Relatório da disciplina de Redes Neurais

Vitor Amorim*, Wesley Pereira Pimentel†
Universidade Federal do Espírito Santo
Vitória, Espírito Santo, Brasil
vitor.amorim@edu.ufes.br, wesley.pereira@edu.ufes.br

***Abstract*—Relatório apresentado na disciplina de Redes Neurais, ministrada pelo professor dr. Thomas Walter Rauber, como forma parcial de avaliação. O presente trabalho tem como finalidade utilizar o método de descida de gradientes para resolução de 4 problemas de minimização.**

***Index Terms*—Inteligência Artificial, Otimização, Redes Neurais**

I. INTRODUÇÃO

O Algoritmo Genético (AG), pode ser definido, por [1], como um método computacional de busca baseado em mecanismos de evolução natural e da genética. Em um AG, uma população de possíveis soluções para um dado problema evolui de acordo com operadores probabilísticos concebidos a partir de conceitos biológicos. De modo que há uma tendência dos indivíduos representarem soluções cada vez melhores à medida que o processo avança.

II. MÉTODO

Foi utilizado o algoritmo genético para resolver os três problemas de otimização descritos. E adotado como parâmetro para o problema do caxeiro viajante e da regressão, o tamanho da população igual a 200 e o número de gerações igual a 1000, e para o do aprendizado de máquina, 100, indivíduos e 250 gerações. O parâmetro de mutação estabelecido foi de 10%. O programa foi implementado em linguagem de programação *Python*, com o compilador *Python* 3.8, num ambiente cuja máquina possuia um processador da Intel Core i5-8250U CPU @ 1.60GHz, Memória RAM de 8GB, no sistema operacional Microsoft Windows.

Executou-se o algoritmo para resolver os problemas 5 vezes com os mesmos parâmetros mencioandos e obteve como retorno do código, as representações gráficas do problema abordado, bem como o menor(Min) e maior(Max) valor da função objetivo, que pode ser chamada de Fitness, no contexto de algoritmos evolutivos. Além desse dado, obteve-se também o valor médio da função objetivo, o desvio padrão e o tempo para execução do código conforme o problema.

III. RESULTADOS

A. Descida de gradientes: unidimensional

A análise realizada levou em consideração a Equação 1. E foi determinado que o valor inicial para a variável, x, seria $x_0 = 1$, e a taxa de aprendizagem, alpha, $\alpha = 0.1$. Fez-se o cálculo da derivada para calcular a descida de gradientes utilizando-se o método de diferenças finitas(MDF), adotando-se o passo, $h = 0.01$. Posteriormente, foi calculada a derivada analiticamente.

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

A Equação 1 também pode ser escrita como:

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

Derivada	MDF	Analítica	$\bar{x}$	$\lvert\sigma\rvert$
x_0	1.0000	1.0000	1.0000	0.0000
x_1	0.9627	0.9632	0.9630	0.0005

Table I: Comparação entre o próximo valor de x, tuilizando-se aproximação da derivada e calculando-a de forma analítica.

Nesta próxima etapa, foi utilizado o método de descida de gradientes, cuja derivadas foram obtidas analiticamente. O valor inicial definido foi de $x_0 = 3$, e taxa de aprendizagem, alpha, $\alpha = 0.1$. As condições de parada implementadas foram o tamanho da derivada, $gmin = 0.1$ e o número máximo de iterações atingidos, $k_{max} = 100$.

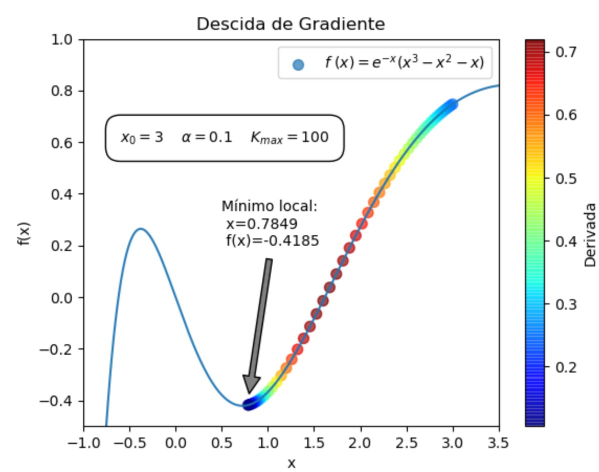

Figure 1: Atualização do valor de x, atualizando-o com base no gradiente.

