ISCTE INSTITUTO UNIVERSITÁRIO DE LISBOA

11 MARÇO DE 2024

PROPOSTA 1.

INTELIGÊNCIA COMPUTACIONAL E OTIMIZAÇÃO

Elaborado por:
António Nanita, 122240
Miguel Valadares, 98345
Gonçalo Botelho, 98893
Rodrigo Alves, 121700

Índice

Introdução

A otimização de funções não lineares é um campo fundamental na matemática aplicada e na ciência da computação, desempenhando um papel crucial na resolução de uma variedade de problemas complexos. O problema em questão reside na minimização de funções não lineares, um desafio que permeia diversas áreas do conhecimento, desde a modelagem de fenómenos físicos até à otimização de parâmetros em aplicações de *machine learning*.

As funções não lineares podem representar uma ampla diversidade de fenómenos e processos da vida real, muitas vezes caracterizados por comportamentos não triviais e múltiplos mínimos locais. A busca pelo mínimo global neste tipo de funções é frequentemente complicada pela presença desses mínimos locais, tornando o problema de otimização não linear altamente desafiador.

A complexidade destas funções e a presença de múltiplos mínimos locais exigem métodos de otimização eficazes que possam encontrar soluções de forma rápida e precisa. Nesse contexto, os métodos de busca em linha emergem como ferramentas poderosas, explorando eficientemente o espaço de busca para encontrar os mínimos desejados.

Aplicações Reais

Uma aplicação da otimização de funções não lineares, e que já trabalhamos, consiste no desenvolvimento e treino de redes neuronais artificiais, mais especificamente no *backpropagation*. Em outras palavras, permite que a rede neuronal "aprenda" consoante um erro obtido, e o objetivo consiste em encontrar o menor erro possível.

Contextualização Teórica

Métodos de Busca em Linha

Os métodos de busca em linha são algoritmos que visam responder ao problema anteriormente identificado. Através de várias iterações, procura-se atingir soluções aproximadas x^* em problemas de minimização ($min_x\, f(x) = ?$).

A partir de um ponto inicial /semente, pretende-se, em cada passo, avançar numa direção de descida, de forma a garantir que o valor seguinte é sempre inferior ao anterior, com o objetivo de atingir o mínimo. Neste trabalho abordar-se-ão alguns métodos de busca diferentes, como o Método de Descida Máxima e o Método de *Newton-Raphson*. Para além disso, será analisada a utilização das condições de *Wolfe*.

Método de Descida Máxima

O Método de Descida Máxima, também conhecido como método do gradiente descendente, é uma abordagem iterativa usada para encontrar o mínimo de uma função. Este método baseia-se na minimização iterativa de uma função, ajustando os parâmetros na direção oposta ao gradiente local.

O gradiente indica a direção de maior inclinação da função, e ao mover-se na direção oposta, procura-se atingir o ponto de mínimo da função. A atualização dos parâmetros ocorre multiplicando o gradiente pela taxa de aprendizagem, controlando assim o tamanho dos passos em cada iteração. Este método revela-se particularmente eficaz em problemas de otimização convexa, proporcionando uma abordagem sistemática para a convergência em direção ao mínimo global.

Método de *Newton-Raphson*

O Método de *Newton-Raphson* é uma técnica iterativa utilizada para encontrar raízes de funções ou extremos (mínimos/máximos). Esta abordagem fundamenta-se na aproximação local da função por uma parábola, utilizando informações da função e da sua segunda derivada (curvatura). Ao determinar o mínimo ou máximo desta parábola, obtém-se uma estimativa mais precisa da solução.

A atualização dos parâmetros leva em consideração tanto o gradiente como a matriz Hessiana da função. Este método permite uma convergência mais rápida em comparação com o Método de Descida Máxima, uma vez que incorpora informações adicionais sobre a curvatura da função. Contudo, o cálculo da matriz Hessiana pode ser computacionalmente exigente, e em alguns casos, a singularidade ou complexidade do problema pode afetar a eficácia do método.

Condições de *Wolfe*

As condições de *Wolfe* são um conjunto de regras utilizadas na determinação do tamanho do passo ideal durante a otimização, sendo compostas por duas partes: a condição de diminuição suficiente e a condição de curvatura.

