Definições e aplicações do Método do Gradiente

Douglas Silva, Gutembergue da Silva, Jeffet da Silva, Roger Conceição
Departamento de Engenharia Elétrica
Universidade Federal do Rio Grande do Norte
Rio Grande do Norte, Brasil

Resumo

Este trabalho tem como objetivo principal mostrar, de forma resumida, a teoria envolvida por trás do Método do Gradiente, que é utilizado para obtermos a resolução de sistemas lineares. Ademais, serão feitas algumas averiguações quanto as suas aplicações, através de uma exemplificação prática, desenvolvida por meio de um algoritmo, desenvolvido em C++.

1 Introdução

O Método do Gradiente consiste em um processo iterativo amplamente utilizado para se obter a solução numérica de sistema lineares, no qual transforma o cálculo do vetor solução em uma obtenção de um ponto de mínimo local de uma função $f(x)$, em que x é um vetor de n dimensões. Tal método é aplicado principalmente em problemas de otimização e no âmbito de *machine learning* e *deep learning* como forma de ajuste dos pesos utilizados em redes neurais. Sua utilização traduz-se em representar um sistema linear qualquer $Ax = b$ na sua forma quadrática, conforme a equação 1 mostrada abaixo e então, aplicar as definições do método no processo iterativo.

$$f(x) = \frac{1}{2}x^T Ax - b^T x + c \tag{1}$$

Neste trabalho serão apresentados conceitos acerca do método e sua aplicação para a resolução de um sistema linear determinado.

2 Método do Gradiente

O Método do Gradiente utiliza o gradiente de uma função para obter um ponto de mínimo ou de máximo local, iterativamente, tendo em vista que o gradiente da função em determinado ponto fornece a direção de maior crescimento ou decrescimento da função a partir deste ponto. Em outras palavras, o gradiente possibilita conhecer a direção na qual ocorre a máxima taxa de variação instantânea da função. A equação a seguir apresenta a definição do gradiente de uma função $f(x, y)$.

$$\nabla f(x, y) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial x}\right) \tag{2}$$

A característica do gradiente supracitada permite que este seja útil para determinar pontos de máximo ou de mínimo locais, já que sua direção é indicada pelo par de coordenadas obtido.

Diante do exposto, o Método do Gradiente obtém pontos críticos a partir da seguinte expressão.

$$x_i^{k+1} = x_i^k \pm \alpha^k \cdot \nabla f(x_i^k) \tag{3}$$

Assim, para um gradiente positivo em 3, teremos $f(x_i^{k+1}) \geq f(x_i^k)$, obtendo um ponto de máximo após a realização das iterações. Por outro lado, para um gradiente negativo em 3, teremos $f(x_i^{k+1}) \leq f(x_i^k)$, obtendo um ponto de mínimo após as iterações.

Para a correta compreensão e interpretação da equação apresentada em 3, vejamos que x_i^k determina o vetor das entradas inciais (variáveis independentes), para uma função de várias variáveis, sendo este o vetor dos chutes iniciais, a partir do qual se inicia o método iterativo. O fator α^k, por sua vez, determina o tamanho do passo dado, na iteração k. Por fim, temos que o operador nabla, aplicado à função f para os valores de iteração k, define a direção para a qual obteremos o vetor posterior x_i^{k+1}. Assim, sendo repetido este processo, obteremos, sucessivamente, as direções de máximo ou mínimo declive da função para novos pontos sobre sua curva. Descritivamente, podemos dizer que, em 3, o vetor x da iteração posterior, x_i^{k+1}, será igual ao vetor x anterior, x_i^k, somado do gradiente de f aplicado em x_i^k, ajustado pelo passo α^k.

