Uma Ferramenta para otimização em Engenharia Mecânica e aplicações na Fundição Eletromagnética de Metais

Alfredo Canelas, Jean R. Roche, José Herskovits

Departamento de Engenharia Mecânica
COPPE – UFRJ

STIC-AMSUD, Novembro de 2009

Conteúdo

Conteúdo

Objetivo

1) Definir e implementar um método numérico para problemas de otimização da Engenharia Mecânica com as seguintes características:
   - O Método dos Elementos de Contorno (MEC) é utilizado para calcular a solução da equação de estado do problema.
   - A formulação SAND é utilizada para definir o problema de otimização.
   - Um algoritmo de ponto interior é utilizado para a solução do problema de otimização discreto.

2) Estudar numericamente o problema inverso de *Fundição Eletromagnética* (EMC).

Objetivo

1) Definir e implementar um método numérico para problemas de otimização da Engenharia Mecânica com as seguintes características:
    - O Método dos Elementos de Contorno (MEC) é utilizado para calcular a solução da equação de estado do problema.
    - A formulação SAND é utilizada para definir o problema de otimização.
    - Um algoritmo de ponto interior é utilizado para a solução do problema de otimização discreto.
2) Estudar numericamente o problema inverso de *Fundição Eletromagnética* (EMC).

Motivação

Esta ferramenta computacional está motivada no seguinte:

1) Análise numérica de problemas não lineares:
   - Problemas de superfície livre:
     Problema direto de Fundição Eletromagnética.
     Problema de infiltração em meio poroso.
   - Problemas com condições de contorno não lineares:
     Problema de contato.
2) Otimização de forma
   - As funções e as suas derivadas podem ser expressadas como uma integral no contorno.
3) Outros problemas de otimização de análise não linear:
   - Problema inverso de Fundição Eletromagnética.

Motivação

Esta ferramenta computacional está motivada no seguinte:

1) Análise numérica de problemas não lineares:
   - Problemas de superfície livre:
     Problema direto de Fundição Eletromagnética.
     Problema de infiltração em meio poroso.
   - Problemas com condições de contorno não lineares:
     Problema de contato.
2) Otimização de forma
   - As funções e as suas derivadas podem ser expressadas como uma integral no contorno.
3) Outros problemas de otimização de análise não linear:
   - Problema inverso de Fundição Eletromagnética.

Motivação

Esta ferramenta computacional está motivada no seguinte:

1) Análise numérica de problemas não lineares:
   - Problemas de superfície livre:
     Problema direto de Fundição Eletromagnética.
     Problema de infiltração em meio poroso.
   - Problemas com condições de contorno não lineares:
     Problema de contato.
2) Otimização de forma
   - As funções e as suas derivadas podem ser expressadas como uma integral no contorno.
3) Outros problemas de otimização de análise não linear:
   - Problema inverso de Fundição Eletromagnética.

Método dos Elementos de Contorno

Por que usar o Método dos Elementos de Contorno?

O MEC utiliza a representação integral:

$$\mathbf{c}(\xi)\mathbf{u}(\xi) = \int_\Gamma \mathbf{u}^*(\xi,\mathbf{x})\mathbf{p}(\mathbf{x})\,d\Gamma - \int_\Gamma \mathbf{p}^*(\xi,\mathbf{x})\mathbf{u}(\mathbf{x})\,d\Gamma\,. \tag{1}$$

- Não precisa da discretização do interior do domínio.
- Define um número menor de variáveis.

Muito indicado para:

- Problemas de domínio infinito.
- Otimização de forma.

Método dos Elementos de Contorno

Por que usar o Método dos Elementos de Contorno?

O MEC utiliza a representação integral:

$$\mathbf{c}(\xi)\mathbf{u}(\xi) = \int_\Gamma \mathbf{u}^*(\xi, \mathbf{x})\mathbf{p}(\mathbf{x})\, d\Gamma - \int_\Gamma \mathbf{p}^*(\xi, \mathbf{x})\mathbf{u}(\mathbf{x})\, d\Gamma \,. \tag{1}$$

- Não precisa da discretização do interior do domínio.
- Define um número menor de variáveis.

Muito indicado para:

- Problemas de domínio infinito.
- Otimização de forma.

Método dos Elementos de Contorno

Por que usar o Método dos Elementos de Contorno?

O MEC utiliza a representação integral:

$$\mathbf{c}(\xi)\mathbf{u}(\xi) = \int_{\Gamma} \mathbf{u}^*(\xi, \mathbf{x})\mathbf{p}(\mathbf{x})\, d\Gamma - \int_{\Gamma} \mathbf{p}^*(\xi, \mathbf{x})\mathbf{u}(\mathbf{x})\, d\Gamma \,. \tag{1}$$

- Não precisa da discretização do interior do domínio.
- Define um número menor de variáveis.

Muito indicado para:

- Problemas de domínio infinito.
- Otimização de forma.

Método dos Elementos de Contorno

Por que usar o Método dos Elementos de Contorno?

O MEC utiliza a representação integral:

$$\mathbf{c}(\xi)\mathbf{u}(\xi) = \int_\Gamma \mathbf{u}^*(\xi,\mathbf{x})\mathbf{p}(\mathbf{x})\,\mathrm{d}\Gamma - \int_\Gamma \mathbf{p}^*(\xi,\mathbf{x})\mathbf{u}(\mathbf{x})\,\mathrm{d}\Gamma\,. \tag{1}$$

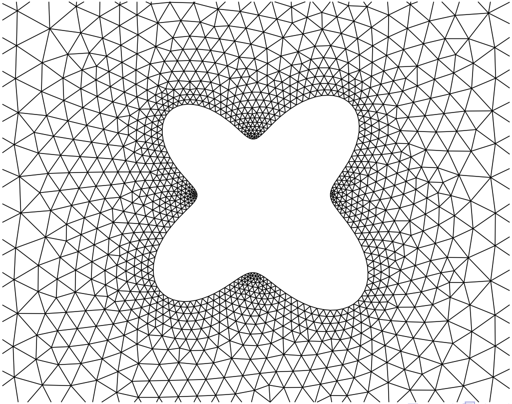
- Não precisa da discretização do interior do domínio.
- Define um número menor de variáveis.

Muito indicado para:

- Problemas de domínio infinito.
- Otimização de forma.

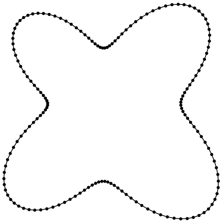
Malha

Malha de Elementos Finitos para simulação de um problema de Fundição Eletromagnética.



Malha

Malha equivalente utilizada pelo Método dos Elementos de Contorno.



Formulação SAND

Por que usar a formulação SAND?

Na formulação clássica do problema de otimização em Engenharia Mecânica o objetivo é encontrar $\mathbf{x} \in \mathbb{R}^n$:

$$\begin{array}{lll} \text{minimize} & f(\mathbf{x}, \tilde{\mathbf{u}}(\mathbf{x})) & \\ \text{sujeito a:} & \mathbf{g}(\mathbf{x}, \tilde{\mathbf{u}}(\mathbf{x})) & \geq 0 \end{array} \tag{2}$$

Com $\tilde{\mathbf{u}}$ definida pela equação de estado:

$$\mathbf{h}(\mathbf{x}, \tilde{\mathbf{u}}(\mathbf{x})) = 0 \tag{3}$$

Formulação SAND

Por que usar a formulação SAND?

Na formulação clássica do problema de otimização em Engenharia Mecânica o objetivo é encontrar $\mathbf{x} \in \mathbb{R}^n$:

$$\begin{array}{lll} \text{minimize} & f(\mathbf{x}, \tilde{\mathbf{u}}(\mathbf{x})) & \\ \text{sujeito a:} & \mathbf{g}(\mathbf{x}, \tilde{\mathbf{u}}(\mathbf{x})) \geq & 0 \end{array} \tag{2}$$

Com $\tilde{\mathbf{u}}$ definida pela equação de estado:

$$\mathbf{h}(\mathbf{x}, \tilde{\mathbf{u}}(\mathbf{x})) = 0 \tag{3}$$

Formulação SAND

Na formulação SAND:

- As variáveis de estados são adicionadas como variáveis de otimização.
- A equação de estado é adicionada como restrição de igualdade.

$$\begin{array}{lrcl} \text{minimize} & f(\mathbf{x},\mathbf{u}) & & \\ \text{sujeito a:} & \mathbf{g}(\mathbf{x},\mathbf{u}) & \geq & 0 \\ & \mathbf{h}(\mathbf{x},\mathbf{u}) & = & 0 \end{array} \tag{4}$$

Vantagens da formulação SAND

1) Não precisa da iteração para a solução da equação de estado.
2) A análise de sensibilidade é mais simples (dependência explícita de **x** e **u**).
3) Em alguns casos permite utilizar algoritmos de otimização muito eficientes (Problemas lineares ou quadráticos).
4) Propriedade de separabilidade parcial das funções. Cálculo eficiente das derivadas $\longrightarrow$ Permite utilizar algoritmos de otimização baseados no método de Newton.

Vantagens da formulação SAND

1) Não precisa da iteração para a solução da equação de estado.
2) A análise de sensibilidade é mais simples (dependência explícita de **x** e **u**).
3) Em alguns casos permite utilizar algoritmos de otimização muito eficientes (Problemas lineares ou quadráticos).
4) Propriedade de separabilidade parcial das funções. Cálculo eficiente das derivadas $\longrightarrow$ Permite utilizar algoritmos de otimização baseados no método de Newton.

Vantagens da formulação SAND

1) Não precisa da iteração para a solução da equação de estado.
2) A análise de sensibilidade é mais simples (dependência explícita de **x** e **u**).
3) Em alguns casos permite utilizar algoritmos de otimização muito eficientes (Problemas lineares ou quadráticos).
4) Propriedade de separabilidade parcial das funções. Cálculo eficiente das derivadas $\rightarrow$ Permite utilizar algoritmos de otimização baseados no método de Newton.

Vantagens da formulação SAND

1) Não precisa da iteração para a solução da equação de estado.
2) A análise de sensibilidade é mais simples (dependência explícita de $\mathbf{x}$ e $\mathbf{u}$).
3) Em alguns casos permite utilizar algoritmos de otimização muito eficientes (Problemas lineares ou quadráticos).
4) Propriedade de separabilidade parcial das funções. Cálculo eficiente das derivadas $\longrightarrow$ Permite utilizar algoritmos de otimização baseados no método de Newton.

Vantagens da formulação SAND

Muito indicada para:

1) Problemas não lineares.
2) Problemas onde seja necessário achar derivadas segundas.
3) Problemas onde seja difícil achar uma configuração inicial conveniente: permite uma maior liberdade na escolha da configuração inicial.

Vantagens da formulação SAND

Muito indicada para:

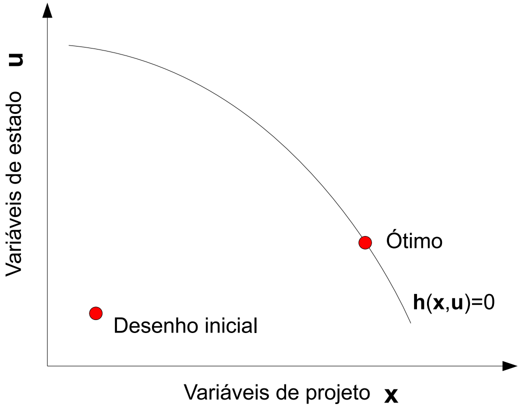
1) Problemas não lineares.
2) Problemas onde seja necessário achar derivadas segundas.
3) Problemas onde seja difícil achar uma configuração inicial conveniente: permite uma maior liberdade na escolha da configuração inicial.

Vantagens da formulação SAND

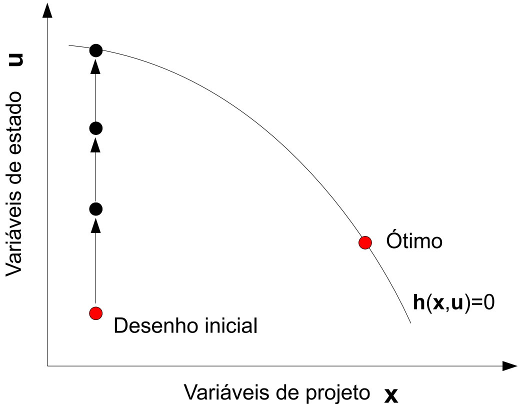
Muito indicada para:

1) Problemas não lineares.
2) Problemas onde seja necessário achar derivadas segundas.
3) Problemas onde seja difícil achar uma configuração inicial conveniente: permite uma maior liberdade na escolha da configuração inicial.

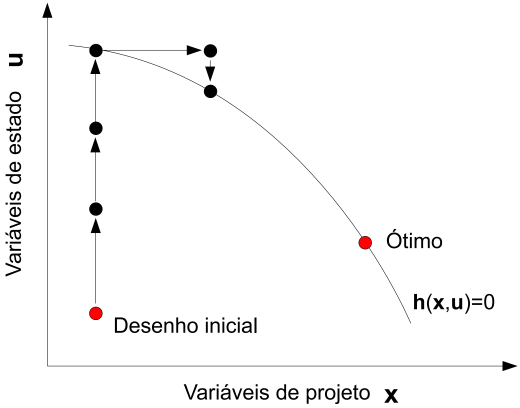
Formulação SAND



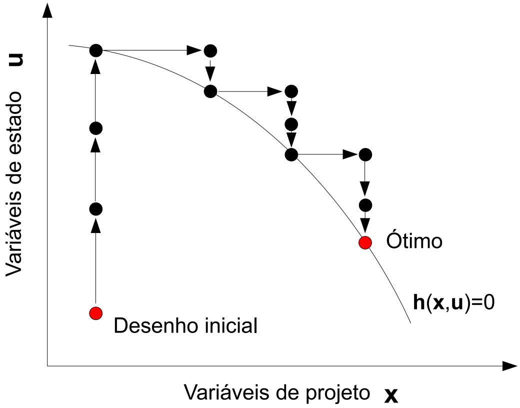
Formulação SAND



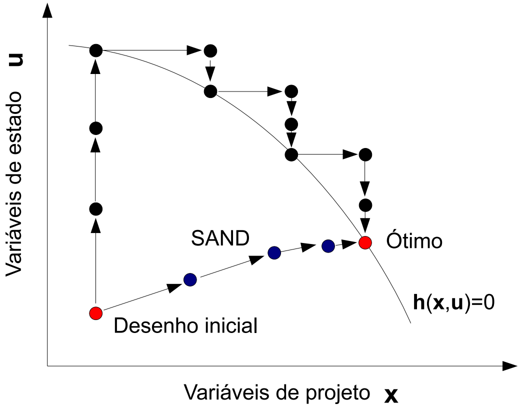
Formulação SAND



Formulação SAND



Formulação SAND



Algoritmo de ponto interior

Por que usar o algoritmo de ponto interior?

