Nemlineáris függvényillesztés

Kormányos Andor

Komplex Rendszerek Fizikája Tanszék

2020. február 20.

Nemlineáris függvényillesztés

A χ^2 definíció szerint:

$$\chi^2 = \sum_i \frac{[y_i - y(x_i|a)]^2}{\sigma_i^2},$$

Ezt szeretnénk minimalizálni$\Rightarrow$

$$\frac{\partial \chi^2}{\partial a_j} = 2 \cdot \sum_{i=1}^{N} \left[\frac{f(x_i|a) - y_i}{\sigma_i^2} \cdot \left. \frac{\partial f(x|a)}{\partial a_j} \right|_{x=x_i} \right] = 0$$

Ha f valamilyen paramétertől nem függő függvények lineárkombinációjaként állt elő, akkor a probléma lineáris volt.

Nemlineáris függvényillesztés

1. példa: mérési hiba nem csak a mért y_i értékekben, hanem a x_i mérési pontokatban is

Tegyük fel, hogy ilyen mérési eredményekre akarunk egyenest illeszteni. A költségfüggvény:

$$\chi^2 = \sum_{i=1}^{N} \frac{(y_i - a - bx_i)^2}{\sigma_{y,i}^2 + b^2 \sigma_{x,i}^2}$$

Nemlineáris függvényillesztés

1. példa: mérési hiba nem csak a mért y_i értékekben, hanem a x_i mérési pontokatban is

Tegyük fel, hogy ilyen mérési eredményekre akarunk egyenest illeszteni. A költségfüggvény:

$$\chi^2 = \sum_{i=1}^{N} \frac{(y_i - a - bx_i)^2}{\sigma_{y,i}^2 + b^2\sigma_{x,i}^2}$$

A $\frac{\partial \chi^2(a,b)}{\partial a} = 0$ egyenlet linearis a-ban, de a $\frac{\partial \chi^2(a,b)}{\partial b} = 0$ egyenlet nemlineáris b-ben

Nemlineáris függvényillesztés

2. példa: Breit-Wigner formula a rezonáns szórás méréseredményeinek kiértékelésére

$$f(E) = \frac{f_r}{(E - E_r)^2 + \Gamma^2/4}$$

E energia, ennek függvényében mérjük a szórást. f_r, E_r, Γ ismeretlen paraméterek, ezeket szeretnénk meghatározni

Másik példa: Fano rezonancia

$$f(E) = \frac{(q\Gamma/2 + (E - E_r +))^2}{(E - E_r)^2 + \Gamma^2/4}$$

q, E_r, Γ paraméterek

$\frac{\partial \chi^2}{\partial (parameterek)} = 0$ nemlineáris egyenletek a paraméterekre

Extrémumok megkeresése

Függvényillesztéskor a χ^2 minimumát keressük. Eddig a deriváltra vonatkozó egyenletrendszert probáltuk megoldani. Ha ez nehézkes vagy nem lehetséges:

- próbáljuk χ^2-et más módon minimalizálni
- ha ismerjük a $\frac{\partial f(x|a)}{\partial a_j}$ deriváltakat, akkor az segíthet

Extrémumok megkeresése

Függvényillesztéskor a χ^2 minimumát keressük. Eddig a deriváltra vonatkozó egyenletrendszert probáltuk megoldani. Ha ez nehézkes vagy nem lehetséges:

- próbáljuk χ^2-et más módon minimalizálni
- ha ismerjük a $\frac{\partial f(x|a)}{\partial a_j}$ deriváltakat, akkor az segíthet

Most a függvényillesztés a kapcsán merült fel, de az extrémumkeresés általánosabb probléma

- extrémum: minimum vagy maximum
- rendszer energiaminimuma
- legkisebb hatás stb.

Extrémumkeresés alapproblémája

Adott egy $f(\mathbf{x})$ függvény

- skalár értékű, de
- a változója lehet vektor is
- általában f valamilyen költségfüggvény, pl. χ^2 illesztésé

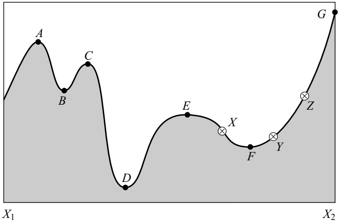
Hol van a függvény minimuma, illetve maximuma?

