ALGORITMOS BASEADOS NO MÉTODO DE NEWTON PARA PROBLEMAS DE OTIMIZAÇÃO NÃO LINEARES

Alfredo Canelas José Herskovits Sandro R. Mazorche

Departamento de Engenharia Mecânica
COPPE – UFRJ

CILAMCE, 2008

Conteúdo

Conteúdo

Problema de Otimização Não Linear

- Encontrar **x** 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 $\mathcal{N}$ tal que $\forall \mathbf{x} \in \Omega \cap \mathcal{N}$

$$f(\mathbf{x}) \geq f(\mathbf{x}^*)$$

Problema de Otimização Não Linear

- Encontrar **x** 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 $\mathcal{N}$ tal que $\forall \mathbf{x} \in \Omega \cap \mathcal{N}$

$$f(\mathbf{x}) \geq f(\mathbf{x}^*)$$

Problema de Otimização Não Linear

- Encontrar **x** 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 $\mathcal{N}$ tal que $\forall \mathbf{x} \in \Omega \cap \mathcal{N}$

$$f(\mathbf{x}) \geq f(\mathbf{x}^*)$$

Problema de Otimização Não Linear

- Regularidade: para os pontos $\mathbf{x}$:

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

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

$$\begin{aligned} \nabla f(\mathbf{x}) - \sum_{i=1}^{m} \lambda_i \nabla \mathbf{g}_i(\mathbf{x}) - \sum_{i=1}^{p} \mu_i \nabla \mathbf{h}_i(\mathbf{x}) &= 0 \\ \mathbf{g}_i(\mathbf{x})\lambda_i &= 0 \\ \mathbf{h}(\mathbf{x}) &= 0 \\ \mathbf{g}(\mathbf{x}) &\geq 0 \\ \lambda &\geq 0 \end{aligned}$$

Problema de Otimização Não Linear

- Regularidade: para os pontos $\mathbf{x}$:

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

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

$$\begin{aligned}
\nabla f(\mathbf{x}) - \sum_{i=1}^{m} \lambda_i \nabla \mathbf{g}_i(\mathbf{x}) - \sum_{i=1}^{p} \mu_i \nabla \mathbf{h}_i(\mathbf{x}) &= 0 \\
\mathbf{g}_i(\mathbf{x})\lambda_i &= 0 \\
\mathbf{h}(\mathbf{x}) &= 0 \\
\mathbf{g}(\mathbf{x}) &\geq 0 \\
\boldsymbol{\lambda} &\geq 0
\end{aligned}$$

Problema de Otimização Não Linear

- Regularidade: para os pontos **x**:

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

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

$$\boxed{\begin{aligned} \nabla f(\mathbf{x}) - \sum_{i=1}^{m} \lambda_i \nabla \mathbf{g}_i(\mathbf{x}) - \sum_{i=1}^{p} \mu_i \nabla \mathbf{h}_i(\mathbf{x}) &= 0 \\ \mathbf{g}_i(\mathbf{x})\lambda_i &= 0 \\ \mathbf{h}(\mathbf{x}) &= 0 \end{aligned}}$$

$$\begin{aligned} \mathbf{g}(\mathbf{x}) &\geq 0 \\ \boldsymbol{\lambda} &\geq 0 \end{aligned}$$

Sistema de Equações (Newton)

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

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

- **NÃO FUNCIONA**

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

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

- **NÃO FUNCIONA**

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

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

- **NÃO FUNCIONA**

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

Sistema singular

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

- **NÃO FUNCIONA**

Direção 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: $(\mathbf{x}_{k+1},\boldsymbol{\lambda}_{k+1},\boldsymbol{\mu}_{k+1}) = (\mathbf{x}_k,\boldsymbol{\lambda}_k,\boldsymbol{\mu}_k) + (\mathbf{d_x},\mathbf{d}_{\boldsymbol{\lambda}},\mathbf{d}_{\boldsymbol{\mu}})$:

Sistema singular

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

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

Conteúdo

Objetivo e Motivação

- Objetivo:
  - Definir um algoritmo de otimização baseado na direção da iteração de Newton.
- Motivação:
  - Existem problemas onde **H** pode ser calculada com um custo pequeno.
  - Maior velocidade de convergência.

