Ministério da Educação
Universidade Federal do Agreste
de Pernambuco

Relatório Final
Métodos de Otimização

$\propto$

Ministério da Educação
Universidade Federal do Agreste
de Pernambuco

Relatório Final
Métodos de Otimização

Mateus Baltazar de Almeida
Matheus Machado Vieira
Orientador: Gersonilo Oliveira da Silva

$\propto$

Sumário

Capítulo 1

Métodos Matemáticos de Otimização

1.1 O Conceito de Otimização

Diz-se otimização, o processo que tem como objetivo encontrar condições que minimizam ou maximizam algo (seja energia, tempo, dinheiro, etc). Sendo este, muitas vezes um trabalho árduo, custoso.

Dessa maneira, na matemática, tal processo é amplamente utilizado quando busca-se valores em conjunto um A (que pode ter restrições), com o objetivo de encontrar uma solução ótima (a melhor resposta para o problema), aplicando os valores de A em uma função objetivo predefinida.

Podendo assim, ser representada da forma como a seguir.

Dada uma função

$$f : A \subseteq \mathbb{R} \to \mathbb{R}. \tag{1.1}$$

A maximização pode ser definida como; a busca pelo elemento $x^* \in A$, que satisfaz:

$$f(x^*) \geq f(x), \tag{1.2}$$

para todo $x \in A$.

A minimização pode ser definida como; a busca pelo elemento $x^* \in A$, que satisfaz:

$$f(x^*) \leq f(x), \tag{1.3}$$

para todo $x \in A$.

Com isso, podemos agora entender como esse processo pode ser custoso. Iniciando com o fato de que existem os pontos máximos e mínimos (pontos críticos[1]), locais e globais, na imagem das funções. Sendo os pontos críticos locais, aqueles que não são os menores ou maiores valores para a minimização e maximização, respectivamente. E os pontos globais, aqueles que representam o menor ou maior valor na imagem das funções, para a minimização e maximização, respectivamente.

[1] Iremos denotar como pontos críticos apenas aqueles onde a função tem derivada e esta se anula.

Criando, desse modo, uma certa incerteza ao encontrar um valor crítico numa função, já que, é estritamente difícil saber se o ponto crítico encontrado é local ou global. Como pode-se perceber na Figura 1.1.

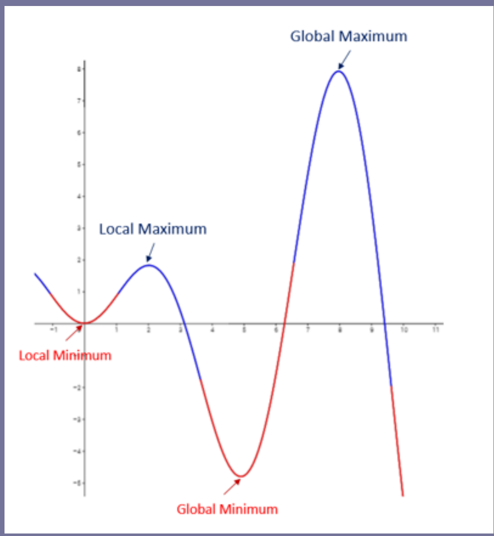

Figura 1.1: Exemplo de pontos críticos locais e globais indicados no gráfico de uma função.

Como na grande maioria dos casos não é possível conhecer antecipadamente todo o espaço gerado pela função, a busca pelo ótimo pode resultar em pontos críticos locais. Donde, dependendo do problema, tais resultados chegam a ser satisfatórios. No entanto, quando o objetivo é estritamente encontrar os máximos ou mínimos globais, o trabalho necessário acaba sendo mais custoso, pelo fato de que há a incerteza da existência de mais pontos na grande maioria dos problemas.

Ademais, o conjunto A, o qual contém os valores aplicáveis no problema, podem respeitar restrições, fazendo com que o campo de respostas do problema, não seja contínuo.

1.2 Otimização de Funções à Uma Variável Real

Evidentemente, as funções possuem as variáveis dependentes (que representam o objeto da otimização) e as variáveis independentes (que suas grandezas podem ser selecionadas), podemos denotar que, para a equação

$$y = f(x), \tag{1.4}$$

quando buscamos otimizá-la, temos como objetivo encontrar valores, x, que quando aplicados à $f(x)$, temos o mínimo ou máximo valor y (seja ele local, ou preferencialmente, global).