A condição de diminuição suficiente, também conhecida como condição de Armijo, garante que o passo dado resulta numa diminuição adequada do valor da função objetivo. A condição de curvatura garante que o tamanho do passo não seja excessivamente grande. Se este for muito grande, pode-se ultrapassar o mínimo pretendido. Estas condições ajudam a garantir a convergência dos algoritmos, sendo amplamente utilizadas.

Lista de Funções

Neste trabalho, iremos aprofundar a exploração dos métodos de busca em linha mencionados, aplicando-os a diversas funções. Além disso, realizaremos uma análise comparativa de desempenho entre esses métodos, considerando tanto os casos em que são aplicadas as condições de *Wolfe* quanto aqueles em que essas condições são desconsideradas.

Ao aplicar os métodos de busca em linha a diferentes funções, procuramos compreender como cada algoritmo se comporta perante diferentes tipologias de função. Desta forma, permitirá avaliar a robustez e a eficácia dos métodos em contextos variados, fornecendo *insights* valiosos sobre as suas características de desempenho em diferentes cenários.

A inclusão da análise com e sem as condições de *Wolfe* amplia a abordagem da avaliação, uma vez que as condições de *Wolfe* são critérios que garantem a convergência dos métodos de otimização.

Função de *Rosenbrock*

A função a ser estudada é uma variação da função de *Rosenbrock, que* é frequentemente utilizada como um desafio significativo para avaliar algoritmos de otimização devido às suas características distintas. A tipologia da função assemelha-se a um “vale alongado e estreito”, o que cria um ambiente desafiador para os métodos de otimização.

O mínimo global desta função ocorre em $x = 1$ e $y = 1$, onde o valor da função atinge o mínimo de zero. Contudo, para descobrir este mínimo é uma tarefa complexa devido à forma peculiar da função, que exige métodos de otimização robustos e eficientes.

A sensibilidade às condições iniciais é uma característica proeminente desta função, pois variações mínimas nas condições iniciais podem resultar em trajetórias de otimização distintas, evidenciando a influência crucial das mesmas no processo de busca pela solução ótima.

Além disso, a função de *Rosenbrock* modificada destaca a importância de estabelecer critérios de paragem adequados ao implementar algoritmos de otimização. Dada a sua complexidade, é essencial definir critérios que evitem convergência prematura ou que garantam uma exploração completa da superfície de otimização.

$$r(x, y) = 100(y - x^2)^2 + (1 - x)^2$$

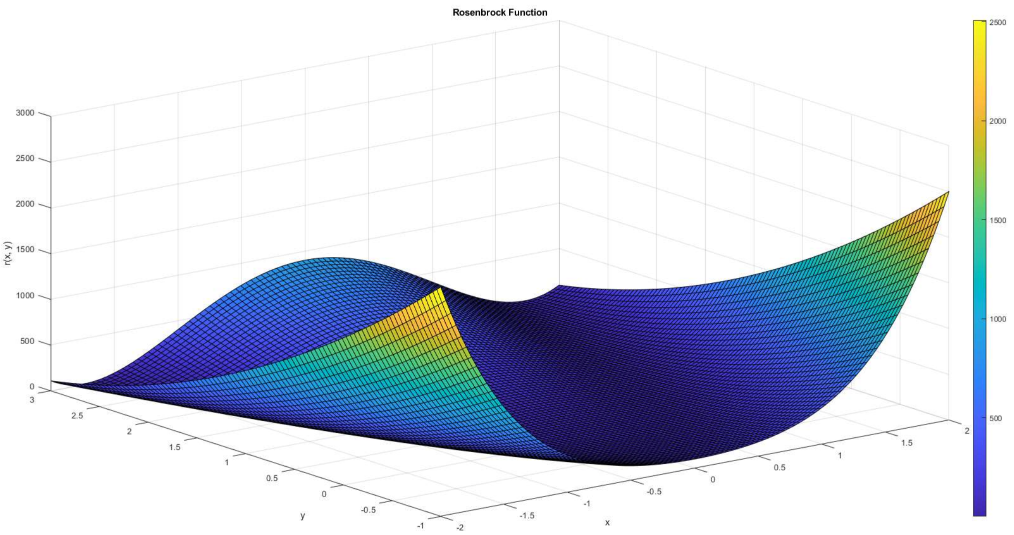

Figura 1 - Função de Rosenbrock

Função Matemática Complexa f(x,y)

A função apresenta uma estrutura matemática complexa e desafiante, tornando-se um campo fascinante para o estudo de algoritmos de otimização. Composta por termos polinomiais, funções trigonométricas e a inclusão de parâmetros fixos como 100, 61, 25 e 16, esta função revela uma superfície de otimização com múltiplos termos não-lineares e não-convexa.

