

Capítulo 5 - Optimização

Carlos Balsa

balsa@ipb.pt

Departamento de Matemática
Escola Superior de Tecnologia e Gestão de Bragança

Matemática Aplicada - Mestrados Eng. Química e Industrial



Outline

- 1 Problemas de Optimiza  o
 - Defini  es
 - Optimiza  o Unidimensional
 - Optimiza  o Multidimensional

Outline

- 1 Problemas de Optimização
 - Definições
 - Optimização Unidimensional
 - Optimização Multidimensional

- 2 Considerações Finais

Optimiza  o

- Dada uma fun   o $f : \mathbb{R}^n \rightarrow \mathbb{R}$ e um conjunto $\mathbf{S} \subseteq \mathbb{R}^n$, encontrar $\mathbf{x}^* \in \mathbf{S}$ tal que $f(\mathbf{x}^*) \leq f(\mathbf{x})$ para todo o $\mathbf{x} \in \mathbf{S}$.

Optimizaão

- Dada uma funo $f : \mathbb{R}^n \rightarrow \mathbb{R}$ e um conjunto $\mathbf{S} \subseteq \mathbb{R}^n$, encontrar $\mathbf{x}^* \in \mathbf{S}$ tal que $f(\mathbf{x}^*) \leq f(\mathbf{x})$ para todo o $\mathbf{x} \in \mathbf{S}$.
- $\mathbf{x}^*$  chamado *minimizador* ou *mnimo* de f

Optimizaão

- Dada uma funo $f : \mathbb{R}^n \rightarrow \mathbb{R}$ e um conjunto $\mathbf{S} \subseteq \mathbb{R}^n$, encontrar $\mathbf{x}^* \in \mathbf{S}$ tal que $f(\mathbf{x}^*) \leq f(\mathbf{x})$ para todo o $\mathbf{x} \in \mathbf{S}$.
- $\mathbf{x}^*$  chamado *minimizador* ou *mnimo* de f
- Basta considerar a minimizao, pois o mximo de f  igual ao mnimo de $-f$

Optimiza  o

- Dada uma fun   o $f : \mathbb{R}^n \rightarrow \mathbb{R}$ e um conjunto $\mathbf{S} \subseteq \mathbb{R}^n$, encontrar $\mathbf{x}^* \in \mathbf{S}$ tal que $f(\mathbf{x}^*) \leq f(\mathbf{x})$ para todo o $\mathbf{x} \in \mathbf{S}$.
- $\mathbf{x}^*$   chamado *minimizador* ou *m nimo* de f
- Basta considerar a minimiza  o, pois o m ximo de f   igual ao m nimo de $-f$
- A *fun   o objectivo* f   normalmente deriv vel, podendo ser linear ao n o linear

Optimiza  o

- Dada uma fun   o $f : \mathbb{R}^n \rightarrow \mathbb{R}$ e um conjunto $\mathbf{S} \subseteq \mathbb{R}^n$, encontrar $\mathbf{x}^* \in \mathbf{S}$ tal que $f(\mathbf{x}^*) \leq f(\mathbf{x})$ para todo o $\mathbf{x} \in \mathbf{S}$.
- $\mathbf{x}^*$   chamado *minimizador* ou *m nimo* de f
- Basta considerar a minimiza  o, pois o m ximo de f   igual ao m nimo de $-f$
- A *fun   o objective* f   normalmente deriv vel, podendo ser linear ao n o linear
- O conjunto restri   o $\mathbf{S}$   definido por um sistema de equa   es e inequa   es que pode ser linear ou n o linear

Optimiza  o

- Dada uma fun  o $f : \mathbb{R}^n \rightarrow \mathbb{R}$ e um conjunto $\mathbf{S} \subseteq \mathbb{R}^n$, encontrar $\mathbf{x}^* \in \mathbf{S}$ tal que $f(\mathbf{x}^*) \leq f(\mathbf{x})$ para todo o $\mathbf{x} \in \mathbf{S}$.
- $\mathbf{x}^*$   chamado *minimizador* ou *m nimo* de f
- Basta considerar a minimiza  o, pois o m ximo de f   igual ao m nimo de $-f$
- A *fun  o objective* f   normalmente deriv vel, podendo ser linear ao n o linear
- O conjunto restri  o $\mathbf{S}$   definido por um sistema de equa  es e inequa  es que pode ser linear ou n o linear
- Os pontos $\mathbf{x} \in \mathbf{S}$ s o chamados pontos *pratic veis*