Partindo dessa perspectiva, acaba surgindo a necessidade de utilizar algum recurso para encontrar os pontos críticos. E nesse sentindo, pode-se utilizar a técnica de **derivação**, que oferece o recurso de identificar tais pontos.

Podemos observar um fator interessante; por exemplo, quando a função está num ponto máximo, existem duas possibilidades, a primeira, a função para de crescer e, em seguinda, torna-se indefinida; e a segunda possibilidade, quando a função para de crescer e começa a decrescer. Com isso, é importante ressaltar que por definição, quando a derivada (taxa de variação) em um ponto é positiva, a função cresce, e quando é negativa a função decresce. Conclui-se que, quando a taxa de variação é zero, a função ou para de crescer ou de decrescer, o que configura um ponto de máximo ou de mínimo. Mas também, é quando podo ocorrer o que denominamos ponto de sela (é um comportamento errático da função, que precisa de uma análise mais refinada). Ou em última instância, quando a função é constante.

Com o uso da derivada, podemos pensar num método de otimização bastante simples, considerando $f(x)$ a função que queremos otimizar e $f'(x)$ sua função derivada, podemos dizer que o conjunto O, abaixo definido, possui todos os ótimos locais e globais de $f(x)$:

$$O := \{f(x)|f'(x) = 0\}. \tag{1.5}$$

Portanto, aplicando um filtro em O para obter o máximo e mínimo do conjunto, acabamos por obter o máximo e mínimo de $f(x)$:

$$max(f(x)) = max(O), \tag{1.6}$$

$$min(f(x)) = min(O). \tag{1.7}$$

Então podemos perceber três problemas:

- Determinar como encontrar os pontos onde a derivada se anule, isto é, os pontos críticos;
- Determinar se temos de fato todos os pontos;
- Classificação dos pontos críticos, que nos revela quais são os máximos e os mínimos, e, ademais, da sua relevância quanto a serem locais ou globais.

Consideraremos, por agora, apenas o problema de encontrar os pontos críticos.

A taxa de variação de uma função, $y = f(x)$, em relação a x, é dada pela relação $\Delta y/\Delta x$. Sendo o cálculo da derivada demonstrado pela Figura 1.2, o que nos leva ao fato de que os pontos críticos provém do conjunto solução da equação $f'(x) = 0$, como definido na equação 1.5. Além disso, caso a função esteja definida num intervalo fechado, temos pelo Teorema do Valor Absoluto, a garantia que ela atingirá o valor dos seus pontos extremos, ou seja, seus valores de máximo e mínimo serão atingidos naquele intervalo. O que culmina nossa busca pelos pontos críticos.

Partindo desse pressuposto, é nítido que a derivada de uma função é também uma função. Portanto, fica evidente que uma função pode ser derivada mais de uma vez, sendo essas derivadas, denominadas de ‘primeira derivada’, ‘segunda derivada’

$$\frac{dy}{dx} = f'(x) = y' = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}$$

Figura 1.2: Derivada: taxa de variação dy/dx.

e por ai em diante. De modo que, a segunda derivada é a derivada da primeira derivada. Concluindo-se que dada a função $f(x)$, sua primeira derivada é $df/dx = p(x)$, e sua segunda derivada $dp/dx = s(x)$.

Agora vamos tratar da classificação dos pontos críticos.

É sabido que a primeira derivada representa a taxa de variação instantânea de um ponto na curva, e, que, a segunda derivada proporciona informações complementares, como por exemplo, se é um ponto de máximo, mínimo ou inflexão, de modo que, ela determina a concavidade da curva naquele ponto. Como pode ser visto na Figura 1.3. E, de modo construtivo, devemos ressaltar que é importante o estudo do gráfico das derivadas de uma função, pois é necessariamente através do uso da intrepretação do comportamento da primeira derivada e/ou da segunda derivada, que obtemos a classificação dos pontos críticos seguindo os critérios que são apontados na Figura 1.3.

