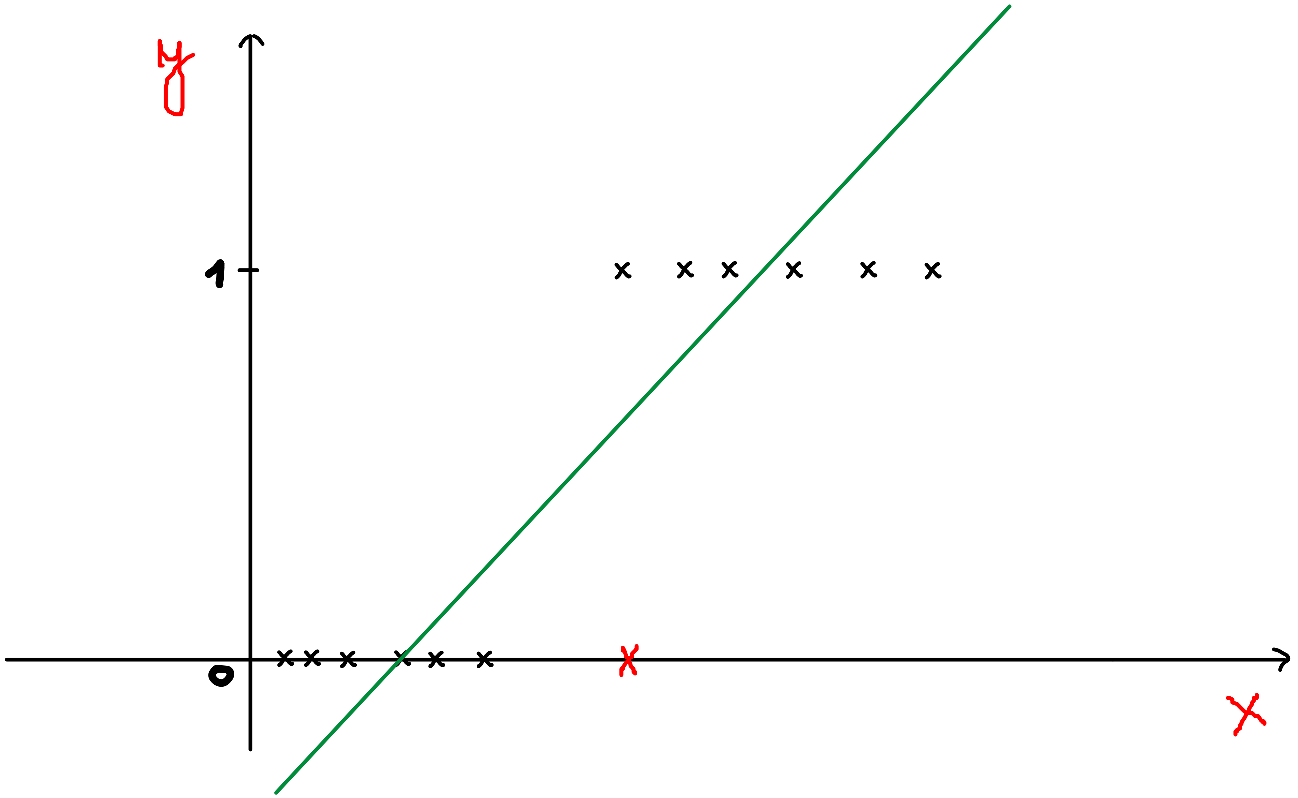
Logistička regresija

Primenjeni algoritmi

Problem klasifikacije

- Primer binarne klasifikacije $y \in \{0,1\}$

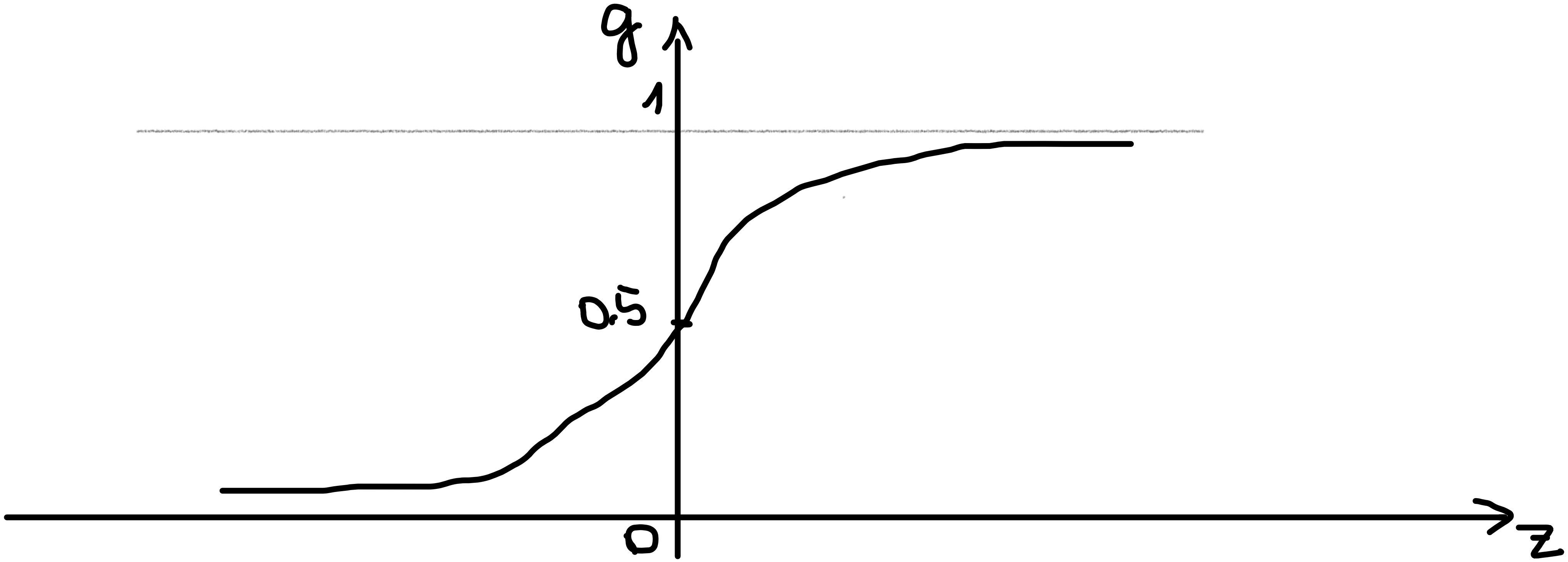


Logistička regresija - hipoteza

- Želja je da hipoteza $h(x) \in [0,1]$

$$h_\theta(x) = g(\theta^T x) = \frac{1}{1 - e^{-\theta^T x}}$$

$$g(z) = \frac{1}{1-e^{-z}}$$



Postupak određivanja parametara

1. Pretpostavlja se probabilistički model

$$P(y = 1 \mid \boldsymbol{x};\ \boldsymbol{\theta}) = h_\theta(\boldsymbol{x})$$

$$P(y = 0 \mid \boldsymbol{x};\ \boldsymbol{\theta}) = 1 - h_\theta(\boldsymbol{x})$$

$$p(y \mid \boldsymbol{x};\ \boldsymbol{\theta}) = (h_\theta(\boldsymbol{x}))^y (1 - h_\theta(\boldsymbol{x}))^{1-y}, \qquad y \in [0,1]$$

2. Primenjuje se princip maksimalne verodostojnosti za ceo obučavajući skup

$$\mathcal{L}(\theta) = p(\boldsymbol{y} \mid X;\ \boldsymbol{\theta}) = \prod_{i=1}^{m} p\left(y^{(i)} \middle| \boldsymbol{x}^{(i)};\ \boldsymbol{\theta}\right) = \prod_{i=1}^{m} \left(h_\theta\left(\boldsymbol{x}^{(i)}\right)\right)^{y^{(i)}} \left(1 - h_\theta\left(\boldsymbol{x}^{(i)}\right)\right)^{1-y^{(i)}}$$

Pojednostavljenje: Maksimizacija log verodostojnosti

