

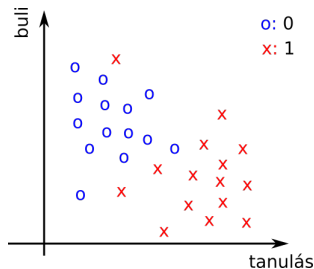
# Logisztikus regresszió: költségfüggvény

Kormányos Andor

Komplex Rendszerek Fizikája Tanszék

2022 október 25.

# Emlékeztető: mennyit muszály tanulni egy vizsgára?



**Kérdés:** ha van egy  $(x^{(i)}, y^{(i)})$ ,  $y^{(i)} = \{0, 1\}$  adathalmaz

- milyen  $h(\mathbf{a}; \mathbf{x})$  hipotézis függvény segítségével adhatjuk meg legjobban az  $y = 0$  és  $y = 1$  kimenetek közötti határvonalat?  $\Rightarrow$  tárgyaltuk
- mi legyen a  $J(\mathbf{a}; x^{(i)}, y^{(i)})$  költségfüggvény?

# Költségfüggvény<sup>1</sup>

Maximum likelihood becsléssel kaphatjuk:

- az a paramétervektor legvalószínűbb értéke az, amelyre

$$J(\mathbf{a}; \mathbf{x}^{(i)}, y^{(i)}) = -\frac{1}{N} \sum_{i=1}^N y^{(i)} \log(g(\mathbf{a} \cdot \mathbf{x}^{(i)})) + (1 - y^{(i)}) \log(1 - g(\mathbf{a} \cdot \mathbf{x}^{(i)}))$$

minimális

Levezetés: lásd “**Kitekintés**” fóliák között

---

<sup>1</sup>Andrew Ng, Coursera alapján

# Költségfüggvény<sup>1</sup>

Maximum likelihood becsléssel kaphatjuk:

- az a paramétervektor legvalószínűbb értéke az, amelyre

$$J(\mathbf{a}; \mathbf{x}^{(i)}, y^{(i)}) = -\frac{1}{N} \sum_{i=1}^N y^{(i)} \log(g(\mathbf{a} \cdot \mathbf{x}^{(i)})) + (1 - y^{(i)}) \log(1 - g(\mathbf{a} \cdot \mathbf{x}^{(i)}))$$

minimális

Levezetés: lásd “**Kitekintés**” fóliák között

**Állítás:** ez tekinthető a logisztikus regresszió költségfüggvényének.

---

<sup>1</sup>Andrew Ng, Coursera alapján

# Költségfüggvény

$J(\mathbf{a}; \mathbf{x}^{(i)}, y^{(i)})$  tekinthető a logisztikus regresszió költségfüggvényének.

$$J(\mathbf{a}; \mathbf{x}^{(i)}, y^{(i)}) = -\frac{1}{N} \sum_{i=1}^N y^{(i)} \log(g(\mathbf{a} \cdot \mathbf{x}^{(i)})) + (1 - y^{(i)}) \log(1 - g(\mathbf{a} \cdot \mathbf{x}^{(i)}))$$

Nézzük meg, hogy a  $J(\mathbf{a}; \mathbf{x}^{(i)}, y^{(i)})$ -hez milyen járulékot ad az  $i$ -ik adatpont!

Tfh  $y^{(i)} = 1$ ! Ekkor csak a  $-\log(g(\mathbf{a} \cdot \mathbf{x}^{(i)}))$  tagot kell tekinteni

# Költségfüggvény

$J(\mathbf{a}; \mathbf{x}^{(i)}, y^{(i)})$  tekinthető a logisztikus regresszió költségfüggvényének.

$$J(\mathbf{a}; \mathbf{x}^{(i)}, y^{(i)}) = -\frac{1}{N} \sum_{i=1}^N y^{(i)} \log(g(\mathbf{a} \cdot \mathbf{x}^{(i)})) + (1 - y^{(i)}) \log(1 - g(\mathbf{a} \cdot \mathbf{x}^{(i)}))$$

Nézzük meg, hogy a  $J(\mathbf{a}; \mathbf{x}^{(i)}, y^{(i)})$ -hez milyen járulékot ad az  $i$ -ik adatpont!

