -\nabla_x \ell(x^k,\lambda^k) \\ A(x^k)d & = & -c(x^k). \end{array}$$

Faça $\mu = \xi + \lambda^k$.

$$\left\{ \begin{array}{rcl} B(x^k,\lambda^k)d + \nabla f(x^k) + A(x^k)^T\mu & = & 0 \\ A(x^k)d + c(x^k) & = & 0, \end{array} \right.$$

que representa as condições de otimalidade do problema quadrático

$$\begin{array}{ll} \text{minimizar} & f(x^k) + \nabla f(x^k)^T d + \frac{1}{2} d^T B(x^k,\lambda^k) d \\ \text{sujeito a} & A(x^k)d + c(x^k) = 0. \end{array}$$

Interpretação: Minimizamos a cada iteração um modelo quadrático da função objetivo, sujeito à linearização das restrições. Mas na Hessiana do modelo incorporamos informações sobre a curvatura das restrições.

Outra interpretação

Condições de otimalidade (KKT) do subproblema

$$
\begin{aligned}
B(x^k,\lambda^k)d + A(x^k)^T\xi &= -\nabla_x \ell(x^k,\lambda^k) \\
A(x^k)d &= -c(x^k).
\end{aligned}
$$

Faça $\mu = \xi + \lambda^k$.

$$
\left\{
\begin{aligned}
B(x^k,\lambda^k)d + \nabla f(x^k) + A(x^k)^T\mu &= 0 \\
A(x^k)d + c(x^k) &= 0,
\end{aligned}
\right.
$$

que representa as condições de otimalidade do problema quadrático

$$
\begin{array}{ll}
\text{minimizar} & f(x^k) + \nabla f(x^k)^T d + \dfrac{1}{2} d^T B(x^k,\lambda^k) d \\
\text{sujeito a} & A(x^k)d + c(x^k) = 0.
\end{array}
$$

Interpretação: Minimizamos a cada iteração um modelo quadrático da função objetivo, sujeito à linearização das restrições. Mas na Hessiana do modelo incorporamos informações sobre a curvatura das restrições.

Outra interpretação

Condições de otimalidade (KKT) do subproblema

$$\begin{array}{rcl} B(x^k,\lambda^k)d + A(x^k)^T\xi & = & -\nabla_x \ell(x^k,\lambda^k) \\ A(x^k)d & = & -c(x^k). \end{array}$$

Faça $\mu = \xi + \lambda^k$.

$$\left\{ \begin{array}{rcl} B(x^k,\lambda^k)d + \nabla f(x^k) + A(x^k)^T\mu & = & 0 \\ A(x^k)d + c(x^k) & = & 0, \end{array} \right.$$

que representa as condições de otimalidade do problema quadrático

$$\begin{array}{ll} \text{minimizar} & f(x^k) + \nabla f(x^k)^T d + \frac{1}{2} d^T B(x^k,\lambda^k) d \\ \text{sujeito a} & A(x^k)d + c(x^k) = 0. \end{array}$$

Interpretação: Minimizamos a cada iteração um modelo quadrático da função objetivo, sujeito à linearização das restrições. Mas na Hessiana do modelo incorporamos informações sobre a curvatura das restrições.

Discussão

Hessiana do modelo precisa incorporar informações sobre a curvatura das restrições.

Subproblema quadrático

$$\begin{array}{ll} \text{minimizar} & f(x^k) + \nabla f(x^k)^T d + \frac{1}{2} d^T B(x^k, \lambda^k) d \\ \text{sujeito a} & A(x^k) d + c(x^k) = 0. \end{array}$$

Discussão

Hessiana do modelo precisa incorporar informações sobre a curvatura das restrições.

Subproblema quadrático

$$\begin{array}{ll} \text{minimizar} & f(x^k) + \nabla f(x^k)^T d + \frac{1}{2} d^T \nabla^2 f(x^k) d \\ \text{sujeito a} & A(x^k) d + c(x^k) = 0. \end{array}$$

O algoritmo pode não funcionar bem como mostra o próximo exemplo.

Exemplo

Problema

$$\begin{array}{ll} \text{minimizar} & f(x) = -\dfrac{x_1^2}{2} + 2x_2 \\ \text{sujeito a} & c(x) = x_1^2 + x_2^2 - 1 = 0, \end{array}$$

Solução: $x^* = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$, com multiplicador $\lambda^* = 1$.