$$\mathcal{L}(\theta) = \prod_{i=1}^{m} \left(h_\theta\left(\boldsymbol{x}^{(i)}\right)\right)^{y^{(i)}} \left(1 - h_\theta\left(\boldsymbol{x}^{(i)}\right)\right)^{1-y^{(i)}}$$

$$\ell(\theta) = \ln\left(\mathcal{L}(\theta)\right)$$

$$= \sum_{i=1}^{m} \left(y^{(i)} \ln\left(h_\theta\left(\boldsymbol{x}^{(i)}\right)\right) + \left(1 - y^{(i)}\right) \ln\left(1 - h_\theta\left(\boldsymbol{x}^{(i)}\right)\right)\right)$$

$$\max \ell(\theta) \qquad \Rightarrow \frac{\partial \ell(\theta)}{\partial \theta_j} = 0, \qquad i = 1, \dots, n$$

$$\frac{\partial \ell(\theta)}{\partial \theta_j} = \sum_{i=1}^{m} \left(y^{(i)} - h_\theta\left(\boldsymbol{x}^{(i)}\right)\right) x_j^{(i)} = 0\,, \qquad i = 1, \dots, n$$

Dobijanje parametara

- Primenom Paketnog gradijentnog postupka dobije se da je

$$\frac{\partial \ell(\theta)}{\partial \theta_j}$$

$$\theta_j = \theta_j + \alpha \underbrace{\sum_{i=1}^{m} \left(y^{(i)} - h_\theta\left(x^{(i)}\right)\right) x_j^{(i)}} \qquad i = 1, \dots, n$$

- Jedna od numeričkih metoda za određivanje parametara je i Njutn-Rapsonov postupak

Podsetnik: Numeričko rešavanje sistema nelinearnih algebarskih jednačina

Problem:

$$f_1(x_1, x_2, \dots, x_n) = 0$$
$$f_2(x_1, x_2, \dots, x_n) = 0$$
$$\dots.$$
$$f_n(x_1, x_2, \dots, x_n) = 0$$