Optimiza  o

- Dada uma fun   o $f : \mathbb{R}^n \rightarrow \mathbb{R}$ e um conjunto $\mathbf{S} \subseteq \mathbb{R}^n$, encontrar $\mathbf{x}^* \in \mathbf{S}$ tal que $f(\mathbf{x}^*) \leq f(\mathbf{x})$ para todo o $\mathbf{x} \in \mathbf{S}$.
- $\mathbf{x}^*$   chamado *minimizador* ou *m nimo* de f
- Basta considerar a minimiza  o, pois o m ximo de f   igual ao m nimo de $-f$
- A *fun   o objective* f   normalmente deriv vel, podendo ser linear ao n o linear
- O conjunto restri   o $\mathbf{S}$   definido por um sistema de equa   es e inequa   es que pode ser linear ou n o linear
- Os pontos $\mathbf{x} \in \mathbf{S}$ s o chamados pontos *pratic veis*
- Se $\mathbf{S} = \mathbb{R}$ o problema n o tem restri   es

Problemas de Optimizaão

- Um problema genérico de optimizaão contínuo:

$$\min f(x) \text{ sujeito a } g(x) = 0 \text{ e } h(x) \leq 0$$

em que $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ e $h : \mathbb{R}^n \rightarrow \mathbb{R}^p$

Problemas de Optimizaão

- Um problema genérico de optimizaão contínuo:

$$\min f(x) \text{ sujeito a } g(x) = 0 \text{ e } h(x) \leq 0$$

em que $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ e $h : \mathbb{R}^n \rightarrow \mathbb{R}^p$

- *Programaão linear*: f , g e h so todas lineares

Problemas de Optimizaão

- Um problema genérico de optimizaão contínuo:

$$\min f(x) \text{ sujeito a } g(x) = 0 \text{ e } h(x) \leq 0$$

em que $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ e $h : \mathbb{R}^n \rightarrow \mathbb{R}^p$

- **Programaão linear:** f , g e h so todas lineares
- **Programaão no linear:** pelo menos uma das funões f , g e h  no linear

M todo de Newton

- *M todo de Newton* para a resolu  o de $f(x) = 0$:

$$x_{k+1} = x_k - f(x_k)/f'(x_k)$$

M todo de Newton

- *M todo de Newton* para a resolu  o de $f(x) = 0$:

$$x_{k+1} = x_k - f(x_k)/f'(x_k)$$

- Num problema de optimiza  o o M todo de Newton   usado para a resolu  o de $f'(x) = 0$:

$$x_{k+1} = x_k - f'(x_k)/f''(x_k)$$

Método de Newton

- *Método de Newton* para a resolução de $f(x) = 0$:

$$x_{k+1} = x_k - f(x_k)/f'(x_k)$$

- Num problema de optimização o Método de Newton é usado para a resolução de $f'(x) = 0$:

$$x_{k+1} = x_k - f'(x_k)/f''(x_k)$$

- Converte quadraticamente (duplica o número de dígitos correctos em cada iteração) para o mínimo desde que o ponto de partida esteja suficientemente próximo da solução

Exemplo: M todo de Newton

- Usar o m todo de Newton para minimizar $f(x) = 0.5 - xe^{-x^2}$

Exemplo: Mtodo de Newton

- Usar o mtodo de Newton para minimizar $f(x) = 0.5 - xe^{-x^2}$
- A primeira e a segunda derivada de f so dada por

$$f' = (2x^2 - 1)e^{-x^2}$$

e

$$f'' = 2x(3 - 2x^2)e^{-x^2}$$

Exemplo: M todo de Newton

- Usar o m todo de Newton para minimizar $f(x) = 0.5 - xe^{-x^2}$
- A primeira e a segunda derivada de f s o dada por

$$f' = (2x^2 - 1)e^{-x^2}$$

e

$$f'' = 2x(3 - 2x^2)e^{-x^2}$$

- Formula de recorr ncia do M todo de Newton para encontrar o zero de f'  