Ponto corrente: $x = \begin{pmatrix} \delta \\ -\sqrt{1-\delta^2} \end{pmatrix}$, com $\delta > 0$ suficientemente pequeno.

Calcule o novo iterando supondo que o passo foi obtido por:

- PQS.
- minimização do subproblema em que o modelo é o polinômio de Taylor de ordem 2. Note que neste caso, o novo ponto se distancia da solução.

CENGAGE Learning

Resolução por Programação Quadrática Sequencial

Subproblema quadrático - PQS

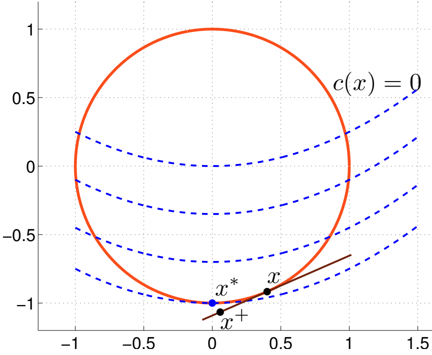
$$\begin{array}{ll} \text{minimizar} & \frac{1}{2}(d_1^2 + 2d_2^2) - \delta d_1 + 2d_2 \\ \text{sujeito a} & \delta d_1 - \sqrt{1-\delta^2}\, d_2 = 0, \end{array}$$

cuja solução é o vetor

$$d_{\text{pqs}} = \frac{(\sqrt{1-\delta^2} - 2)\delta}{1+\delta^2} \begin{pmatrix} \sqrt{1-\delta^2} \\ \delta \end{pmatrix}.$$

Neste caso,

$$\frac{\|x + d_{\text{pqs}} - x^*\|}{\|x - x^*\|^2} \approx \frac{1}{2}.$$



$x^+ = x + d_{\text{pqs}}$

Resolução com modelo obtido por Taylor

Subproblema quadrático

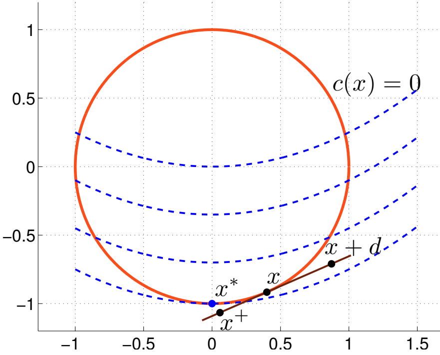
$$\begin{array}{ll} \text{minimizar} & -\dfrac{d_1^2}{2} - \delta d_1 + 2d_2 \\ \text{sujeito a} & \delta d_1 - \sqrt{1-\delta^2}\, d_2 = 0 \end{array}$$

Pelas condições de KKT do problema, temos

$$d = \begin{pmatrix} \dfrac{2\delta}{\sqrt{1-\delta^2}} - \delta \\ \dfrac{2\delta^2}{1-\delta^2} - \dfrac{\delta^2}{\sqrt{1-\delta^2}} \end{pmatrix}$$

Para δ suficientemente pequeno o ponto x fica muito próximo da solução x^*. No entanto,

$$\|x+d-x^*\| \approx 2\|x-x^*\|.$$



O problema geral

$$
\begin{array}{ll}
\text{minimizar} & f(x) \\
\text{sujeito a} & c_{\mathcal{E}}(x) = 0 \\
 & c_{\mathcal{I}}(x) \leq 0
\end{array}
$$

- $f, c_i : \mathbb{R}^n \to \mathbb{R} \in \mathcal{C}^2,\ i \in \mathcal{E} \cup \mathcal{I}$
- $m = \text{card}(\mathcal{E} \cup \mathcal{I})$

CENGAGE Learning

Medida de inviabilidade

Considere a função $c^+ : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ dada por

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

e a medida de inviabilidade $h : \mathbb{R}^n \longrightarrow \mathbb{R}$ dada por

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

Penalidade Exata

$h(x) = 0 \iff x$ é viável

Medida de inviabilidade

Considere a função $c^+ : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ dada por

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

e a medida de inviabilidade $h : \mathbb{R}^n \longrightarrow \mathbb{R}$ dada por