- a két probléma azonos $f(\mathbf{x}) \Rightarrow -f(\mathbf{x})$ felcseréléssel
- együttesen *extrémumkeresésnek* hívjuk.
- általában minimumkeresésről beszélünk

Feladat: találjuk meg az extrémumot

- minél kevesebb lépésben
- minél pontosabban
- minél kevesebb függvénykiértékeléssel

Lokális és globális minimumok



Fő probléma: lokális minimumok, mi a globális minimumot keressük

- nemlineáris esetben nincs jó globális algoritmus
- emiatt rossz helyről indulva rossz minimumot találnak
- "bennragadnak" a lokális minimumban

Közel a minimumhoz

χ^2 minimális egy $\mathbf{a}_{min}$ parametervektor esetén. Ha van valami sejtésünk arról, hogy $\mathbf{a}_{min}$ értékek hol helyezkednek el a paramétertérben, akkor ezt kihasználhatjuk:

- meghatározunk egy $\mathbf{a}_0$ kezdeti paramétervektort, ami már "közel" van $\mathbf{a}_{min}$-hez
- a χ^2 gradiensét és a Hessian mátrixát is ki tudjuk értékelni $\mathbf{a}_0$-ra, mert az $f(x|\mathbf{a})$ modell explicit adott

Közel a minimumhoz

$\chi^2(\mathbf{a})$ Taylor sorfejtéséből:

$$\chi^2(\mathbf{a}) \approx \chi_0^2 + \left.\frac{\partial \chi^2}{\partial a_k}\right|_{\mathbf{a}_0} (\mathbf{a} - \mathbf{a}_0) + (\mathbf{a} - \mathbf{a}_0)^\mathsf{T} \, \frac{1}{2} \left.\frac{\partial^2 \chi^2}{\partial a_k \partial a_l}\right|_{\mathbf{a}_0} (\mathbf{a} - \mathbf{a}_0)$$

Tehát

$$\nabla \chi^2(\mathbf{a}) = \nabla \chi^2(\mathbf{a}_0) + \mathbf{D} \cdot (\mathbf{a} - \mathbf{a}_0)$$

Jelölések:

$$\nabla \chi^2(\mathbf{a}_0) = \left.\frac{\partial \chi^2}{\partial a_k}\right|_{\mathbf{a}_0}$$

$$D_{kl} = \left.\frac{\partial^2 \chi^2}{\partial a_k \partial a_l}\right|_{\mathbf{a}_0} \qquad \text{Hessian mátrix}$$

Figyelem, most másképp használjuk a Hessian mátrixot, mint korábban!

Mivel a model explicit adott, $\nabla \chi^2(\mathbf{a}_0)$ és $\mathbf{D}$ kiszámolható

Közel a minimumhoz

Ha $\mathbf{a} = \mathbf{a}_{min}$, akkor $\nabla\chi^2(\mathbf{a}_{min}) = 0$. $\Rightarrow$

$$(\mathbf{a}_{min} - \mathbf{a}_0) = -\mathbf{D}^{-1}[\nabla\chi^2(\mathbf{a}_0)]$$

Elvileg egy lépésben megtalálhatjuk $\mathbf{a}_{min}$-t:

$$\mathbf{a}_{min} = \mathbf{a}_0 - \mathbf{D}^{-1}[\nabla\chi^2(\mathbf{a}_0)]$$

Közel a minimumhoz

Ha $\mathbf{a} = \mathbf{a}_{min}$, akkor $\nabla\chi^2(\mathbf{a}_{min}) = 0$. $\Rightarrow$

$$(\mathbf{a}_{min} - \mathbf{a}_0) = -\mathbf{D}^{-1}[\nabla\chi^2(\mathbf{a}_0)]$$

Elvileg egy lépésben megtalálhatjuk $\mathbf{a}_{min}$-t:

$$\mathbf{a}_{min} = \mathbf{a}_0 - \mathbf{D}^{-1}[\nabla\chi^2(\mathbf{a}_0)]$$

Hogyan számolhatjuk $\mathbf{D}$-t?

$$\frac{\partial^2\chi^2}{\partial a_k \partial a_l} = 2\sum_{i=1}^{N} \frac{1}{\sigma_i^2}\left[\frac{f(x_i|\mathbf{a})}{\partial a_k}\frac{f(x_i|\mathbf{a})}{\partial a_l} - (y_i - f(x_i|\mathbf{a}))\frac{\partial^2 f(x_i|\mathbf{a})}{\partial a_k \partial a_l}\right]$$

Az $(y_i - f(x_i|\mathbf{a}))$ lehet positív vagy negatív is. Gyakran jó közelítés:

$$\sum_{i=1}^{N} \frac{1}{\sigma_i^2}(y_i - f(x_i|\mathbf{a}))\frac{\partial^2 f(x_i|\mathbf{a})}{\partial a_k \partial a_l} \approx 0$$

Általános esetben

Ha nem lehetünk biztosak abban, hogy $\mathbf{a}_0$ közel van $\mathbf{a}_{min}$-hez, akkor függvény minimalizációs technikákat lehet alkalmazni, pl *konjugált gradiens módszer* (később)