A introdução de termos trigonométricos, como cos(5y), adiciona uma dimensão cíclica à problemática, exigindo métodos robustos para lidar com oscilações. Os coeficientes diferenciados associados a cada termo refletem-se na convergência dos algoritmos, evidenciando a importância de ajustes precisos.

$$f(x, y) = (x^2 - 100)(x^2 - 61) + (y^2 - 25)(y^2 - 16) - 15\cos\left(x - \frac{7}{5}\right)\cos\left(y - \frac{9}{z}\right)$$

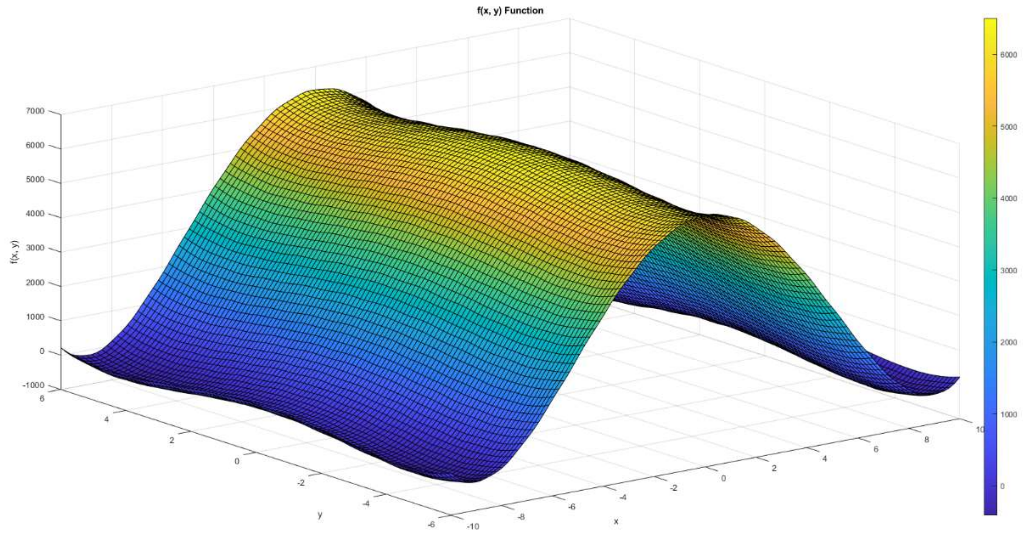

Figura 2 - Função Matemática Complexa f(x,y)

Função Matemática Complexa g(x,y)

A função apresenta uma configuração matemática desafiante que incorpora termos polinomiais, constituindo um estímulo interessante para análise e otimização. Composta por fatores quadráticos *(x + 3) (x - 3) (x - 5)* e a presença da variável linear *y*, esta função revela uma superfície de otimização com características não-lineares.

Os fatores quadráticos indicam raízes nos pontos *x = -3, x = 3 e x = 5*, sugerindo possíveis pontos críticos. A inclusão da variável *y* nos termos lineares adiciona uma dimensão ao problema, aumentando a complexidade da otimização.

A análise atenta da tipologia desta função, tendo em conta as raízes dos fatores quadráticos e a influência da variável *y*, é crucial para compreender a busca pelo mínimo global. A implementação de métodos de otimização eficazes torna-se essencial para explorar de forma eficiente esta complexa superfície de otimização.