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

Penalidade Exata

$h(x) = 0 \Longleftrightarrow x$ é viável

Critério de aceitação do passo

- Duas funções a serem minimizadas:
  - função objetivo f
  - medida de inviabilidade h
- Algoritmo iterativo: gera uma sequência de pontos

$$\text{ponto corrente } x$$
$$\downarrow$$
$$\text{ponto tentativo } x^+$$
$$\downarrow$$
$$\text{Como decidir se } x^+ \text{ é melhor que } x?$$

Critério de aceitação do passo

- Duas funções a serem minimizadas:
  - função objetivo f
  - medida de inviabilidade h
- Algoritmo iterativo: gera uma sequência de pontos

ponto corrente x

$\downarrow$

ponto tentativo x^+

$\downarrow$

Como decidir se x^+ é melhor que x?

Critério de aceitação do passo

- Duas funções a serem minimizadas:
  - função objetivo f
  - medida de inviabilidade h
- Algoritmo iterativo: gera uma sequência de pontos

ponto corrente x

$\downarrow$

ponto tentativo x^+

$\downarrow$

Como decidir se x^+ é melhor que x?

CENGAGE Learning

Critério de aceitação do passo

- Duas funções a serem minimizadas:
  - função objetivo f
  - medida de inviabilidade h
- Algoritmo iterativo: gera uma sequência de pontos

ponto corrente x

$\downarrow$

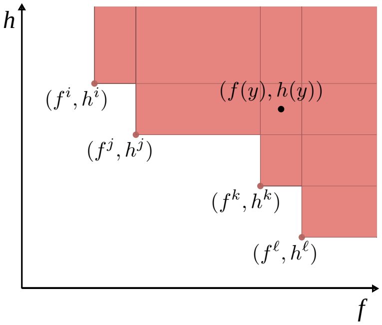
ponto tentativo x^+

$\downarrow$

Como decidir se $(f(x^+), h(x^+))$ é melhor que $(f(x), h(x))$?

Métodos de Filtro

- Fletcher e Leyffer (2002)
- Conjunto de pares
  $F = \{(f^j, h^j),\ \ j = 1, \dots, n_F\}$
- y é proibido pelo filtro se o par $(f(y), h(y))$ for dominado por algum elemento de F



Definição

$(f(y), h(y))$ é dominado por $(f(x), h(x)) \Leftrightarrow f(y) \geq f(x)$ e $h(y) \geq h(x)$.

Diferentes filtros

Filtro Original (Fletcher e Leyffer, 2002)

$$\mathcal{R}_j = \left\{ x \in \mathbb{R}^n \mid f(x) \geq f(x^j) - \alpha h(x^j) \text{ e } h(x) \geq (1-\alpha) h(x^j) \right\}$$

Filtro Inclinado (Chin, 2003)

$$\mathcal{R}_j = \left\{ x \in \mathbb{R}^n \mid f(x) \geq f(x^j) - \alpha h(x) \text{ e } h(x) \geq (1-\alpha) h(x^j) \right\}.$$

Diferentes filtros

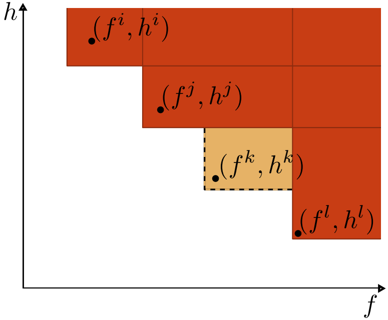
Filtro Original (Fletcher e Leyffer, 2002)

$$\mathcal{R}_j = \left\{x \in \mathbb{R}^n \mid f(x) \geq f(x^j) - \alpha h(x^j) \text{ e } h(x) \geq (1-\alpha)h(x^j)\right\}$$

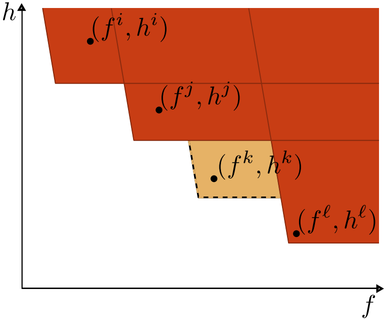
Filtro Inclinado (Chin, 2003)

$$\mathcal{R}_j = \left\{x \in \mathbb{R}^n \mid f(x) \geq f(x^j) - \alpha h(x) \text{ e } h(x) \geq (1-\alpha)h(x^j)\right\}.$$