Vantagens:

1) Conveniente para problemas de grande porte.
2) Problemas esparsos ou com outro tipo de estrutura.
3) Conveniente para executar em computadores de arquitetura paralela.

- Muito indicado para utilizar junto à formulação SAND.

Algoritmo de ponto interior

Por que usar o algoritmo de ponto interior?

Vantagens:

1) Conveniente para problemas de grande porte.
2) Problemas esparsos ou com outro tipo de estrutura.
3) Conveniente para executar em computadores de arquitetura paralela.

- Muito indicado para utilizar junto à formulação SAND.

Algoritmo de ponto interior

Por que usar o algoritmo de ponto interior?

Vantagens:

1) Conveniente para problemas de grande porte.
2) Problemas esparsos ou com outro tipo de estrutura.
3) Conveniente para executar em computadores de arquitetura paralela.

- Muito indicado para utilizar junto à formulação SAND.

Algoritmo de ponto interior

Por que usar o algoritmo de ponto interior?

Vantagens:

1) Conveniente para problemas de grande porte.
2) Problemas esparsos ou com outro tipo de estrutura.
3) Conveniente para executar em computadores de arquitetura paralela.

- Muito indicado para utilizar junto à formulação SAND.

Algoritmo de ponto interior

Por que usar o algoritmo de ponto interior?

Vantagens:

1) Conveniente para problemas de grande porte.
2) Problemas esparsos ou com outro tipo de estrutura.
3) Conveniente para executar em computadores de arquitetura paralela.

- Muito indicado para utilizar junto à formulação SAND.

Vantagens do método proposto

1) MEC: Redução do número de variáveis.
2) MEC: Problemas de domínio infinito.

3) SAND: Problemas não lineares.
4) SAND: Análise de sensibilidade simples.

5) Ponto interior: Conveniente para problemas de grande porte.

Vantagens do método proposto

1) MEC: Redução do número de variáveis.
2) MEC: Problemas de domínio infinito.
3) SAND: Problemas não lineares.
4) SAND: Análise de sensibilidade simples.
5) Ponto interior: Conveniente para problemas de grande porte.

Vantagens do método proposto

1) MEC: Redução do número de variáveis.
2) MEC: Problemas de domínio infinito.
3) SAND: Problemas não lineares.
4) SAND: Análise de sensibilidade simples.
5) Ponto interior: Conveniente para problemas de grande porte.

Conteúdo

Problema de Otimização Não Linear

- Encontrar $\mathbf{x} \in \mathbb{R}^n$ tal que:

$$\begin{array}{ll} \text{minimize} & f(\mathbf{x}) \\ \text{sujeito a:} & \mathbf{g}(\mathbf{x}) \geq 0 \\ & \mathbf{h}(\mathbf{x}) = 0 \end{array}$$

- Região viável:

$$\Omega = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{g}(\mathbf{x}) \geq 0,\ \mathbf{h}(\mathbf{x}) = 0\}$$

- $\mathbf{x}^*$ é mínimo local se existe $\mathcal{N}(\mathbf{x}^*)$ tal que:

$$f(\mathbf{x}) \geq f(\mathbf{x}^*)\,, \quad \forall \mathbf{x} \in \Omega \cap \mathcal{N}(\mathbf{x}^*) \tag{5}$$

Karush-Kuhn-Tucker

- Regularidade das restrições: para todo $\mathbf{x} \in \Omega$:

$$\{\nabla \mathbf{g}_i(\mathbf{x}) \mid g_i(\mathbf{x}) = 0, \nabla \mathbf{h}_i(\mathbf{x}) \mid i \in \{1,..,p\}\} \quad \text{é l.i.}$$

- **TEOREMA** de Karush-Kuhn-Tucker:

$$\nabla f(\mathbf{x}^*) - \sum_{i=1}^{m} \boldsymbol{\lambda}_i \nabla \mathbf{g}_i(\mathbf{x}^*) - \sum_{i=1}^{p} \boldsymbol{\mu}_i \nabla \mathbf{h}_i(\mathbf{x}^*) = 0 \qquad (6)$$

$$\mathbf{g}_i(\mathbf{x}^*)\boldsymbol{\lambda}_i = 0 \qquad (7)$$

$$\mathbf{h}(\mathbf{x}^*) = 0 \qquad (8)$$

$$\mathbf{g}(\mathbf{x}^*) \geq 0 \qquad (9)$$

$$\boldsymbol{\lambda} \geq 0 \qquad (10)$$

Karush-Kuhn-Tucker

- Regularidade das restrições: para todo $\mathbf{x} \in \Omega$:

$$\{\nabla \mathbf{g}_i(\mathbf{x}) \mid g_i(\mathbf{x}) = 0, \nabla \mathbf{h}_i(\mathbf{x}) \mid i \in \{1,..,p\}\} \quad \text{é l.i.}$$

- **TEOREMA** de Karush-Kuhn-Tucker:

Sistema de Equações (Newton)

$$\boxed{\begin{aligned} \nabla f(\mathbf{x}^*) - \sum_{i=1}^{m} \boldsymbol{\lambda}_i \nabla \mathbf{g}_i(\mathbf{x}^*) - \sum_{i=1}^{p} \boldsymbol{\mu}_i \nabla \mathbf{h}_i(\mathbf{x}^*) &= 0 \\ \mathbf{g}_i(\mathbf{x}^*)\boldsymbol{\lambda}_i &= 0 \\ \mathbf{h}(\mathbf{x}^*) &= 0 \end{aligned}} \qquad \begin{aligned} &(6) \\ &(7) \\ &(8) \end{aligned}$$

$$\mathbf{g}(\mathbf{x}^*) \geq 0 \qquad (9)$$

$$\boldsymbol{\lambda} \geq 0 \qquad (10)$$

Iteração de Newton

- Função lagrangiana:

$$L(\mathbf{x},\boldsymbol{\lambda},\boldsymbol{\mu}) = f(\mathbf{x}) - \sum_{i=1}^{m} \lambda_i \mathbf{g}_i(\mathbf{x}) - \sum_{i=1}^{p} \mu_i \mathbf{h}_i(\mathbf{x})$$

$$\mathbf{H}(\mathbf{x},\boldsymbol{\lambda},\boldsymbol{\mu}) = \nabla^2_{\mathbf{xx}} L(\mathbf{x},\boldsymbol{\lambda},\boldsymbol{\mu}) = \nabla^2 f(\mathbf{x}) - \sum_{i=1}^{m} \lambda_i \nabla^2 \mathbf{g}_i(\mathbf{x}) - \sum_{i=1}^{p} \mu_i \nabla^2 \mathbf{h}_i(\mathbf{x})$$

- Iteração de Newton:

$$(\mathbf{x}_{k+1},\boldsymbol{\lambda}_{k+1},\boldsymbol{\mu}_{k+1}) = (\mathbf{x}_k,\boldsymbol{\lambda}_k,\boldsymbol{\mu}_k) + (\mathbf{d}_\mathbf{x},\mathbf{d}_\lambda,\mathbf{d}_\mu) \qquad (11)$$

$$\begin{pmatrix} \mathbf{H}_k & -\nabla \mathbf{g}_k^T & -\nabla \mathbf{h}_k^T \\ \boldsymbol{\Lambda}_k \nabla \mathbf{g}_k & \mathbf{G}_k & 0 \\ \nabla \mathbf{h}_k & 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{d}_\mathbf{x} \\ \mathbf{d}_\lambda \\ \mathbf{d}_\mu \end{pmatrix} = \begin{pmatrix} -\nabla_\mathbf{x} L_k^T \\ -\mathbf{G}_k \boldsymbol{\lambda}_k \\ -\mathbf{h}_k \end{pmatrix} \qquad (12)$$

- NÃO FUNCIONA

Iteração de Newton

- Função lagrangiana:

$$L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = f(\mathbf{x}) - \sum_{i=1}^{m} \lambda_i \mathbf{g}_i(\mathbf{x}) - \sum_{i=1}^{p} \mu_i \mathbf{h}_i(\mathbf{x})$$

$$\mathbf{H}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = \nabla^2_{\mathbf{xx}} L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = \nabla^2 f(\mathbf{x}) - \sum_{i=1}^{m} \lambda_i \nabla^2 \mathbf{g}_i(\mathbf{x}) - \sum_{i=1}^{p} \mu_i \nabla^2 \mathbf{h}_i(\mathbf{x})$$

- Iteração de Newton:

$$(\mathbf{x}_{k+1}, \boldsymbol{\lambda}_{k+1}, \boldsymbol{\mu}_{k+1}) = (\mathbf{x}_k, \boldsymbol{\lambda}_k, \boldsymbol{\mu}_k) + (\mathbf{d}_{\mathbf{x}}, \mathbf{d}_{\boldsymbol{\lambda}}, \mathbf{d}_{\boldsymbol{\mu}}) \qquad (11)$$

$$\begin{pmatrix} \mathbf{H}_k & -\nabla \mathbf{g}_k^T & -\nabla \mathbf{h}_k^T \\ \boldsymbol{\Lambda}_k \nabla \mathbf{g}_k & \mathbf{G}_k & 0 \\ \nabla \mathbf{h}_k & 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{d}_{\mathbf{x}} \\ \mathbf{d}_{\boldsymbol{\lambda}} \\ \mathbf{d}_{\boldsymbol{\mu}} \end{pmatrix} = \begin{pmatrix} -\nabla_{\mathbf{x}} L_k^T \\ -\mathbf{G}_k \boldsymbol{\lambda}_k \\ -\mathbf{h}_k \end{pmatrix} \qquad (12)$$

- **NÃO FUNCIONA**

Iteração de Newton

- Função lagrangiana:

$$L(\mathbf{x},\boldsymbol{\lambda},\boldsymbol{\mu}) = f(\mathbf{x}) - \sum_{i=1}^{m} \lambda_i \mathbf{g}_i(\mathbf{x}) - \sum_{i=1}^{p} \mu_i \mathbf{h}_i(\mathbf{x})$$

$$\mathbf{H}(\mathbf{x},\boldsymbol{\lambda},\boldsymbol{\mu}) = \nabla^2_{\mathbf{xx}} L(\mathbf{x},\boldsymbol{\lambda},\boldsymbol{\mu}) = \nabla^2 f(\mathbf{x}) - \sum_{i=1}^{m} \lambda_i \nabla^2 \mathbf{g}_i(\mathbf{x}) - \sum_{i=1}^{p} \mu_i \nabla^2 \mathbf{h}_i(\mathbf{x})$$

- Iteração de Newton:

$$(\mathbf{x}_{k+1},\boldsymbol{\lambda}_{k+1},\boldsymbol{\mu}_{k+1}) = (\mathbf{x}_k,\boldsymbol{\lambda}_k,\boldsymbol{\mu}_k) + (\mathbf{d}_\mathbf{x},\mathbf{d}_\lambda,\mathbf{d}_\mu) \qquad (11)$$

$$\begin{pmatrix} \mathbf{H}_k & -\nabla\mathbf{g}_k^T & -\nabla\mathbf{h}_k^T \\ \boldsymbol{\Lambda}_k\nabla\mathbf{g}_k & \mathbf{G}_k & 0 \\ \nabla\mathbf{h}_k & 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{d}_\mathbf{x} \\ \mathbf{d}_\lambda \\ \mathbf{d}_\mu \end{pmatrix} = \begin{pmatrix} -\nabla_\mathbf{x} L_k^T \\ -\mathbf{G}_k\boldsymbol{\lambda}_k \\ -\mathbf{h}_k \end{pmatrix} \qquad (12)$$

- **NÃO FUNCIONA**

Iteração de Newton

- Função lagrangiana:

$$L(\mathbf{x},\boldsymbol{\lambda},\boldsymbol{\mu}) = f(\mathbf{x}) - \sum_{i=1}^{m} \lambda_i \mathbf{g}_i(\mathbf{x}) - \sum_{i=1}^{p} \mu_i \mathbf{h}_i(\mathbf{x})$$

$$\mathbf{H}(\mathbf{x},\boldsymbol{\lambda},\boldsymbol{\mu}) = \nabla^2_{\mathbf{xx}} L(\mathbf{x},\boldsymbol{\lambda},\boldsymbol{\mu}) = \nabla^2 f(\mathbf{x}) - \sum_{i=1}^{m} \lambda_i \nabla^2 \mathbf{g}_i(\mathbf{x}) - \sum_{i=1}^{p} \mu_i \nabla^2 \mathbf{h}_i(\mathbf{x})$$

- Iteração de Newton:

$$(\mathbf{x}_{k+1},\boldsymbol{\lambda}_{k+1},\boldsymbol{\mu}_{k+1}) = (\mathbf{x}_k,\boldsymbol{\lambda}_k,\boldsymbol{\mu}_k) + (\mathbf{d}_\mathbf{x},\mathbf{d}_\lambda,\mathbf{d}_\mu) \quad (11)$$

Sistema singular

$$\boxed{\begin{pmatrix} \mathbf{H}_k & -\nabla \mathbf{g}_k^T & -\nabla \mathbf{h}_k^T \\ \boldsymbol{\Lambda}_k \nabla \mathbf{g}_k & \mathbf{G}_k & 0 \\ \nabla \mathbf{h}_k & 0 & 0 \end{pmatrix}} \begin{pmatrix} \mathbf{d}_\mathbf{x} \\ \mathbf{d}_\lambda \\ \mathbf{d}_\mu \end{pmatrix} = \begin{pmatrix} -\nabla_\mathbf{x} L_k^T \\ -\mathbf{G}_k \boldsymbol{\lambda}_k \\ -\mathbf{h}_k \end{pmatrix} \quad (12)$$

