

# Logisztikus regresszió (logistic regression)

Kormányos Andor

Komplex Rendszerek Fizikája Tanszék

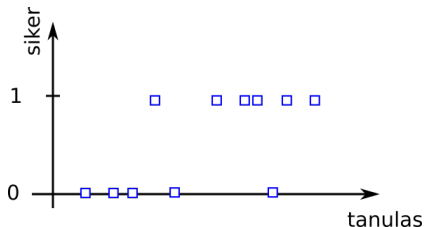
2022 október 25.

# Klasszifikáció

- a klasszifikációs problémák hasonlóak a regresszióhoz
- de az  $y^{(i)}$  csak néhány diszkrét értéket vehet fel
- most csak bináris klasszifikációt tekintünk ( $y^{(i)} = \{0, 1\}$ )

# Klasszifikáció

- a klasszifikációs problémák hasonlóak a regresszióhoz
- de az  $y^{(i)}$  csak néhány diszkrét értéket vehet fel
- most csak bináris klasszifikációt tekintünk ( $y^{(i)} = \{0, 1\}$ )



- ha a tanulással  $x_{sajat}$  időt töltünk, akkor a korábbi adatok alapján az  $y = 1$  (sikeres vizsga) vagy a  $y = 0$  (sikertelen vizsga) a valószínűbb?

Klasszifikációs probléma esetén:

- mi legyen a  $h(\mathbf{x}; \mathbf{a})$  hipotézis függvény?
- milyen költségfüggvényt használjunk?

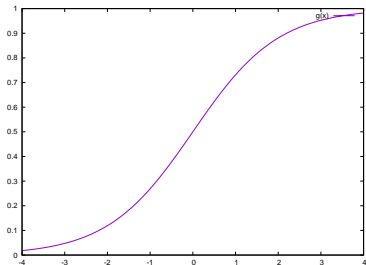
# Mi legyen a hipotézis?

- válasszunk olyan  $h(\mathbf{x}; \mathbf{a})$  függényt, amely a valós számokat leképezi a  $(0, 1)$  intervallumra
- ilyen pl az ún *sigmoid függvény* (*logisztikus függvény*):

$$h(\mathbf{x}; \mathbf{a}) := g(\mathbf{a} \cdot \mathbf{x}) = \frac{1}{1 + e^{-(\mathbf{a} \cdot \mathbf{x})}}$$

- Jelölés:  $\mathbf{a} \cdot \mathbf{x}$  az  $\mathbf{a}$  és  $\mathbf{x}$  vektorok skalárszorzata
- **Konvenció:**  $\mathbf{x}$  vektor 0-ik eleme:  $x_0 = 1$
- legegyszerűbb eset:  $\mathbf{a} \cdot \mathbf{x} = a_0 + a_1 x$

# Sigmoid függvény



- nagy pozitív argumentum értékekre  $\rightarrow 1$ , nagy negatív argumentum értékekre  $\rightarrow 0$

# Klasszifikáció, döntési határ (Decision boundary)

Hogyan használhatjuk a  $g(\mathbf{a} \cdot \mathbf{x})$  függvényt klasszifikációra?

## Valószínűségi értelmezés:

- adott  $\mathbf{a}$  paramétervektor mellett  $g(\mathbf{a} \cdot \mathbf{x})$  adja meg az  $y = 1$  kimenet valószínűségét:

$$P(y = 1|\mathbf{x}; \mathbf{a}) = g(\mathbf{a} \cdot \mathbf{x})$$

# Klasszifikáció, döntési határ (Decision boundary)

Hogyan használhatjuk a  $g(\mathbf{a} \cdot \mathbf{x})$  függvényt klasszifikációra?

