

# Machine Learning For Robotics

## Assignment 2

Maha Qaiser 22i-2348

→ **GitHub Link:**

[https://github.com/mahaqj/Machine-Learning-Assignment-2/blob/main/ML\\_Assignment\\_2.ipynb](https://github.com/mahaqj/Machine-Learning-Assignment-2/blob/main/ML_Assignment_2.ipynb)

→ **Gradient Descent**

1. Stochastic:

- Updates weights after every individual data point.
- Fast updates, but very noisy because each data point affects the direction.
- Can escape local minima due to randomness but may take longer to converge.

2. Batch:

- Updates weights after processing the entire dataset.
- More stable updates and less noise, but computationally expensive.
- Not practical for large datasets.

3. Mini-Batch:

- Updates weights after a small batch of data points have been processed (mix of SGD and BGD).
- Reduces noise while still being computationally efficient.
- Balance of speed and stability.

4. Stochastic with Momentum:

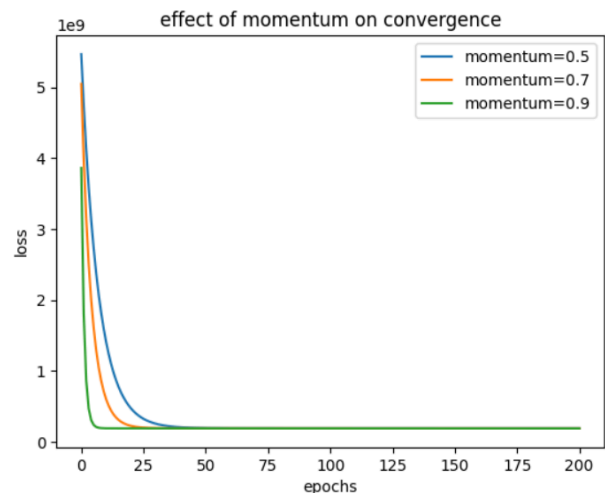
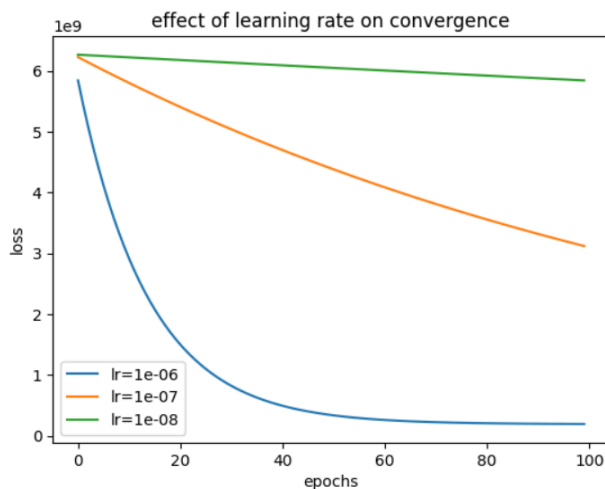
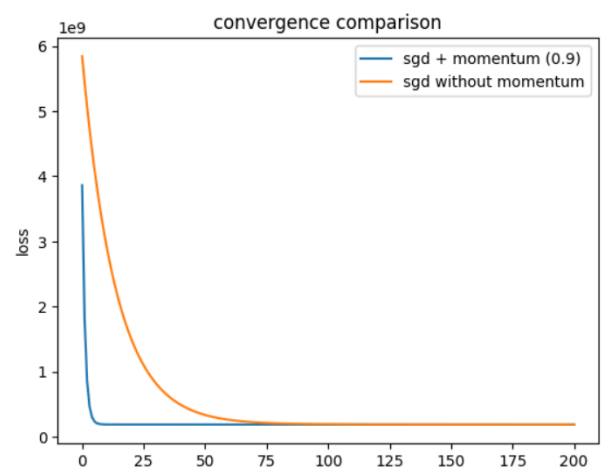
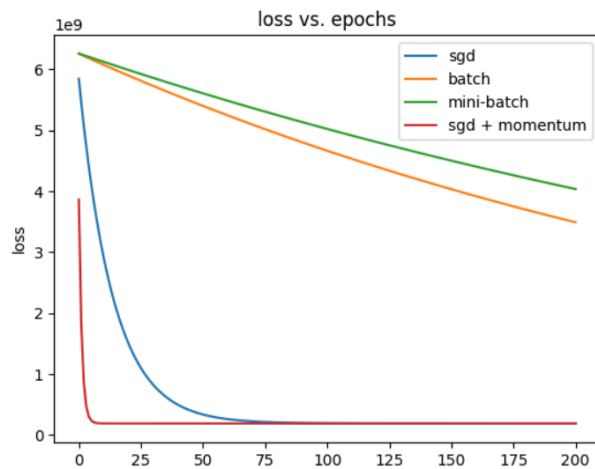
- Like SGD, but incorporates past updates to smooth out the learning process.
- Helps accelerate convergence and reduces oscillations.
- Instead of stopping at every step, it builds momentum, allowing it to avoid getting stuck in local minima.

→ **Observations**

● Loss vs. Epochs

- SGD: Fast initial drop, but fluctuates before settling.
- Batch: Slower, smooth decline but takes longer.

- Mini-Batch: Balanced between speed and stability.
  - SGD + Momentum: Reaches optimal loss much faster and smoother.
- Effect of Learning Rate
  - High learning rate ( $1e-6$ ): Fast convergence.
  - Medium learning rate ( $1e-7$ ): Converges slower but still reasonable.
  - Low learning rate ( $1e-8$ ): Takes too long, barely making progress.
- Effect of Momentum
  - Higher momentum (0.9): Faster convergence, minimal oscillation.
  - Medium momentum (0.7): Still fast, but slightly more oscillation.
  - Lower momentum (0.5): Converges slower with more fluctuation.



- Choosing the right values for learning rate, momentum, and batch size took some trial and error.

→ **Conclusions**

- SGD is fast but unstable, and works well when you need quick updates.
- Batch is slow but steady, great when stability is a priority.
- Mini-Batch is the best compromise, making it the most commonly used.
- Momentum speeds up convergence and reduces unnecessary movement.
- Learning rate matters, too high can overshoot, too low can stall.
- Higher momentum means faster, smoother learning.