Az optimalizálási folyamat futtatását a gradiens módszer esetében több kritérium alapján lehet befejezni:

* **Maximális Iterációszám:**  
  Az optimalizálási algoritmus egy előre meghatározott maximális iterációszámot ér el. Ha ez a szám teljesült, az algoritmus megáll.
* **Konvergencia Kritérium:**  
  A költségfüggvény változása két egymást követő iteráció között elég kicsi. Azaz, ha a költségfüggvény értéke egy előre meghatározott kis értékkel (ϵ) csökken, akkor az algoritmus megáll.
* **Gradiens Nagysága:**  
  Az algoritmus megáll, ha a gradiens nagysága egy előre meghatározott kis értéknél (ϵ) kisebb.

A gradiens módszer nem mindig találja meg a globális minimumot több okból:

* **Nem Konvex Költségfüggvény:**  
  Sok gépi tanulási probléma nem konvex költségfüggvénnyel rendelkezik, amely több lokális minimumot tartalmazhat. A gradiens módszer hajlamos a legközelebbi lokális minimumhoz konvergálni, különösen akkor, ha az inicializálás nem optimális.
* **Gradiens Iránya:**  
  A gradiens irány mindig a legnagyobb meredekség irányába mutat, ami nem garantálja, hogy a globális minimumhoz vezet, különösen nem konvex függvények esetén.
* **Tanulási Ráta:**  
  Ha a tanulási ráta túl nagy, az optimalizálás ugrálhat a minimum körül és esetleg nem talál stabil minimumot. Ha túl kicsi, az optimalizálás nagyon lassan konvergálhat, és fennáll a veszélye, hogy lokális minimumokba ragad.

A gradiens egy többváltozós függvény irányított származéka, amely megmutatja a függvény növekedésének legmeredekebb irányát az aktuális pontban. Az optimalizálás során a gradiens irányába történő mozgás csökkenti a költségfüggvény értékét. Az aktuális paraméterértékek frissítése a negatív gradiens irányába történik. A negatív gradiens irányába való mozgás tehát az aktuális pozíciót a költségfüggvény legmeredekebb csökkenésének irányába viszi, ami segít a minimum megtalálásában.