$$x_{k+1} = x_k - (2x^2 - 1) / (2x(3 - 2x^2))$$

Exemplo: M todo de Newton

- Usar o m todo de Newton para minimizar $f(x) = 0.5 - xe^{-x^2}$
- A primeira e a segunda derivada de f s o dada por

$$f' = (2x^2 - 1)e^{-x^2}$$

e

$$f'' = 2x(3 - 2x^2)e^{-x^2}$$

- Formula de recorr ncia do M todo de Newton para encontrar o zero de f'  

$$x_{k+1} = x_k - (2x^2 - 1) / (2x(3 - 2x^2))$$

- Usando a estimativa inicial $x_0 = 1$, obtemos

x_k	$f(x_k)$
1.000	0.132
0.500	0.111
0.700	0.071
0.707	0.071

M todo de Newton para Optimiza  o Multidimensional

- M todo de Newton para minimizar uma fun   o unidimensional, procura o zero de $f(x)$ atrav s da recorr ncia:

$$x_{k+1} = x_k - f'(x_k)/f''(x_k)$$

M todo de Newton para Optimiza  o Multidimensional

- M todo de Newton para minimizar uma fun  o unidimensional, procura o zero de $f(x)$ atrav s da recorr ncia:

$$x_{k+1} = x_k - f'(x_k)/f''(x_k)$$

- No caso multidimensional o m todo de Newton   usado para procurar o zero do gradiente da fun  o, $\nabla f(\mathbf{x}) = 0$, atrav s da formula de recorr ncia:

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \mathbf{H}_f^{-1}(\mathbf{x}_k) \nabla f(\mathbf{x}_k)$$

em que $\mathbf{H}(\mathbf{x})$   a matriz Hessiana das segundas derivadas parciais

$$\{\mathbf{H}(\mathbf{x})\}_{ij} = \frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j}$$

Método de Newton, continuação

- A matriz Hessiana não é invertida explicitamente, em vez disso resolve-se o sistema linear

$$\mathbf{H}_f(\mathbf{x}_k)\delta_k = -\nabla f(\mathbf{x})$$

em ordem a δ_k , e depois actualiza-se a solução aproximada

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \delta_k$$

Método de Newton, continuação

- A matriz Hessiana não é invertida explicitamente, em vez disso resolve-se o sistema linear

$$\mathbf{H}_f(\mathbf{x}_k)\delta_k = -\nabla f(\mathbf{x})$$

em ordem a δ_k , e depois actualiza-se a solução aproximada

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \delta_k$$

- Convergência quadrática mas necessita de uma boa estimativa inicial da solução

Método de Newton, continuação

- A matriz Hessiana não é invertida explicitamente, em vez disso resolve-se o sistema linear

$$\mathbf{H}_f(\mathbf{x}_k)\delta_k = -\nabla f(\mathbf{x})$$

em ordem a δ_k , e depois actualiza-se a solução aproximada

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \delta_k$$

- Convergência quadrática mas necessita de uma boa estimativa inicial da solução
- Necessita da segunda derivada da função

Exemplo: M todo de Newton

- Usar o m todo de Newton para minimizar

$$f(\mathbf{x}) = 0.5x_1^2 + 2.5x_2^2$$

Exemplo: M todo de Newton

- Usar o m todo de Newton para minimizar

$$f(\mathbf{x}) = 0.5x_1^2 + 2.5x_2^2$$

- Gradiente e matriz Hessiana s o dados por

$$\nabla f(\mathbf{x}) = \begin{bmatrix} x_1 \\ 5x_2 \end{bmatrix} \quad \text{e} \quad \mathbf{H}_f(\mathbf{x}) = \begin{bmatrix} 1 & 0 \\ 0 & 5 \end{bmatrix}$$

Exemplo: M todo de Newton

- Usar o m todo de Newton para minimizar

$$f(\mathbf{x}) = 0.5x_1^2 + 2.5x_2^2$$

- Gradiente e matriz Hessiana s o dados por

$$\nabla f(\mathbf{x}) = \begin{bmatrix} x_1 \\ 5x_2 \end{bmatrix} \quad \text{e} \quad \mathbf{H}_f(\mathbf{x}) = \begin{bmatrix} 1 & 0 \\ 0 & 5 \end{bmatrix}$$

