

Všeobecná lineárna regresia

Máme stĺpcové vektory.

Ku každému vstupu pridáme atribút s hodnotou 1.

Máme model:

$$y = \vec{x}^T \cdot \vec{\theta}$$

Všeobecná lineárna regresia

Máme stĺpcové vektory.

Ku každému vstupu pridáme atribút s hodnotou 1.

Máme model:

$$y = \vec{x}^T \cdot \vec{\theta}$$

Vstupy zoradíme do matice po riadkoch a výstupy do stĺpcového vektora:

$$X = \begin{pmatrix} (\vec{x}^{(1)})^T \\ (\vec{x}^{(2)})^T \\ \dots \\ (\vec{x}^{(n)})^T \end{pmatrix}$$

$$\vec{y} = \begin{pmatrix} y^{(1)} \\ y^{(2)} \\ \dots \\ y^{(n)} \end{pmatrix}$$

$$X\vec{\theta} - \vec{y} = \begin{pmatrix} (\vec{x}^{(1)})^T \vec{\theta} - y^{(1)} \\ (\vec{x}^{(2)})^T \vec{\theta} - y^{(2)} \\ \vdots \\ (\vec{x}^{(n)})^T \vec{\theta} - y^{(n)} \end{pmatrix}$$

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

Gradient definícia:

$$\nabla_{\vec{\theta}} E = \left(\frac{\partial E}{\partial \theta_1}, \frac{\partial E}{\partial \theta_2}, \dots, \frac{\partial E}{\partial \theta_n} \right)$$

Vyjadruje smer nahor.

Gradient chyby

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

Jedna zložka gradientu:

$$\frac{\partial E}{\partial \theta_j} = \sum_{i=0}^n 2((\vec{x}^{(i)})^T \vec{\theta} - y^{(i)}) x_j^{(i)}$$

Maticovým násobením sa to všetko zapíše takto:

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

Chceme:

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

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

$$\vec{\theta} = (X^T X)^{-1} X^T \vec{y}$$

Tomuto sa hovorí aj normálne rovnice lineárnej regresie.

```
import numpy as np

X = [[122, 3], [39, 1], [67, 3]]
y = [400000, 76000, 175000]

X = np.hstack([np.array(X, float),
                np.ones(shape=(len(y), 1))])
y = np.array(y, float)

XXi = np.linalg.inv(X.T.dot(X))
theta = XXi.dot(X.T).dot(y)
print theta

print np.linalg.solve(X.T.dot(X), X.T.dot(y))
```

- $X^T X - O(m^2 n)$
- Invertovanie matice / riešenie sústavy – $O(m^3)$.

Numerické metódy - gradient descent

Iterujeme:

$$\vec{\theta} = \vec{\theta} - \alpha \nabla_{\vec{\theta}} E$$

Po dosadení (faktor 2 sa skryje do α):

$$\vec{\theta} = \vec{\theta} - \alpha X^T (X\vec{\theta} - \vec{y})$$

Stochastic gradient descent

Nespočítame chybu zo všetkých príkladov, ale urobíme update po každom príklade (spočítame deriváciu parametrov z jedného príkladu):

- while (mení sa chyba):
 - ▶ for i in range(n):
 - ★ $\theta = \theta - \alpha \vec{x}^{(i)} ((\vec{x}^{(i)})^T \theta - y^{(i)})$

Veľmi často konverguje rýchlejšie ako klasický gradient descent.