- Filtro permanente: $F_k = \left\{ (f^i, h^i), (f^j, h^j), (f^l, h^l) \right\}$
- Filtro temporário: $\bar{F}_k = F_k \cup \left\{ (f^k, h^k) \right\}$



Original



Inclinado

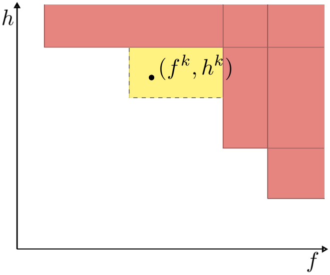
CENGAGE Learning

Dados: $x^0 \in \mathbb{R}^n$, $F_0 = \emptyset$, $\mathcal{F}_0 = \emptyset$, $\alpha \in (0,1)$.
$k = 0$
REPITA
Defina $\bar{F}_k = F_k \bigcup \{(f^k, h^k)\}$ e
$\bar{\mathcal{F}}_k = \mathcal{F}_k \bigcup \mathcal{R}_k$, com $\mathcal{R}_k$ original ou inclinado
Passo:
SE x^k é estacionário, pare com sucesso
SENÃO, calcule $x^{k+1} \notin \bar{\mathcal{F}}_k$.
Atualização do filtro:
SE $f(x^{k+1}) < f(x^k)$,
$F_{k+1} = F_k, \quad \mathcal{F}_{k+1} = \mathcal{F}_k$ (iteração f)
SENÃO,
$F_{k+1} = \bar{F}_k, \quad \mathcal{F}_{k+1} = \bar{\mathcal{F}}_k$ (iteração h)
$k = k + 1$.

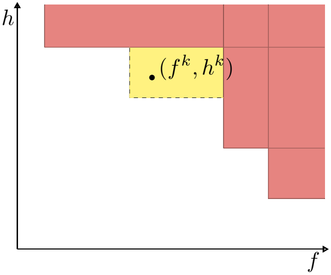
Caso em que o ponto corrente é viável

Se x^k é viável, então qualquer x não proibido deve satisfazer $f(x) < f(x^k)$.

Tipos de iteração

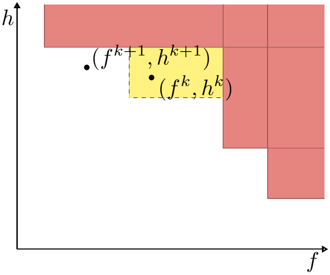


Iteração do tipo f

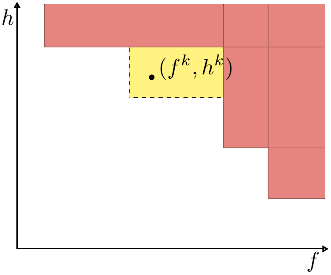


Iteração do tipo h

Tipos de iteração

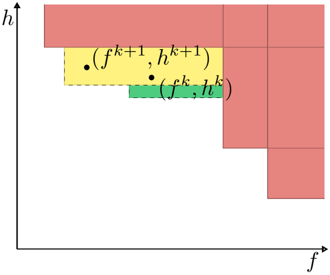


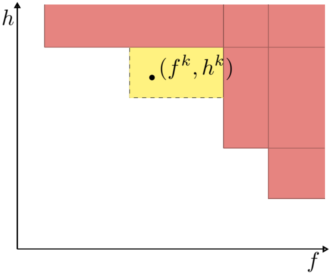
Iteração do tipo f
O filtro permanente
não muda



Iteração do tipo h

Tipos de iteração

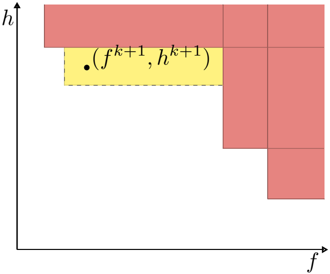




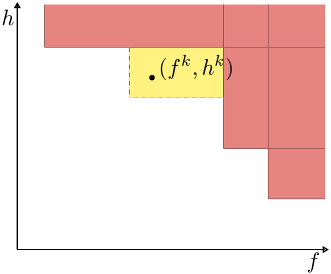
Iteração do tipo f
O filtro permanente
não muda