- Escolhendo $\mathbf{x}_0 = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$ obtemos $\nabla f(\mathbf{x}_0) = \begin{bmatrix} 5 \\ 5 \end{bmatrix}$

Exemplo: Mtodo de Newton

- Usar o mtodo de Newton para minimizar

$$f(\mathbf{x}) = 0.5x_1^2 + 2.5x_2^2$$

- Gradiente e matriz Hessiana so dados por

$$\nabla f(\mathbf{x}) = \begin{bmatrix} x_1 \\ 5x_2 \end{bmatrix} \quad \text{e} \quad \mathbf{H}_f(\mathbf{x}) = \begin{bmatrix} 1 & 0 \\ 0 & 5 \end{bmatrix}$$

- Escolhendo $\mathbf{x}_0 = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$ obtemos $\nabla f(\mathbf{x}_0) = \begin{bmatrix} 5 \\ 5 \end{bmatrix}$
- Resolvendo o sistema linear $\mathbf{H}_f(\mathbf{x}_k)\delta_k = -\nabla f(\mathbf{x}_k)$, obtemos o passo $\delta_k = \begin{bmatrix} -5 \\ -1 \end{bmatrix}$, e $\mathbf{x}_1 = \mathbf{x}_0 + \delta_0 = \begin{bmatrix} 5 \\ 1 \end{bmatrix} + \begin{bmatrix} -5 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ que  a soluo exacta deste problema

M todo do Gradiente Conjugado

- Um outro m todo que n o requer explicitamente as segundas derivadas   o m todo do *gradiente conjugado* (CG)

Método do Gradiente Conjugado

- Um outro método que não requer explicitamente as segundas derivadas é o método do *gradiente conjugado* (CG)
- CG gera uma sequência de direcções de procura conjugadas (ortogonais) entre elas

M  todo do Gradiente Conjugado

- Um outro m  todo que n  o requer explicitamente as segundas derivadas    o m  todo do *gradiente conjugado* (CG)
- CG gera uma sequ  ncia de direc   es de procura conjugadas (ortogonais) entre elas
- Teoricamente, para fun   es objectivo quadr  ticas o CG converge para a solu   o exacta num numero m  ximo de n itera   es, em que n    a dimens  o do problema

M todo do Gradiente Conjugado

- Um outro m todo que n o requer explicitamente as segundas derivadas   o m todo do *gradiente conjugado* (CG)
- CG gera uma sequ ncia de direc   es de procura conjugadas (ortogonais) entre elas
- Teoricamente, para fun   es objectivo quadr ticas o CG converge para a solu   o exacta num numero m ximo de n itera   es, em que n   a dimens  o do problema
- Tamb m   eficiente para os outros tipos de problemas de minimiza   o sem restri   es

M todo do Gradiente Conjugado, continua  o

ALGORITMO DO GRADIENTE CONJUGADO

$\mathbf{x}_0$ = aproxima  o inicial

$\mathbf{g}_0 = \nabla f(\mathbf{x}_0)$

$\mathbf{s}_0 = -\mathbf{g}_0$

for $k = 0, 1, 2, \dots$

Escolher o α_k que minimiza $f(\mathbf{x}_k + \alpha_k \mathbf{s}_k)$

$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{s}_k$

$\mathbf{g}_{k+1} = \nabla f(\mathbf{x}_{k+1})$

$\beta_{k+1} = (\mathbf{g}_{k+1}^T \mathbf{g}_{k+1}) / (\mathbf{g}_k^T \mathbf{g}_k)$

$\mathbf{s}_{k+1} = -\mathbf{g}_{k+1} + \beta_{k+1} \mathbf{s}_k$

end

Exemplo: M todo do Gradiente Conjugado

- Usar o CG para minimizar

$$f(\mathbf{x}) = 0.5x_1^2 + 2.5x_2^2$$

Exemplo: M todo do Gradiente Conjugado

- Usar o CG para minimizar

$$f(\mathbf{x}) = 0.5x_1^2 + 2.5x_2^2$$

- Gradiente   igual a $\nabla f(\mathbf{x}) = \begin{bmatrix} x_1 \\ 5x_2 \end{bmatrix}$

