Labwork 1: Gradient Descent

April 29, 2025

1 Introduction

Gradient descent: definition
Gradient descent is an optimization algorithm used to minimize a function by iteratively moving in the direction of the steepest descent, as defined by the negative of the gradient.

2 Implementation

The implementation of gradient descent in this lab minimizes the function $f(x)$. The algorithm starts with an initial value of x, a learning rate r, and a convergence threshold ϵ. It iteratively updates x using the formula $x = x - r \cdot f'(x)$, where $f'(x)$ is the gradient of $f(x)$. The process continues until $f(x)$ becomes less than or equal to ϵ, at which point the algorithm returns the final values of x and $f(x)$.

3 Evaluation

Test with $y = x^4$
Output of my program:

Enter initial x: 2
Enter learning rate r: 0.05
Enter convergence threshold epsilon: 0.005
Final x: 0.3999999999999999, Final f(x): 0.025599999999999977

The x and f(x) steps: Current x: 0.3999999999999999, f(x): 0.025599999999999977
Current x: 0.38719999999999993, f(x): 0.022477157800345586
Current x: 0.3755898978303999, f(x): 0.019900116114290756
Current x: 0.3649931718645291, f(x): 0.017747572531879738
Current x: 0.35526829266132753, f(x): 0.015930367528179314
Current x: 0.3463002153759361, f(x): 0.014381727201132847
Current x: 0.3379942851093202, f(x): 0.013050808848442437

Current x: 0.3302717824378793, f(x): 0.011898326472229414
Current x: 0.3230666095439659, f(x): 0.010893521523512159
Current x: 0.3163227856997814, f(x): 0.010012024563155282
Current x: 0.30999252748521344, f(x): 0.009234319577443924
Current x: 0.3040347583400302, f(x): 0.008544623793691794
Current x: 0.2984139379836053, f(x): 0.007930058940763027
Current x: 0.2930991333237783, f(x): 0.00738003016633253
Current x: 0.28806327389789593, f(x): 0.006885750438280605
Current x: 0.2832825498919133, f(x): 0.006439902494252792
Current x: 0.27873592144853765, f(x): 0.006036313180470917
Current x: 0.27440471566128116, f(x): 0.005669780969569371
Current x: 0.27027229326037167, f(x): 0.005335880645204224
Current x: 0.2663237711378589, f(x): 0.005030830985947597
Current x: 0.2625457899501838, f(x): 0.0047513851304683925
Final x: 0.2625457899501838, Final f(x): 0.0047513851304683925

How ever, for the learning rate of 0.2, the program gives:
OverflowError: (34, 'Numerical result out of range')

The reason for this is that the learning rate is too high, causing the algorithm to overshoot the minimum and produce an overflow error.

4 Conclusion

In this labwork, I implemented the gradient descent algorithm to minimize the function $f(x) = x^4$. The result is that the algorithm converged to a minimum value of $f(x)$ at $x \approx 0.4$ with a final function value of approximately 0.0256.

An interesting finding is that the choice of learning rate significantly affects the convergence speed and stability of the algorithm. A smaller learning rate leads to slower convergence, while a larger one may cause divergence.