Iteração do tipo h

Tipos de iteração

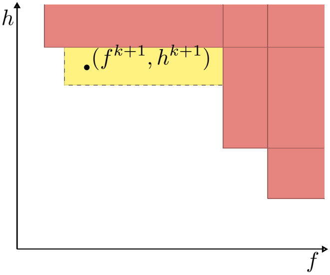


Iteração do tipo f
O filtro permanente
não muda

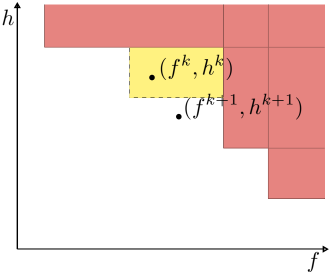


Iteração do tipo h

Tipos de iteração

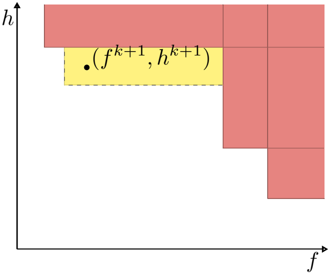


Iteração do tipo f
O filtro permanente
não muda

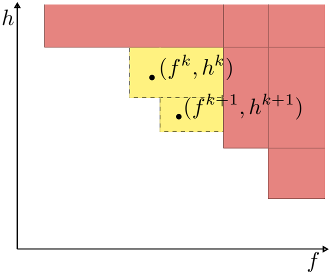


Iteração do tipo h
O filtro temporário
torna-se permanente

Tipos de iteração

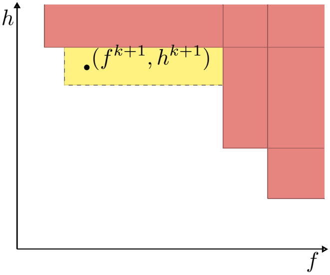


Iteração do tipo f
O filtro permanente
não muda

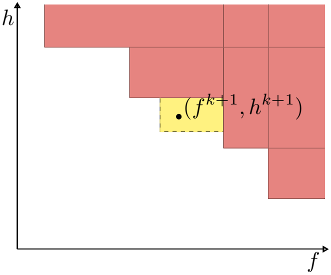


Iteração do tipo h
O filtro temporário
torna-se permanente

Tipos de iteração



Iteração do tipo f
O filtro permanente
não muda



Iteração do tipo h
O filtro temporário
torna-se permanente

O algoritmo está bem definido

Lema

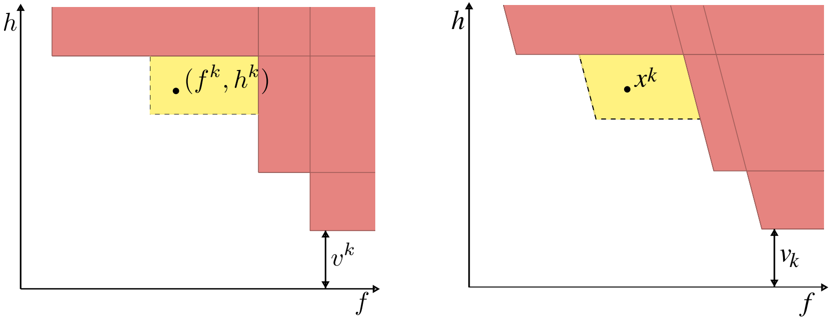
Para todo $k \in \mathbb{N}$ tal que x^k é não estacionário, as seguintes afirmações são válidas:

(i) $h^j > 0$, para todo $j \in \mathbb{N}$ tal que $(f^j, h^j) \in F_k$;

(ii) Existe $x^{k+1} \notin \bar{\mathcal{F}}_k$.