3 Aplicação em Problema Numérico

Para a comprovação teórica apresentada, neste trabalho será implementado o método do gradiente para obtenção da solução numérica de um sistema linear apresentando o resultado e as configurações para a obtenção do mesmo. O sistema proposto consiste em uma matriz simples 2x2 que multiplicado por um vetor x resultará em um vetor 1x2. A configuração explorada está descrita abaixo.

$$\begin{bmatrix} 4 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 5 \\ 4 \end{bmatrix}$$

Porém, alguns parâmetros são necessários durante a implementação do método iterativo, eles são:

$$r^k = b - Ax^k \tag{4}$$

$$\alpha = \frac{(r^k)^T r^k}{(r^k)^T A r^k} \tag{5}$$

$$x^{k+1} = x^k + \alpha r^k \tag{6}$$

Em que r^k é o resíduo a cada iteração e x^k é o chute inicial para a primeira iteração. A escolha do sistema relativamente simples se deu para que ao resolver um sistema com apenas duas variáveis seja possível visualizar de forma mais efetiva as plotagens dos pontos para a verificação da ortogonalidade gerada pelo gradiente.

Ao realizar a implementação do algoritmo em um código C++, pôde-se realizar as iterações para obtenção da solução numérica do sistema. O desenvolvimento se dá basicamente como um *loop* no qual a cada volta serão calculados os parâmetros informados acima em que cada novo vetor solução deve ser perpendicular ao anterior, seguindo a definição do gradiente. A discussão acerca dos resultados obtidos será realizada no tópico a seguir.

4 Resultados

Com a implementação do programa, foi possível obter a solução desse sistema, já conhecida $(1, 1)$, para verificar como se deram as aproximações a cada iteração realizada. A tabela apresentada abaixo mostra a solução para cada iteração usando um chute inicial distante das raízes com $x_1 = 15$ e $x_2 = 15$ para acompanhar a resposta do método que com apenas 3 iterações já atingiu uma precisão de 10^{-8}.

n	x_1	x_2
0	15	15
1	-0.27	2.79
2	1.09	1.09
3	0.99	1.01

Tabela 1: Solução numérica.

Além da tabela apresentada, o gráfico abaixo apresenta o caminho seguido pelo método para encontrar a solução desejada, mantendo sempre o novo vetor x^{k+1} perpendicular ao vetor x^k anterior, verificando o desenvolvimento conceitual apresentado anteriormente.

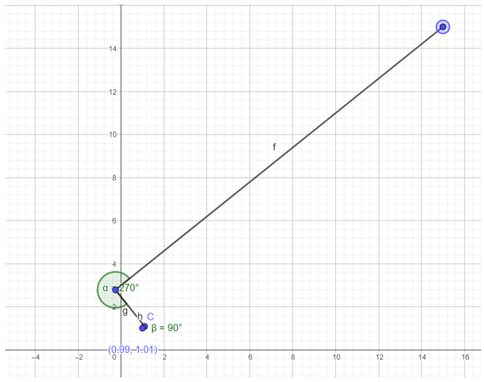

Figura 1: Configuração das iterações.

5 Conclusões

Conclui-se, por fim, que o presente trabalho, que enfatizou os conceitos e aplicações do Método do Gradiente, obteve êxito em seus resultados finais, visto que o que foi proposto para o seu desenvolvimento, a resolução de um sistema linear qualquer, foi cumprido. Além disso, é importante destacar, também, que foi possível realizar a verificação dos conhecimentos teóricos através de uma exemplificação prática, desenvolvida por meio de um algoritmo, implementado na linguagem de programação C++.

6 Referências

[1] GUERRA, Rodrigo, Método da Descida do Gradiente. Disponível em: Link 1

[2] GRADIENTE Descendente: Laboratório de Estatística e Geoinformação - LEG/UFPR. Paraná: UFPR - Universidade Federal do Paraná, [202-]. Disponível em: Link 2. Acesso em: 8 dez. 2022.

[3] MICROTONAL. Otimização Método de Busca pelo Gradiente – Engenharia Elétrica. Disponível em: Link 3

[4] VELASQUEZ, Ester. O que é o vetor gradiente? Disponível em: Link 4

[5] 11. Método do Gradiente. Produção: Thadeu Penna. [S. l.: s. n.], 2020. Disponível em: Link 5. Acesso em: 8 dez. 2022.