

# Iterativne numerične metode v posplošenih linearnih modelih

Mitja Mandić

Mentor: izred. prof. dr. Jaka Smrekar

20. november 2020

# Posplošeni linearni modeli

- Slučajni del, sistematični del, povezovalna funkcija

# Posplošeni linearni modeli

- Slučajni del, sistematični del, povezovalna funkcija
- Linearna regresija:

$$Y = x^T \beta$$

- Problem - ni najboljša. Rešitev?  
Transformacija Y

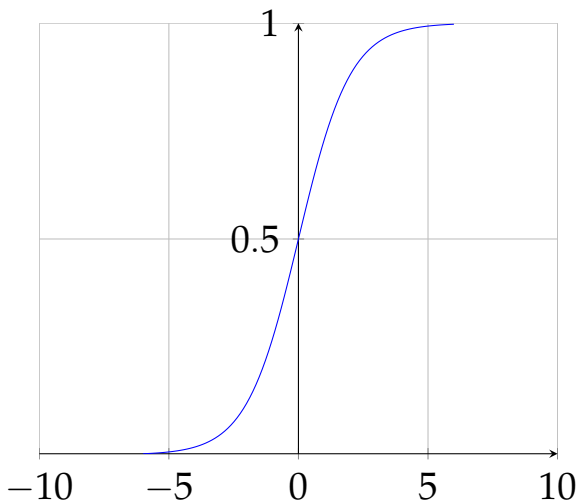
# Logistični model

- Za kategorične podatke  $\rightarrow$  binomska porazdelitev

# Logistični model

- Za kategorične podatke  $\rightarrow$  binomska porazdelitev
- $\text{logit}(p_i) = \log\left(\frac{p_i}{1-p_i}\right) = x^T \beta$

$$p = \frac{e^{x^T \beta}}{1 + e^{x^T \beta}}$$



# Numerične metode

- Za ocenjevanje parametrov  $\beta$  običajno rešujemo sistem enačb največjega verjetja

# Numerične metode

- Za ocenjevanje parametrov  $\beta$  običajno rešujemo sistem enačb največjega verjetja
- Newtonova metoda še vedno zelo aktualna:

$$x_{i+1} = x_i - \frac{f'(x_i)}{f''(x_i)}$$

- Izboljšava: Fisher-scoring



# Fisher scoring

$$\beta_{i+1} = \beta_i + \frac{\dot{l}(\beta_i)}{E(\ddot{l}(\beta_i))}$$

- za logistično regresijo sovpadata z Newtonovo metodo
- Informacijska matrika je pozitivno definitna  
→ imamo naraščajoč algoritem

# Fisher scoring za logistični model

$$\begin{aligned}\hat{\beta}_{i+1} &= \hat{\beta}_i + (X^T v(\hat{\beta}_i) X)^{-1} X^T (y - \mu(\hat{\beta}_i)) \\ &= \hat{\beta}_i + (\text{inverz info})(\text{score})\end{aligned}$$

# Fisher scoring za logistični model

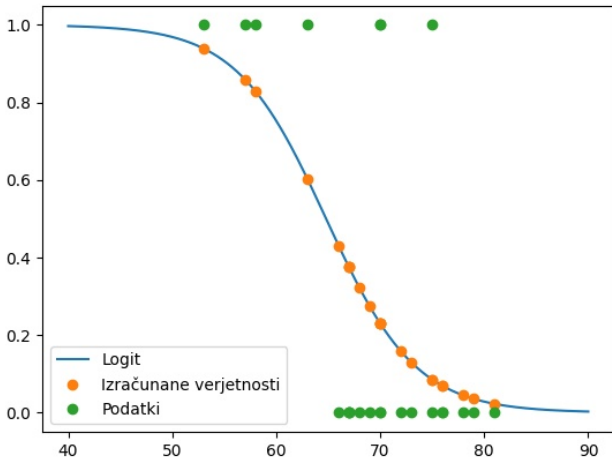
$$\begin{aligned}\hat{\beta}_{i+1} &= \hat{\beta}_i + (X^T v(\hat{\beta}_i) X)^{-1} X^T (y - \mu(\hat{\beta}_i)) \\ &= \hat{\beta}_i + (\text{inverz info})(\text{score})\end{aligned}$$

Računanje inverza je lahko problematično. To rešimo takole:

$$\begin{aligned}h &= \hat{\beta}_{i+1} - \hat{\beta}_i \\ X^T v(\hat{\beta}_i) X &= h * X^T (y - \mu(\hat{\beta}_i))\end{aligned}$$

# Challenger podatki

$$p = \frac{e^{15.04290 - 0.23216x}}{1 + e^{15.04290 - 0.23216x}}$$



# Kaj še bom naredil

- Analiziral časovno zahtevnost že postavljenega algoritma
- Izvedel podobno še za kak drugačen model
- Teorijo razvil še za splošen posplošen linearen model

# References I