Altura do filtro

$$v_k = \min\left\{1, \min\left\{(1-\alpha)h^j \mid \left(f^j, h^j\right) \in F_k\right\}\right\}$$



Hipóteses

- A sequência (x^k) permanece em um conjunto convexo e compacto $X \subset \mathbb{R}^n$.
- As funções f, c_i, $i \in \mathcal{E} \cup \mathcal{I}$, são duas vezes continuamente diferenciáveis.
- Dado um ponto viável não estacionário $\bar{x} \in X$, existem $M > 0$ e uma vizinhança V de $\bar{x}$ tal que se $x^k \in V$, então

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

onde $v_k = \min\left\{1, \min\left\{(1-\alpha)h^j \mid \left(f^j, h^j\right) \in F_k\right\}\right\}$ é a altura do filtro.

Convergência global

Hipóteses

- A sequência (x^k) permanece em um conjunto convexo e compacto $X \subset \mathbb{R}^n$.
- As funções f, c_i, $i \in \mathcal{E} \cup I$, são duas vezes continuamente diferenciáveis.
- Dado um ponto viável não estacionário $\bar{x} \in X$, existem $M > 0$ e uma vizinhança V de $\bar{x}$ tal que se $x^k \in V$, então

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

onde $v_k = \min\left\{1, \min\left\{(1-\alpha)h^j \mid \left(f^j, h^j\right) \in F_k\right\}\right\}$ é a altura do filtro.

Convergência global

Hipóteses

- A sequência (x^k) permanece em um conjunto convexo e compacto $X \subset \mathbb{R}^n$.
- As funções f, c_i, $i \in \mathcal{E} \cup \mathcal{I}$, são duas vezes continuamente diferenciáveis.
- Dado um ponto viável não estacionário $\bar{x} \in X$, existem $M > 0$ e uma vizinhança V de $\bar{x}$ tal que se $x^k \in V$, então

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

onde $v_k = \min\left\{1, \min\left\{(1-\alpha)h^j \mid \left(f^j, h^j\right) \in F_k\right\}\right\}$ é a altura do filtro.

Convergência global

Resultados sobre viabilidade

- Existe um conjunto infinito $\mathbb{N}' \subset \mathbb{N}$ tal que $h(x^k) \overset{\mathbb{N}'}{\longrightarrow} 0$.
- Se o filtro inclinado for usado, então $h(x^k) \to 0$, ou seja, qualquer ponto de acumulação da sequência (x^k) é viável.

Convergência global

Resultados sobre viabilidade

- Existe um conjunto infinito $\mathbb{N}' \subset \mathbb{N}$ tal que $h(x^k) \xrightarrow{\mathbb{N}'} 0$.
- Se o filtro inclinado for usado, então $h(x^k) \to 0$, ou seja, qualquer ponto de acumulação da sequência (x^k) é viável.

Convergência global

Considere o conjunto das iterações h dado por

$$\mathcal{K}_a = \left\{ k \in \mathbb{N} \mid \left(f^k, \, h^k \right) \text{ é adicionado ao filtro} \right\}.$$

Lema

Se $\bar{x} \in X$ é um ponto não estacionário, então nenhuma subsequência de $(x^k)_{k \in \mathcal{K}_a}$ converge para $\bar{x}$.

Interpretação: em uma vizinhança de $\bar{x}$, toda iteração k é do tipo f.

Convergência global

Considere o conjunto das iterações h dado por

$$\mathcal{K}_a = \left\{ k \in \mathbb{N} \mid \left(f^k,\ h^k \right) \text{ é adicionado ao filtro} \right\}.$$

Lema

Se $\bar{x} \in X$ é um ponto não estacionário, então nenhuma subsequência de $(x^k)_{k \in \mathcal{K}_a}$ converge para $\bar{x}$.

Interpretação: em uma vizinhança de $\bar{x}$, toda iteração k é do tipo f.

CENGAGE Learning

Convergência global

Teorema

A sequência (x^k) gerada pelo algoritmo de filtro tem um ponto de acumulação estacionário.

Principais Referências Bibliográficas

R. Fletcher e S. Leyffer.
Nonlinear programming without a penalty function.
Mathematical Programming - Ser. A, 91(2):239–269, 2002.

C. C. Gonzaga, E. W. Karas e M. Vanti.
A globally convergent filter method for nonlinear programming.
SIAM J. Optimization, 14(3):646–669, 2003.

G. A. Periçaro, A. A. Ribeiro e E. W. Karas.
Global Convergence of a General Filter Algorithm Based on an Efficiency Condition of the Step.
Applied Mathematics and Computation, v. 219, (17), pp. 9581-9597, 2013.