Objetivo e Motivação

- Objetivo:
  - Definir um algoritmo de otimização baseado na direção da iteração de Newton.
- Motivação:
  - Existem problemas onde **H** pode ser calculada com um custo pequeno.
  - Maior velocidade de convergência.

Conteúdo

Algoritmo FDIPA

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

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

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

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

- Atualização: $\mathbf{x}_{k+1} = \mathbf{x}_k + t\,\mathbf{d}_{\mathbf{x}}$:
- A direção $\mathbf{d}_{\mathbf{x}} = \mathbf{d}_{\mathbf{x}}^{\alpha} + \rho \mathbf{d}_{\mathbf{x}}^{\beta}$:

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

Algoritmo FDIPA

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

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

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

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

- Atualização: $\mathbf{x}_{k+1} = \mathbf{x}_k + t\,\mathbf{d}_{\mathbf{x}}$:
- A direção $\mathbf{d}_{\mathbf{x}} = \mathbf{d}_{\mathbf{x}}^{\alpha} + \rho \mathbf{d}_{\mathbf{x}}^{\beta}$:

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

Algoritmo FDIPA

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

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

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

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

- Atualização: $\mathbf{x}_{k+1} = \mathbf{x}_k + t\,\mathbf{d}_{\mathbf{x}}$:
- A direção $\mathbf{d}_{\mathbf{x}} = \mathbf{d}_{\mathbf{x}}^{\alpha} + \rho \mathbf{d}_{\mathbf{x}}^{\beta}$:

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

Algoritmo FDIPA

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

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

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

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

- Atualização: $\mathbf{x}_{k+1} = \mathbf{x}_k + t\,\mathbf{d}_{\mathbf{x}}$:
- A direção $\mathbf{d}_{\mathbf{x}} = \mathbf{d}_{\mathbf{x}}^{\alpha} + \rho\mathbf{d}_{\mathbf{x}}^{\beta}$:

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

Algoritmo FDIPA

- FDIPA gera seqüência $\{\mathbf{x}_k\}_{k\in\mathbb{N}} \subset \Delta$:
$$\Delta = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{g}(\mathbf{x}) \geq 0,\ \mathbf{h}(\mathbf{x}) \geq 0\}$$
- A seqüência reduz em cada iteração o valor da função potencial $\phi_{\mathbf{c}}(\mathbf{x})$:
$$\phi_{\mathbf{c}}(\mathbf{x}) = f(\mathbf{x}) + \sum_{i=1}^{p} \mathbf{c}_i |\mathbf{h}_i(\mathbf{x})|$$
- Atualização: $\mathbf{x}_{k+1} = \mathbf{x}_k + t\,\mathbf{d}_\mathbf{x}$:
- A direção $\mathbf{d}_\mathbf{x} = \mathbf{d}_\mathbf{x}^\alpha + \rho\mathbf{d}_\mathbf{x}^\beta$:
$$\begin{pmatrix} \boxed{\mathbf{B}_k} & -\nabla\mathbf{g}_k^T & -\nabla\mathbf{h}_k^T \\ -\boldsymbol{\Lambda}_k\nabla\mathbf{g}_k & -\mathbf{G}_k & 0 \\ -\nabla\mathbf{h}_k & 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{d}_\mathbf{x}^\alpha & \mathbf{d}_\mathbf{x}^\beta \\ \boldsymbol{\lambda}^\alpha & \boldsymbol{\lambda}^\beta \\ \boldsymbol{\mu}^\alpha & \boldsymbol{\mu}^\beta \end{pmatrix} = \begin{pmatrix} -\nabla f_k^T & 0 \\ 0 & -\boldsymbol{\Lambda}_k\,\boldsymbol{\omega}^I \\ \mathbf{h}_k & -\boldsymbol{\omega}^E \end{pmatrix}$$