Vektorki zapis

$$\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{0}$$

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix}$$

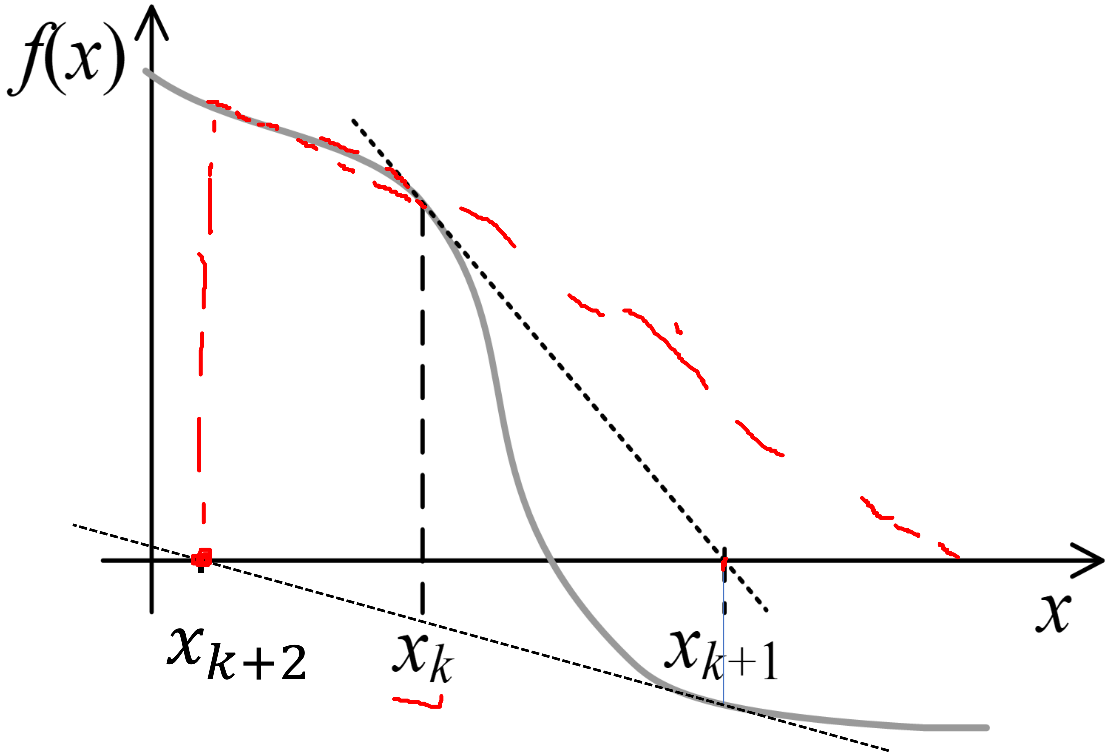
Primenjuje se za određivanje $\max \ell(\theta)$ $\quad \Rightarrow \dfrac{\partial \ell(\theta)}{\partial \theta_j} = 0, \quad i = 1, \dots, n$

f

Njutn-Rapsonov postupak – grafička predstava

Pojednostavljen problem jedne promenljive $f(x) = 0$

- u tekućoj tački x_k se odredi tangenta na krivu f i nova tačka x_{k+1} se nalazi na preseku tangente i x-ose.



Njutn-Rapsonov postupak

$$f'(x_k) = \frac{f(x_k)}{\Delta x_k}$$

Njutnov metod (poznat i kao Njutn-Rapsonov postupak):

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}, \qquad k = 0,1,\ldots$$

$$\Delta x_k = \frac{f(x_k)}{f'(x_k)}$$

$$x_{k+1} + \Delta x_k = x_k$$

$$x_{k+1} = x_k - \Delta x_k$$

- Za razmatrani problem

$$\ell'(\theta) = 0$$

$$\theta^{(k+1)} = \theta^{(k)} - \frac{\ell'(\theta^{(k)})}{\ell''(\theta^{(k)})}, \qquad k = 0,1,\ldots \;\; ; \quad k - iteracija$$

Njutn-Rapsonov postupak

- Generalizacija Njutnovog metoda za višedimenzioni problem:

$$\frac{\partial \ell(\theta)}{\partial \theta_j} = \underline{\sum_{i=1}^{m} \left(y^{(i)} - h_\theta\left(\boldsymbol{x}^{(\boldsymbol{i})}\right) \right) x_j^{(i)} = 0}\,, \qquad i = 1, \dots, n$$

- Za razmatrani problem

$$\theta^{(k+1)} = \theta^{(k)} - H^{-1} \nabla_\theta \ell'(\theta^{(k)}), \qquad k = 0,1,\dots \;\; ; \quad k - iteracija$$

Gde je

- $\nabla_\theta \ell'(\theta^{(k)})$ – vektor parcijalnih izvoda
- H Hesijanova matrica sa elementima $H_{ij} = \dfrac{\partial^2 \ell(\theta)}{\partial \theta_i \partial \theta_j}$