## Valószínűségi értelmezés:

- adott  $\mathbf{a}$  paramétervektor mellett  $g(\mathbf{a} \cdot \mathbf{x})$  adja meg az  $y = 1$  kimenet valószínűségét:

$$P(y = 1 | \mathbf{x}; \mathbf{a}) = g(\mathbf{a} \cdot \mathbf{x})$$

**Klasszifikáció:** adott  $\mathbf{a}$  és  $\mathbf{x}$  mellett

ha  $g(\mathbf{a} \cdot \mathbf{x}) \geq 0.5 \Rightarrow$  kimenet "y = 1"

ha  $g(\mathbf{a} \cdot \mathbf{x}) < 0.5 \Rightarrow$  kimenet "y = 0"



# Példa egyszerű klasszifikációra

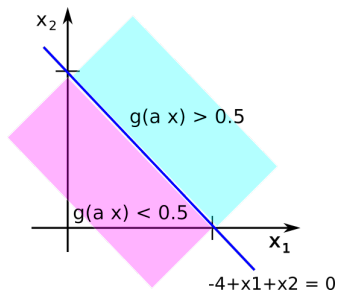
A  $g(\mathbf{a} \cdot \mathbf{x})$  függvényt akarjuk arra használni, hogy megjósoljuk, bukunk vagy átmegyünk a vizsgán:  $y = \{0, 1\}$

Tfh  $\mathbf{a} \cdot \mathbf{x} = -4.0 + 1 \cdot x$

- ha kevesebb, mint 4 órát tanultunk, i.e.,  $x < 4.0$ :
  - $-4.0 + 1 \cdot x < 0 \Rightarrow g(\mathbf{a} \cdot \mathbf{x}) < 0.5$
  - a kimenet “ $y = 0$ ”, bukás a vizsgán
- ha  $x \geq 4$  órát tanultunk:
  - $-4.0 + 1 \cdot x > 0 \Rightarrow g(\mathbf{a} \cdot \mathbf{x}) \geq 0.5$
  - a kimenet “ $y = 1$ ”, vagyis sikeres vizsga

## Kétváltozós példa: döntési határ

$$\text{Tfh } \mathbf{a} \cdot \mathbf{x} = -4.0 + 1 \cdot x_1 + 1 \cdot x_2$$



- $-4.0 + 1 \cdot x_1 + 1 \cdot x_2 = 0$  kijelöl egy határt
- egy nevezzük döntési határnak (decision boundary)
- felső félsíkban  $-4.0 + 1 \cdot x_1 + 1 \cdot x_2 \geq 0$ , vagyis a klasszifikáció értéke “ $y = 1$ ”
- az alsóban  $-4.0 + 1 \cdot x_1 + 1 \cdot x_2 < 0$ , tehát “ $y = 0$ ”

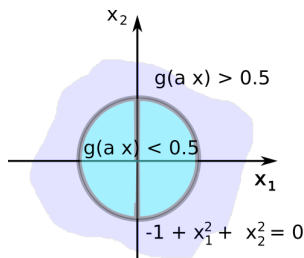
# Egy bonyolultabb döntési határ

- az  $\mathbf{x}$  vektort általánosan értjük: pl  $\mathbf{x} = (x_1^2, x_2^2 \dots x_j^2)^T$
- tehát  $\mathbf{x}$  tartalmazhatja az eredeti  $x_j$ -k valamilyen függvényét is

# Egy bonyolultabb döntési határ

- az  $\mathbf{x}$  vektort általánosan értjük: pl  $\mathbf{x} = (x_1^2, x_2^2 \dots x_j^2)^T$
- tehát  $\mathbf{x}$  tartalmazhatja az eredeti  $x_j$ -k valamilyen függvényét is

Tfh  $\mathbf{a} \cdot \mathbf{x} = -1.0 + x_1^2 + x_2^2$



- $-1 + x_1^2 + x_2^2 = 0$  kijelöl egy határt
- ha  $-1 + x_1^2 + x_2^2 < 0$  akkor a klasszifikáció értéke “ $y = 0$ ”
- ha  $-1 + x_1^2 + x_2^2 \geq 0$  akkor “ $y = 1$ ”

$\mathbf{a} \cdot \mathbf{x}$ -től **függően bonyolult döntési határokat tudunk leírni.**