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

Algoritmo FDIPA

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

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

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

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

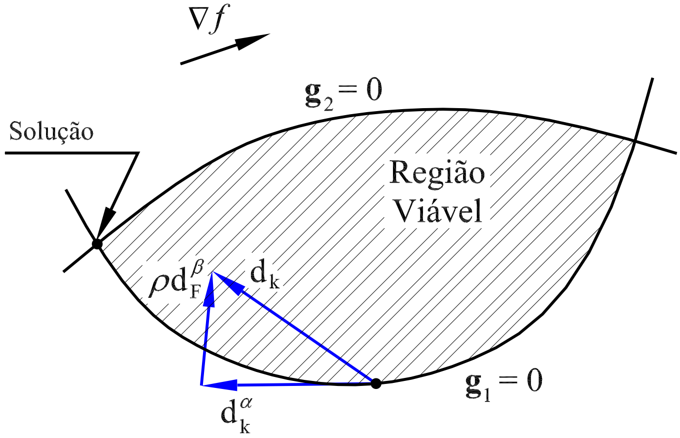
- Atualização: $\mathbf{x}_{k+1} = \mathbf{x}_k + t\,\mathbf{d}_{\mathbf{x}}$:
- A direção $\mathbf{d}_{\mathbf{x}} = \mathbf{d}_{\mathbf{x}}^{\alpha} + \boxed{\rho \mathbf{d}_{\mathbf{x}}^{\beta}}$:

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

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

$\mathbf{d}_{\mathbf{x}}^{\beta}$: direção de restauração

Direção de restauração



Convergência global do FDIPA

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

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

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

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

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

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

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

Convergência global do FDIPA

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

  $\rightarrow$ Sistema 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 + $\mathbf{B}_k$ positiva definida + $\boldsymbol{\lambda} > 0$ + $\mathbf{g}(\mathbf{x}_x) \geq 0$:

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

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

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

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

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

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

Convergência global do FDIPA

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

  $\rightarrow$ Sistema não singular

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

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

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

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

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

Convergência global do FDIPA

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

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

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

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

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

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

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

Conteúdo

Direção de Newton

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

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

- Espaço tangente:

$$\mathcal{T} = \{\mathbf{v} \mid \mathbf{v} \perp \{\nabla \mathbf{g}_i(\mathbf{x}) \mid g_i(\mathbf{x}) = 0, \nabla \mathbf{h}_i(\mathbf{x}) \mid i \in \{1, .., p\}\}\}$$

- Matriz $\mathbf{M}$:

$$\mathbf{M}(\mathbf{x}, \lambda, \mu) = \mathbf{H}(\mathbf{x}, \lambda, \mu) + \nabla \mathbf{g}^{I^T}(\mathbf{x}) \mathbf{G}^I(\mathbf{x})^{-1} \boldsymbol{\Lambda}^I \nabla \mathbf{g}^I(\mathbf{x})$$

- Função potencial:

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

Direção de Newton

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

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

- Espaço tangente:

$$\mathcal{T} = \{\mathbf{v} \mid \mathbf{v} \perp \{\nabla \mathbf{g}_i(\mathbf{x}) \mid g_i(\mathbf{x}) = 0, \nabla \mathbf{h}_i(\mathbf{x}) \mid i \in \{1, .., p\}\}\}$$

- Matriz $\mathbf{M}$:

$$\mathbf{M}(\mathbf{x}, \lambda, \mu) = \mathbf{H}(\mathbf{x}, \lambda, \mu) + \nabla \mathbf{g}^{I^T}(\mathbf{x}) \mathbf{G}^I(\mathbf{x})^{-1} \mathbf{\Lambda}^I \nabla \mathbf{g}^I(\mathbf{x})$$

- Função potencial:

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

Direção de Newton

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

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

- Espaço tangente:

$$\mathcal{T} = \{\mathbf{v} \mid \mathbf{v} \perp \{\nabla \mathbf{g}_i(\mathbf{x}) \mid g_i(\mathbf{x}) = 0, \nabla \mathbf{h}_i(\mathbf{x}) \mid i \in \{1, .., p\}\}\}$$

