

Költségfüggvény

Kormányos Andor

Komplex Rendszerek Fizikája Tanszék

2023 november 14.

Kérdés:

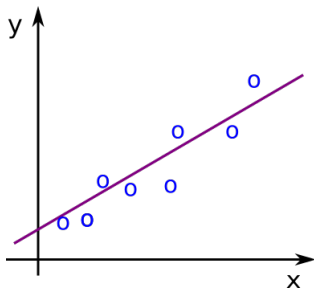
- adott egy $\{(\mathbf{x}^{(i)}, y^{(i)})\}$ adathalmaz
- erre szeretnénk illeszteni egy $h(\mathbf{x}; \mathbf{a})$ függényt
- hogyan határozhatjuk meg a $\mathbf{a} = (a_0, a_1 \dots a_k)$ paramétervektort?

Példa: egyenesillesztés $\Rightarrow h(x; \mathbf{a}) = a_0 + a_1x$

Költségfüggvény

Intuíció:

- az a $h(x; \mathbf{a})$ írja le legjobban az adatokat, amelyekre a mért $y^{(i)}$ adat és a hipotézis $h(x^{(i)}; \mathbf{a})$ értéke **átlagosan** a legkevésbe tér el
- $h(x; \mathbf{a})$ nem kell, hogy minden $y^{(i)}$ ponton átmenjen, csak az, hogy átlagosan kicsi legyen a különbség



Stratégia:

- definiálunk egy $\text{cost}(y^{(i)}, h(\mathbf{x}^{(i)}; \mathbf{a}))$ függvényt
- ez adja meg, hogy mennyire tér el az $y^{(i)}$ adattól a $h(\mathbf{x}^{(i)}; \mathbf{a})$ hipotézis
- definiáljuk a $J(\mathbf{a}; \mathbf{x}^{(i)}, y^{(i)})$ költségfüggvényt:

$$J(\mathbf{a}; \mathbf{x}^{(i)}, y^{(i)}) = \frac{1}{N} \sum_{i=1}^N \text{cost}(y^{(i)}, h(\mathbf{x}^{(i)}; \mathbf{a}))$$

- adott $h(\mathbf{x}; \mathbf{a})$ hipotézis (függvényalak) mellett az eltérés átlagosan akkor lesz a legkisebb, ha megkeressük a $J(\mathbf{a}; \mathbf{x}^{(i)}, y^{(i)})$ minimumát a $a_0, a_1 \dots a_k$ függvényében

$$\min_{\mathbf{a}} [J(\mathbf{a}; \mathbf{x}^{(i)}, y^{(i)})] = ?$$

- \Rightarrow függvény minimalizációs probléma

Legkisebb négyzetek módszer

Példa: egyenesillesztés legkisebb négyzetek módszerrel

- $h(x; \mathbf{a}) = a_0 + a_1x$
- költségfüggvény

$$J(a_0, a_1; x^{(i)}, y^{(i)}) = \frac{1}{2N} \sum_i^N \left(y^{(i)} - a_0 - a_1x^{(i)} \right)^2$$

- az \mathbf{a} paramétervektor meghatározása:

$$\min_{a_0, a_1} [J(a_0, a_1; x^{(i)}, y^{(i)})]$$

- milyen $\text{cost}(y^{(i)}, h(\mathbf{x}^{(i)}; \mathbf{a}))$ -t használjunk?
 - a legkisebb négyzetek módszer gyakran alkalmazott
 - egyenesillesztés esetén, az adatokra tett bizonyos feltételezések mellett, a legkisebb négyzetek módszernek adható egy valószínűségelméleti megalapozás (maximum likelihood becslés)
 - később még látunk példát másfajta költségfüggvényre is
- milyen algoritmussal tudjuk megtalálni egy függvény minimumát vagy maximumát?