- **NÃO FUNCIONA**

Iteração de Newton

- Função lagrangiana:

$$L(\mathbf{x},\boldsymbol{\lambda},\boldsymbol{\mu}) = f(\mathbf{x}) - \sum_{i=1}^{m} \lambda_i \mathbf{g}_i(\mathbf{x}) - \sum_{i=1}^{p} \mu_i \mathbf{h}_i(\mathbf{x})$$

$$\mathbf{H}(\mathbf{x},\boldsymbol{\lambda},\boldsymbol{\mu}) = \nabla^2_{\mathbf{xx}} L(\mathbf{x},\boldsymbol{\lambda},\boldsymbol{\mu}) = \nabla^2 f(\mathbf{x}) - \sum_{i=1}^{m} \lambda_i \nabla^2 \mathbf{g}_i(\mathbf{x}) - \sum_{i=1}^{p} \mu_i \nabla^2 \mathbf{h}_i(\mathbf{x})$$

- Iteração de Newton:

$$(\mathbf{x}_{k+1},\boldsymbol{\lambda}_{k+1},\boldsymbol{\mu}_{k+1}) = (\mathbf{x}_k,\boldsymbol{\lambda}_k,\boldsymbol{\mu}_k) + (\mathbf{d}_\mathbf{x},\mathbf{d}_\lambda,\mathbf{d}_\mu) \qquad (11)$$

Sistema singular

$$\boxed{\begin{pmatrix} \mathbf{H}_k & -\nabla \mathbf{g}_k^T & -\nabla \mathbf{h}_k^T \\ \boldsymbol{\Lambda}_k \nabla \mathbf{g}_k & \mathbf{G}_k & 0 \\ \nabla \mathbf{h}_k & 0 & 0 \end{pmatrix}} \begin{pmatrix} \mathbf{d}_\mathbf{x} \\ \mathbf{d}_\lambda \\ \mathbf{d}_\mu \end{pmatrix} = \begin{pmatrix} -\nabla_\mathbf{x} L_k^T \\ -\mathbf{G}_k \boldsymbol{\lambda}_k \\ -\mathbf{h}_k \end{pmatrix} \qquad (12)$$

- **NÃO FUNCIONA** $\mathbf{g}(\mathbf{x}) \geq 0,\ \boldsymbol{\lambda} \geq 0$?

Iteração de Newton

Idéia:

- Definir um algoritmo de otimização baseado na direção da iteração de Newton.

Motivação: problemas com custo de **H** pequeno

- Problemas de programação linear ou quadrática.
- Problemas da coleção CUTE.
  - Economia.
  - Controle ótimo.
  - Otimização de redes.
- Na Engenharia Mecânica:
  - Utilizando a formulação SAND.
  - Utilizando modelos aproximados.

Iteração de Newton

Idéia:

- Definir um algoritmo de otimização baseado na direção da iteração de Newton.

Motivação: problemas com custo de **H** pequeno

- Problemas de programação linear ou quadrática.
- Problemas da coleção CUTE.
  - Economia.
  - Controle ótimo.
  - Otimização de redes.
- Na Engenharia Mecânica:
  - Utilizando a formulação SAND.
  - Utilizando modelos aproximados.

Iteração de Newton

Idéia:

- Definir um algoritmo de otimização baseado na direção da iteração de Newton.

Motivação: problemas com custo de **H** pequeno

- Problemas de programação linear ou quadrática.
- Problemas da coleção CUTE.
  - Economia.
  - Controle ótimo.
  - Otimização de redes.
- Na Engenharia Mecânica:
  - Utilizando a formulação SAND.
  - Utilizando modelos aproximados.

Iteração de Newton

Idéia:

- Definir um algoritmo de otimização baseado na direção da iteração de Newton.

Motivação: problemas com custo de **H** pequeno

- Problemas de programação linear ou quadrática.
- Problemas da coleção CUTE.
  - Economia.
  - Controle ótimo.
  - Otimização de redes.
- Na Engenharia Mecânica:
  - Utilizando a formulação SAND.
  - Utilizando modelos aproximados.

Iteração de Newton

Idéia:

- Definir um algoritmo de otimização baseado na direção da iteração de Newton.

Motivação: problemas com custo de **H** pequeno

- Problemas de programação linear ou quadrática.
- Problemas da coleção CUTE.
  - Economia.
  - Controle ótimo.
  - Otimização de redes.
- Na Engenharia Mecânica:
  - Utilizando a formulação SAND.
  - Utilizando modelos aproximados.

Conteúdo

Algoritmo FDIPA

Herskovits (1998).

- FDIPA gera seqüência $\{\mathbf{x}_k\}_{k\in\mathbb{N}} \subset \Delta$:

$$\Delta = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{g}(\mathbf{x}) \geq 0,\ \mathbf{h}(\mathbf{x}) \geq 0\} \tag{13}$$

- A seqüência reduz em cada iteração o valor da função potencial $\phi_{\mathbf{c}}(\mathbf{x})$:

$$\phi_{\mathbf{c}}(\mathbf{x}) = f(\mathbf{x}) + \sum_{i=1}^{p} \mathbf{c}_i |\mathbf{h}_i(\mathbf{x})| \tag{14}$$

- A direção $\mathbf{d_x} = \mathbf{d_x^\alpha} + \rho \mathbf{d_x^\beta}$:

$$\begin{pmatrix} \mathbf{B}_k & -\nabla \mathbf{g}_k^T & -\nabla \mathbf{h}_k^T \\ -\mathbf{\Lambda}_k \nabla \mathbf{g}_k & -\mathbf{G}_k & 0 \\ -\nabla \mathbf{h}_k & 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{d_x^\alpha} & \mathbf{d_x^\beta} \\ \boldsymbol{\lambda}^\alpha & \boldsymbol{\lambda}^\beta \\ \boldsymbol{\mu}^\alpha & \boldsymbol{\mu}^\beta \end{pmatrix} = \begin{pmatrix} -\nabla f_k^T & 0 \\ 0 & -\mathbf{\Lambda}_k\, \boldsymbol{\omega}^I \\ \mathbf{h}_k & -\boldsymbol{\omega}^E \end{pmatrix} \tag{15}$$

Algoritmo FDIPA

Herskovits (1998).

- FDIPA gera seqüência $\{\mathbf{x}_k\}_{k\in\mathbb{N}} \subset \Delta$:

$$\Delta = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{g}(\mathbf{x}) \geq 0,\ \mathbf{h}(\mathbf{x}) \geq 0\} \tag{13}$$

- A seqüência reduz em cada iteração o valor da função potencial $\phi_{\mathbf{c}}(\mathbf{x})$:

$$\phi_{\mathbf{c}}(\mathbf{x}) = f(\mathbf{x}) + \sum_{i=1}^{p} \mathbf{c}_i |\mathbf{h}_i(\mathbf{x})| \tag{14}$$

- A direção $\mathbf{d_x} = \mathbf{d_x^\alpha} + \rho \mathbf{d_x^\beta}$:

$$\begin{pmatrix} \mathbf{B}_k & -\nabla \mathbf{g}_k^T & -\nabla \mathbf{h}_k^T \\ -\mathbf{\Lambda}_k \nabla \mathbf{g}_k & -\mathbf{G}_k & 0 \\ -\nabla \mathbf{h}_k & 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{d_x^\alpha} & \mathbf{d_x^\beta} \\ \boldsymbol{\lambda}^\alpha & \boldsymbol{\lambda}^\beta \\ \boldsymbol{\mu}^\alpha & \boldsymbol{\mu}^\beta \end{pmatrix} = \begin{pmatrix} -\nabla f_k^T & 0 \\ 0 & -\mathbf{\Lambda}_k\, \boldsymbol{\omega}^I \\ \mathbf{h}_k & -\boldsymbol{\omega}^E \end{pmatrix} \tag{15}$$

Algoritmo FDIPA

Herskovits (1998).

- FDIPA gera seqüência $\{\mathbf{x}_k\}_{k\in\mathbb{N}} \subset \Delta$:

$$\Delta = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{g}(\mathbf{x}) \geq 0,\ \mathbf{h}(\mathbf{x}) \geq 0\} \tag{13}$$

- A seqüência reduz em cada iteração o valor da função potencial $\phi_{\mathbf{c}}(\mathbf{x})$:

$$\phi_{\mathbf{c}}(\mathbf{x}) = f(\mathbf{x}) + \sum_{i=1}^{p} \mathbf{c}_i |\mathbf{h}_i(\mathbf{x})| \tag{14}$$

- A direção $\mathbf{d_x} = \mathbf{d_x}^\alpha + \rho \mathbf{d_x}^\beta$:

$$\begin{pmatrix} \mathbf{B}_k & -\nabla \mathbf{g}_k^T & -\nabla \mathbf{h}_k^T \\ -\mathbf{\Lambda}_k \nabla \mathbf{g}_k & -\mathbf{G}_k & 0 \\ -\nabla \mathbf{h}_k & 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{d_x}^\alpha & \mathbf{d_x}^\beta \\ \boldsymbol{\lambda}^\alpha & \boldsymbol{\lambda}^\beta \\ \boldsymbol{\mu}^\alpha & \boldsymbol{\mu}^\beta \end{pmatrix} = \begin{pmatrix} -\nabla f_k^T & 0 \\ 0 & -\mathbf{\Lambda}_k\, \omega^I \\ \mathbf{h}_k & -\omega^E \end{pmatrix} \tag{15}$$

Algoritmo FDIPA

Herskovits (1998).

- FDIPA gera seqüência $\{\mathbf{x}_k\}_{k\in\mathbb{N}} \subset \Delta$:

$$\Delta = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{g}(\mathbf{x}) \geq 0,\ \mathbf{h}(\mathbf{x}) \geq 0\} \tag{13}$$

- A seqüência reduz em cada iteração o valor da função potencial $\phi_{\mathbf{c}}(\mathbf{x})$:

$$\phi_{\mathbf{c}}(\mathbf{x}) = f(\mathbf{x}) + \sum_{i=1}^{p} \mathbf{c}_i |\mathbf{h}_i(\mathbf{x})| \tag{14}$$

- A direção $\mathbf{d_x} = \mathbf{d_x^\alpha} + \rho \mathbf{d_x^\beta}$:

$$\begin{pmatrix} \mathbf{B}_k & -\nabla \mathbf{g}_k^T & -\nabla \mathbf{h}_k^T \\ -\boldsymbol{\Lambda}_k \nabla \mathbf{g}_k & -\mathbf{G}_k & 0 \\ -\nabla \mathbf{h}_k & 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{d_x^\alpha} & \mathbf{d_x^\beta} \\ \boldsymbol{\lambda}^\alpha & \boldsymbol{\lambda}^\beta \\ \boldsymbol{\mu}^\alpha & \boldsymbol{\mu}^\beta \end{pmatrix} = \begin{pmatrix} -\nabla f_k^T & 0 \\ 0 & -\boldsymbol{\Lambda}_k\, \boldsymbol{\omega}^I \\ \mathbf{h}_k & -\boldsymbol{\omega}^E \end{pmatrix} \tag{15}$$

Algoritmo FDIPA

Herskovits (1998).

- FDIPA gera seqüência $\{\mathbf{x}_k\}_{k\in\mathbb{N}} \subset \Delta$:

$$\Delta = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{g}(\mathbf{x}) \geq 0,\ \mathbf{h}(\mathbf{x}) \geq 0\} \tag{13}$$

- A seqüência reduz em cada iteração o valor da função potencial $\phi_{\mathbf{c}}(\mathbf{x})$:

$$\phi_{\mathbf{c}}(\mathbf{x}) = f(\mathbf{x}) + \sum_{i=1}^{p} \mathbf{c}_i |\mathbf{h}_i(\mathbf{x})| \tag{14}$$

- A direção $\mathbf{d_x} = \mathbf{d_x^\alpha} + \rho \mathbf{d_x^\beta}$:

$$\begin{pmatrix} \boxed{\mathbf{B}_k} & -\nabla \mathbf{g}_k^T & -\nabla \mathbf{h}_k^T \\ -\boldsymbol{\Lambda}_k \nabla \mathbf{g}_k & -\mathbf{G}_k & 0 \\ -\nabla \mathbf{h}_k & 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{d_x^\alpha} & \mathbf{d_x^\beta} \\ \boldsymbol{\lambda}^\alpha & \boldsymbol{\lambda}^\beta \\ \boldsymbol{\mu}^\alpha & \boldsymbol{\mu}^\beta \end{pmatrix} = \begin{pmatrix} -\nabla f_k^T & 0 \\ 0 & -\boldsymbol{\Lambda}_k\, \boldsymbol{\omega}^I \\ \mathbf{h}_k & -\boldsymbol{\omega}^E \end{pmatrix} \tag{15}$$

$\mathbf{B}_k$ positiva definida em vez da Hessiana

Algoritmo FDIPA

Herskovits (1998).