Tfh  $y^{(i)} = 1$ ! Ekkor csak a  $-\log(g(\mathbf{a} \cdot \mathbf{x}^{(i)}))$  tagot kell tekinteni

- ha a hipotézis értéke  $g(\mathbf{a} \cdot \mathbf{x}^{(i)}) \rightarrow 0$  akkor  $-\log(g(\mathbf{a} \cdot \mathbf{x}^{(i)})) \rightarrow \infty$

# Költségfüggvény

$J(\mathbf{a}; \mathbf{x}^{(i)}, y^{(i)})$  tekinthető a logisztikus regresszió költségfüggvényének.

$$J(\mathbf{a}; \mathbf{x}^{(i)}, y^{(i)}) = -\frac{1}{N} \sum_{i=1}^N y^{(i)} \log(g(\mathbf{a} \cdot \mathbf{x}^{(i)})) + (1 - y^{(i)}) \log(1 - g(\mathbf{a} \cdot \mathbf{x}^{(i)}))$$

Nézzük meg, hogy a  $J(\mathbf{a}; \mathbf{x}^{(i)}, y^{(i)})$ -hez milyen járulékot ad az  $i$ -ik adatpont!

Tfh  $y^{(i)} = 1!$  Ekkor csak a  $-\log(g(\mathbf{a} \cdot \mathbf{x}^{(i)}))$  tagot kell tekinteni

- ha a hipotézis értéke  $g(\mathbf{a} \cdot \mathbf{x}^{(i)}) \rightarrow 0$  akkor  $-\log(g(\mathbf{a} \cdot \mathbf{x}^{(i)})) \rightarrow \infty$

Megfelel az intuíciónak: ha  $g(\mathbf{a} \cdot \mathbf{x}^{(i)}) = 0$ , vagyis  $P(y = 1 | \mathbf{x}^{(i)}; \mathbf{a}) = 0$ , de a mért érték  $y^{(i)} = 1$ , akkor a hipotézis és a mérés nagyon eltér  $\rightarrow$  nagy "költség"

# Költségfüggvény

$J(\mathbf{a}; \mathbf{x}^{(i)}, y^{(i)})$  tekinthető a logisztikus regresszió költségfüggvényének.

$$J(\mathbf{a}; \mathbf{x}^{(i)}, y^{(i)}) = -\frac{1}{N} \sum_{i=1}^N y^{(i)} \log(g(\mathbf{a} \cdot \mathbf{x}^{(i)})) + (1 - y^{(i)}) \log(1 - g(\mathbf{a} \cdot \mathbf{x}^{(i)}))$$

Nézzük meg, hogy a  $J(\mathbf{a}; \mathbf{x}^{(i)}, y^{(i)})$ -hez milyen járulékot ad az  $i$ -ik adatpont!

Tfh  $y^{(i)} = 1$ ! Ekkor csak a  $-\log(g(\mathbf{a} \cdot \mathbf{x}^{(i)}))$  tagot kell tekinteni

- ha a hipotézis értéke  $g(\mathbf{a} \cdot \mathbf{x}^{(i)}) \rightarrow 0$  akkor  $-\log(g(\mathbf{a} \cdot \mathbf{x}^{(i)})) \rightarrow \infty$

Megfelel az intuíciónak: ha  $g(\mathbf{a} \cdot \mathbf{x}^{(i)}) = 0$ , vagyis  $P(y = 1 | \mathbf{x}^{(i)}; \mathbf{a}) = 0$ , de a mért érték  $y^{(i)} = 1$ , akkor a hipotézis és a mérés nagyon eltér  $\rightarrow$  nagy "költség"

- ha a hipotézis értéke  $g(\mathbf{a} \cdot \mathbf{x}^{(i)}) \rightarrow 1$  akkor  $-\log(g(\mathbf{a} \cdot \mathbf{x}^{(i)})) \rightarrow 0$
- tehát ha  $y^{(i)} = 1$  és a hipotézis  $g(\mathbf{a} \cdot \mathbf{x}^{(i)}) \rightarrow 1 \Rightarrow$  kis költség



$J(\mathbf{a}; \mathbf{x}^{(i)}, y^{(i)})$  tekinthető a logisztikus regresszió költségfüggvényének.

$$J(\mathbf{a}; \mathbf{x}^{(i)}, y^{(i)}) = -\frac{1}{N} \sum_{i=1}^N y^{(i)} \log(g(\mathbf{a} \cdot \mathbf{x}^{(i)})) + (1 - y^{(i)}) \log(1 - g(\mathbf{a} \cdot \mathbf{x}^{(i)}))$$

$J(\mathbf{a}; \mathbf{x}^{(i)}, y^{(i)})$  tekinthető a logisztikus regresszió költségfüggvényének.

