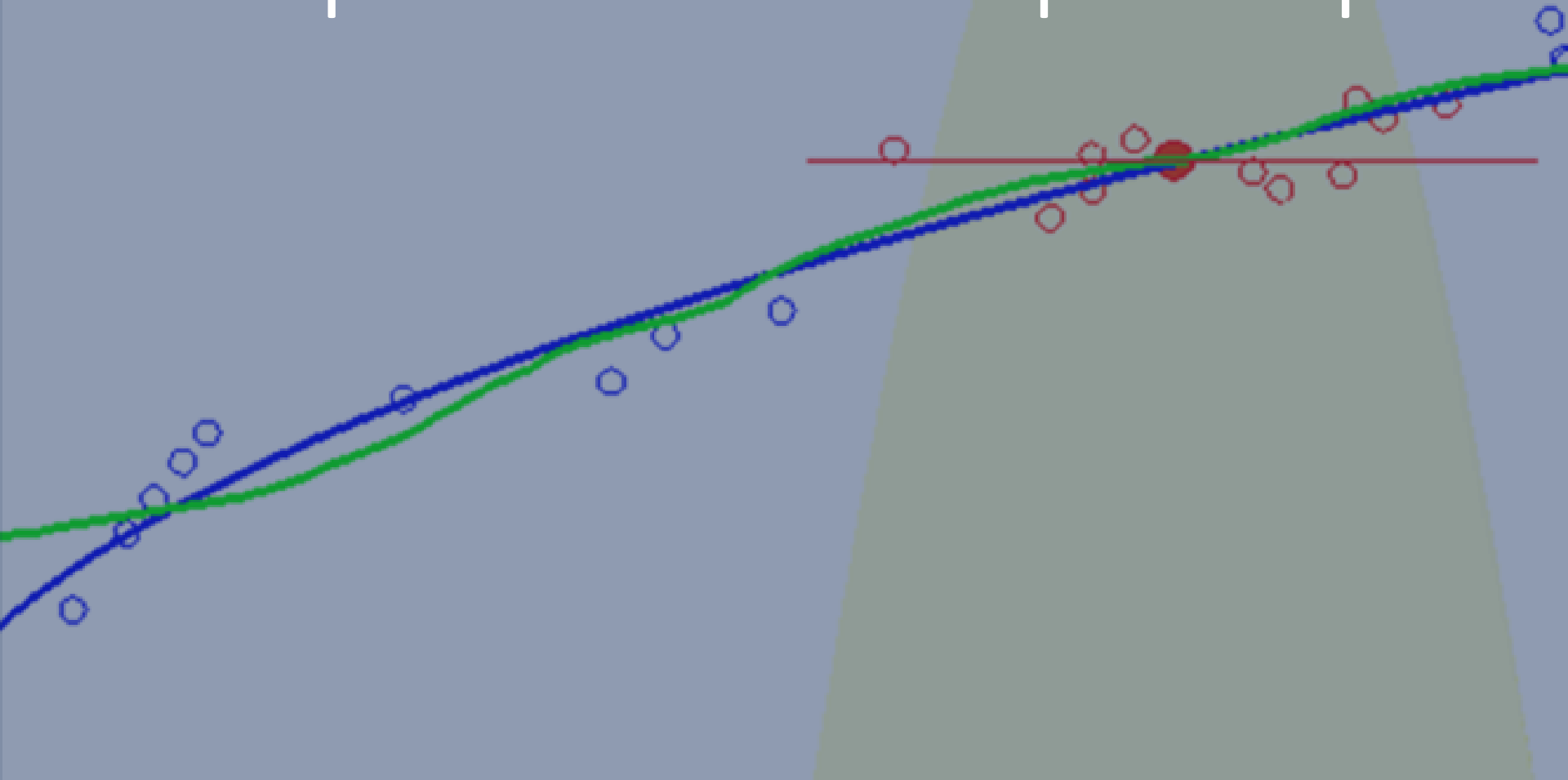


Neparametarski pristup



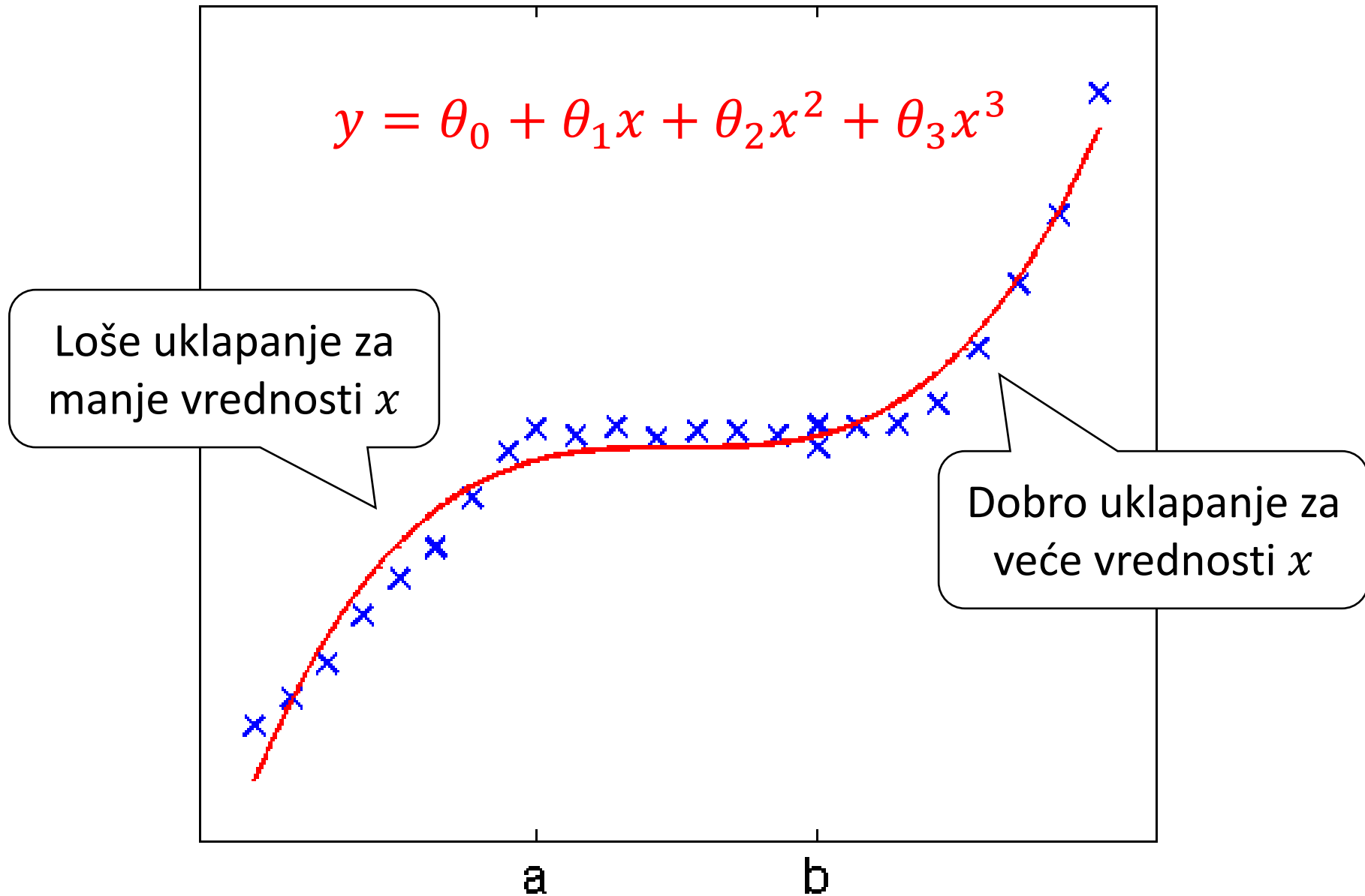
Regresija: parametarski pristup

- Parametarski pristup: $h_{\theta}(x)$ je u potpunosti definisana (konačnim brojem) fiksnih parametara θ
- Videli smo da je izbor obeležja od velike važnosti za dobre preformanse algoritma
- Za fiksno θ postoji granica do koje model može biti fleksibilan

Neparametarski pristup

- Postoje i *neparametarski pristupi*
 - Izbor obeležja nije kritičan problem
 - Jednostavni su za implementaciju
 - Ovi modeli su ekstremno fleksibilni
 - Sa povećanjem podataka se povećava i kompleksnost modela
 - Jednostavan pristup koji je u praksi zapanjujuće teško pobediti
- Međutim, sve ovo zavisi od toga da li imamo dovoljno podataka da koristimo ovakav pristup

