

# Görbeillesztés

Lineáris regresszió, legkisebb négyzetek módszere

Illés Gergő és Sarkadi Balázs

PTE TTK Fizikai Intézet

2023. március 2.

# Tartalom

## Lineáris regresszió

### Általános ismertetés

# Általános ismertetés

- ▶ A lineáris regresszió a leggyakrabban használt görbeillesztési módszer.
- ▶ Lineáris regresszió használatakor lineáris kapcsolatot feltételezünk a független- és függő változó között.
- ▶ Nemlineáris kapcsolatokra is alkalmazható a függő- és független változók közti kapcsolat linearizálásával.

# Általános ismertetés

Tételezzük fel, hogy rendelkezünk egy adathalmazzal amely  $n$  darab statisztikai egységet tartalmaz. Ezt mátrix formájában a következőképp írhatjuk le.

$$\begin{bmatrix} \{y_1, [x_{11}, \dots, x_{1p}]\} \\ \vdots \\ \{y_n, [x_{n1}, \dots, x_{np}]\} \end{bmatrix}$$

Ebben az írásmódban  $y$  a függő változó,  $\vec{x}$   $p$  hosszúságú vektor pedig az úgynevezett regresszor ami a független változókat tartalmazza.

# Általános ismertetés

A becslésünk jóságára vezessünk be egy hibaváltozót, ez legyen  $\epsilon$ .  
Lineáris függést, valamint  $\epsilon$  hibát feltételezve az egyes  $y_i$ -k a következőképp írhatók fel:

$$y_i = \beta_0 + \sum_{j=1}^p \beta_j x_{ij} + \epsilon_i = \vec{x}_i \vec{\beta}^T + \epsilon_i, \quad i = 1, 2, \dots, n$$

Az egyenletből látszik, hogy itt feltételezünk egy 0-ik  $x_{i0}$  elemet, ami minden  $\vec{x}$  esetén 1-nek adódik.