$$J(\mathbf{a}; \mathbf{x}^{(i)}, y^{(i)}) = -\frac{1}{N} \sum_{i=1}^N y^{(i)} \log(g(\mathbf{a} \cdot \mathbf{x}^{(i)})) + (1 - y^{(i)}) \log(1 - g(\mathbf{a} \cdot \mathbf{x}^{(i)}))$$

Tfh  $y^{(i)} = 0$  ! Ekkor csak a  $-\log(1 - g(\mathbf{a} \cdot \mathbf{x}^{(i)}))$  tagot kell tekinteni

# Költségfüggvény

$J(\mathbf{a}; \mathbf{x}^{(i)}, y^{(i)})$  tekinthető a logisztikus regresszió költségfüggvényének.

$$J(\mathbf{a}; \mathbf{x}^{(i)}, y^{(i)}) = -\frac{1}{N} \sum_{i=1}^N y^{(i)} \log(g(\mathbf{a} \cdot \mathbf{x}^{(i)})) + (1 - y^{(i)}) \log(1 - g(\mathbf{a} \cdot \mathbf{x}^{(i)}))$$

Tfh  $y^{(i)} = 0$  ! Ekkor csak a  $-\log(1 - g(\mathbf{a} \cdot \mathbf{x}^{(i)}))$  tagot kell tekinteni

- ha a hipotézis értéke  $g(\mathbf{a} \cdot \mathbf{x}^{(i)}) \rightarrow 0$  akkor  $-\log(1 - g(\mathbf{a} \cdot \mathbf{x}^{(i)})) \rightarrow 0$
- tehát ha  $y^{(i)} = 0$  és a hipotézis  $g(\mathbf{a} \cdot \mathbf{x}^{(i)}) \rightarrow 0$  értéket ad  $\Rightarrow$  kis költség

# Költségfüggvény

$J(\mathbf{a}; \mathbf{x}^{(i)}, y^{(i)})$  tekinthető a logisztikus regresszió költségfüggvényének.

$$J(\mathbf{a}; \mathbf{x}^{(i)}, y^{(i)}) = -\frac{1}{N} \sum_{i=1}^N y^{(i)} \log(g(\mathbf{a} \cdot \mathbf{x}^{(i)})) + (1 - y^{(i)}) \log(1 - g(\mathbf{a} \cdot \mathbf{x}^{(i)}))$$

Tfh  $y^{(i)} = 0$  ! Ekkor csak a  $-\log(1 - g(\mathbf{a} \cdot \mathbf{x}^{(i)}))$  tagot kell tekinteni

- ha a hipotézis értéke  $g(\mathbf{a} \cdot \mathbf{x}^{(i)}) \rightarrow 0$  akkor  $-\log(1 - g(\mathbf{a} \cdot \mathbf{x}^{(i)})) \rightarrow 0$
- tehát ha  $y^{(i)} = 0$  és a hipotézis  $g(\mathbf{a} \cdot \mathbf{x}^{(i)}) \rightarrow 0$  értéket ad  $\Rightarrow$  kis költség
- ha a hipotézis értéke  $g(\mathbf{a} \cdot \mathbf{x}^{(i)}) \rightarrow 1$  akkor  $-\log(1 - g(\mathbf{a} \cdot \mathbf{x}^{(i)})) \rightarrow \infty$
- tehát ha  $y^{(i)} = 0$  és a hipotézis  $g(\mathbf{a} \cdot \mathbf{x}^{(i)}) \rightarrow 1 \Rightarrow$  nagy költség

Összefoglalva:

$$\text{cost}(y^{(i)}, h(\mathbf{x}^{(i)}; \mathbf{a})) = \begin{cases} -\log(h(\mathbf{x}^{(i)}; \mathbf{a})) & \text{ha } y^{(i)} = 1 \\ -\log(1 - h(\mathbf{x}^{(i)}; \mathbf{a})) & \text{ha } y^{(i)} = 0 \end{cases}$$

Költségfüggvény

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

# Költségfüggvény

Összefoglalva:

$$\text{cost}(y^{(i)}, h(\mathbf{x}^{(i)}; \mathbf{a})) = \begin{cases} -\log(h(\mathbf{x}^{(i)}; \mathbf{a})) & \text{ha } y^{(i)} = 1 \\ -\log(1 - h(\mathbf{x}^{(i)}; \mathbf{a})) & \text{ha } y^{(i)} = 0 \end{cases}$$

Költségfüggvény

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

Hogyan határozhatjuk meg  $\mathbf{a}$ -t?

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