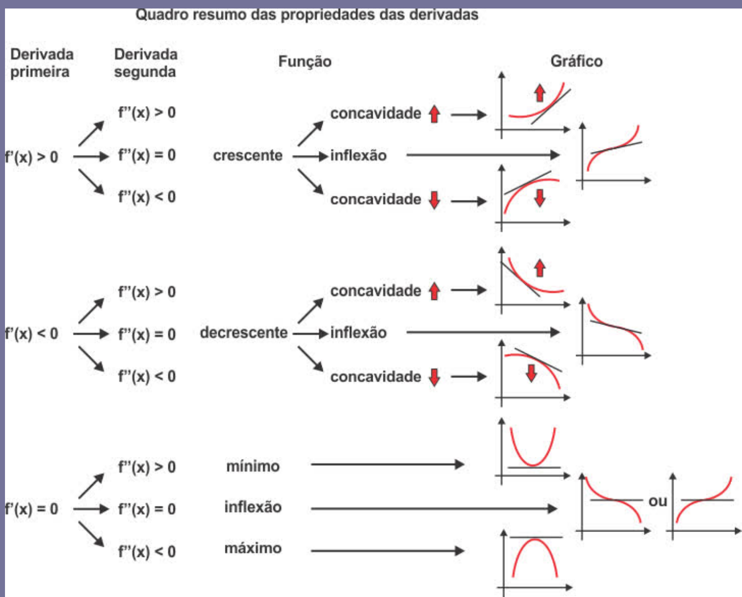

Figura 1.3: Propriedades e relação da primeira e segunda derivada.

1.3 Programando o Método

A própria definição da derivada já nos oferece uma visão simples de como ela pode ser implementada num programa de computador. Mas, com alguns nuances que devem ser levados em consideração, segue a tal implementação:

```
pub fn derive1x1_v1<F>(f: &F, x: f64) -> f64
where
    F: Fn(f64) -> f64,
{
    (f(x + h) - f(x)) / h
}
```

Essa implementação é ingênua no que diz respeito a precisão da operação. O uso, depende da aplicação, caso se queira prezar por velocidade de cálculo e muito pouco sobre precisão, então provavelmente, essa solução seja boa o suficiente. O problema com a precisão desta implementação é que não se é levado em consideração o aspecto infinitesimal da variação (Δx na Figura 1.2, ou a variável h na implementação em Rust), já que é por este aspecto que definimos a derivada. Não temos como ter uma variável com tal propriedade num programa de computador. Como consequência, $f(x+h)-f(x)$ já é calculada de forma falha, entregando uma variação diferente da que seria quando se tem uma variável infinitesimal. O que de fato é entregue, é a variação um pouco mais à frente do ponto esperado.

A partir daí podemos reformular, considerando esse deslocamento, redefinindo da seguinte forma:

$$f'(x) = \frac{f(x+h) - f(x-h)}{2h}. \tag{1.8}$$

Que é equivalente à:

$$f'(x) = \frac{f(x+h) - f(x-h)}{(x+h) - (x-h)}. \tag{1.9}$$

Usaremos essa segunda formulação por motivos de custo de multiplicação de de números reais se comparado a somas e subtrações, além da possível perda de precisão. Assim, conseguimos compensar o tal deslocamento numa nova função insignificantemente mais custosa, e mais precisa. Como vemos no código a seguir:

```
pub fn derive1x1_v2<F>(f: &F, x: f64) -> f64
where
    F: Fn(f64) -> f64,
{
    let x1 = x - h;
    let x2 = x + h;

    let y1 = f(x1);
    let y2 = f(x2);
    return (y2 - y1) / (x2 - x1);
}
```

1.4 Otimização de Funções à Várias Variáveis

Capítulo 2

Métodos Clássicos de Otimização

2.1 O Método de Newton

O Método de Newton nos diz que uma sequência $\{x_k\}$ converge para o mínimo de $f(x)$:

$$x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)} \tag{2.1}$$

O que este método faz de fato é encontrar zeros de uma função $g(x)$ a partir da mesma sequência $\{x_k\}$, afim de que x_k converga para uma entrada x^* de $g(x)$, tal que a mesma se anule:

$$x_{k+1} = x_k - \frac{g(x_k)}{g'(x_k)} \tag{2.2}$$

$$g(x^*) = 0 \tag{2.3}$$

Com isso temos um método que minimiza uma função que pode ser derivada duas vezes (caso não possa, aproximações de suas primeiras e segundas derivadas podem ser boas o suficiente) a partir de um valor de entrada. O método é simples, entrega muitas vezes ótimos locais próximos ao ponto inicial, mas tem seu destaque pode ser curto e facilmente computável.