- FDIPA gera seqüência $\{\mathbf{x}_k\}_{k\in\mathbb{N}} \subset \Delta$:

$$\Delta = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{g}(\mathbf{x}) \geq 0,\ \mathbf{h}(\mathbf{x}) \geq 0\} \tag{13}$$

- A seqüência reduz em cada iteração o valor da função potencial $\phi_{\mathbf{c}}(\mathbf{x})$:

$$\phi_{\mathbf{c}}(\mathbf{x}) = f(\mathbf{x}) + \sum_{i=1}^{p} \mathbf{c}_i |\mathbf{h}_i(\mathbf{x})| \tag{14}$$

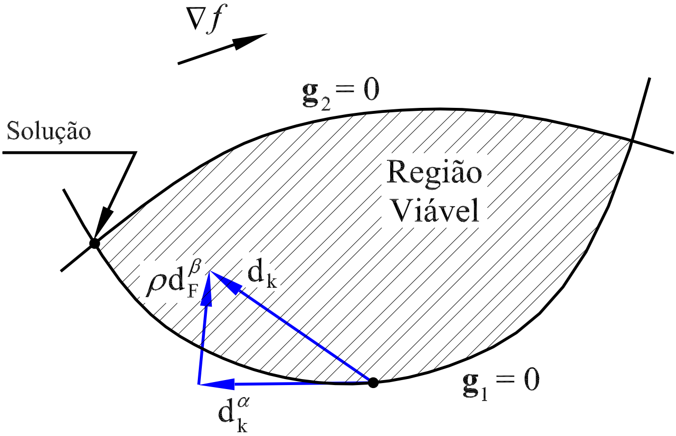
- A direção $\mathbf{d_x} = \mathbf{d_x^\alpha} + \boxed{\rho \mathbf{d_x^\beta}}$:

$$\begin{pmatrix} \boxed{\mathbf{B}_k} & -\nabla \mathbf{g}_k^T & -\nabla \mathbf{h}_k^T \\ -\boldsymbol{\Lambda}_k \nabla \mathbf{g}_k & -\mathbf{G}_k & 0 \\ -\nabla \mathbf{h}_k & 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{d_x^\alpha} & \boxed{\mathbf{d_x^\beta}} \\ \boldsymbol{\lambda}^\alpha & \boxed{\boldsymbol{\lambda}^\beta} \\ \boldsymbol{\mu}^\alpha & \boxed{\boldsymbol{\mu}^\beta} \end{pmatrix} = \begin{pmatrix} -\nabla f_k^T & \boxed{0} \\ 0 & \boxed{-\boldsymbol{\Lambda}_k \boldsymbol{\omega}^I} \\ \mathbf{h}_k & \boxed{-\boldsymbol{\omega}^E} \end{pmatrix} \tag{15}$$

$\mathbf{B}_k$ positiva definida em vez da Hessiana

$\mathbf{d_x^\beta}$: direção de desvio

Direção de desvio



Convergência global do FDIPA

- Regularidade das restrições $+$ **B** positiva definida $+$ $\boldsymbol{\lambda} > 0$ $+$ $\mathbf{g}(\mathbf{x}) \geq 0$:

$\rightarrow$ Sistema linear é não singular

- **x** não é ponto estacionário: $\rightarrow \|\mathbf{d_x}\| > 0$
- Viabilidade uniforme: Existe $\theta_1 > 0$ tal que $\forall t \in [0, \theta_1]$

$$\rightarrow \mathbf{x}_k + t\,\mathbf{d_x} \in \Delta$$

- Descida uniforme: Existe $\theta_2 > 0$ tal que $\forall t \in [0, \theta_2]$

$$\phi_{\mathbf{c}_k}(\mathbf{x}_k + t_k\,\mathbf{d_x}) \leq \phi_{\mathbf{c}_k}(\mathbf{x}_k) + t_k \eta \nabla\phi_{\mathbf{c}_k}\,\mathbf{d_x}$$

- Busca linear: Achar t (Armijo, Wolfe, etc.).

Convergência global do FDIPA

- Regularidade das restrições + $\mathbf{B}$ positiva definida + $\boldsymbol{\lambda} > 0$ + $\mathbf{g}(\mathbf{x}) \geq 0$:

$$\rightarrow \text{Sistema linear é não singular}$$

- $\mathbf{x}$ não é ponto estacionário: $\rightarrow \|\mathbf{d_x}\| > 0$
- Viabilidade uniforme: Existe $\theta_1 > 0$ tal que $\forall t \in [0, \theta_1]$

$$\rightarrow \mathbf{x}_k + t\,\mathbf{d_x} \in \Delta$$

- Descida uniforme: Existe $\theta_2 > 0$ tal que $\forall t \in [0, \theta_2]$

$$\phi_{\mathbf{c}_k}(\mathbf{x}_k + t_k\,\mathbf{d_x}) \leq \phi_{\mathbf{c}_k}(\mathbf{x}_k) + t_k \eta \nabla \phi_{\mathbf{c}_k}\,\mathbf{d_x}$$

- Busca linear: Achar t (Armijo, Wolfe, etc.).

Convergência global do FDIPA

- Regularidade das restrições + $\mathbf{B}$ positiva definida + $\boldsymbol{\lambda} > 0$ + $\mathbf{g}(\mathbf{x}) \geq 0$:

$$\rightarrow \text{Sistema linear é não singular}$$

- $\mathbf{x}$ não é ponto estacionário: $\rightarrow \|\mathbf{d_x}\| > 0$
- Viabilidade uniforme: Existe $\theta_1 > 0$ tal que $\forall t \in [0, \theta_1]$

$$\rightarrow \mathbf{x}_k + t\,\mathbf{d_x} \in \Delta$$

- Descida uniforme: Existe $\theta_2 > 0$ tal que $\forall t \in [0, \theta_2]$

$$\phi_{\mathbf{c}_k}(\mathbf{x}_k + t_k\,\mathbf{d_x}) \leq \phi_{\mathbf{c}_k}(\mathbf{x}_k) + t_k \eta \nabla \phi_{\mathbf{c}_k}\,\mathbf{d_x}$$

- Busca linear: Achar t (Armijo, Wolfe, etc.).

Convergência global do FDIPA

- Regularidade das restrições + $\mathbf{B}$ positiva definida + $\boldsymbol{\lambda} > 0$ + $\mathbf{g}(\mathbf{x}) \geq 0$:

$$\rightarrow \text{Sistema linear é não singular}$$

- $\mathbf{x}$ não é ponto estacionário: $\rightarrow \|\mathbf{d_x}\| > 0$
- Viabilidade uniforme: Existe $\theta_1 > 0$ tal que $\forall t \in [0, \theta_1]$

$$\rightarrow \mathbf{x}_k + t\,\mathbf{d_x} \in \Delta$$

- Descida uniforme: Existe $\theta_2 > 0$ tal que $\forall t \in [0, \theta_2]$

$$\phi_{\mathbf{c}_k}(\mathbf{x}_k + t_k\,\mathbf{d_x}) \leq \phi_{\mathbf{c}_k}(\mathbf{x}_k) + t_k \eta\, \nabla\phi_{\mathbf{c}_k}\,\mathbf{d_x}$$

- Busca linear: Achar t (Armijo, Wolfe, etc.).

Convergência global do FDIPA

- Regularidade das restrições $+$ **B** positiva definida $+ \boldsymbol{\lambda} > 0 + \mathbf{g}(\mathbf{x}) \geq 0$:

$\rightarrow$ Sistema linear é não singular

- **x** não é ponto estacionário: $\rightarrow \|\mathbf{d_x}\| > 0$
- Viabilidade uniforme: Existe $\theta_1 > 0$ tal que $\forall t \in [0, \theta_1]$

$$\rightarrow \mathbf{x}_k + t\,\mathbf{d_x} \in \Delta$$

- Descida uniforme: Existe $\theta_2 > 0$ tal que $\forall t \in [0, \theta_2]$

$$\phi_{\mathbf{c}_k}(\mathbf{x}_k + t_k\,\mathbf{d_x}) \leq \phi_{\mathbf{c}_k}(\mathbf{x}_k) + t_k \eta\, \nabla\phi_{\mathbf{c}_k}\,\mathbf{d_x}$$

- Busca linear: Achar t (Armijo, Wolfe, etc.).

Conteúdo

Algoritmo FDIPA-H

- Manter a hessiana $\mathbf{H}_k$ no sistema:

$$\begin{pmatrix} \mathbf{H}_k & -\nabla \mathbf{g}_k^T & -\nabla \mathbf{h}_k^T \\ -\mathbf{\Lambda}_k \nabla \mathbf{g}_k & -\mathbf{G}_k & 0 \\ -\nabla \mathbf{h}_k & 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{d}_{\mathbf{x}}^{\alpha} \\ \boldsymbol{\lambda}^{\alpha} \\ \boldsymbol{\mu}^{\alpha} \end{pmatrix} = \begin{pmatrix} -\nabla f_k^T \\ 0 \\ \mathbf{h}_k \end{pmatrix}$$

- $A = \{i \mid \mathbf{g}_i(\mathbf{x}) = 0\}, \quad I = \{i \mid \mathbf{g}_i(\mathbf{x}) > 0\}.$
- Espaço tangente:

$$\mathcal{T} = \{\mathbf{v} \in \mathbb{R}^n \mid \mathbf{v} \perp \{\nabla \mathbf{g}_i(\mathbf{x}) \mid i \in A, \nabla \mathbf{h}_i(\mathbf{x}) \mid i \in \{1, .., p\}\}\} \qquad (16)$$

- Matriz $\mathbf{M}$:

$$\mathbf{M}(\mathbf{x}, \lambda, \mu) = \mathbf{H}(\mathbf{x}, \lambda, \mu) + \sum_{i \in I} \mathbf{g}_i(\mathbf{x})^{-1} \lambda_i \nabla \mathbf{g}_i^T(\mathbf{x}) \nabla \mathbf{g}_i(\mathbf{x}) \qquad (17)$$

- Função potencial:

$$\phi_c(\mathbf{x}) = f(\mathbf{x}) + c\, \Psi(\mathbf{h}(\mathbf{x})) \qquad (18)$$

Algoritmo FDIPA-H

- Manter a hessiana $\mathbf{H}_k$ no sistema:

$$\begin{pmatrix} \mathbf{H}_k & -\nabla \mathbf{g}_k^T & -\nabla \mathbf{h}_k^T \\ -\mathbf{\Lambda}_k \nabla \mathbf{g}_k & -\mathbf{G}_k & 0 \\ -\nabla \mathbf{h}_k & 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{d}_\mathbf{x}^\alpha \\ \boldsymbol{\lambda}^\alpha \\ \boldsymbol{\mu}^\alpha \end{pmatrix} = \begin{pmatrix} -\nabla f_k^T \\ 0 \\ \mathbf{h}_k \end{pmatrix}$$

- $A = \{i \mid \mathbf{g}_i(\mathbf{x}) = 0\}, \quad I = \{i \mid \mathbf{g}_i(\mathbf{x}) > 0\}$.
- Espaço tangente:

$$\mathcal{T} = \{\mathbf{v} \in \mathbb{R}^n \mid \mathbf{v} \perp \{\nabla \mathbf{g}_i(\mathbf{x}) \mid i \in A, \nabla \mathbf{h}_i(\mathbf{x}) \mid i \in \{1,..,p\}\}\} \tag{16}$$

- Matriz $\mathbf{M}$:

$$\mathbf{M}(\mathbf{x},\lambda,\mu) = \mathbf{H}(\mathbf{x},\lambda,\mu) + \sum_{i \in I} \mathbf{g}_i(\mathbf{x})^{-1} \lambda_i \nabla \mathbf{g}_i^T(\mathbf{x}) \nabla \mathbf{g}_i(\mathbf{x}) \tag{17}$$

- Função potencial:

$$\phi_c(\mathbf{x}) = f(\mathbf{x}) + c\,\Psi(\mathbf{h}(\mathbf{x})) \tag{18}$$

Algoritmo FDIPA-H

- Manter a hessiana $\mathbf{H}_k$ no sistema:

$$\begin{pmatrix} \mathbf{H}_k & -\nabla \mathbf{g}_k^T & -\nabla \mathbf{h}_k^T \\ -\boldsymbol{\Lambda}_k \nabla \mathbf{g}_k & -\mathbf{G}_k & 0 \\ -\nabla \mathbf{h}_k & 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{d}_{\mathbf{x}}^{\alpha} \\ \boldsymbol{\lambda}^{\alpha} \\ \boldsymbol{\mu}^{\alpha} \end{pmatrix} = \begin{pmatrix} -\nabla f_k^T \\ 0 \\ \mathbf{h}_k \end{pmatrix}$$

- $A = \{i \mid \mathbf{g}_i(\mathbf{x}) = 0\}, \quad I = \{i \mid \mathbf{g}_i(\mathbf{x}) > 0\}.$
- Espaço tangente:

$$\mathcal{T} = \{\mathbf{v} \in \mathbb{R}^n \mid \mathbf{v} \perp \{\nabla \mathbf{g}_i(\mathbf{x}) \mid i \in A, \nabla \mathbf{h}_i(\mathbf{x}) \mid i \in \{1,..,p\}\}\} \qquad (16)$$

- Matriz **M**:

$$\mathbf{M}(\mathbf{x},\boldsymbol{\lambda},\boldsymbol{\mu}) = \mathbf{H}(\mathbf{x},\boldsymbol{\lambda},\boldsymbol{\mu}) + \sum_{i \in I} \mathbf{g}_i(\mathbf{x})^{-1} \lambda_i \nabla \mathbf{g}_i^T(\mathbf{x}) \nabla \mathbf{g}_i(\mathbf{x}) \qquad (17)$$

- Função potencial:

$$\phi_c(\mathbf{x}) = f(\mathbf{x}) + c\,\Psi(\mathbf{h}(\mathbf{x})) \qquad (18)$$

Algoritmo FDIPA-H

- Manter a hessiana $\mathbf{H}_k$ no sistema:

$$\begin{pmatrix} \mathbf{H}_k & -\nabla \mathbf{g}_k^T & -\nabla \mathbf{h}_k^T \\ -\mathbf{\Lambda}_k \nabla \mathbf{g}_k & -\mathbf{G}_k & 0 \\ -\nabla \mathbf{h}_k & 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{d}_{\mathbf{x}}^{\alpha} \\ \boldsymbol{\lambda}^{\alpha} \\ \boldsymbol{\mu}^{\alpha} \end{pmatrix} = \begin{pmatrix} -\nabla f_k^T \\ 0 \\ \mathbf{h}_k \end{pmatrix}$$

- $A = \{i \mid \mathbf{g}_i(\mathbf{x}) = 0\}, \quad I = \{i \mid \mathbf{g}_i(\mathbf{x}) > 0\}$.
- Espaço tangente:

$$\mathcal{T} = \{\mathbf{v} \in \mathbb{R}^n \mid \mathbf{v} \perp \{\nabla \mathbf{g}_i(\mathbf{x}) \mid i \in A, \nabla \mathbf{h}_i(\mathbf{x}) \mid i \in \{1, .., p\}\}\} \quad (16)$$

- Matriz **M**:

$$\mathbf{M}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = \mathbf{H}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) + \sum_{i \in I} \mathbf{g}_i(\mathbf{x})^{-1} \boldsymbol{\lambda}_i \nabla \mathbf{g}_i^T(\mathbf{x}) \nabla \mathbf{g}_i(\mathbf{x}) \quad (17)$$

- Função potencial:

$$\phi_c(\mathbf{x}) = f(\mathbf{x}) + c\, \Psi(\mathbf{h}(\mathbf{x})) \quad (18)$$

