Függvényextrémum keresés

Kormányos Andor

Komplex Rendszerek Fizikája Tanszék

2021. november 17.

Függvényextrémumok keresése, optimalizáció

Alapkérdés

- adott egy $f(\mathbf{x})$ függvény, hol van a minimuma, illetve maximuma?
- jelölés: $\mathbf{x}$ azon változók v paraméterek, amelyek függvényében a minimumot keressük
- pl. költségfüggvény $J(\mathbf{a}; \mathbf{x}^{(i)}, y^{(i)})$ esetén az $\mathbf{a}$ paraméterek
- extrémum: minimum vagy maximum
- pl. rendszer energiaminimuma, legkisebb hatás stb.

A továbbiakban egy $f(\mathbf{x})$ függvény minimumkeresését tekintjük

- maximumkeresés: $f(\mathbf{x}) \to -f(\mathbf{x})$

Függvényextrémumok keresése, optimalizáció

Alapkérdés

- adott egy $f(\mathbf{x})$ függvény, hol van a minimuma, illetve maximuma?
- jelölés: $\mathbf{x}$ azon változók v paraméterek, amelyek függvényében a minimumot keressük
- pl. költségfüggvény $J(\mathbf{a}; \mathbf{x}^{(i)}, y^{(i)})$ esetén az $\mathbf{a}$ paraméterek
- extrémum: minimum vagy maximum
- pl. rendszer energiaminimuma, legkisebb hatás stb.

A továbbiakban egy $f(\mathbf{x})$ függvény minimumkeresését tekintjük

- maximumkeresés: $f(\mathbf{x}) \to -f(\mathbf{x})$

Optimalizáció: egy általánosabb feladat

- keressük egy $f(\mathbf{x})$ függvény extrémumát adott határfeltételek mellett
- pl. milyen $\mathbf{x}$-re lesz $f(\mathbf{x})$ maximális, ha megköveteljük, hogy $g(\mathbf{x}) > 0$

Függvényextrémumok keresése

Feladat: találjuk meg az extrémumot

- minél kevesebb lépésben
- minél kevesebb függvénykiértékeléssel
- minél pontosabban

Függvényextrémumok keresése

Feladat: találjuk meg az extrémumot

- minél kevesebb lépésben
- minél kevesebb függvénykiértékeléssel
- minél pontosabban

Két fő csoportra oszthatóak a módszerek:

- csak az $f(\mathbf{x})$ függvényértékeit használják
- $f(\mathbf{x})$ deriváltját is használják

Fő probléma: lokális és globális minimumok

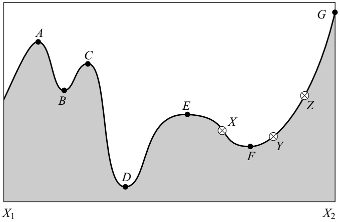


Figure: Egy függvény lokális és globális minimumai. ©Numerical Recipies

- az algoritmusok általában lokálisan működnek
- emiatt rossz helyről indulva rossz minimumot találnak
- "bennragadnak" a lokális minimumban
- univerzálisan jó globális algoritmus nincs

Minimum bekeretezése egy dimenzióban

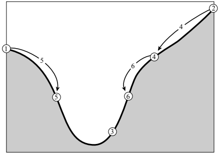
Egyszerű példa: egyváltozós $f(x)$ függvény minimumkeresése:

- ha $a < b < c$, továbbá $f(b) < f(a)$ és $f(b) < f(c)$
- akkor a függvénynek *lokális minimuma* van a és c között
- a minimum bekeretezéséhez tehát három pontot kell használni

Minimum iteratív keresése egy dimenzóban

Kiindulás:

- bekereteztük a minimumot az a, b és c számokkal



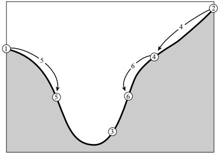
[1]

[1] ábra: Numerical Recipies

Minimum iteratív keresése egy dimenzóban

Kiindulás:

- bekereteztük a minimumot az a, b és c számokkal



[1]

Iterációs lépés:

- választunk egy új x pontot az $[a, c]$ intervallumon belül, pl. $b < x < c$
- ha $f(x) < f(b)$, akkor az új bekeretező ponthármas (b, x, c)
- ha $f(x) > f(b)$, akkor az új bekeretező ponthármas (a, b, x)
- addig folytatjuk, míg a két szélső pont közötti távolság elég kicsi nem lesz

[1] ábra: Numerical Recipies

Deriváltat is használó módszer egy dimenzióban

Ha nem csak az $f(x)$ függvény, de a deriváltja is ismert:

- ha a derivált gyorsan számolható
- a minimumot itt is a, b és c pontok keretezik: $f(b) < f(a)$ és $f(b) < f(c)$
- kiszámoljuk $f'(b)$-t, az előjele megadja, hogy merre lépünk
- többváltozós esetre is kidolgozható hasonló módszer
- néhány változós föggvények esetén a deriváltat használó módszerek konvergenciája sokkal gyorsabb lehet, mint a csak függvénykiértékelést használók konvergenciája

Deriváltat is használó módszer egy dimenzióban

Ha nem csak az $f(x)$ függvény, de a deriváltja is ismert:

- ha a derivált gyorsan számolható
- a minimumot itt is a, b és c pontok keretezik: $f(b) < f(a)$ és $f(b) < f(c)$
- kiszámoljuk $f'(b)$-t, az előjele megadja, hogy merre lépünk
- többváltozós esetre is kidolgozható hasonló módszer
- néhány változós föggvények esetén a deriváltat használó módszerek konvergenciája sokkal gyorsabb lehet, mint a csak függvénykiértékelést használók konvergenciája

Másik lehetőség:

- keressük a derivált zérushelyeit valamely gyökkereső módszerrel (pl. Newton-Ralphson)
- minden lépésben ki kell értékelni $f(x)$-t is, hogy meg tudjuk megmondani, maximum vagy minimum felé megyünk