- Matriz **M**:

$$\mathbf{M}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = \mathbf{H}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) + \nabla \mathbf{g}^{IT}(\mathbf{x}) \mathbf{G}^{I}(\mathbf{x})^{-1} \mathbf{\Lambda}^{I} \nabla \mathbf{g}^{I}(\mathbf{x})$$

- Função potencial:

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

Direção de Newton

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

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

- Espaço tangente:

$$\mathcal{T} = \{\mathbf{v} \mid \mathbf{v} \perp \{\nabla \mathbf{g}_i(\mathbf{x}) \mid g_i(\mathbf{x}) = 0, \nabla \mathbf{h}_i(\mathbf{x}) \mid i \in \{1, .., p\}\}\}$$

- Matriz **M**:

$$\mathbf{M}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = \mathbf{H}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) + \nabla \mathbf{g}^{IT}(\mathbf{x}) \mathbf{G}^{I}(\mathbf{x})^{-1} \boldsymbol{\Lambda}^{I} \nabla \mathbf{g}^{I}(\mathbf{x})$$

- Função potencial:

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

Conteúdo

Direção de Newton

- **Assumindo**: Regularidade $+$ $\mathbf{M}_k$ positiva definida no espaço $\mathcal{T}$ $+$ $+$ $\boldsymbol{\lambda} > 0$ $+$ $\mathbf{g}(\mathbf{x}_k) \geq 0$:
- **LEMA**: $\rightarrow$ Sistema não singular
- **LEMA**: $\mathbf{x}$ não é ponto estacionário: $\rightarrow \|\mathbf{d}_\mathbf{x}^\alpha\| > 0$
- **TEOREMA**: $\mathbf{d}_\mathbf{x}^\alpha = \mathbf{d}_\mathbf{x}^1 + \mathbf{d}_\mathbf{x}^2$
  - $\mathbf{d}_\mathbf{x}^1$ 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 $\phi(\mathbf{x}) = c\,\Psi(\mathbf{h}(\mathbf{x}))$

Direção de Newton

- **Assumindo**: Regularidade $+$ $\mathbf{M}_k$ positiva definida no espaço $\mathcal{T}$ $+$ $+\ \boldsymbol{\lambda} > 0 + \mathbf{g}(\mathbf{x}_k) \geq 0$:
- **LEMA**: $\rightarrow$ Sistema não singular
- **LEMA**: $\mathbf{x}$ não é ponto estacionário: $\rightarrow \|\mathbf{d}_{\mathbf{x}}^{\alpha}\| > 0$
- **TEOREMA**: $\mathbf{d}_{\mathbf{x}}^{\alpha} = \mathbf{d}_{\mathbf{x}}^{1} + \mathbf{d}_{\mathbf{x}}^{2}$
  - $\mathbf{d}_{\mathbf{x}}^{1}$ 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 $\phi(\mathbf{x}) = c\,\Psi(\mathbf{h}(\mathbf{x}))$

Direção de Newton

- **Assumindo**: Regularidade $+$ $\mathbf{M}_k$ positiva definida no espaço $\mathcal{T}$ $+$ $+ \boldsymbol{\lambda} > 0 + \mathbf{g}(\mathbf{x}_k) \geq 0$:
- **LEMA**: $\rightarrow$ Sistema não singular
- **LEMA**: $\mathbf{x}$ não é ponto estacionário: $\rightarrow \|\mathbf{d}_{\mathbf{x}}^{\alpha}\| > 0$
- **TEOREMA**: $\mathbf{d}_{\mathbf{x}}^{\alpha} = \mathbf{d}_{\mathbf{x}}^{1} + \mathbf{d}_{\mathbf{x}}^{2}$
  - $\mathbf{d}_{\mathbf{x}}^{1}$ 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 $\phi(\mathbf{x}) = c\,\Psi(\mathbf{h}(\mathbf{x}))$