Problemas de maximização podem ser vistos sob o seguinte olhar:

$$max(f(x)) = min(-1 * f(x)) \tag{2.4}$$

Que com isto podem ser otimizados pelo Método de Newton também.

O movimento de x_k dentro da sequência, é determinado pela relação das quantidades e propriedades que tanto a primeira quanto a segunda derivada oferecem. As quantidades determinam a velocidade do motivento, e os sinais indicam a direção do movimento. De certa forma podemos ver esse movimento da sequência $\{x_k\}$ como instantes do movimento de uma bola numa ladeira, que no começo de sua descida é acelerada, e, conforme chega ao plano no fim da ladeira, começa a reduzir sua velocidade, até supostamente chegar no ponto mais baixo.

2.2 Outros Métodos

2.3 Programando os Métodos

Método de Newton

A forma mais simples e mais util de implementar o metodo de Newton é na forma de busca das raizes, que uma vez implementada, só precisamos por como entrada a priemira e a segunda derivada da função que desejamos minimizar, já que o método não precisa saber qual a função de fato. A seguir temos a implementação na liguagem de programação Rust:

Listing 2.1: Implementação do Método de Newton para uma variável em Rust.

```
pub fn newton1x1<F>(funcao_derivada: &F, x: f64) -> (usize, f64)
where
    F: Fn(f64) -> f64,
{
    let mut entrada_atual = x;
    let maximo_iteracoes = 100;

    for iteracao_atual in 1..=maximo_iteracoes {
        let diferenca: f64 =
            funcao_derivada(entrada_atual.clone())
            /
            derive1x1(&funcao_derivada, &entrada_atual);

        println!("diferenca: {}", diferenca);
        entrada_atual -= diferenca;

        if diferenca.abs() < 0.0000001 {
            return (iteracao_atual, entrada_atual);
        }
    }

    return (maximo_iteracoes, entrada_atual);
}
```

Os parâmetros da função são:

- Uma função $f : \mathbb{R} \to \mathbb{R}$
- Uma entrada x sendo o chute inicial do otimo.

A função derive1x1 recebe como parametro uma função e um ponto, tendo como retorno a derivada da função entregue, no ponto especificado. Restingindo-se à funções do tipo $f : \mathbb{R} \to \mathbb{R}$, donde deve ser escrita pelo úsuario como bem entender.

2.4 O Método de Newton para Várias Variáveis

Capítulo 3

Os Métodos Modernos de Otimização

3.1 Breve Relato Histórico

3.2 Métodos de Um

3.2.1 O Método - Uma breve descrição

3.2.2 Exemplos Aplicações

3.2.3 Possíveis Aplicações

3.3 Métodos de Dois

3.3.1 O Método - Uma breve descrição

3.3.2 Exemplos Aplicações

3.3.3 Possíveis Aplicações

3.4 Um com o outro

Capítulo 4

Aplicações à Mecânica Celeste

4.1 Entendendo o Problema de N Corpos

4.2 A Otimização na Mecânica

4.3 Resultados Numéricos

Capítulo 5

Demais Resultados

5.1 Outros Resultados

5.1.1

Referências Bibliográficas

[1] *Garcia, Arnaldo. Elementos de Álgebra. Rio de Janeiro. Associação instituto Nacional de Matemática Pura e Aplicada, 2003.*

[2] *Lima, Elon Lajes Álgebra Linear. 7º Edição. Rio de Janeiro; IMPA, 2008.*

[3] *Halmos, Paul R. Espaços Vetoriais de Dimensão Finita. Tradução [de] Guilherme de la Penha. Rio de Janeiro: Campus, 1978.*