# Általános lineáris függvényillesztés

Kormányos Andor

Komplex Rendszerek Fizikája Tanszék

2023. március 2.

# Általános lineáris függvényillesztés

Az $\{(x^{(i)}, y^{(i)})\}$ adathalmazra szeretnénk illeszteni egy modellt.

- $h(x; \mathbf{a})$ alakja:

$$h(x; \mathbf{a}) = \sum_{j=1}^{M} a_j f_j(x),$$

  - $f_j(x)$ tetszőleges ún. bázisfüggvény
  - $f_j(x)$ nem függ az $a_j$-ktől
- példa bázisfüggvényre: $f_j(x) = x^{j-1}$
- ekkor a lineáris függvényillesztés lényegében egy polinomillesztés

# Általános lineáris függvényillesztés

Költségfüggvény:

$$\chi^2(\mathbf{a}; x^{(i)}, y^{(i)}) = \sum_{i=1}^{N} \left[ \frac{y^{(i)} - h(x^{(i)}; \mathbf{a})}{\sigma_i} \right]^2$$

A minimumban

$$\frac{\partial \chi^2(\mathbf{a})}{\partial a_j} = 0$$

Behelyettesítve $h(x; \mathbf{a})$-t:

$$\frac{\partial \chi^2(\mathbf{a})}{\partial a_j} = 2 \cdot \sum_{i=1}^{N} \left[ \frac{1}{\sigma_i^2} \cdot \left( \sum_{k=1}^{M} a_k f_k(x^{(i)}) - y^{(i)} \right) \cdot f_j(x^{(i)}) \right] = 0$$

Ez egy lineáris egyenletrendszer az $a_k$ együtthatókra!

# A tervmátrix[1]

Bevezetjük a következő jelöléseket:

$$X_{ij} = \frac{f_j(x^{(i)})}{\sigma_i} \qquad\qquad b_i = \frac{y^{(i)}}{\sigma_i}$$

$X_{ij}$ az úgynevezett *tervmátrix*:

- az $M$ oszlopa a bázisfüggvényeknek felel meg
- az $N$ sora a mérési pontoknak
- a mátrixelemek a $j$. bázisfüggvény $x^{(i)}$ helyeken vett értékei

[1]design matrix

# A tervmátrix és a $b_i$ vektor felépítése

$$X_{ij} = \begin{pmatrix} \frac{f_1(x^{(1)})}{\sigma_1} & \frac{f_2(x^{(1)})}{\sigma_1} & \dots & \frac{f_M(x^{(1)})}{\sigma_1} \\ \frac{f_1(x^{(2)})}{\sigma_2} & \frac{f_2(x^{(2)})}{\sigma_2} & \dots & \frac{f_M(x^{(2)})}{\sigma_2} \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ \frac{f_1(x^{(N)})}{\sigma_N} & \frac{f_2(x^{(N)})}{\sigma_N} & \dots & \frac{f_M(x^{(N)})}{\sigma_N} \end{pmatrix} \qquad b_i = \begin{pmatrix} \frac{y^{(1)}}{\sigma_1} \\ \frac{y^{(2)}}{\sigma_2} \\ \vdots \\ \vdots \\ \frac{y^{(N)}}{\sigma_N} \end{pmatrix}$$

# A lineáris illesztés normálegyenletei

A $\frac{\partial \chi^2(\mathbf{a})}{\partial a_j}$-ra felírt egyenletek ezzel:

$$\sum_{i=1}^{N} \left[ \left( \sum_{k=1}^{M} a_k X_{ik} - b_i \right) \cdot X_{ij} \right] = 0,$$

# A lineáris illesztés normálegyenletei

A $\frac{\partial \chi^2(\mathbf{a})}{\partial a_j}$-ra felírt egyenletek ezzel:

$$\sum_{i=1}^{N} \left[ \left( \sum_{k=1}^{M} a_k X_{ik} - b_i \right) \cdot X_{ij} \right] = 0,$$

Indexes írásmóddal:

$$X_{ik} a_k X_{ij} = X_{ij} b_i$$

Mátrixos írásmódban:

$$\mathbf{X}^T \mathbf{X} \mathbf{a} = \mathbf{X}^T \mathbf{b}$$

- $\mathbf{X}^T \mathbf{X}$ egy $M \times M$-es mátrix
- **a** tartalmazza az ismeretleneket
- ez egy $M \times M$ **lineáris egyenletrendszer**: az illesztés normálegyenletei

# Normálegyenletek megoldása

Hogyan oldjuk meg?