Exemplo: M todo do Gradiente Conjugado

- Usar o CG para minimizar

$$f(\mathbf{x}) = 0.5x_1^2 + 2.5x_2^2$$

- Gradiente   igual a $\nabla f(\mathbf{x}) = \begin{bmatrix} x_1 \\ 5x_2 \end{bmatrix}$
- Escolhendo $\mathbf{x}_0 = \begin{bmatrix} 5 & 1 \end{bmatrix}$ obtemos a direc  o de procura inicial (corresponde ao gradiente negativo)

$$\mathbf{s}_0 = -\mathbf{g}_0 = -\nabla f(\mathbf{x}_0) = \begin{bmatrix} -5 \\ -5 \end{bmatrix}$$

Exemplo: Método do Gradiente Conjugado

- Usar o CG para minimizar

$$f(\mathbf{x}) = 0.5x_1^2 + 2.5x_2^2$$

- Gradiente é igual a $\nabla f(\mathbf{x}) = \begin{bmatrix} x_1 \\ 5x_2 \end{bmatrix}$
- Escolhendo $\mathbf{x}_0 = \begin{bmatrix} 5 & 1 \end{bmatrix}$ obtemos a direcção de procura inicial (corresponde ao gradiente negativo)

$$\mathbf{s}_0 = -\mathbf{g}_0 = -\nabla f(\mathbf{x}_0) = \begin{bmatrix} -5 \\ -5 \end{bmatrix}$$

- O mínimo exacto ao longo da linha de procura é $\alpha_0 = 1/3$, pelo que a próxima solução aproximada é $\mathbf{x}_1 = \begin{bmatrix} 3.333 & -0.667 \end{bmatrix}^T$ com o qual calculamos o novo gradiente

$$\mathbf{g}_1 = \nabla f(\mathbf{x}_1) = \begin{bmatrix} 3.333 \\ -3.333 \end{bmatrix}$$

Exemplo, continuação

- Neste ponto, em vez de procurar na direcção do gradiente negativo calcula-se

$$\beta_1 = (\mathbf{g}_1^T \mathbf{g}_1) / (\mathbf{g}_0^T \mathbf{g}_0) = 0.444$$

que origina a nova direcção de procura

$$\mathbf{s}_1 = -\mathbf{g}_1 + \beta_1 \mathbf{s}_0 = \begin{bmatrix} -3.333 \\ 3.333 \end{bmatrix} + 0.444 \begin{bmatrix} -5 \\ -5 \end{bmatrix} = \begin{bmatrix} -5.556 \\ 1.111 \end{bmatrix}$$

Exemplo, continuação

- Neste ponto, em vez de procurar na direcção do gradiente negativo calcula-se

$$\beta_1 = (\mathbf{g}_1^T \mathbf{g}_1) / (\mathbf{g}_0^T \mathbf{g}_0) = 0.444$$

que origina a nova direcção de procura

$$\mathbf{s}_1 = -\mathbf{g}_1 + \beta_1 \mathbf{s}_0 = \begin{bmatrix} -3.333 \\ 3.333 \end{bmatrix} + 0.444 \begin{bmatrix} -5 \\ -5 \end{bmatrix} = \begin{bmatrix} -5.556 \\ 1.111 \end{bmatrix}$$

- A deslocação ao longo desta linha que minimiza a função é $\alpha_1 = 0.6$, que origina a solução exacta $\mathbf{x}_2 = \begin{bmatrix} 0 & 0 \end{bmatrix}^T$, tal como seria de esperar para uma função quadrática

Métodos Disponíveis na NMLibforOctave

- Método de Newton: `[] = OPT_NEWTON()`

Métodos Disponíveis na NMLibforOctave

- Método de Newton: `[] = OPT_NEWTON()`
- Mét. Gradiente Conjugado: `[] = OPT_CG()`

Bibliografia

Exposi  o baseada essencialmente no cap  tulo 6 de

- Michael T. Heath. "Scientific Computing an Introductory Survey". McGraw-Hill, 2002, New York.