

Javított optimumkereső módszerek

Kovács Bálint

2023. november 20.

AdaMax:

AdaMax egy optimalizációs algoritmus, amely a népszerű Adam optimalizációs módszert bővíti ki. Az AdaMax algoritmus nem a múltbeli négyzetes gradiensek mozgó átlagát használja, hanem a múltbeli gradiensek végtelen normáját (maximumát). Ez a módosítás segít abban, hogy AdaMax jól teljesítsen olyan helyzetekben, ahol más algoritmusok belassulnának, különösen a mély neurális hálók képzésekor.

MaxProp:

MaxProp, azaz Maximum Proposition, egy optimalizációs algoritmus, amely a sztochasztikus optimalizációs módszer konvergenciasebességét javítja. Ez azzal érhető el, hogy bevezet egy új, adaptív tanulási ráta mechanizmust, amely az átvitt információ maximális értékén alapul. A MaxProp az egyes paraméterek tanulási rátáját adaptálja, figyelembe véve a gradienskomponensek maximális értékét. Ez az alkalmazkodóképesség segít a MaxPropnak kezelni a változó gradienseket és felgyorsítja az optimalizációs algoritmus konvergenciáját.

Nadam:

Nadam egy olyan optimalizációs algoritmus, amely a Nesterov Accelerated Gradient (NAG) és az Adam optimalizátor kombinációjából származik. Ötvözi mindkét algoritmus előnyeit, kihasználva a NAG momentum tagot a gyorsabb konvergencia és az Adam adaptív tanulási ráta előnyeért. A Nadamról ismert, hogy hatékonyan képez mély neurális hálózatokat, jó általánosítható teljesítményt nyújtva. Az adaptív tanulási ráta és a momentum hozzájárul a Nadam képességéhez, hogy hatékonyan mozogjon a bonyolult veszteséges tájakon.

NadaMax:

A NadaMax az AdaMax kiterjesztése, amely ötvözi a Nesterov Accelerated Gradient (NAG) és az AdaMax előnyeit. Hasonlóan a Nadamhoz, a NadaMax is alkalmazza a NAG momentum tagot a konvergencia fokozása érdekében,

és az AdaMax múltbeli gradienseinek végtelen normáját. Ez a kombináció eredményez egy olyan optimalizációs algoritmust, amely jól teljesít különböző optimalizációs problémáknál, különösen a mély neurális hálók képzésekor.

Eve:

Az Eve, vagyis „Exponential Variance Reduction for Efficient Optimization”, egy optimalizátor, amelyet a nem-konvex és nagy dimenziójú optimalizációs problémák kihívásaira terveztek. Bevezet egy adaptív tanulási ráta sémát, amely figyelembe veszi a gradiensek történelmi variációját. Az Eve dinamikusan állítja be a tanulási rátákat az exponenciális mozgóátlag négyzetes gradiens alapján, lehetővé téve számára, hogy különböző típusú optimalizációs tájakon hatékonyan konvergáljon, például mély tanulási helyzetekben.