Direção de Newton

- **Assumindo**: Regularidade das restrições $+$ **M** positiva definida no espaço $\mathcal{T}$ $+ \boldsymbol{\lambda} > 0 + \mathbf{g}(\mathbf{x}_k) \geq 0$:

1) **LEMA 3.44**: $\longrightarrow$ Sistema linear de Newton é não singular
2) **LEMA 3.45**: $\mathbf{x}$ não é ponto estacionário: $\longrightarrow \|\mathbf{d}_{\mathbf{x}}^{\alpha}\| > 0$
3) **TEOREMA 3.46**: $\mathbf{d}_{\mathbf{x}}^{\alpha} = \mathbf{d}_{\mathbf{x}}^{1} + \mathbf{d}_{\mathbf{x}}^{2}$
   - $\mathbf{d}_{\mathbf{x}}^{1}$ é de descida para a função f: Direção de otimalidade
   - $\mathbf{d}_{\mathbf{x}}^{2}$ é de descida para a função $\Psi(\mathbf{h}(\mathbf{x}))$: Direção de viabilidade
   - Existe c_0 tal que $\forall c \geq c_0$, $\mathbf{d}_{\mathbf{x}}^{\alpha}$ é de descida para a função potencial $\phi(\mathbf{x}) = f(\mathbf{x}) + c\,\Psi(\mathbf{h}(\mathbf{x}))$

Direção de Newton

- **Assumindo**: Regularidade das restrições $+$ **M** positiva definida no espaço $\mathcal{T} + \boldsymbol{\lambda} > 0 + \mathbf{g}(\mathbf{x}_k) \geq 0$:

1) **LEMA 3.44**: $\rightarrow$ Sistema linear de Newton é não singular
2) **LEMA 3.45**: **x** não é ponto estacionário: $\rightarrow \|\mathbf{d}_\mathbf{x}^\alpha\| > 0$
3) **TEOREMA 3.46**: $\mathbf{d}_\mathbf{x}^\alpha = \mathbf{d}_\mathbf{x}^1 + \mathbf{d}_\mathbf{x}^2$
   - $\mathbf{d}_\mathbf{x}^1$ é de descida para a função f: Direção de otimalidade
   - $\mathbf{d}_\mathbf{x}^2$ é de descida para a função $\Psi(\mathbf{h}(\mathbf{x}))$: Direção de viabilidade
   - Existe c_0 tal que $\forall c \geq c_0$, $\mathbf{d}_\mathbf{x}^\alpha$ é de descida para a função potencial $\phi(\mathbf{x}) = f(\mathbf{x}) + c\,\Psi(\mathbf{h}(\mathbf{x}))$

Direção de Newton

- **Assumindo**: Regularidade das restrições $+$ $\mathbf{M}$ positiva definida no espaço $\mathcal{T}$ $+$ $\boldsymbol{\lambda} > 0 + \mathbf{g}(\mathbf{x}_k) \geq 0$:

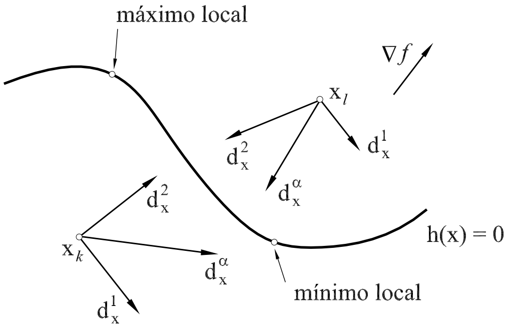
1) **LEMA 3.44**: $\rightarrow$ Sistema linear de Newton é não singular
2) **LEMA 3.45**: $\mathbf{x}$ não é ponto estacionário: $\rightarrow \|\mathbf{d}_\mathbf{x}^\alpha\| > 0$
3) **TEOREMA 3.46**: $\mathbf{d}_\mathbf{x}^\alpha = \mathbf{d}_\mathbf{x}^1 + \mathbf{d}_\mathbf{x}^2$
   - $\mathbf{d}_\mathbf{x}^1$ é de descida para a função f: Direção de otimalidade
   - $\mathbf{d}_\mathbf{x}^2$ é de descida para a função $\Psi(\mathbf{h}(\mathbf{x}))$: Direção de viabilidade
   - Existe c_0 tal que $\forall c \geq c_0$, $\mathbf{d}_\mathbf{x}^\alpha$ é de descida para a função potencial $\phi(\mathbf{x}) = f(\mathbf{x}) + c\,\Psi(\mathbf{h}(\mathbf{x}))$

Direção de Newton

- **Assumindo**: Regularidade das restrições $+$ $\mathbf{M}$ positiva definida no espaço $\mathcal{T}$ $+$ $\boldsymbol{\lambda} > 0$ $+$ $\mathbf{g}(\mathbf{x}_k) \geq 0$:

1) **LEMA 3.44**: $\rightarrow$ Sistema linear de Newton é não singular
2) **LEMA 3.45**: $\mathbf{x}$ não é ponto estacionário: $\rightarrow \|\mathbf{d}_{\mathbf{x}}^{\alpha}\| > 0$
3) **TEOREMA 3.46**: $\mathbf{d}_{\mathbf{x}}^{\alpha} = \mathbf{d}_{\mathbf{x}}^{1} + \mathbf{d}_{\mathbf{x}}^{2}$
    - $\mathbf{d}_{\mathbf{x}}^{1}$ é de descida para a função f: Direção de otimalidade
    - $\mathbf{d}_{\mathbf{x}}^{2}$ é de descida para a função $\Psi(\mathbf{h}(\mathbf{x}))$: Direção de viabilidade
    - Existe c_0 tal que $\forall c \geq c_0$, $\mathbf{d}_{\mathbf{x}}^{\alpha}$ é de descida para a função potencial $\phi(\mathbf{x}) = f(\mathbf{x}) + c\,\Psi(\mathbf{h}(\mathbf{x}))$

Direção de Newton



- **M** deve ser positiva definida no espaço $\mathcal{T}$

$\mathbf{M}_k$ não é positiva definida em $\mathcal{T}$

- Duas perguntas:
  1) Como saber se **M** é positiva definida no espaço $\mathcal{T}$?
  2) O que fazer se não é ?
- **TEOREMA 3.40**: **M** é positiva definida no espaço tangente $\mathcal{T} \Leftrightarrow$ a matriz **A** do sistema linear tem inércia $\{n, m+p, 0\}$
  - n é a dimensão de **x**
  - m é a dimensão de $\mathbf{g}(\mathbf{x})$
  - p é a dimensão de $\mathbf{h}(\mathbf{x})$
- inércia$(\mathbf{A}) = \{i_+, i_-, i_0\}$
  - i_+ é o número de valores próprios positivos
  - i_- é o número de valores próprios negativos
  - i_0 é o número de valores próprios nulos

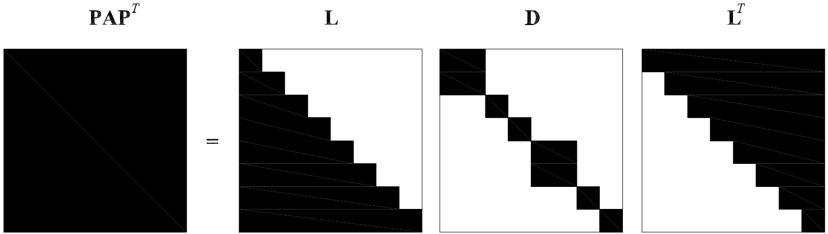
$\mathbf{M}_k$ não é positiva definida em $\mathcal{T}$

- Duas perguntas:
  1) Como saber se **M** é positiva definida no espaço $\mathcal{T}$?
  2) O que fazer se não é ?
- **TEOREMA 3.40**: **M** é positiva definida no espaço tangente $\mathcal{T} \Leftrightarrow$ a matriz **A** do sistema linear tem inércia $\{n, m+p, 0\}$
  - n é a dimensão de $\mathbf{x}$
  - m é a dimensão de $\mathbf{g}(\mathbf{x})$
  - p é a dimensão de $\mathbf{h}(\mathbf{x})$
- inércia$(\mathbf{A}) = \{i_+, i_-, i_0\}$
  - i_+ é o número de valores próprios positivos
  - i_- é o número de valores próprios negativos
  - i_0 é o número de valores próprios nulos

Decomposição $\mathbf{LDL}^T$

- Para saber a inércia ou resolver o sistema linear usamos a decomposição $\mathbf{LDL}^T$:

$$\mathbf{PAP}^T = \mathbf{LDL}^T$$

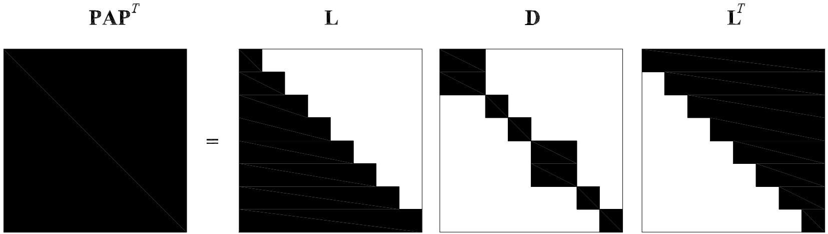


- **TEOREMA** (Sylvester): inércia($\mathbf{A}$) $=$ inércia($\mathbf{D}$)

Decomposição $\mathbf{LDL}^T$

- Para saber a inércia ou resolver o sistema linear usamos a decomposição $\mathbf{LDL}^T$:

$$\mathbf{PAP}^T = \mathbf{LDL}^T$$



- **TEOREMA** (Sylvester): inércia($\mathbf{A}$) $=$ inércia($\mathbf{D}$)

$\mathbf{M}$ não é positiva definida em $\mathcal{T}$

(1) Substituir $\mathbf{d}_\mathbf{x}^1$ por $-\mathbf{d}_\mathbf{x}^1$, se $\mathbf{d}_\mathbf{x}^1$ não for de descida para a função objetivo.

(2) Na próxima iteração substituir $\mathbf{H}_k$ por $\mathbf{H}_k + \gamma\mathbf{I}$

- Sempre existe γ conveniente
- Preserva esparsidade de $\mathbf{H}_k$

- Exemplo: utilizar $\mathbf{H}_k + \gamma\mathbf{I}$ positiva definida:

$$\gamma = \max_{1\le i\le n}\left\{\max\left(1.2\sum_{j\neq i}|(\mathbf{H}_k)_{ij}| - (\mathbf{H}_k)_{ii}\,, 0\right)\right\}. \qquad (19)$$

- **TEOREMA 3.50** Toda subseqüência convergente da seqüência gerada pelo algoritmo FDIPA-H converge para um ponto de Karush-Kuhn-Tucker do problema de otimização.

$\mathbf{M}$ não é positiva definida em $\mathcal{T}$

(1) Substituir $\mathbf{d}_\mathbf{x}^1$ por $-\mathbf{d}_\mathbf{x}^1$, se $\mathbf{d}_\mathbf{x}^1$ não for de descida para a função objetivo.

(2) Na próxima iteração substituir $\mathbf{H}_k$ por $\mathbf{H}_k + \gamma \mathbf{I}$

- Sempre existe γ conveniente
- Preserva esparsidade de $\mathbf{H}_k$

- Exemplo: utilizar $\mathbf{H}_k + \gamma \mathbf{I}$ positiva definida:

$$\gamma = \max_{1 \leq i \leq n} \left\{ \max \left(1.2 \sum_{j \neq i} |(\mathbf{H}_k)_{ij}| - (\mathbf{H}_k)_{ii}\, , 0 \right) \right\} . \tag{19}$$

- **TEOREMA 3.50** Toda subseqüência convergente da seqüência gerada pelo algoritmo FDIPA-H converge para um ponto de Karush-Kuhn-Tucker do problema de otimização.

Algoritmo de Otimização FDIPA-H

Dados: $\mathbf{x}_0 \in \Delta^0$, $\boldsymbol{\lambda}_0 \in \mathbb{R}^m$ positivo, $\boldsymbol{\mu}_0 \in \mathbb{R}^p$
$\omega^I \in \mathbb{R}^m$ positivo, $\omega^E \in \mathbb{R}^p$ positivo e c_0 positivo.

Passo 1: Teste de convergência

Passo 2: Cálculo da direção de busca

2.1 Escolha o parâmetro γ_k e solução do sistema:

$$\begin{pmatrix} \mathbf{H}_k + \gamma_k \mathbf{I} & -\nabla \mathbf{g}_k^T & -\nabla \mathbf{h}_k^T \\ -\boldsymbol{\Lambda}_k \nabla \mathbf{g}_k & -\mathbf{G}_k & 0 \\ -\nabla \mathbf{h}_k & 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{d}_\mathbf{x}^\alpha & \mathbf{d}_\mathbf{x}^\beta \\ \boldsymbol{\lambda}^\alpha & \boldsymbol{\lambda}^\beta \\ \boldsymbol{\mu}^\alpha & \boldsymbol{\mu}^\beta \end{pmatrix} = \begin{pmatrix} -\nabla f_k^T & 0 \\ 0 & -\boldsymbol{\Lambda}_k \, \omega^I \\ \mathbf{h}_k & -\omega^E \end{pmatrix}.$$

2.2 Se $\mathbf{d}_\mathbf{x}^\alpha = 0$ pare

2.4 Atualize c_k

2.5 Calcule ρ e a direção $\mathbf{d}_\mathbf{x} = \mathbf{d}_\mathbf{x}^\alpha + \rho \mathbf{d}_\mathbf{x}^\beta$

Passo 3: Busca linear

Passo 4: Atualização: $\mathbf{x}_{k+1} = \mathbf{x}_k + t_k \, \mathbf{d}_\mathbf{x}$

4.2 Retorne ao Passo 1.

Algoritmo de Otimização FDIPA-H

Dados: $\mathbf{x}_0 \in \Delta^0$, $\boldsymbol{\lambda}_0 \in \mathbb{R}^m$ positivo, $\boldsymbol{\mu}_0 \in \mathbb{R}^p$
$\boldsymbol{\omega}^I \in \mathbb{R}^m$ positivo, $\boldsymbol{\omega}^E \in \mathbb{R}^p$ positivo e c_0 positivo.

