Programming Assignment 2

EE-5121 Convex Optimization

Keshaw Choudhary *[EE21B069]*

26 May, 2025

Objective

Implement and compare the performance of **Mirror Descent (MD)** and **Accelerated Mirror Descent (AMD)** for convex problems including:

- Logistic Regression
- Quadratic Function
- Log-Sum-Exp Function

Problem 1: Mirror Descent on Logistic Regression

1.1 Derivation with Entropy Mirror Map

Mirror Descent update:

$$\nabla\phi(\theta^{t+1}) = \nabla\phi(\theta^t) - \eta_t \nabla L(\theta^t)$$

For entropy: $\phi(\theta) = \sum_j \theta_j \log \theta_j$, we have:

$$\nabla\phi(\theta) = \log(\theta) + 1, \quad \theta^{t+1} = \exp\left(\log(\theta^t) + 1 - \eta\nabla L(\theta^t)\right)/Z$$

1.2 Sigmoid Function Plot

The logistic function is:

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

It maps real numbers to the interval (0, 1). See Figure 1.

1.3 Logistic Regression with GD

Gradient descent updates:

$$\theta^{t+1} = \theta^t - \eta\nabla L(\theta^t)$$

1.4 Mirror Descent Convergence

MD was applied using entropy mirror map. Figure 2 shows loss per iteration.

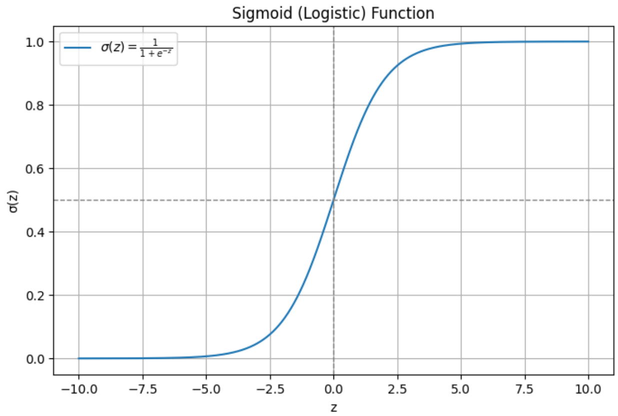

Figure 1: Sigmoid function

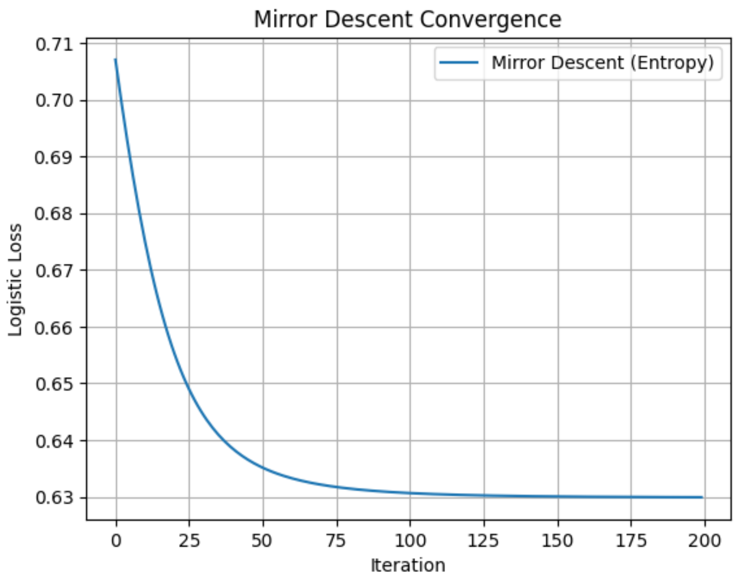

Figure 2: Mirror Descent convergence on logistic loss

1.5 MD vs GD Comparison

Gradient Descent and Mirror Descent were compared. AMD generally showed faster convergence. See Figure 3.

Problem 2: Accelerated Mirror Descent

2.1 Quadratic Objective

$$f(x) = \frac{1}{2}x^T Q x + b^T x$$

AMD and MD applied. AMD showed faster convergence. See Figure 4.