Direção de Newton

- **Assumindo**: Regularidade $+$ $\mathbf{M}_k$ positiva definida no espaço $\mathcal{T}$ $+$ $+\boldsymbol{\lambda} > 0 + \mathbf{g}(\mathbf{x}_k) \geq 0$:
- **LEMA**: $\rightarrow$ Sistema não singular
- **LEMA**: $\mathbf{x}$ não é ponto estacionário: $\rightarrow \|\mathbf{d}_\mathbf{x}^\alpha\| > 0$
- **TEOREMA**: $\mathbf{d}_\mathbf{x}^\alpha = \mathbf{d}_\mathbf{x}^1 + \mathbf{d}_\mathbf{x}^2$
  - $\mathbf{d}_\mathbf{x}^1$ 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 $\phi(\mathbf{x}) = c\,\Psi(\mathbf{h}(\mathbf{x}))$

Direção de Newton

- **Assumindo**: Regularidade $+$ $\mathbf{M}_k$ positiva definida no espaço $\mathcal{T}$ $+$ $+\boldsymbol{\lambda} > 0 + \mathbf{g}(\mathbf{x}_k) \geq 0$:
- **LEMA**: $\longrightarrow$ Sistema não singular
- **LEMA**: $\mathbf{x}$ não é ponto estacionário: $\longrightarrow \|\mathbf{d}_{\mathbf{x}}^{\alpha}\| > 0$
- **TEOREMA**: $\mathbf{d}_{\mathbf{x}}^{\alpha} = \mathbf{d}_{\mathbf{x}}^{1} + \mathbf{d}_{\mathbf{x}}^{2}$
  - $\mathbf{d}_{\mathbf{x}}^{1}$ 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 $\phi(\mathbf{x}) = c\,\Psi(\mathbf{h}(\mathbf{x}))$

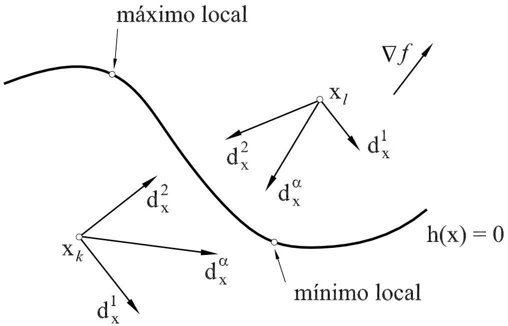
Direção de Newton

- **Assumindo**: Regularidade $+$ $\mathbf{M}_k$ positiva definida no espaço $\mathcal{T}$ $+$ $+\ \boldsymbol{\lambda} > 0 + \mathbf{g}(\mathbf{x}_k) \geq 0$:
- **LEMA**: $\rightarrow$ Sistema não singular
- **LEMA**: $\mathbf{x}$ não é ponto estacionário: $\rightarrow \|\mathbf{d}_\mathbf{x}^\alpha\| > 0$
- **TEOREMA**: $\mathbf{d}_\mathbf{x}^\alpha = \mathbf{d}_\mathbf{x}^1 + \mathbf{d}_\mathbf{x}^2$
  - $\mathbf{d}_\mathbf{x}^1$ 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 $\phi(\mathbf{x}) = c\,\Psi(\mathbf{h}(\mathbf{x}))$

Direção de Newton

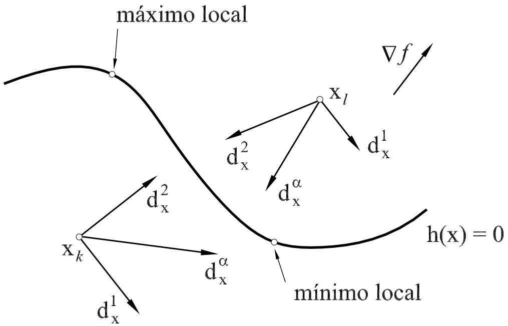
- **Assumindo**: Regularidade $+$ $\mathbf{M}_k$ positiva definida no espaço $\mathcal{T}$ $+$ $+$ $\boldsymbol{\lambda} > 0$ $+$ $\mathbf{g}(\mathbf{x}_k) \geq 0$:
- **LEMA**: $\rightarrow$ Sistema não singular
- **LEMA**: $\mathbf{x}$ não é ponto estacionário: $\rightarrow$ $\|\mathbf{d}_\mathbf{x}^\alpha\| > 0$
- **TEOREMA**: $\mathbf{d}_\mathbf{x}^\alpha = \mathbf{d}_\mathbf{x}^1 + \mathbf{d}_\mathbf{x}^2$
  - $\mathbf{d}_\mathbf{x}^1$ 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 $\phi(\mathbf{x}) = c\,\Psi(\mathbf{h}(\mathbf{x}))$

