Elsheikh Mohamed

UTA ID: 1001486517

Linear Regression

Q: What are the pros and cons of using the normal equation to solve for the weights in linear regression as opposed to using gradient descent?

Answer:

- **Pros of the Normal Equation:**
  - It gives a direct, closed-form solution — no need to iterate.
  - It works well for small to moderately sized datasets.
- **Cons of the Normal Equation:**
  - It is computationally expensive for large datasets because it involves matrix inversion, which is **$O(n^3)$** in time.
  - It does not scale well with a large number of features.
- **Gradient Descent Advantages:**
  - It handles large datasets efficiently, especially with many features.
  - It is more flexible, allowing for optimization with regularization and different loss functions.
- **Gradient Descent Drawbacks:**
  - Requires tuning of hyperparameters like learning rate and number of iterations.
  - May take longer to converge or get stuck in local minima without proper configuration.

Logistic Regression

Q: Why is the softmax function used in multi-class logistic regression?

Answer:

The softmax function is used because it converts the raw output scores (logits) from the model into probabilities for each class. It ensures that:

- All predicted probabilities are between 0 and 1.

- The sum of all class probabilities equals 1.
- It highlights the most likely class while still assigning nonzero probabilities to others.

This makes softmax ideal for multi-class classification problems, where we need a probability distribution across more than two classes.