B. Descida de gradientes: bidimensional

O objetivo abordado neste problema é utlizar o método de diferenças finitas, bem como o método analítico, para o cálculo da derivada de uma função de duas variáveis. A função analisada é representada pela Equação 3.

$$f(x_1, x_2) = (4 - 2.1x_1^2 + \frac{x_1^3}{3})x_1^3 + x_1x_2 + (-4 + 4x_2^2)x_2^2 \quad (3)$$

$$f(x_1, x_2) = 4x_1^3 - 2.1x_1^5 + \frac{x_1^6}{3} + x_1x_2 - 4x_2^2 + 4x_2^4 \quad (4)$$

O gradiente das funções calculados analiticamente, por meio do algoritmo estão representados pela Equação 5.

$$\nabla f(x_1, x_2) = \begin{bmatrix} \dfrac{\partial f(x_1, x_2)}{\partial x_1} \\ \\ \dfrac{\partial f(x_1, x_2)}{\partial x_2} \end{bmatrix} \quad (5)$$

$$\nabla f(x_1, x_2) = \begin{bmatrix} 2x_1^5 - 10.5x_1^4 + 12x_1^2 + x_2 \\ \\ 16x_2^3 - 8x_2 + x_1 \end{bmatrix}$$

Executou-se o algoritmo elaborado para estimar mínimos da Equação 4, e foi implementado o método das diferenças finitas para duas vairáveis, bem como o calculo do gradiente de forma analítica. A Figura 2 representa a curva de nível da Equação 3.

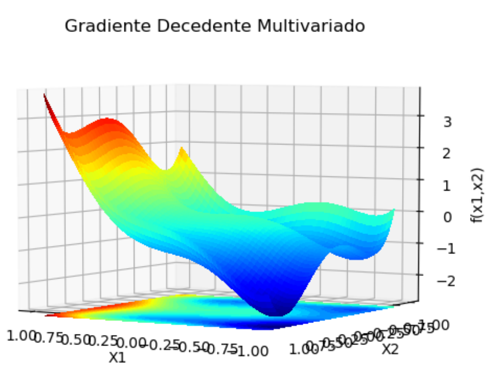

Figure 2: Superfície representado no domínio $-1 <= x <= 1$.

C. Aprendizagem de Máquinas

Este problema consistiu em realizar a busca de hiper-parâmetros e *feature selection*. E o algoritmo para realizar a classificação foi o *k-nearest neighbors* (K-NN).

Problema	**Min.**	**Max.**	$\bar{x}$	σ	**Tempo(s)**
TSP	118.462	147.385	133.097	9.829	33.156
Regressão	1.814	2.175	1.907	0.135	253.481
AM	-0.816	-0.810	-0.814	0.002	380.070

Table II: Resultados Obtidos em 5 rodadads de teste.

IV. Discussion

Para o caso o Problema do Caxeiro Viajante, o algoritmo genético neste contexto, para os parâmetros adotados, não apresentou um resultado que se assemelhe ao ótimo. Fato que fica evidente ao observar a Figura 1 e 2, ao perceber que algumas rotas se cruzam. O valor da função objetivo converge muito rapidamente, ficando o valor, restrito a possíveis ótimos locais da função. É muito provável que ao rodar o algoritmo por um número maior de gerações, obtenha-se melhores resultados.

No problema de regressão, é possível perceber que foram obtidos resultados satisfatórios, com um erro quadrático que apresenta indícios de estar próximo do ótimo, tal como é observado nas Figuras 5, 6 e 7. Em todas as simulações, o valor da função objetivo converge.

O problema da busca de hiperparâmetros também apresenta uma boa convergência, para valores bem próximos. Indicando que a função objetivo está numa região de ótimo ou que está próxima a uma.

References

[1] L. H. M. Costa, M. A. H. d. Castro, and H. Ramos, "Utilização de um algoritmo genético híbrido para operação ótima de sistemas de abastecimento de água," *Engenharia Sanitaria e Ambiental*, vol. 15, pp. 187–196, 2010.