Direção de Newton



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

Direção de Newton



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

Conteúdo

$\mathbf{M}_k$?

- Duas perguntas:
  - Como saber se $\mathbf{M}_k$ positiva definida no espaço $\mathcal{T}$?
  - O que fazer se não é ?
- **TEOREMA**: $\mathbf{M}_k$ é positiva definida no espaço tangente $\mathcal{T} \Leftrightarrow$ a matriz $\mathbf{A}$ 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

$\mathbf{M}_k$?

- Duas perguntas:
  - Como saber se $\mathbf{M}_k$ positiva definida no espaço $\mathcal{T}$?
  - O que fazer se não é ?
- **TEOREMA**: $\mathbf{M}_k$ é positiva definida no espaço tangente $\mathcal{T} \Leftrightarrow$ a matriz $\mathbf{A}$ 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

$\mathbf{M}_k$?

- Duas perguntas:
  - Como saber se $\mathbf{M}_k$ positiva definida no espaço $\mathcal{T}$?
  - O que fazer se não é ?
- **TEOREMA**: $\mathbf{M}_k$ é positiva definida no espaço tangente $\mathcal{T} \Leftrightarrow$ a matriz $\mathbf{A}$ 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

$\mathbf{M}_k$?

- Duas perguntas:
  - Como saber se $\mathbf{M}_k$ positiva definida no espaço $\mathcal{T}$?
  - O que fazer se não é ?
- **TEOREMA**: $\mathbf{M}_k$ 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

$\mathbf{M}_k$?

- Duas perguntas:
  - Como saber se $\mathbf{M}_k$ positiva definida no espaço $\mathcal{T}$?
  - O que fazer se não é ?
- **TEOREMA**: $\mathbf{M}_k$ 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

$\mathbf{M}_k$?

- Duas perguntas:
  - Como saber se $\mathbf{M}_k$ positiva definida no espaço $\mathcal{T}$?
  - O que fazer se não é ?
- **TEOREMA**: $\mathbf{M}_k$ é positiva definida no espaço tangente $\mathcal{T} \Leftrightarrow$ a matriz $\mathbf{A}$ 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 $\mathbf{LDL}^T$:

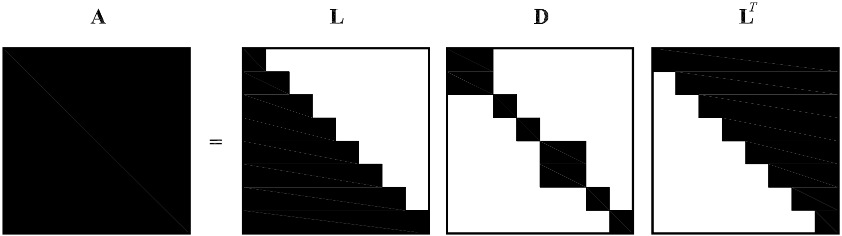
$$\mathbf{PAP}^T = \mathbf{LDL}^T$$

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

Decomposição $\mathbf{LDL}^T$

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

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

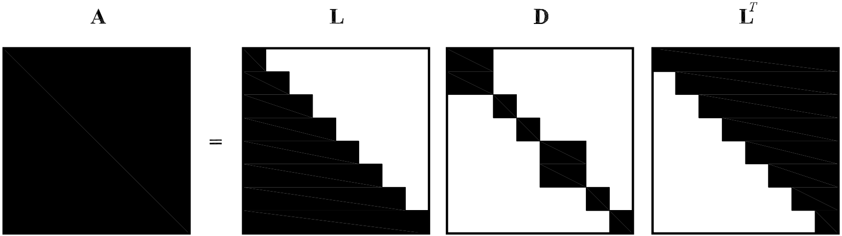


- **TEOREMA** (Sylvester): inércia($\mathbf{A}$) = inércia($\mathbf{D}$)