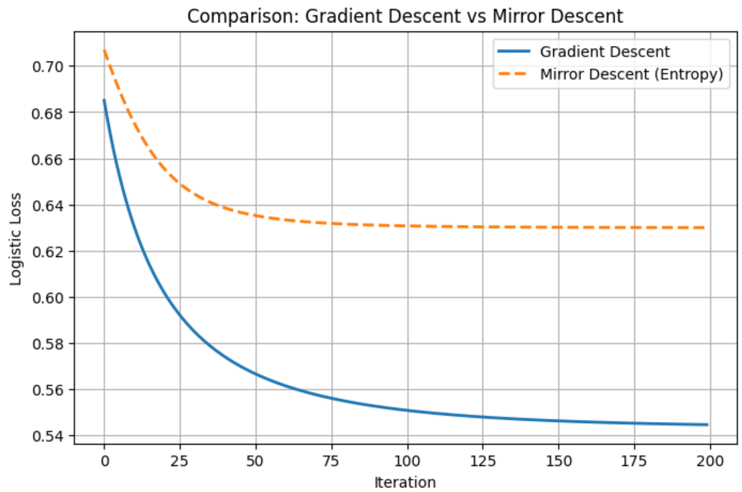

Figure 3: Loss comparison: GD vs MD

2.2 Log-Sum-Exp Objective

$$f(x) = \log \left(\sum_{i=1}^{d} e^{a_i^T x + b_i} \right)$$

Gradient:

$$\nabla f(x) = A^T p, \quad p_i = \frac{e^{a_i^T x + b_i}}{\sum_j e^{a_j^T x + b_j}}$$

2.3 Comparison

- MD converges slowly but stably.
- AMD improves convergence speed.
- On both objectives, AMD outperforms MD.

Conclusion

Accelerated mirror descent significantly improves optimization performance over standard MD, especially for smooth convex functions like logistic and log-sum-exp losses.

Note: All plots were generated using Python and matplotlib in Jupyter.

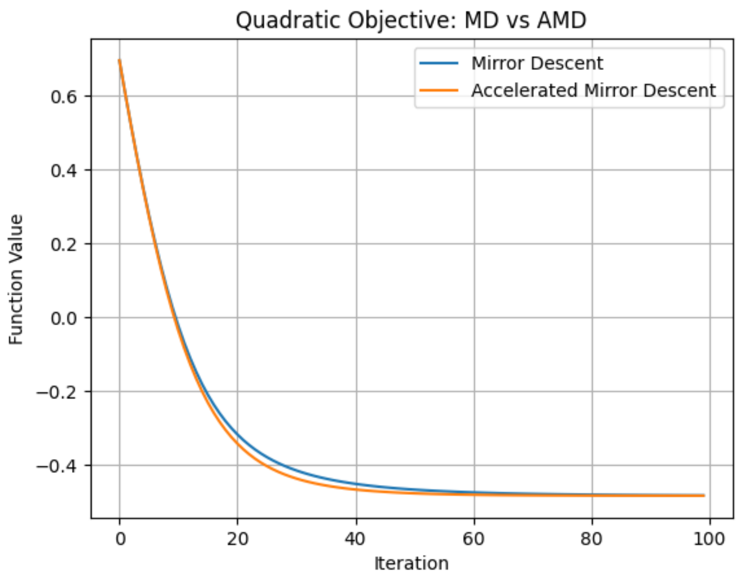

Figure 4: Quadratic objective: MD vs AMD

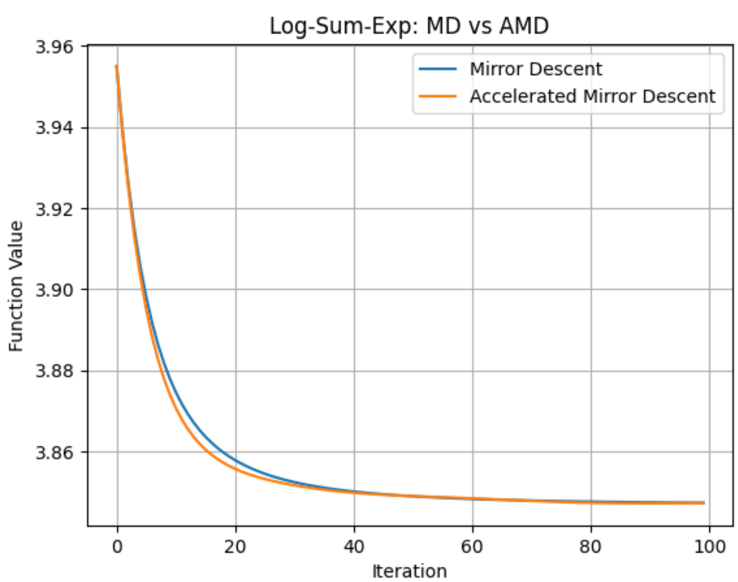

Figure 5: Log-Sum-Exp: MD vs AMD