Trabalho 3 - Avaliação de PNL

Otimização 2025.1

Pedro Saito 122149392

Marcos Silva 122133854

Milton Salgado 122169279

1. Questão 1

Enunciado: Para os seguintes problemas de otimização não linear irrestrita, estude de forma detalhada e conceitual , as condições de otimalidade de primeira e segunda orde, apresentando todos os cálculos . Ache os pontos críticos, extremo local e global se existirem.

$\mathbf{1.1}$) Minimize $x_1 x_2, (x_1, x_2 \in \mathbb{R})$.

As condições de otimalidade de primeira ordem garantem que, para que um ponto x^* seja um mínimo ou máximo local de f, é necessário que o gradiente em x^* seja nulo, isto é,

$$\nabla f(x^*) = 0.$$

Desse modo, como o gradiente da função f é dado por (x_2, x_1), seu único ponto crítico é $(0, 0)$. Com isso, vamos analisar as condições de otimalidade de segunda ordem para a matriz Hessiana de f.

Dada uma função de duas variáveis $f(\boldsymbol{x}) = f(x_1, x_2)$ com os pontos críticos $(0, 0)$, calcularemos as derivadas de segunda ordem para cada componente. Assim, teremos:

$$H_1 = \left[\frac{\partial^2 f}{\partial x_1^2}(x^*)\right] = [0] \quad \text{e} \quad H_2 = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(\boldsymbol{x}) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(\boldsymbol{x}) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(\boldsymbol{x}) & \frac{\partial^2 f}{\partial x_2^2}(\boldsymbol{x}) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Pelo critério de menores principais em $\mathbb{R}^2$, vamos analisar os determinantes:

$$\det(H_1) = 0 \quad \text{e} \quad \det(H_2) = -1$$

Como podemos observar, o primeiro determinante $\det(H_1)$ é zero e, portanto, o teste é *inconclusivo*. Realizando uma análise da função nos quadrantes de f temos:

- Para $x_1 > 0$ e $x_2 > 0 : f > 0$.
- Para $x_1 < 0$ e $x_2 < 0 : f > 0$.
- Para $x_1 > 0$ e $x_2 < 0 : f < 0$.
- Para $x_1 < 0$ e $x_2 > 0 : f < 0$.

Com efeito, a origem $(0, 0)$ é um **ponto de sela**, pois a função assume apenas valores positivos em alguns quadrantes e negativos em outros, indicando a presença de direções de crescimento e decrescimento próximas ao ponto crítico. Assim, como não tem pontos de mínimo, a função f é **ilimitada.** ■

1.2) Minimize $(x_1 - 1)^2 - x_1 x_2 + (x_2 - 1)^2, (x_1, x_2 \in \mathbb{R}^2)$.

Assim como na questão anterior, calcularemos as condições de otimalidade de primeira e segunda ordem. Nesse sentido, o gradiente da função é $(2x_1 - x_2 - 2, 2x_2 - 2 - x_1)$. Portanto, igualando cada componente a zero encontramos seus pontos críticos:

$$\begin{cases} 2x_1 - x_2 - 2 = 0 \\ 2x_2 - 2 - x_1 = 0. \end{cases}$$

Obtemos o ponto crítico $(2, 2)$. Observando a matriz Hessiana e calculando seus menores principais, descobrimos que todos os seus menores principais são positivos.

$$|H_1| = \left| \begin{bmatrix} \frac{\partial f^2}{\partial x_1^2} \end{bmatrix} \right| = |[2]| = 2$$

$$|H_2| = \left| \begin{bmatrix} \frac{\partial f^2}{\partial x_1^2} & \frac{\partial f^2}{\partial x_1 \partial x_2} \\ \frac{\partial f^2}{\partial x_2 \partial x_1} & \frac{\partial f^2}{\partial x_2^2} \end{bmatrix} \right| = \left| \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \right| = 3$$

Logo, pelo teste de Sylvester, $f(2, 2) = -2$ é mínimo local. Além disso, por ser o único ponto crítico da função, também é mínimo global. ■

2. Questão 2

Enunciado: Faça uma síntese sobre um dos tópicos de otimização não linear irrestrita, exemplifique:

- Método de Newton (Escolhemos esse).
- Métodos de Máxima Descida

2.1. Método de Newton Clássico

Em análise numérica, o método de Newton-Raphson é um dos métodos mais eficientes para obtenção das raízes de uma função $f(x) = 0$. Começamos escolhendo a entrada inicial x_0 da função e calculamos a reta tangente ao ponto atual juntamente de sua interseção com o eixo das abcissas. Assim, repetimos esse processo até obter a interseção que corresponde à raiz da função.

O método de Newton é descrito pela seguinte sequência recursiva:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, n \in \mathbb{N}$$

onde x_n é a n-ésima iteração do algoritmo e $f'(x_n)$ é a derivada em x_n. Começamos a partir de uma estimativa x_0 e então iteramos até que o erro $\varepsilon_n = x_{n+1} - x_n$ seja menor que alguma tolerância aceitável que definimos.

2.2. Método de Newton em Otimização Não Linear Irrestrita

No contexto de otimização não linear irrestrita, o método de Newton é uma técnica iterativa que utiliza informações da função e de suas derivadas para encontrar os pontos críticos onde o gradiente da função é zero, podendo este ser um ponto de mínimo, máximo ou ponto de sela. Este método corresponde a uma extensão do método de Newton clássico para encontrar as raízes.