Passo 1: Teste de convergência

Passo 2: Cálculo da direção de busca

2.1 Escolha o parâmetro γ_k e solução do sistema:

$$\begin{pmatrix} \mathbf{H}_k + \gamma_k \mathbf{I} & -\nabla \mathbf{g}_k^T & -\nabla \mathbf{h}_k^T \\ -\boldsymbol{\Lambda}_k \nabla \mathbf{g}_k & -\mathbf{G}_k & 0 \\ -\nabla \mathbf{h}_k & 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{d}_\mathbf{x}^\alpha & \mathbf{d}_\mathbf{x}^\beta \\ \boldsymbol{\lambda}^\alpha & \boldsymbol{\lambda}^\beta \\ \boldsymbol{\mu}^\alpha & \boldsymbol{\mu}^\beta \end{pmatrix} = \begin{pmatrix} -\nabla f_k^T & 0 \\ 0 & -\boldsymbol{\Lambda}_k \,\boldsymbol{\omega}^I \\ \mathbf{h}_k & -\boldsymbol{\omega}^E \end{pmatrix}.$$

2.2 Se $\mathbf{d}_\mathbf{x}^\alpha = 0$ pare

2.4 Atualize c_k

2.5 Calcule ρ e a direção $\mathbf{d}_\mathbf{x} = \mathbf{d}_\mathbf{x}^\alpha + \rho \mathbf{d}_\mathbf{x}^\beta$

Passo 3: Busca linear

Passo 4: Atualização: $\mathbf{x}_{k+1} = \mathbf{x}_k + t_k \, \mathbf{d}_\mathbf{x}$

4.2 Retorne ao Passo 1.

Algoritmo de Otimização FDIPA-H

Dados: $\mathbf{x}_0 \in \Delta^0$, $\boldsymbol{\lambda}_0 \in \mathbb{R}^m$ positivo, $\boldsymbol{\mu}_0 \in \mathbb{R}^p$
$\boldsymbol{\omega}^I \in \mathbb{R}^m$ positivo, $\boldsymbol{\omega}^E \in \mathbb{R}^p$ positivo e c_0 positivo.

Passo 1: Teste de convergência

Passo 2: Cálculo da direção de busca

2.1 Escolha o parâmetro γ_k e solução do sistema:

$$\begin{pmatrix} \mathbf{H}_k + \gamma_k \mathbf{I} & -\nabla \mathbf{g}_k^T & -\nabla \mathbf{h}_k^T \\ -\boldsymbol{\Lambda}_k \nabla \mathbf{g}_k & -\mathbf{G}_k & 0 \\ -\nabla \mathbf{h}_k & 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{d}_\mathbf{x}^\alpha & \mathbf{d}_\mathbf{x}^\beta \\ \boldsymbol{\lambda}^\alpha & \boldsymbol{\lambda}^\beta \\ \boldsymbol{\mu}^\alpha & \boldsymbol{\mu}^\beta \end{pmatrix} = \begin{pmatrix} -\nabla f_k^T & 0 \\ 0 & -\boldsymbol{\Lambda}_k \, \omega^I \\ \mathbf{h}_k & -\omega^E \end{pmatrix}.$$

2.2 Se $\mathbf{d}_\mathbf{x}^\alpha = 0$ pare

2.4 Atualize c_k

2.5 Calcule ρ e a direção $\mathbf{d}_\mathbf{x} = \mathbf{d}_\mathbf{x}^\alpha + \rho \mathbf{d}_\mathbf{x}^\beta$

Passo 3: Busca linear

Passo 4: Atualização: $\mathbf{x}_{k+1} = \mathbf{x}_k + t_k \, \mathbf{d}_\mathbf{x}$

4.2 Retorne ao Passo 1.

Algoritmo de Otimização FDIPA-H

Dados: $\mathbf{x}_0 \in \Delta^0$, $\boldsymbol{\lambda}_0 \in \mathbb{R}^m$ positivo, $\boldsymbol{\mu}_0 \in \mathbb{R}^p$
$\boldsymbol{\omega}^I \in \mathbb{R}^m$ positivo, $\boldsymbol{\omega}^E \in \mathbb{R}^p$ positivo e c_0 positivo.

Passo 1: Teste de convergência

Passo 2: Cálculo da direção de busca

2.1 Escolha o parâmetro γ_k e solução do sistema:

$$\begin{pmatrix} \mathbf{H}_k + \gamma_k \mathbf{I} & -\nabla \mathbf{g}_k^T & -\nabla \mathbf{h}_k^T \\ -\boldsymbol{\Lambda}_k \nabla \mathbf{g}_k & -\mathbf{G}_k & 0 \\ -\nabla \mathbf{h}_k & 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{d}_\mathbf{x}^\alpha & \mathbf{d}_\mathbf{x}^\beta \\ \boldsymbol{\lambda}^\alpha & \boldsymbol{\lambda}^\beta \\ \boldsymbol{\mu}^\alpha & \boldsymbol{\mu}^\beta \end{pmatrix} = \begin{pmatrix} -\nabla f_k^T & 0 \\ 0 & -\boldsymbol{\Lambda}_k \, \boldsymbol{\omega}^I \\ \mathbf{h}_k & -\boldsymbol{\omega}^E \end{pmatrix}.$$

2.2 Se $\mathbf{d}_\mathbf{x}^\alpha = 0$ pare
2.4 Atualize c_k
2.5 Calcule ρ e a direção $\mathbf{d}_\mathbf{x} = \mathbf{d}_\mathbf{x}^\alpha + \rho \mathbf{d}_\mathbf{x}^\beta$

Passo 3: Busca linear

Passo 4: Atualização: $\mathbf{x}_{k+1} = \mathbf{x}_k + t_k \, \mathbf{d}_\mathbf{x}$

4.2 Retorne ao Passo 1.

Algoritmo de Otimização FDIPA-H

Dados: $\mathbf{x}_0 \in \Delta^0$, $\boldsymbol{\lambda}_0 \in \mathbb{R}^m$ positivo, $\boldsymbol{\mu}_0 \in \mathbb{R}^p$
$\boldsymbol{\omega}^I \in \mathbb{R}^m$ positivo, $\boldsymbol{\omega}^E \in \mathbb{R}^p$ positivo e c_0 positivo.

Passo 1: Teste de convergência

Passo 2: Cálculo da direção de busca

2.1 Escolha o parâmetro γ_k e solução do sistema:

$$\begin{pmatrix} \mathbf{H}_k + \gamma_k \mathbf{I} & -\nabla \mathbf{g}_k^T & -\nabla \mathbf{h}_k^T \\ -\boldsymbol{\Lambda}_k \nabla \mathbf{g}_k & -\mathbf{G}_k & 0 \\ -\nabla \mathbf{h}_k & 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{d}_\mathbf{x}^\alpha & \mathbf{d}_\mathbf{x}^\beta \\ \boldsymbol{\lambda}^\alpha & \boldsymbol{\lambda}^\beta \\ \boldsymbol{\mu}^\alpha & \boldsymbol{\mu}^\beta \end{pmatrix} = \begin{pmatrix} -\nabla f_k^T & 0 \\ 0 & -\boldsymbol{\Lambda}_k \, \boldsymbol{\omega}^I \\ \mathbf{h}_k & -\boldsymbol{\omega}^E \end{pmatrix}.$$

2.2 Se $\mathbf{d}_\mathbf{x}^\alpha = 0$ pare
2.4 Atualize c_k
2.5 Calcule ρ e a direção $\mathbf{d}_\mathbf{x} = \mathbf{d}_\mathbf{x}^\alpha + \rho \mathbf{d}_\mathbf{x}^\beta$

Passo 3: Busca linear

Passo 4: Atualização: $\mathbf{x}_{k+1} = \mathbf{x}_k + t_k \, \mathbf{d}_\mathbf{x}$

4.2 Retorne ao Passo 1.

Algoritmo de Otimização FDIPA-H

Dados: $\mathbf{x}_0 \in \Delta^0$, $\boldsymbol{\lambda}_0 \in \mathbb{R}^m$ positivo, $\boldsymbol{\mu}_0 \in \mathbb{R}^p$
$\boldsymbol{\omega}^I \in \mathbb{R}^m$ positivo, $\boldsymbol{\omega}^E \in \mathbb{R}^p$ positivo e c_0 positivo.

Passo 1: Teste de convergência

Passo 2: Cálculo da direção de busca

2.1 Escolha o parâmetro γ_k e solução do sistema:

$$\begin{pmatrix} \mathbf{H}_k + \gamma_k \mathbf{I} & -\nabla \mathbf{g}_k^T & -\nabla \mathbf{h}_k^T \\ -\boldsymbol{\Lambda}_k \nabla \mathbf{g}_k & -\mathbf{G}_k & 0 \\ -\nabla \mathbf{h}_k & 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{d}_\mathbf{x}^\alpha & \mathbf{d}_\mathbf{x}^\beta \\ \boldsymbol{\lambda}^\alpha & \boldsymbol{\lambda}^\beta \\ \boldsymbol{\mu}^\alpha & \boldsymbol{\mu}^\beta \end{pmatrix} = \begin{pmatrix} -\nabla f_k^T & 0 \\ 0 & -\boldsymbol{\Lambda}_k\, \boldsymbol{\omega}^I \\ \mathbf{h}_k & -\boldsymbol{\omega}^E \end{pmatrix}.$$

2.2 Se $\mathbf{d}_\mathbf{x}^\alpha = 0$ pare
2.4 Atualize c_k
2.5 Calcule ρ e a direção $\mathbf{d}_\mathbf{x} = \mathbf{d}_\mathbf{x}^\alpha + \rho \mathbf{d}_\mathbf{x}^\beta$

Passo 3: Busca linear

Passo 4: Atualização: $\mathbf{x}_{k+1} = \mathbf{x}_k + t_k\, \mathbf{d}_\mathbf{x}$

4.2 Retorne ao Passo 1.

Exemplos: Coleção de Hock e Shitkowski

Comparação do número de iterações

112 problemas de pequeno tamanho $\approx 2 - 16$ variáveis.

	Resolveu	Falhou*	Foi o melhor**	Foi o pior***
FDIPA	104 (92.9%)	8 (7.1%)	13	71
FAIPA	103 (92.0%)	9 (8.0%)	24	62
FDIPA-H	104 (92.9%)	8 (7.1%)	84	19

(*) Falhou: Erro de execução ou mais de 100 iterações para atingir a tolerância exigida.
(**) Melhor: realizou un número de iterações menor ou igual aos outros dois.
(***) Pior: realizou un número de iterações maior ou igual aos outros dois.

Exemplos: Coleção do CUTE

- 150 problemas de grande porte: 1000 – 50000 variáveis.
  - Economia.
  - Controle ótimo.
  - Otimização de redes.
  - Engenharia Mecânica.
- Formulações SAND: problemas grandes e esparsos.
- As funções têm a propriedade de separabilidade parcial.
- O custo de cálculo da matriz hessiana é pequeno.
- É um conjunto de problemas muito importante para os quais os algoritmos de ponto interior baseados na iteração de Newton são ideais.

Exemplos: Coleção do CUTE

- 150 problemas de grande porte: 1000 – 50000 variáveis.
  - Economia.
  - Controle ótimo.
  - Otimização de redes.
  - Engenharia Mecânica.
- Formulações SAND: problemas grandes e esparsos.
- As funções têm a propriedade de separabilidade parcial.
- O custo de cálculo da matriz hessiana é pequeno.
- É um conjunto de problemas muito importante para os quais os algoritmos de ponto interior baseados na iteração de Newton são ideais.

Exemplos: Coleção do CUTE

Comparação do número de iterações

	Resolveu	Falhou*	Foi o melhor**	Foi o pior***
KNITRO	93 (62.0%)	57 (38.0%)	31	76
LOQO	138 (92.0%)	12 (8.0%)	39	43
FDIPA-H	114 (76.0%)	36 (24.0%)	81	40
SNOPT	62 (41.3%)	88 (58.7%)	-	-
FDIPA	61 (40.6%)	89 (59.3%)	-	-
FAIPA	59 (39.3%)	91 (60.7%)	-	-

(*) A comparação compreende os algoritmos KNITRO, LOQO e FDIPA-H.

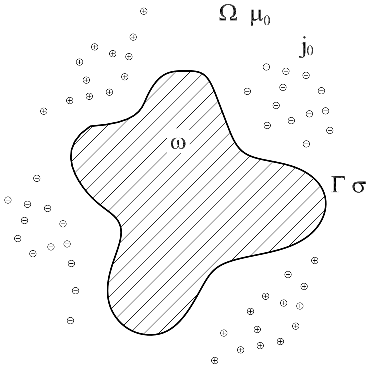
(*) Falhou: Erro de execução, mais de 3000 iterações ou mais de 20 min de execução.
(**) Melhor: realizou un número de iterações menor ou igual aos outros dois.
(***) Pior: realizou un número de iterações maior ou igual aos outros dois.

Conclusões sobre o FDIPA-H

1) O algoritmo tem convergência global (Teorema 3.50)
2) É muito eficiente para os problemas de grande porte da coleção CUTE.

Conteúdo

Problema de Fundição Eletromagnética



O modelo aqui considerado assume que a freqüência da corrente elétrica é muito alta e, portanto, o campo magnético penetra uma distância desprezível no interior do metal líquido.

Problema de Fundição Eletromagnética

Equações do campo magnético

Michel Pierre, Jean R. Roche (1991)

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{j}_0 \quad \text{em } \Omega \tag{20}$$

$$\nabla \cdot \mathbf{B} = 0 \quad \text{em } \Omega \tag{21}$$

$$\mathbf{B} \cdot \nu = 0 \quad \text{em } \Gamma \tag{22}$$

$$\|\mathbf{B}\| = O(\|x\|^{-1}) \quad \text{quando } \|x\| \to \infty \text{ em } \Omega \tag{23}$$

ω: domínio ocupado pelo metal líquido.
Γ: contorno de ω.
$\Omega = \mathbb{R} \setminus \omega$ o exterior do metal líquido.
$\mathbf{j}_0 = (0, 0, j_0)$ e o vetor de densidade de corrente elétrica.
$\mathbf{B} = (B_1, B_2, 0)$ vetor de campo magnético.
μ_0: permeabilidade magnética do vácuo.
ν: vetor unitário normal a superfície Γ.

