Legkisebb négyzetek illesztés

Kormányos Andor

Komplex Rendszerek Fizikája Tanszék

2023. március 2.

A legkisebb négyzetek költségfüggvény

Milyen költségfüggvényt használjunk?

A legkisebb négyzetek költségfüggvény

Milyen költségfüggvényt használjunk?

Definiáljuk az ún. khi-négyzet költségfüggvényt:

$$J(\mathbf{a}; \mathbf{x}^{(i)}, y^{(i)}) \equiv \chi^2(\mathbf{a}; \mathbf{x}^{(i)}, y^{(i)}) = \sum_i \frac{\left[y^{(i)} - h(\mathbf{x}^{(i)}; \mathbf{a})\right]^2}{\sigma_i^2}$$

- a σ_i mérési hiba minden mérési pontban különbözhet
- azok a mérési pontok, ahol σ_i nagy, kisebb súllyal szerepelnek

A legkisebb négyzetek költségfüggvény

Miért pont a χ^2 költségfüggvényt használjuk?

- ha $P(y^{(i)})$ normális eloszlású σ_i szórással és a hipotézis $h(\mathbf{x}^{(i)}; \mathbf{a}) = a + b \cdot x$
- $\Rightarrow$ megmutatható, hogy a maximum likelihood becslés a χ^2 költségfüggvényhez vezet
- $\chi^2(\mathbf{a}; \mathbf{x}^{(i)}, y^{(i)})$ költségfüggvény általában akkor is eredményre vezet, ha $P(y^{(i)})$ nem normális eloszlású és/vagy $h(\mathbf{x}^{(i)}; \mathbf{a})$ nem lineáris függvény

A legkisebb négyzetek költségfüggvény

Miért pont a χ^2 költségfüggvényt használjuk?

- ha $P(y^{(i)})$ normális eloszlású σ_i szórással és a hipotézis $h(\mathbf{x}^{(i)}; \mathbf{a}) = a + b \cdot x$
- $\Rightarrow$ megmutatható, hogy a maximum likelihood becslés a χ^2 költségfüggvényhez vezet
- $\chi^2(\mathbf{a}; \mathbf{x}^{(i)}, y^{(i)})$ költségfüggvény általában akkor is eredményre vezet, ha $P(y^{(i)})$ nem normális eloszlású és/vagy $h(\mathbf{x}^{(i)}; \mathbf{a})$ nem lineáris függvény

Eddig a $h(\mathbf{x}^{(i)}; \mathbf{a})$-ről nem tettünk fel semmit

- lehet matematikai formula
- lehet algoritmus **a** bemenő paraméterekkel

A legkisebb négyzetek költségfüggvény

Miért pont a χ^2 költségfüggvényt használjuk?

- ha $P(y^{(i)})$ normális eloszlású σ_i szórással és a hipotézis $h(\mathbf{x}^{(i)}; \mathbf{a}) = a + b \cdot x$
- $\Rightarrow$ megmutatható, hogy a maximum likelihood becslés a χ^2 költségfüggvényhez vezet
- $\chi^2(\mathbf{a}; \mathbf{x}^{(i)}, y^{(i)})$ költségfüggvény általában akkor is eredményre vezet, ha $P(y^{(i)})$ nem normális eloszlású és/vagy $h(\mathbf{x}^{(i)}; \mathbf{a})$ nem lineáris függvény

Eddig a $h(\mathbf{x}^{(i)}; \mathbf{a})$-ről nem tettünk fel semmit

- lehet matematikai formula
- lehet algoritmus **a** bemenő paraméterekkel

Milyen más költségfüggvény létezik?

- lásd pl *Logisztikus regresszió*

A legkisebb négyzetek módszere

Ha a $h(\mathbf{x}^{(i)}; \mathbf{a})$ egy matematikai formula, akkor számolhatóak a $\frac{\partial \chi^2(\mathbf{a})}{\partial a_k}$ deriváltak.

A minimumban $\frac{\partial \chi^2(\mathbf{a})}{\partial a_k} = 0$, vagyis a χ^2 minimalizációs feltétel:

$$0 = -2 \sum_i \frac{\left[y^{(i)} - h(\mathbf{x}^{(i)}; \mathbf{a}) \right]}{\sigma_i^2} \frac{\partial h(\mathbf{x}^{(i)}; ...a_k...)}{\partial a_k}$$

minden k-ra, ahol k a paraméterek számáig futó index.

A legkisebb négyzetek módszere

Ha a $h(\mathbf{x}^{(i)};\mathbf{a})$ egy matematikai formula, akkor számolhatóak a $\frac{\partial\chi^2(\mathbf{a})}{\partial a_k}$ deriváltak.

A minimumban $\frac{\partial\chi^2(\mathbf{a})}{\partial a_k}=0$, vagyis a χ^2 minimalizációs feltétel:

$$0=-2\sum_i \frac{\left[y^{(i)}-h(\mathbf{x}^{(i)};\mathbf{a})\right]}{\sigma_i^2}\frac{\partial h(\mathbf{x}^{(i)};...a_k...)}{\partial a_k}$$

minden k-ra, ahol k a paraméterek számáig futó index.

Általános esetben:

- a fenti egyenletrendszer bonyolult
- pl a parciális deriváltakat nem feltétlenül könnyű kiszámolni
- $\chi^2(\mathbf{a};\mathbf{x}^{(i)},y^{(i)})$ abszolút minimumát nehéz megtalálni (lokális minimumok)

Példa: egyenes illesztése

Egyváltozós adathalmaz

- $x^{(1)}, x^{(2)}, ..., x^{(i)}$ mérési pontok, ezeknek nincs hibájuk
- $y^{(1)}, y^{(2)}, ..., y^{(i)}$ mért értékek
- $\sigma_1, \sigma_2, ..., \sigma_i$ becsült hibák

Feladat: illesszünk a pontokra egyenest $\chi^2(\mathbf{a}; x^{(i)}, y^{(i)})$ költségfüggvény segítségével

- modell: $h(x; \mathbf{a}) = a + b \cdot x$
- a költségfüggvény:

$$\chi^2(\mathbf{a}; x^{(i)}, y^{(i)}) = \sum_i \left(\frac{y^{(i)} - a - bx^{(i)}}{\sigma_i} \right)^2$$

Milyen a és b mellett lesz χ^2 minimális?

A legkisebb négyzetek módszere

Önálló tanulás:

- Egyenesillesztés
- Kilógó pontok kezelése

fóliák

További megjegyzések:

- a számolás során $\mathbf{x}^{(i)}$-ket (mérési pontok) végig ismertnek vettük, de általános esetben ezeknek is lehet hibájuk
- ha a modell egy algoritmus **a** bemenő paraméterekkel, akkor a minimum keresés is csak algoritmikusan kezelhető