$$g(x, y) = x(x+3)(x-3)(x-5) + y(y+3)(y-3)(y-5)$$

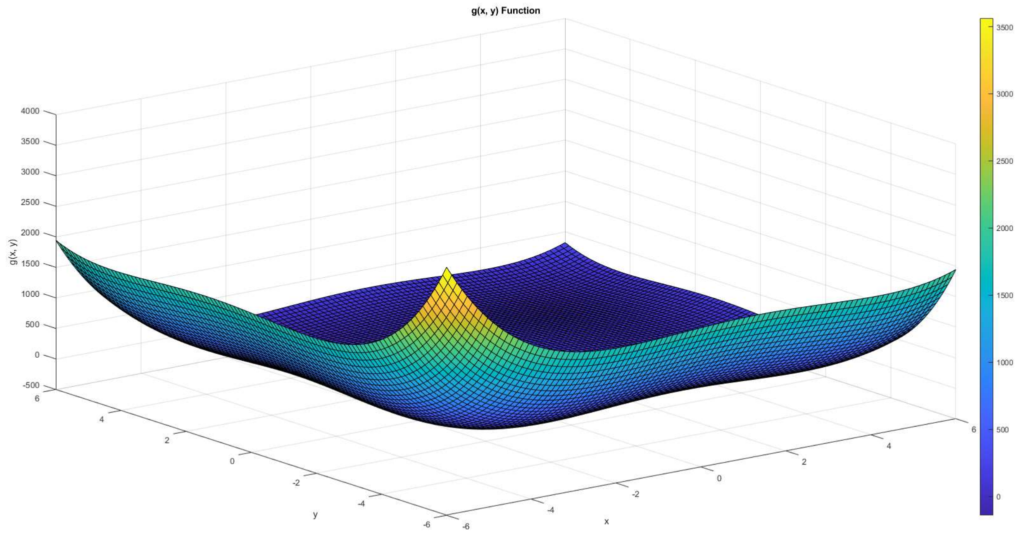

Figura 3 - Função Matemática Complexa g(x,y)

Metodologia

Tendo em conta os objetivos do trabalho, podemos fazer uma exploração de cada método, consoante cada função de teste, utilizados diferentes abordagens.

Diferentes Tolerâncias de Erro:

- Para cada função de teste, ajustaremos a tolerância de erro do método de busca em linha para avaliar como isso afeta a convergência e a precisão da solução.
- Exploraremos uma gama de tolerâncias de erro para identificar o impacto na quantidade de iterações necessárias para alcançar a convergência.

Diferentes Números de Iterações:

- Realizaremos experimentos com diferentes números de iterações para observar como a convergência dos métodos de busca em linha é influenciada pela quantidade de iterações permitidas.
- Analisaremos a relação entre o número de iterações e a precisão da solução, bem como o tempo computacional necessário para atingir a convergência.

Diferentes Pontos Iniciais:

- Exploraremos o efeito dos pontos iniciais na convergência e na qualidade da solução obtida pelos métodos de busca em linha.
- Investigaremos como diferentes pontos iniciais podem influenciar a “trajetória da busca” e se podem levar a minimizantes locais ou globais distintos.

Após a realização das experiências, apresentaremos os resultados obtidos para cada função de teste, incluindo os mínimos e minimizantes locais ou globais encontrados, os vetores gradientes correspondentes e os parâmetros utilizados em cada caso.

Por fim, iremos analisar e interpretar os resultados obtidos, destacando padrões observados, tendências identificadas e insights relevantes sobre o desempenho e comportamento dos métodos de busca em linha em diferentes cenários de otimização.

Bibliografia

Bertolazzi, E. (2011). Unconstrained minimization. Lectures for PHD course on Numerical optimization

Humpherys, J., & Jarvis, T. J. (Eds.). (2011). Labs for Foundations of Applied Mathematics - Algorithm Design and Optimization (Vol. 2) Lectures for PHD course on Numerical optimization

Parkinson, A. R., Balling, R. J., & Hedengren, J. D. (2013). Optimization Methods for Engineering Design - Applications and Theory. Brigham Young University.

Bazzet, D. T. (14 de feveiro de 2023). Intro to Gradient Descent || Optimizing High-Dimensional Equations. Obtido de Youtube: https://www.youtube.com/watch?v=fXQXE96r4AY

Bierlaire, M. (04 de maio de 2019). 11 Descent methods - Playlist. Obtido de Youtube: https://www.youtube.com/playlist?list=PL10NOnsbP5Q7wNrYItE2GhKq05cVov97e

Visually, E. (08 de outubro de 2021). Gradient Descent in 3 minutes. Obtido de Youtube: https://www.youtube.com/watch?v=qg4PchTECck