Closed-form solution (Normal equation)

Normal Equation

Analitičko pronalaženje rešenja

Pronalaženje ekstrema (minimuma) → izjednačavanje parcijalnih izvoda sa 0

$$\mathrm{J}(\theta) = \frac{1}{2N}\sum_{i=1}^{N}\left(\theta_1 x^{(i)} + \theta_0 - y^{(i)}\right)^2$$

$$\frac{\partial J}{\partial \theta_1} = 0$$

$$\frac{\partial J}{\partial \theta_1} = 0$$

Normal Equation

$$\mathrm{J}(\theta) = \frac{1}{2N}\sum_{i=1}^{N}\left(\theta_1 x^{(i)} + \theta_0 - y^{(i)}\right)^2$$

$$\frac{\partial J}{\partial \theta_d} = \frac{1}{N}\sum_{i=1}^{N}\left(h_\theta\left(x^{(i)}\right) - y^{(i)}\right)x_d^{(i)} = 0$$

$$X = \begin{bmatrix} 1 & x^{(1)} \\ 1 & x^{(2)} \\ \cdots & \cdots \\ 1 & x^{(N)} \end{bmatrix} \quad y = \begin{bmatrix} y^{(1)} \\ y^{(2)} \\ \cdots \\ y^{(N)} \end{bmatrix} \Rightarrow \theta = (X^T X)^{-1} X^T \mathrm{y}$$

Normal Equation

Redudantna obeležja

$N > D$

Šta ako je matrica singularna?

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

$X_{N \times D}$

$(X^T)_{D \times N}$

$(X^T X)_{D \times D}$

kompleksnost inverzije matrice je $O(D^3)$

Normal Equation

- Nema parametara koje moramo birati
- Jedna iteracija
- $(X^T X)^{-1}$ je sporo za veliko D
- $X^T X$ može biti singularna matrica

Gradient descent

- Moramo birati α
- Mnogo iteracija
- Vreme ne zavisi od D