Para uma função $f : \mathbb{R}^n \to \mathbb{R}$, a ideia central do método é utilizar uma aproximação quadrática da função para iterativamente obter um ponto crítico. A estimativa da solução a cada iteração é calculada da seguinte forma:

$$x_{k+1} = x_k - H_f(x_k)^{-1} \nabla f(x_k)$$

Onde:

- x_k é o ponto atual na iteração k.
- $\nabla f(x_k)$ é o gradiente de f avaliado em x_k.
- $H_f(x_k)$ é a matriz Hessiana de f avaliada em x_k, ou seja, a matriz de segunda derivada de f.

2.2.1. Passo-a-Passo

1. **Gradiente**: Calcular o gradiente $\nabla f(x)$, que é um vetor de primeiras derivadas parciais da função f em relação a cada nível.
2. **Hessiana**: Calcular a matriz Hessiana $H_f(x)$, que é uma matriz de segundas derivadas parciais da função f definida na questão anterior.
3. **Resolução do Sistema Linear**: Resolver o sistema linear $H_f(x)d = -\nabla f(x_k)$ para encontrar a direção d. Onde $H_f(x_k)^{-1} \nabla f(x_k)$ é a direção do passo de Newton.
4. **Atualização da Solução**: Atualizar a solução x_k usando $x_{k+1} = x_k + d$.
5. **Repetição**: Repetir os passos anteriores até que a norma do gradiente $\|\nabla f(x_k)\|$ seja suficientemente pequena, indicando que x_k está próximo de um ponto crítico.

2.2.2. Prova

A demonstração do Método de Newton no contexto de otimização não linear irrestrita envolve a seguinte demonstração:

Convergência Local Quadrática: Se a aproximação inicial x_0 estiver suficientemente próxima ao ponto crítico x^* onde o gradiente $\nabla f(x^*) = 0$ e a Hessiana $H_f(x^*)$ é positiva definida, então a sequência gerada pelo Método de Newton convergirá para x^* com taxa quadrática.

Além disso, são necessárias algumas condições a mais para a convergência, tal como:

- **Matriz Hessiana Positiva Definida**: A matriz Hessiana $H_f(x^*)$ no ponto ótimo deve ser positiva definida. Isso nos garante que x^* seja um mínimo local e que o método converge.
- **Proximidade Inicial**: A aproximação x_0 deve ser próxima o bastante do ponto crítico x^*.
- **Classe C^2**: A função deve ser duas vezes continuamente diferenciável. Isso permite a utilização de expansão de Taylor para aproximar a função.

Primeiramente, considere as seguintes suposições:

- x^* é um ponto crítico tal que $H_f(x^*)$ é positiva definida.
- A função f pertence à classe C^2, ou seja, possui derivadas contínuas até a segunda ordem.

Como havíamos mencionado, a função f pode ser aproximada ao redor de x_k usando a expansão de Taylor de segunda ordem:

$$f(x) \approx f(x_k) + \nabla f(x_k)^T (x - x_k) + \frac{1}{2}(x - x_k)^T H_f(x_k)(x - x_k)$$

No ponto x_k, o gradiente é dado por $\nabla f(x_k)$. Assim, a atualização corresponde a;

$$x_{k+1} = x_k - H_f(x_k)^{-1} \nabla f(x_k).$$

Usando a expansão de Taylor para o gradiente $\nabla f(x)$ em torno de x_k:

$$\nabla f(x_{k+1}) \approx \nabla f(x_k) + H_f(x_k)\left(-H_f(x_k)^{-1} \nabla f(x_k)\right) = 0$$

Com base nisso, concluímos que o gradiente em x_{k+1} é zero, indicando que x_{k+1} é um ponto crítico. Para mostrar a convergência quadrática, vamos considerar a diferença $e_{k+1} = x_{k+1} - x^*$. Usando a definição de x_{k+1} e considerando que $\nabla f(x^*) = 0$:

$$e_{k+1} = x_{k+1} - x^* = x_k - H_f(x_k)^{-1} \nabla f(x_k) - x^*$$

Como x_k está próximo de x^*, podemos usar a expansão de Taylor para $\nabla f(x_k)$:

$$\nabla f(x_k) = H_f(x^*)(x_k - x^*) + \mathcal{O}\left(\|x_k - x^*\|^2\right)$$

Simplificando e usando $e_k = x_k - x^*$:

$$e_{k+1} = \left(I - H_f(x_k)^{-1} H_{f(x^*)}\right) e_k - H_f(x_k)^{-1} \mathcal{O}\left(\|e_k\|^2\right)$$

Se x_k está suficientemente próximo de x^*, podemos aproximar $H_f(x_k) \approx H_f(x^*)$, então:

$$
\begin{aligned}
e_{k+1} &\approx -H_f(x^*)^{-1}\mathcal{O}\left(\|e_k\|^2\right) \\
\|e_{k+1}\| &\approx C\|e_k\|^2
\end{aligned}
$$

Onde C é uma constante, demonstrando que o erro decresce quadraticamente.

Bibliografia

(1) Wikipedia. Hessiano, 2024.

(2) Wikipedia. Método de Newton em otimização, 2024.

(3) Wikipedia. Otimização, 2024.

(4) Andretta, M. Gradientes, 2011.

(5) Wikipedia. Método de Newton–Raphson, 2024.