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

Optimização

- Dada uma função $f : \mathbb{R}^n \to \mathbb{R}$ 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 $\mathbf{x} \in \mathbf{S}$.
- $\mathbf{x}^*$ é chamado *minimizador* ou *mínimo* de f
- Basta considerar a minimização, pois o máximo de f é igual ao mínimo de $-f$
- A *função objectivo* f é normalmente derivável, podendo ser linear ao não linear
- O conjunto restrição $\mathbf{S}$ é definido por um sistema de equações e inequações que pode ser linear ou não linear
- Os pontos $\mathbf{x} \in \mathbf{S}$ são chamados pontos *praticáveis*
- Se $\mathbf{S} = \mathbb{R}$ o problema não tem restrições

Problemas de Optimização

- Um problema genérico de optimização contínuo:

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

  em que $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}^m$ e $h : \mathbb{R}^n \to \mathbb{R}^p$
- *Programação linear*: f, g e h são todas lineares
- *Programação não linear*: pelo menos uma das funções f, g e h é não linear

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 $f'(x) = 0$:

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

- Converge 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: Método de Newton

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

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

e

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

- Formula de recorrência do Método de Newton para encontrar o zero de f' é

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

- 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

Método de Newton para Optimização Multidimensional

- Método de Newton para minimizar uma função unidimensional, procura o zero de $f(x)$ através da recorrência:

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

- No caso multidimensional o método de Newton é usado para procurar o zero do gradiente da função, $\nabla f(\mathbf{x}) = 0$, através da formula de recorrência:

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

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

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

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

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

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

Exemplo: Método de Newton

- Usar o método de Newton para minimizar

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

- Gradiente e matriz Hessiana são 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 $\mathbf{H}_f(\mathbf{x}_k)\delta_k = -\nabla f(\mathbf{x})$, 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 é a solução exacta deste problema

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
- Teoricamente, para funções objectivo quadráticas o CG converge para a solução exacta num numero máximo de n iterações, em que n é a dimensão do problema
- Também é eficiente para os outros tipos de problemas de minimização sem restrições

Método do Gradiente Conjugado, continuação

ALGORITMO DO GRADIENTE CONJUGADO

$\mathbf{x}_0 =$ aproximação inicial

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

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

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

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

$$
\begin{aligned}
\mathbf{x}_{k+1} &= \mathbf{x}_k + \alpha_k \mathbf{s}_k \\
\mathbf{g}_{k+1} &= \nabla f(\mathbf{x}_{k+1}) \\
\beta_{k+1} &= \left(\mathbf{g}_{k+1}^T \mathbf{g}_{k+1}\right) / \left(\mathbf{g}_k^T \mathbf{g}_k\right) \\
\mathbf{s}_{k+1} &= -\mathbf{g}_{k+1} + \beta_{k+1} \mathbf{s}_k
\end{aligned}
$$

end

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

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

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