**Convex**

**A graph of a graph of a person

AI-generated content may be incorrect.**

A graph of a function

AI-generated content may be incorrect.

On this convex function, gradient descent (α = 0.1) converged in 34 iterations. Adagrad (α = 5) oscillated at first but eventually converged within 9 iterations. Adam (α = 0.1) showed the poorest performance, requiring 106 iterations to converge. In contrast, Newton’s method reached the minimum in just a single iteration. This result is expected because the function is quadratic (a second-degree polynomial), and the formulation of Newton’s method guarantees convergence to the minimum of quadratic functions in one step.

**Rosenbrock**

**A graph of a positive and negative reaction

AI-generated content may be incorrect.**

A graph of a function

AI-generated content may be incorrect.

On this function, both gradient descent and Adagrad failed to converge (exceeding 2000 iterations). Adam (α = 0.1) successfully converged to (1, 1) in 543 iterations. Once again, Newton’s method outperformed all other optimization methods, reaching the point (1, 1) in just two iterations.