Globalno fitovanje modela



Lokalno fitovanje modela

Dozvoljavamo da model ima lokalnu strukturu – dozvoljavamo veću fleksibilnost

$$y = \theta_0 + \theta_1 x + \theta_2 x^2 + \theta_3 x^3$$

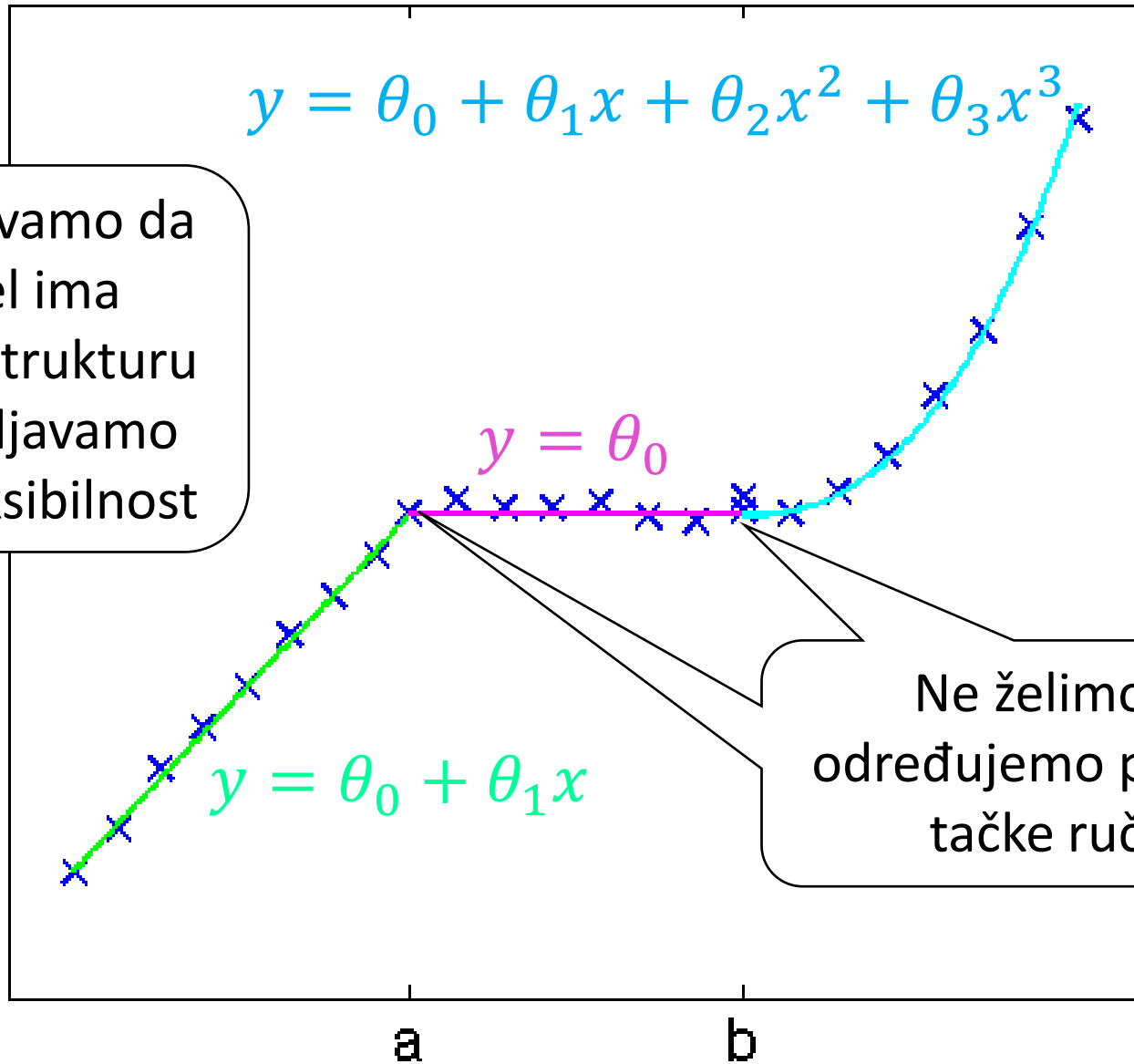
$$y = \theta_0$$

$$y = \theta_0 + \theta_1 x$$

Ne želimo da određujemo prelazne tačke ručno

a

b



Regresiona funkcija

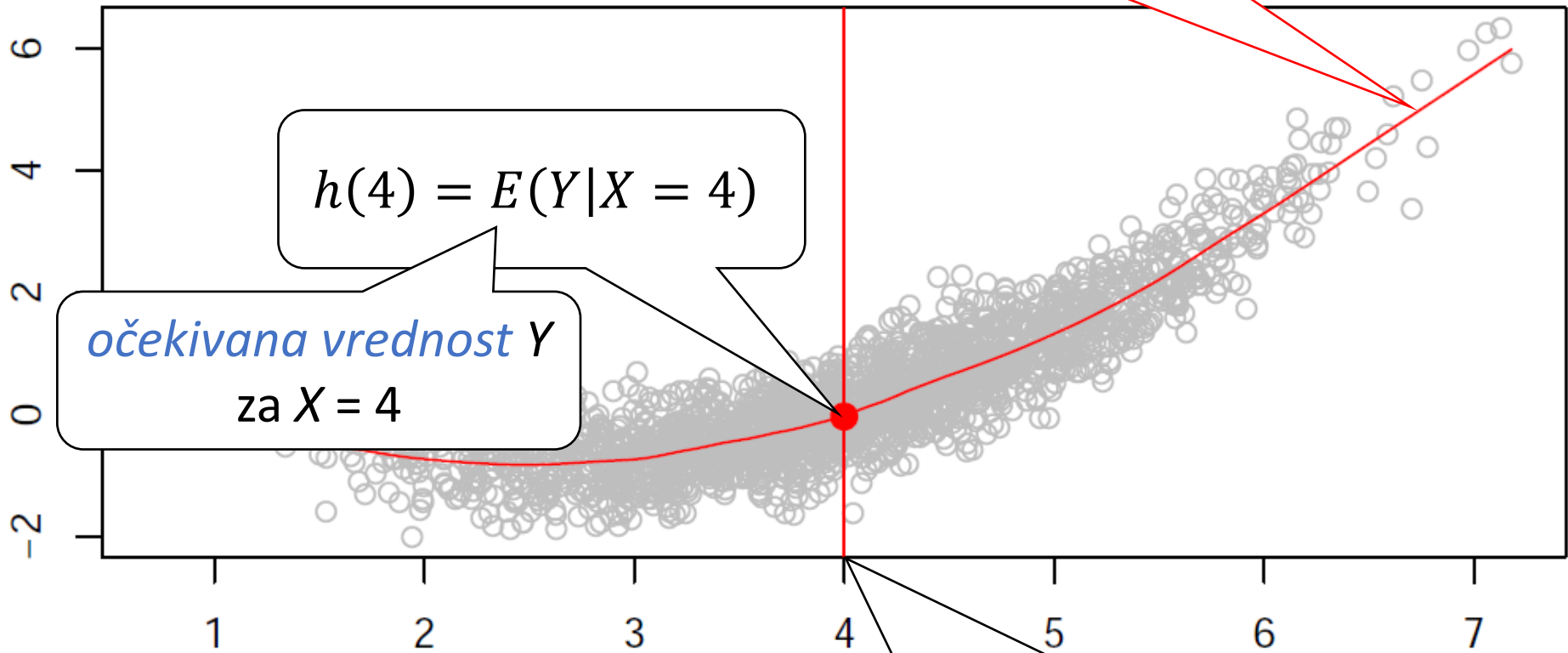
Šta bi bila idealna funkcija $h(x)$?

$h(x) = E(Y|X = x)$
se zove **regresiona funkcija**

$$h(4) = E(Y|X = 4)$$

očekivana vrednost Y
za $X = 4$

Šta bi bila dobra vrednost funkcije u tački 4?



Mana procene $h(x) = E(Y|X = x)$

- Najčešće, za određenu tačku x imamo malo (ili čak 0) opservacija za koje važi $X = x$
- Posledica je da ne možemo izračunati $E(Y|X = x)$
- Zato ćemo malo relaksirati definiciju:

$$\hat{h}(x) = Ave(Y|X \in \mathcal{N}(x))$$

gde \mathcal{N} označava susedstvo (*neighborhood*) tačke x

Najbliži susedi (lokalno uprosečavanje)

Ovo se zove *nearest neighbors* ili *local averaging*

