Root Mean Square Prop

AdaGrad Scene

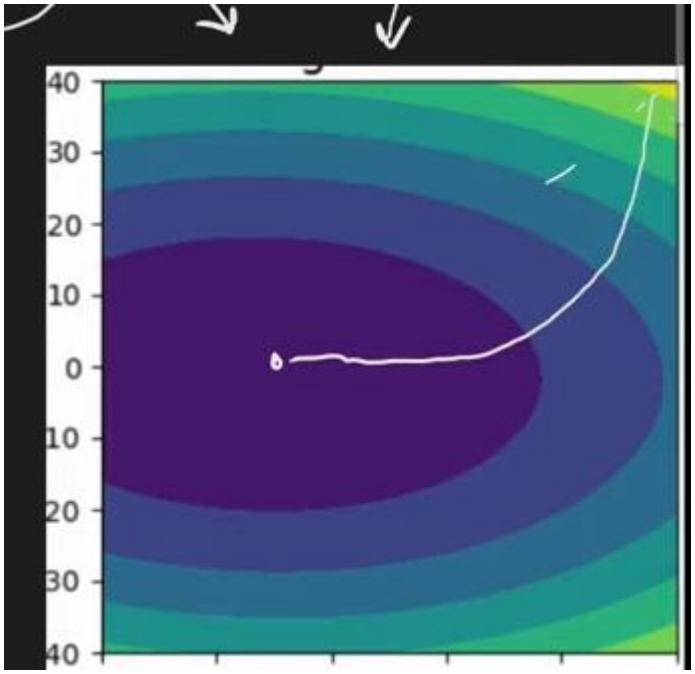
We use Adagrad when there is a sparse Data,

Sparse Data means Columns mai bhaut saare value hai vo 0 hai

Becoz of Sparse Data Cost function Elongated hota hai

Optimize karke Jbh

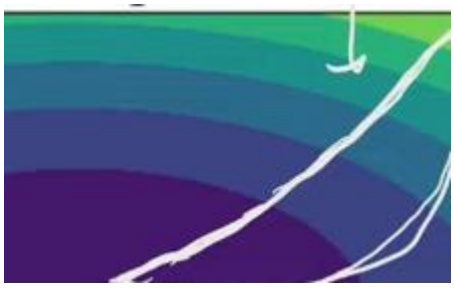
Minima tak pochte hai mtlb lowest Cost or Loss function tak



Minima tak ese hum lowest Minima taka ate hai Batch Gradient & SGD with Momentum

Adagrad yeah problem solve krta hai :

Even Ideally esa hona chaye:



Adagrad ka bhi 1 bhaut bada disadvantage tha ki

In adagrad we reduce the Learning Rate

But at one point LR itna zyaada reduce hota hai

Ki Converge hie nahi hopata hai becoz LR itna chota hai ki Update sbh Small hojata hai

Yeahi Problem Ko Solve karta Hai RMSProp

This RMSProp optimization Technique is Improvement Version of AdaGrad

RMSprop, which stands for Root Mean Square Propagation, is an optimization algorithm commonly used in the training of deep neural networks. It is designed to address some of the limitations of AdaGrad, particularly the issue of diminishing learning rates during training. RMSprop dynamically adjusts the learning rates for each parameter based on the magnitude of their recent gradients.

1. **Benefits:**
    - RMSprop mitigates the issue of diminishing learning rates by using a moving average of squared gradients.
    - It adapts the learning rates individually for each parameter based on recent gradients, allowing for faster convergence.
2. **Adaptability to Different Scales:**
    - RMSprop is effective in dealing with features that have different scales. The adaptive learning rates help in normalizing the updates for parameters with large or small gradients.
3. **Automatic Learning Rate Tuning:**
    - Unlike AdaGrad, RMSprop does not accumulate squared gradients indefinitely, which prevents overly small learning rates later in training.

While RMSprop has been successful in addressing some of the issues of AdaGrad, it is worth noting that the choice of hyperparameters, such as the learning rate and decay factor, can still impact the performance of the algorithm.

Disadvantages of RMSProp

No, it doesn't have much Disadvnatages

Adam ke aanese pehle RMSProp hie zyaada use hota tha

Even in today's time RMSProp competes with Adam

From GPT:

While RMSprop has been successful in addressing some of the issues of AdaGrad, it is worth noting that the choice of hyperparameters, such as the learning rate and decay factor, can still impact the performance of the algorithm. Additionally, later optimization algorithms like Adam further extend and improve upon the ideas introduced by RMSprop.