

Z minula

Lineárna regresia

Vstupy: riadky v matici X .

Očakávané výstupy: vektor $\vec{y}$.

Hľadáme parametre $\vec{\theta}$, aby $E = (X\vec{\theta} - \vec{y})^T (X\vec{\theta} - \vec{y})$ bolo, čo najmenšie.

Trénovanie

Možnosť 1: vyriešiť sústavu $X^T X \vec{\theta} = X^T \vec{y}$

Možnosť 2: (stochastic) gradient descent: $\vec{\theta} = \vec{\theta} - \alpha X^T (X\vec{\theta} - \vec{y})$

E je konvexná funkcia, má najviac jedno lokálne minimum, ktoré je aj globálne a obe metódy nájdu to isté riešenie (až na numerickú chybu).

Predikcia z nového vstupu

$$y_{new} = \vec{x}_{new}^T \cdot \vec{\theta}$$

Stále regresia (jedno reálne číslo ako výstup).
Niekedý sa dáta nesprávajú lineárne.

- Lokálne váhovaná lineárna regresia
- Polynomiálna regresia a jej redukcia na lineárnu
- Neurónové siete

Váňovaná lineárna regresia

Minimalizujeme:

$$E = \sum_{i=1}^n w_i ((\vec{x}^{(i)})^T \vec{\theta} - y^{(i)})^2 = (X\vec{\theta} - \vec{y})^T I_w (X\vec{\theta} - \vec{y})$$

$$I_w = \begin{pmatrix} w_1 & 0 & \dots & 0 \\ 0 & w_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & w_n \end{pmatrix}$$

Váňovaná lineárna regresia

Minimalizujeme:

$$E = \sum_{i=1}^n w_i ((\vec{x}^{(i)})^T \vec{\theta} - y^{(i)})^2 = (X\vec{\theta} - \vec{y})^T I_w (X\vec{\theta} - \vec{y})$$

$$I_w = \begin{pmatrix} w_1 & 0 & \dots & 0 \\ 0 & w_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & w_n \end{pmatrix}$$

Chceme gradient rovný nule (alebo použiť gradient descent):

$$\nabla_{\vec{\theta}} E = 2X^T I_w (X\vec{\theta} - \vec{y}) = \vec{0}$$

Riešenie:

$$X^T I_w X \vec{\theta} = X^T I_w \vec{y}$$

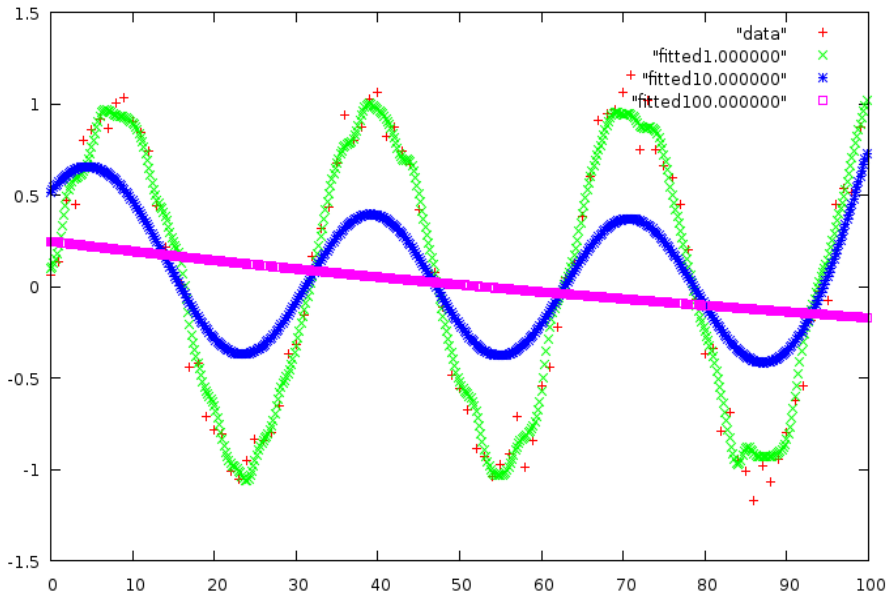
Lokálne váhovaná lineárna regresia

Chceme predikovať hodnotu v $\vec{x}_{new}$.

Položíme: $w_i = e^{\frac{-\|\vec{x}_{new} - \vec{x}^{(i)}\|^2}{\sigma^2}}$ (σ je hyperparameter, ktorý treba dobre nastaviť).

Zbehneme váhovanú lineárnu regresiu a odpredikujeme hodnotu (áno naozaj robíme nové tréovanie regresie pre každý nový výstup).

LVLR výsledok



Polynomiálna regresia

Jeden vstup x , model (do stupňa 2):

$$y = \theta_0 + \theta_1 x + \theta_2 x^2$$

Polynomiálna regresia

Jeden vstup x , model (do stupňa 2):

$$y = \theta_0 + \theta_1 x + \theta_2 x^2$$

Dva vstupy x_1, x_2 , model (do stupňa 2):

$$y = \theta_0 + \theta_{10}x_1 + \theta_{01}x_2 + \theta_{11}x_1x_2 + \theta_{20}x_1^2 + \theta_{02}x_2^2$$

Môžeme zopakovať postup z minula a určiť hodnoty θ . Alebo redukovať na lineárnu regresiu.

Redukcia polynomiálnej regresie

Pre dva vstupy

Vstup: $(1, x_1, x_2)$ zmeníme na vstup:

$(1, x_1, x_2, x_1x_2, x_1^2, x_2^2)$

A môžeme riešiť lineárnu regresiu (očakávané výstupy nemeníme).

Redukcia polynomiálnej regresie

Pre dva vstupy

Vstup: $(1, x_1, x_2)$ zmeníme na vstup:

$$(1, x_1, x_2, x_1 x_2, x_1^2, x_2^2)$$

A môžeme riešiť lineárnu regresiu (očakávané výstupy nemeníme).

Všeobecne

Máme p báзовých funkcií: $\phi_1(\vec{x}), \phi_2(\vec{x}), \dots, \phi_p(\vec{x})$, kde $\phi_i \in \mathbb{R}^m \rightarrow \mathbb{R}$.

Vstupnú maticu X predspracujeme na maticu Φ :

$$\begin{pmatrix} \phi_1(\vec{x}^{(1)}) & \phi_2(\vec{x}^{(1)}) & \dots & \phi_p(\vec{x}^{(1)}) \\ \phi_1(\vec{x}^{(2)}) & \phi_2(\vec{x}^{(2)}) & \dots & \phi_p(\vec{x}^{(2)}) \\ & & \vdots & \\ \phi_1(\vec{x}^{(n)}) & \phi_2(\vec{x}^{(n)}) & \dots & \phi_p(\vec{x}^{(n)}) \end{pmatrix}$$

A riešime lineárnu regresiu, napríklad sústavu: $\Phi^T \Phi \vec{\theta} = \Phi^T \vec{y}$

Bázové funkcie - príklady

Nemusia byť len polynómy.

- $\phi(\vec{x}) = x_4 x_7$, $\phi(\vec{x}) = x_2$
- 0-1 funkcie: $\phi(\vec{x}) = x_6 > 0$
- Rôzne predspracovania: $\phi(\vec{x}) = \log(x_5 + 1)$
- Kernelové funkcie: $\phi(\vec{x}) = e^{\frac{-\|\vec{z} - \vec{x}\|^2}{\sigma^2}}$

Lineárna regresia s predspracovaním

