

Gradient descent: egyszerű példa

Kormányos Andor

Komplex Rendszerek Fizikája Tanszék

2023. november 14.

Egyenesillesztés legkisebb négyzetek módszerrel

- adathalmaz: $\{(x^{(i)}, y^{(i)})\}$, $i = 0 \dots N$ számpárok
- erre szeretnénk egy $h(x; \mathbf{a}) = a_0 + a_1x$ függvényt illeszteni
- most tehát a paramétervektor $\mathbf{a} = (a_0, a_1)$

Egyenesillesztés legkisebb négyzetek módszerrel

- adathalmaz: $\{(x^{(i)}, y^{(i)})\}$, $i = 0 \dots N$ számpárok
- erre szeretnénk egy $h(x; \mathbf{a}) = a_0 + a_1x$ függvényt illeszteni
- most tehát a paramétervektor $\mathbf{a} = (a_0, a_1)$
- legkisebb négyzetek módszere: minimalizáljuk a

$$J(\mathbf{a}; x^{(i)}, y^{(i)}) = \frac{1}{2N} \sum_{i=1}^N \left(y^{(i)} - (a_0 + a_1 x^{(i)}) \right)^2$$

költségfüggvényt!

Egyenesillesztés legkisebb négyzetek módszerrel

- adathalmaz: $\{(x^{(i)}, y^{(i)})\}$, $i = 0 \dots N$ számpárok
- erre szeretnénk egy $h(x; \mathbf{a}) = a_0 + a_1x$ függvényt illeszteni
- most tehát a paramétervektor $\mathbf{a} = (a_0, a_1)$
- legkisebb négyzetek módszere: minimalizáljuk a

$$J(\mathbf{a}; x^{(i)}, y^{(i)}) = \frac{1}{2N} \sum_{i=1}^N \left(y^{(i)} - (a_0 + a_1x^{(i)}) \right)^2$$

költségfüggvényt!

- a parciális deriváltak:

$$\frac{\partial}{\partial a_0} J(\mathbf{a}) = \frac{1}{N} \sum_{i=1}^N \left(y^{(i)} - (a_0 + a_1x^{(i)}) \right)$$

$$\frac{\partial}{\partial a_1} J(\mathbf{a}) = \frac{1}{N} \sum_{i=1}^N \left(y^{(i)} - (a_0 + a_1x^{(i)}) \right) x^{(i)}$$

Egyenesillesztés legkisebb négyzetek módszerrel

Tehát az algoritmus:

$$a_0 = a_0^{(init)}, \quad a_1 = a_1^{(init)} \quad /*kezdőérték adás*/$$

```
repeat until convergence
{
 $a_0 := a_0 - \alpha \frac{1}{N} \sum_{i=1}^N (y^{(i)} - (a_0 + a_1 x^{(i)}))$ 
 $a_1 := a_1 - \alpha \frac{1}{N} \sum_{i=1}^N (y^{(i)} - (a_0 + a_1 x^{(i)})) x^{(i)}$ 
}
```

Megjegyzés

- megmutatható, hogy egyenesillesztés esetén csak egy minimum van, az algoritmus oda konvergál

Legkisebb négyzetek illesztés: kitekintés

- az egyenesillesztés esetén az a_0 , a_1 kiszámítására létezik egy egzakt képlet is
- az egyenesillesztés problémája általánosítható, pl polinomok illesztésére: pl $a_0 + a_1x + a_2x^2 + a_3x^3 + \dots$
- ennek létezik egy nem iteratív megoldása is: **normál egyenletek**
- ez egy lineáris egyenletrendszer megoldására vezet vissza
- ha olyan a probléma, hogy sok paramétert kell illeszteni vagy nagyon sok az adat, akkor a **gradient descent** módszer előnyösebb lehet (jobb skálázódás)