Equilíbrio e restrições

Além disso temos a equação de equilíbrio:

$$\frac{1}{2\mu_0}\|\mathbf{B}\|^2 + \sigma\mathcal{C} = \bar{p} \quad \text{constante em } \Gamma \tag{24}$$

E a equação do volume:

$$\int_\omega \mathrm{d}\Omega = S_0 \tag{25}$$

Assume-se também que j_0 tem suporte compacto e verifica:

$$\int_\Omega j_0 \,\mathrm{d}\Omega = 0 \tag{26}$$

Função fluxo magnético

Com isso, existe a função fluxo magnético $\varphi : \Omega \to \mathbb{R}$ tal que $\mathbf{B} = (\frac{\partial \varphi}{\partial x_2}, -\frac{\partial \varphi}{\partial x_1}, 0)$ e φ é solução das equações de estado:

$$-\Delta\varphi = \mu_0 j_0 \quad \text{em } \Omega \tag{27}$$

$$\varphi = 0 \quad \text{em } \Gamma \tag{28}$$

$$\varphi(x) = O(1) \quad \text{quando } \|x\| \to \infty \tag{29}$$

O equilíbrio em termos do fluxo φ fica:

$$\frac{1}{2\mu_0}\|\nabla\varphi\|^2 + \sigma\mathcal{C} = \bar{p} \quad \text{constante em } \Gamma \tag{30}$$

Problema variacional

A formulação variacional consiste em encontrar ω como um ponto crítico do funcional Energia Total:

$$E(\omega) = -\frac{1}{2\mu_0} \int_\Omega \|\nabla \varphi_\omega\|^2 \, \mathrm{d}\Omega + \sigma \int_\Gamma \mathrm{d}\Gamma \,, \tag{31}$$

sujeito à restrição de igualdade na área de ω:

$$\int_\omega \mathrm{d}\Omega = S_0 \,. \tag{32}$$

onde φ_ω verifica:

$$-\Delta\varphi_\omega = \mu_0 j_0 \quad \text{em } \Omega \tag{33}$$

$$\varphi_\omega = 0 \quad \text{em } \Gamma \tag{34}$$

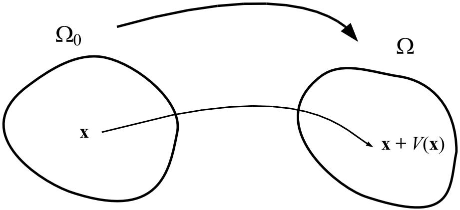
$$\varphi_\omega(x) = O(1) \quad \text{quando } \|x\| \to \infty \tag{35}$$

Diferenciação em relação à forma

Para caracterizar os pontos críticos utilizamos o conceito de diferenciação em relação à forma.

Para um domínio de referência Ω_0, são consideradas as transformações:

$$T = Id + V\,, \quad \text{com} \quad V \in W^{1,\infty}(\mathbb{R}^n, \mathbb{R}^n)\,, \quad \|V\|_{W^{1,\infty}(\mathbb{R}^n,\mathbb{R}^n)} < 1\,, \qquad (36)$$



Definição do domínio transformado pelo campo vetorial V.

Condição de equilíbrio

Função lagrangiana:

$$L(\omega,\bar{p}) = E(\omega) - \bar{p}(S(\omega) - S_0)\,, \tag{37}$$

onde $\bar{p}$ é o multiplicador de Lagrange associado à restrição de área.

Condição de ponto crítico:

$$L'(\omega,\bar{p})(V) = 0\,, \quad \forall V \in W^{1,\infty}(\mathbb{R}^2,\mathbb{R}^2)\,. \tag{38}$$

Teorema 5.1 Condição de equilíbrio do problema variacional:

$$\int_\Gamma \left(\frac{1}{2\mu_0} \|\nabla\varphi\|^2 + \sigma\mathcal{C} - \bar{p} \right) (V\cdot\nu)\, \mathrm{d}\Gamma = 0 \quad \forall V \text{ in } W^{1,\infty}(\mathbb{R}^2,\mathbb{R}^2)\,. \tag{39}$$

Solução da equação de estado

Para achar a solução é considerada uma solução particular φ_1:

$$\varphi_1(x) = -\frac{\mu_0}{2\pi} \int_{\mathbb{R}^2} \ln \|x - y\| j_0(y) \, d\Omega \qquad (40)$$

Então, a função φ pode ser calculada como:

$$\varphi(x) = v(x) + \varphi_1(x) \qquad (41)$$

onde a função v é solução de:

$$\begin{aligned} -\Delta v(x) &= 0 && \text{em } \Omega && (42) \\ v(x) &= -\varphi_1(x) && \text{em } \Gamma && (43) \\ v(x) &= O(1) && \text{quando } \|x\| \to \infty && (44) \end{aligned}$$

Solução da equação homogênea

Uma representação integral de v é dada por:

$$v(x) = -\frac{1}{2\pi}\int_\Gamma q(y)\ln\|x-y\|\,\mathrm{d}\Gamma + c \tag{45}$$

onde c é o valor no infinito, e $q \in H^{-1/2}(\Gamma)$ deve verificar:

$$\int_\Gamma q(x)\,\mathrm{d}\Gamma = 0 \tag{46}$$

As condições de contorno em Γ são impostas de forma fraca:

$$\begin{aligned} -\frac{1}{2\pi}\int_\Gamma g(x)\int_\Gamma q(y)\ln\|x-y\|\,\mathrm{d}\Gamma\,\mathrm{d}\Gamma + c\int_\Gamma g(x)\,\mathrm{d}\Gamma \\ = -\int_\Gamma \varphi_1(x) g(x)\,\mathrm{d}\Gamma \quad \forall g \in H^{-1/2}(\Gamma) \end{aligned} \tag{47}$$

Resumo das equações do problema direto

- Em resumo, temos as equações de estado:

$$-\frac{1}{2\pi}\int_\Gamma g(x)\int_\Gamma q(y)\ln\|x-y\|\,\mathrm{d}\Gamma\,\mathrm{d}\Gamma + \\ + c\int_\Gamma g(x)\,\mathrm{d}\Gamma = -\int_\Gamma \varphi_1(x)g(x)\,\mathrm{d}\Gamma \quad \forall g \in H^{-1/2}(\Gamma) \tag{48}$$

$$\int_\Gamma q(x)\,\mathrm{d}\Gamma = 0 \tag{49}$$

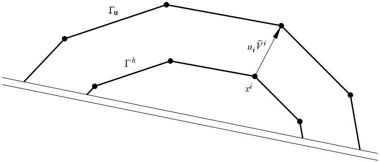
- A restrição de igualdade na área de ω:

$$\int_\omega \mathrm{d}\Omega = S_0 \tag{50}$$

- A equação de equilíbrio no contorno:

$$\int_\Gamma \left(\frac{1}{2\mu_0}\|\nabla\varphi\|^2 + \sigma\mathcal{C} - \bar{p}\right)(V\cdot\nu)\,\mathrm{d}\Gamma = 0 \quad \forall V \text{ in } W^{1,\infty}(\mathbb{R}^n,\mathbb{R}^n) \tag{51}$$

Discretização



Define-se a transformação paramétrica $T_{\mathbf{u}}$ como:

$$T_{\mathbf{u}}(x) = x + V_{\mathbf{u}}(x) \tag{52}$$

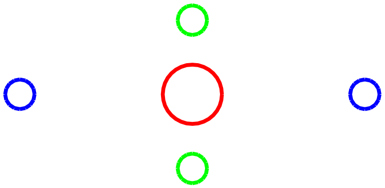
$$V_{\mathbf{u}}(x) = \sum_{i=1}^{n} u_i V^i(x) \tag{53}$$

onde $\mathbf{u}^T = (u_1, \ldots, u_n) \in \mathbb{R}^n$ é o vetor das incógnitas que determinam a evolução do contorno. Então, o contorno atualizado $\Gamma_{\mathbf{u}}$ é dado por:

$$\Gamma_{\mathbf{u}} = \left\{ X \mid X = x + V_{\mathbf{u}}(x);\ x \in \Gamma^h \right\} \tag{54}$$

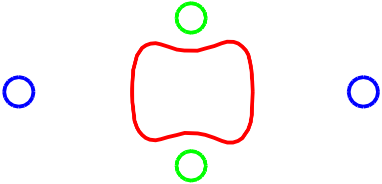
Exemplo

Iter 0



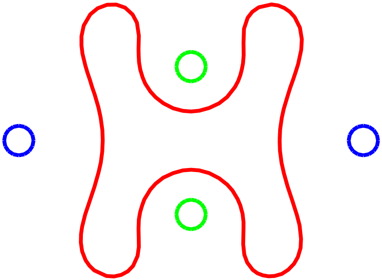
Exemplo

Iter 4



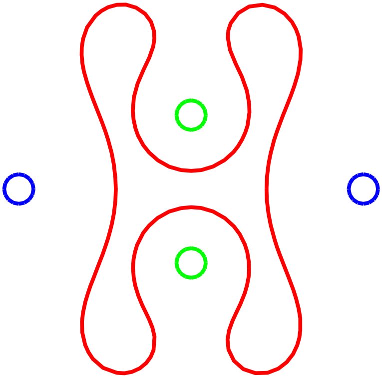
Exemplo

Iter 10

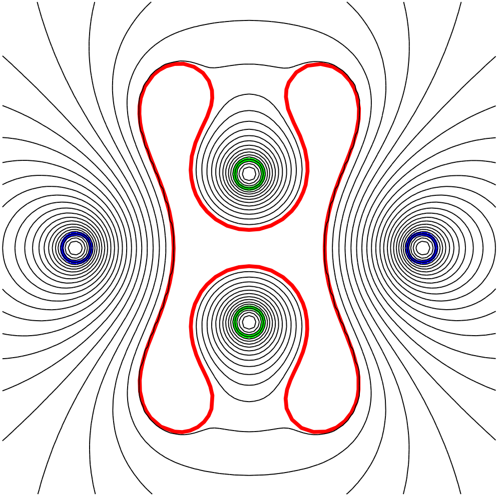


Exemplo

Iter 29



Exemplo



Conteúdo

Problema Inverso

- Encontrar a configuração de indutores para ter ω aproximadamente igual a uma forma objetivo ω^*
- Duas formulações são propostas:
  1) Minimização da "distância" entre a forma objetivo e a forma em equilíbrio.
  2) Minimização da pressão fictícia que equilibra a forma objetivo.
- Otimização da posição dos indutores.
- Otimização da forma dos indutores.

Problema Inverso

- Encontrar a configuração de indutores para ter ω aproximadamente igual a uma forma objetivo ω^*
- Duas formulações são propostas:
  1) Minimização da "distância" entre a forma objetivo e a forma em equilíbrio.
  2) Minimização da pressão fictícia que equilibra a forma objetivo.
- Otimização da posição dos indutores.
- Otimização da forma dos indutores.

Problema Inverso

- Encontrar a configuração de indutores para ter ω aproximadamente igual a uma forma objetivo ω^*
- Duas formulações são propostas:
  1) Minimização da "distância" entre a forma objetivo e a forma em equilíbrio.
  2) Minimização da pressão fictícia que equilibra a forma objetivo.
- Otimização da posição dos indutores.
- Otimização da forma dos indutores.

Problema Inverso - Primeira formulação

A primeira formulação considera uma deformação do domínio ω^* definida pelo mapeamento seguinte:

$$T_Z(x) = (Id + Z)(x), \quad \forall x \in \mathbb{R}^2 \tag{55}$$

onde Z é regular e tem suporte compacto em $\mathbb{R}^2$. Definindo:

$$\omega_Z = T_Z(\omega^*) \tag{56}$$

$$\Gamma_Z = T_Z(\Gamma^*) \tag{57}$$

Primeira formulação do problema inverso:

$$\begin{aligned} &\min_{j_0, Z} \|Z\|^2_{L^2(\Gamma^*)} \\ &\text{sujeito a:} \\ &\omega_Z \text{ é equilibrado sob } j_0 \end{aligned} \tag{58}$$

Problema Inverso - Segunda formulação

Considerando uma função de folga $p(x) : \Gamma^* \to \mathbb{R}$ de forma que a equação de equilíbrio seja verificada para o domínio objetivo:

$$\int_{\Gamma^*} \left(\frac{1}{2\mu_0} \|\nabla \varphi_\omega\|^2 + \sigma \mathcal{C} - \bar{p} + p \right) (V \cdot \nu) \, d\Gamma = 0 \;\; \forall V \text{ in } C^1(\mathbb{R}^2, \mathbb{R}^2) \qquad (59)$$

A função p pode ser interpretada como sendo uma pressão adicional atuando na interface.

Segunda formulação do problema inverso:

$$\begin{aligned} & \min_{j_0, p} \|p\|^2_{L^2(\Gamma^*)} \\ & \text{sujeito a:} \\ & \omega^* \text{ é equilibrado sob a ação de } j_0 \text{ e } p \end{aligned} \qquad (60)$$

Comparação das formulações

Primeira formulação:

1) Acha a forma em equilíbrio.
2) Em geral encontra soluções mais aproximadas à forma objetivo.

Segunda formulação:

1) Não têm variáveis relacionadas à forma do metal líquido.
2) Análise de sensibilidade fica mais simples.
3) O custo computacional do processo de otimização é muito menor para esta formulação.
4) Si for necessário, o resultado pode ser utilizado como ponto inicial da primeira formulação.

Otimização da posição indutores

A densidade de corrente elétrica j_0 é assumida da forma:

$$j_0 = I \sum_{i=1}^{n_c} \alpha_i \delta_{x_i} \ , \tag{61}$$

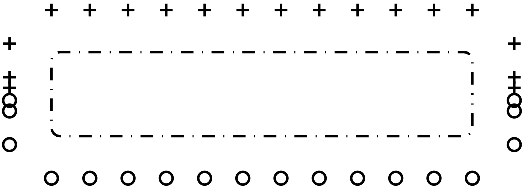
Neste caso a expressão para a função φ_1 é:

$$\varphi_1(x) = -\frac{\mu_0 I}{2\pi} \sum_{i=1}^{n_c} \alpha_i \ln \|x - x_i\| \ . \tag{62}$$

x_i variáveis de projeto do problema.