- lineáris egyenletrendszer megoldó numerikus rutinok
- viszont az egyenletrendszer lehet rosszul kondicionált $\Rightarrow$ pontatlan megoldás
- ilyenkor lehet hasznos az ún szinguláris érték felbontás (singular value decomposition, SVD)
- SVD-re is elérhetőek jó numerikus rutinok

# Példa: parabola illesztése 5 pontra

A mérési adatok

$$
\begin{array}{l}
x^{(i)} = \{ \quad -2 \quad , \quad -1 \quad , \quad 0 \quad , \quad 1 \quad , \quad 2 \quad \} \\
y^{(i)} = \{ \quad\quad 5.1, \quad\quad 1.9, \quad 1.1, \quad 2.1, \quad 4.9\}
\end{array}
$$

- a mérési hiba az egyszerűség kedvéért $\sigma_i = \sigma = 0.1$
- a modell: $f(x) = a_0 + a_1 x + a_2 x^2$, azaz $f_1(x) = 1$, $f_2(x) = x$ és $f_3(x) = x^2$.

$$
\mathbf{X} = 10 \begin{bmatrix} 1 & -2 & 4 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix}
$$

$$
\mathbf{X}^T\mathbf{X} = 10^2 \begin{bmatrix} 5 & 0 & 10 \\ 0 & 10 & 0 \\ 10 & 0 & 34 \end{bmatrix}
$$

$$
\mathbf{X}^T \frac{\mathbf{y}}{\sigma} = 10^2 \begin{bmatrix} 15.1 \\ -0.2 \\ 44.0 \end{bmatrix}
$$

# Példa: parabola illesztése 5 pontra

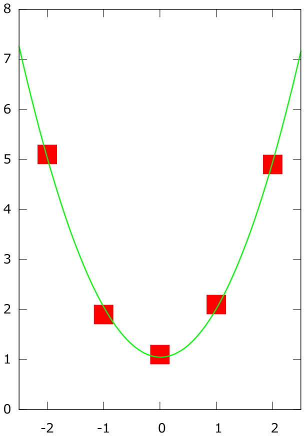
A függvényillesztés eredménye:

Az $\mathbf{X}^T\mathbf{X}\mathbf{a} = \mathbf{X}^T\mathbf{y}$ egyenletet $\mathbf{a}$-ra megoldva:

$$\mathbf{a}_{fit} = \begin{bmatrix} 1.049 \\ -0.020 \\ 0.986 \end{bmatrix}$$

Az illesztett modell:

$$f(x) = 1.049 - 0.020x + 0.986x^2$$



# Többváltozós polinomillesztés

További példa: $y^{(i)}$ több változótól függ: $\mathbf{x}^{(i)}$ K-dimenziós vektor

Ekkor definiáljuk a következő polinomot

$$h(\mathbf{x};\mathbf{a}) = a_0 + \sum_{k=1}^{K}\sum_{j=1}^{M} a_j^{(k)}[x_k]^j,$$

- a probléma immár összesen $M \cdot K + 1$ ismeretlent fog tartalmazni
- ebben még nincsenek vegyes tagok!
- a probléma ugyanúgy oldható meg, mint az előző

# Normálegyenletek vs iteratív módszerek

A költségfüggvény minimalizálására **iteratív módszerek** is léteznek:

- pl legmeredekebb ereszkedés (gradient descent)
- és olyan módszerek, amelyekhez nem szükséges a deriváltak kiszámolása (pl. szimplex módszer)

Hogyan válasszunk a normálegyenletek és az iteratív módszerek között?

- ha nagyon sok ismeretlen paraméter van $\Rightarrow$ a tervmátrix nagyon nagy is lehet
- ekkor gyorsabb lehet valamilyen iteratív módszer
- ha nemlineáris függvényillesztés kell csinálni $\Rightarrow$ iteratív módszer könnyebb lehet

# További fontos kérdések

A paraméterek meghatározásával még nem ért véget a feladat:

- mekkora a meghatározott paraméterek hibája?
- egyáltalán mennyire jó a modell? Hiába kicsi a meghatározott paraméterek hibája, ha rossz a modell, amit használunk