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($\mathbf{D}$)

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

- 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:

$$\gamma_k = \max_{1\leq i\leq n}\left\{\max\left\{(\mathbf{H}_k)_{ii}\ ,\ 1.2\sum_{j\neq i}|(\mathbf{H}_k)_{ij}|\right\}\right\}$$

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

- 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:

$$\gamma_k = \max_{1 \leq i \leq n} \left\{ \max \left\{ (\mathbf{H}_k)_{ii} \ , 1.2 \sum_{j \neq i} |(\mathbf{H}_k)_{ij}| \right\} \right\}$$

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

- 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:

$$\gamma_k = \max_{1\leq i\leq n}\left\{\max\left\{(\mathbf{H}_k)_{ii}\ ,\ 1.2\sum_{j\neq i}\left|(\mathbf{H}_k)_{ij}\right|\right\}\right\}$$

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

- 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:

$$\gamma_k = \max_{1 \leq i \leq n} \left\{ \max \left\{ (\mathbf{H}_k)_{ii} \ , 1.2 \sum_{j \neq i} |(\mathbf{H}_k)_{ij}| \right\} \right\}$$

Conteúdo

Algoritmo de Otimização

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

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:

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

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

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

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

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

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.

Conclusões

- Se **M** positiva definida em $\mathcal{T}$
  $\longrightarrow$ direção de Newton é de descida.
- **M** positiva definida em $\mathcal{T}$
  $\Leftrightarrow$ inércia(**A**) $= \{n, m+p, 0\}$.
- Um algoritmo de otimização que utiliza **H** foi apresentado.

- Trabalhos futuros
  - Provar convergência global.
  - Ver eficiência na solução de exemplos.

FIM

Conclusões

- Se **M** positiva definida em $\mathcal{T}$
  $\rightarrow$ direção de Newton é de descida.
- **M** positiva definida em $\mathcal{T}$
  $\Leftrightarrow$ inércia(**A**) $= \{n, m+p, 0\}$.
- Um algoritmo de otimização que utiliza **H** foi apresentado.
- Trabalhos futuros
  - Provar convergência global.
  - Ver eficiência na solução de exemplos.

FIM

Conclusões

- Se **M** positiva definida em $\mathcal{T}$
  $\rightarrow$ direção de Newton é de descida.
- **M** positiva definida em $\mathcal{T}$
  $\Leftrightarrow$ inércia($\mathbf{A}$) $= \{n, m+p, 0\}$.
- Um algoritmo de otimização que utiliza **H** foi apresentado.
- Trabalhos futuros
  - Provar convergência global.
  - Ver eficiência na solução de exemplos.

FIM

Conclusões

- Se **M** positiva definida em $\mathcal{T}$
  $\longrightarrow$ direção de Newton é de descida.
- **M** positiva definida em $\mathcal{T}$
  $\Leftrightarrow$ inércia(**A**) $= \{n, m+p, 0\}$.
- Um algoritmo de otimização que utiliza **H** foi apresentado.

- Trabalhos futuros
  - Provar convergência global.
  - Ver eficiência na solução de exemplos.

FIM

Conclusões

- Se **M** positiva definida em $\mathcal{T}$
  $\rightarrow$ direção de Newton é de descida.
- **M** positiva definida em $\mathcal{T}$
  $\Leftrightarrow$ inércia(**A**) $= \{n, m+p, 0\}$.
- Um algoritmo de otimização que utiliza **H** foi apresentado.

- Trabalhos futuros
  - Provar convergência global.
  - Ver eficiência na solução de exemplos.

FIM