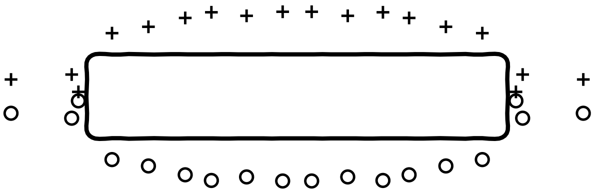
Exemplo 3

Configuração Inicial



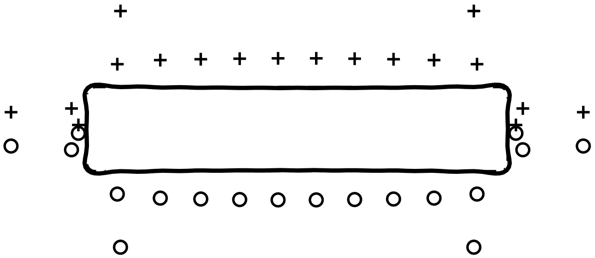
Exemplo 3

Primeira formulação



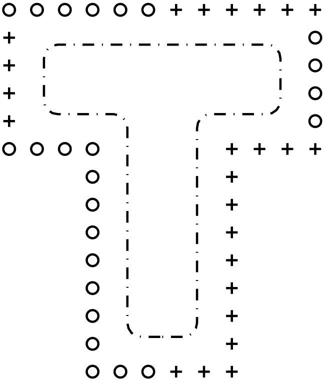
Exemplo 3

Segunda formulação



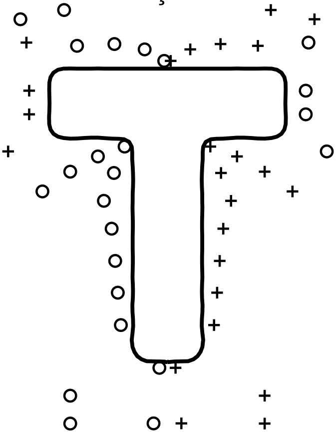
Exemplo 4

Configuração Inicial



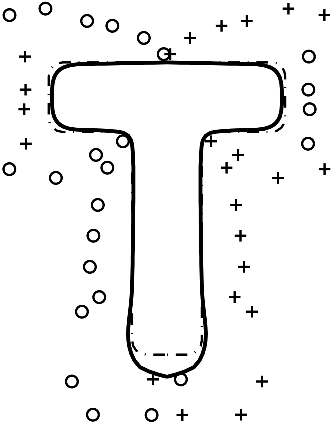
Exemplo 4

Resultado da Primeira formulação

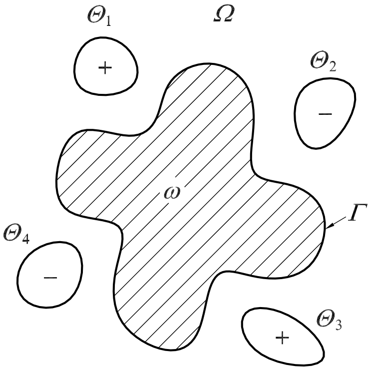


Exemplo 4

Resultado da Segunda formulação



Otimização de forma dos indutores



A hipótese de densidade uniforme é válida para o caso em que o indutor é composto por filamentos entrelaçados e isolados uns dos outros (Litz-Wire).

Otimização de forma dos indutores

A densidade de corrente elétrica j_0 é assumida da forma:

$$j_0 = I \sum_{i=1}^{n_c} \alpha_i \chi_{\Theta_i} \, , \tag{63}$$

Neste caso a expressão para a função φ_1 é:

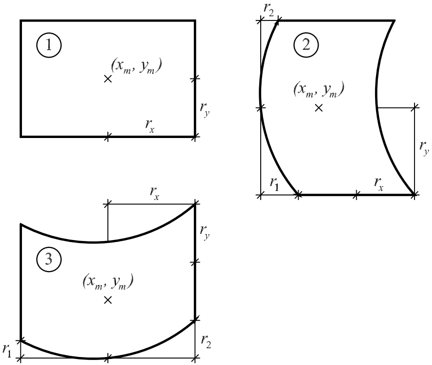
$$\varphi_1(x) = -\frac{\mu_0 I}{2\pi} \sum_{i=1}^{n_c} \alpha_i \int_{\Theta_i} \ln \|x - y\| \, \mathrm{d}\Omega_y \, . \tag{64}$$

Seja $w : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2$:

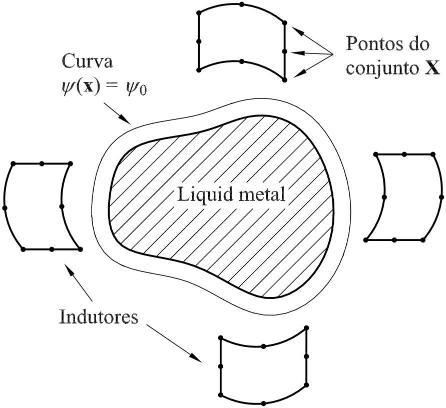
$$w(x, y) = (1/4)(1 - 2 \ln \|x - y\|)(x - y) \, . \tag{65}$$

$$\varphi_1(x) = -\frac{\mu_0 I}{2\pi} \sum_{i=1}^{n_c} \alpha_i \int_{\Gamma_i} w(x, y) \cdot \nu \, \mathrm{d}\Gamma_y \, . \tag{66}$$

Indutores



Restrições geométricas



Restrições geométricas

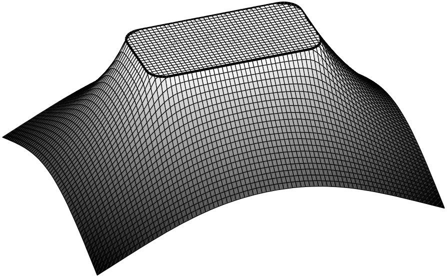
A função ψ proposta é definida pela solução de:

$$\begin{aligned} \Delta\psi(x) &= 0 \qquad \text{in } \Omega^*\,, \\ \psi(x) &= 0 \qquad \text{on } \Gamma^*\,, \\ \int_{\Gamma^*} \nabla\psi(x)\cdot\nu\,\mathrm{d}\Gamma &= -1\,. \end{aligned} \tag{67}$$

Definindo $\psi_j(\mathbf{u}_c) = \psi(x_j(\mathbf{u}_c)) - \psi_0$, as restrições geométricas ficam:

$$\boldsymbol{\psi}(\mathbf{u}_c) \leq 0\,. \tag{68}$$

Restrições geométricas

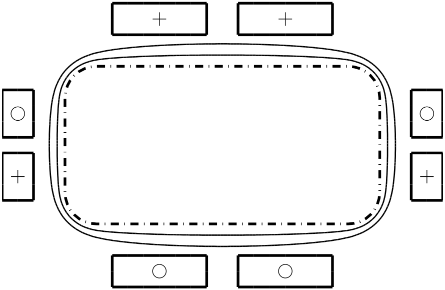


Função ψ

Conteúdo

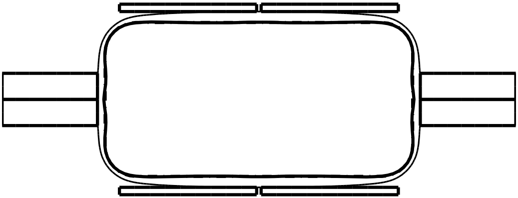
Exemplo 3

Configuração Inicial



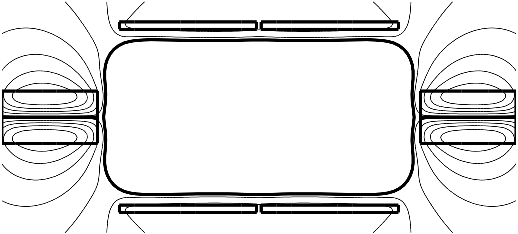
Exemplo 3

Primeira formulação



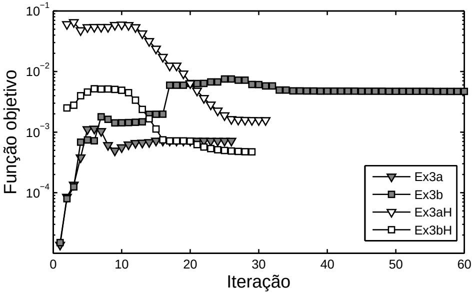
Exemplo 3

Primeira formulação



Exemplo 3

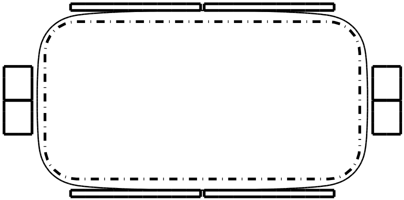
Primeira formulação



Evolução da função objetivo

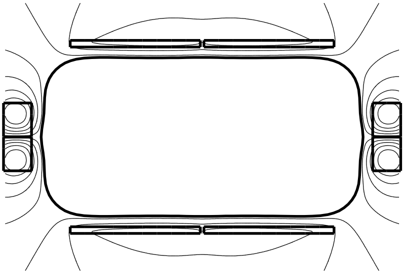
Exemplo 3

Segunda formulação



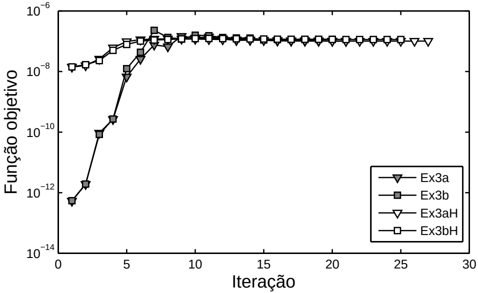
Exemplo 3

Segunda formulação



Exemplo 3

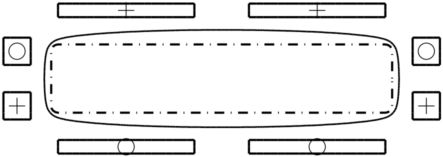
Segunda formulação



Evolução da função objetivo

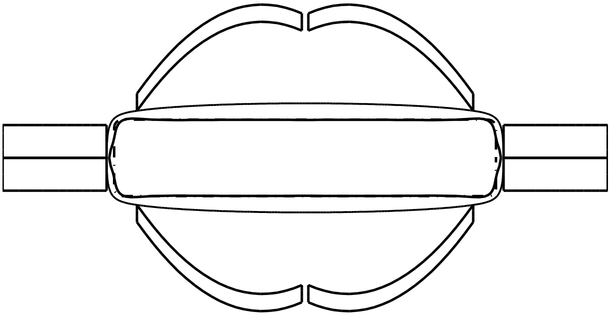
Exemplo 5

Configuração Inicial



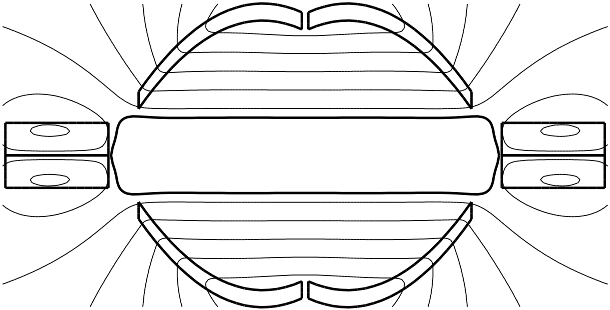
Exemplo 5

Primeira formulação



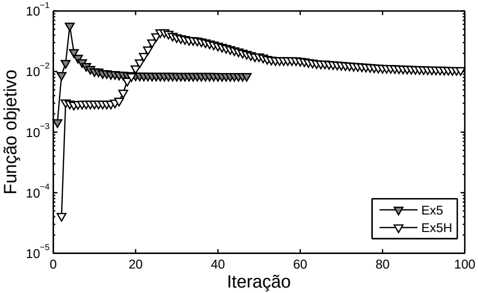
Exemplo 5

Primeira formulação



Exemplo 5

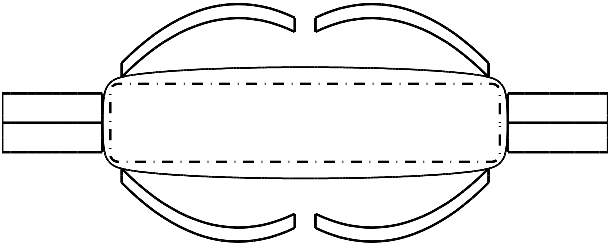
Primeira formulação



Evolução da função objetivo

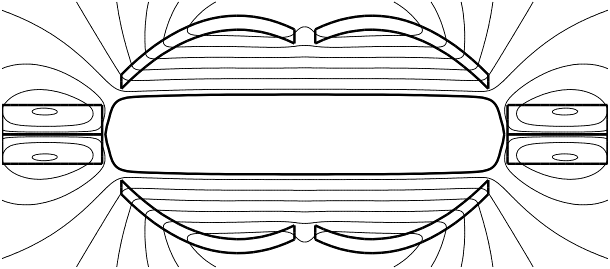
Exemplo 5

Segunda formulação



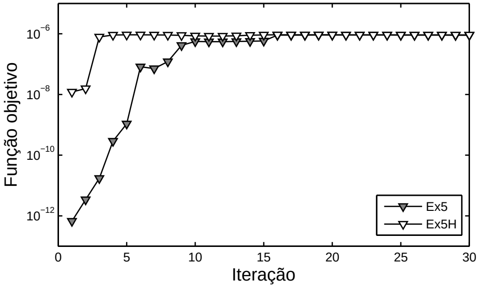
Exemplo 5

Segunda formulação



Exemplo 5

Segunda formulação



Evolução da função objetivo

Conclusões

1) Neste trabalho foi proposto um método para otimização na Engenharia Mecânica baseado na utilização do MEC, na formulação SAND e na utilização de um algoritmo de ponto interior.
2) No Capítulo 3 da tese é apresentado o algoritmo FDIPA-H.
   - Utiliza a matriz hessiana do problema de otimização nos sistemas lineares.
   - É muito eficiente para os problemas de grande porte da coleção CUTE.

Trabalhos futuros

- Ver convergência assintótica.

Conclusões

3) Uma aplicação do método proposto ao problema de Fundição Eletromagnética foi apresentada no Capítulo 5.
   - Duas formulações para a solução do problema inverso.
   - Técnica geral para satisfazer as restrições geométricas de não penetração.

   Trabalhos futuros
   - Ver o problema de otimização topológica dos indutores.
   - Considerar modelos para baixas freqüências da corrente elétrica.