# Fitovanje polinomijalne krive

- Posmatraćemo probelem fitovanja polinomijalne krive iz probablističke perspektive

- Ovo će nam dati uvid u dizajn algoritma – odabranu funkciju greške i regularizaciju

- Dat je uzorak $x = (x_1, x_2, \dots, x_N)$ i odgovarajuće vrednosti ciljne varijable $y = (y_1, y_2, \dots, y_N)$

# Pretpostavke

- Pretpostavka 1: $y$ i $x$ su povezani na sledeći način:

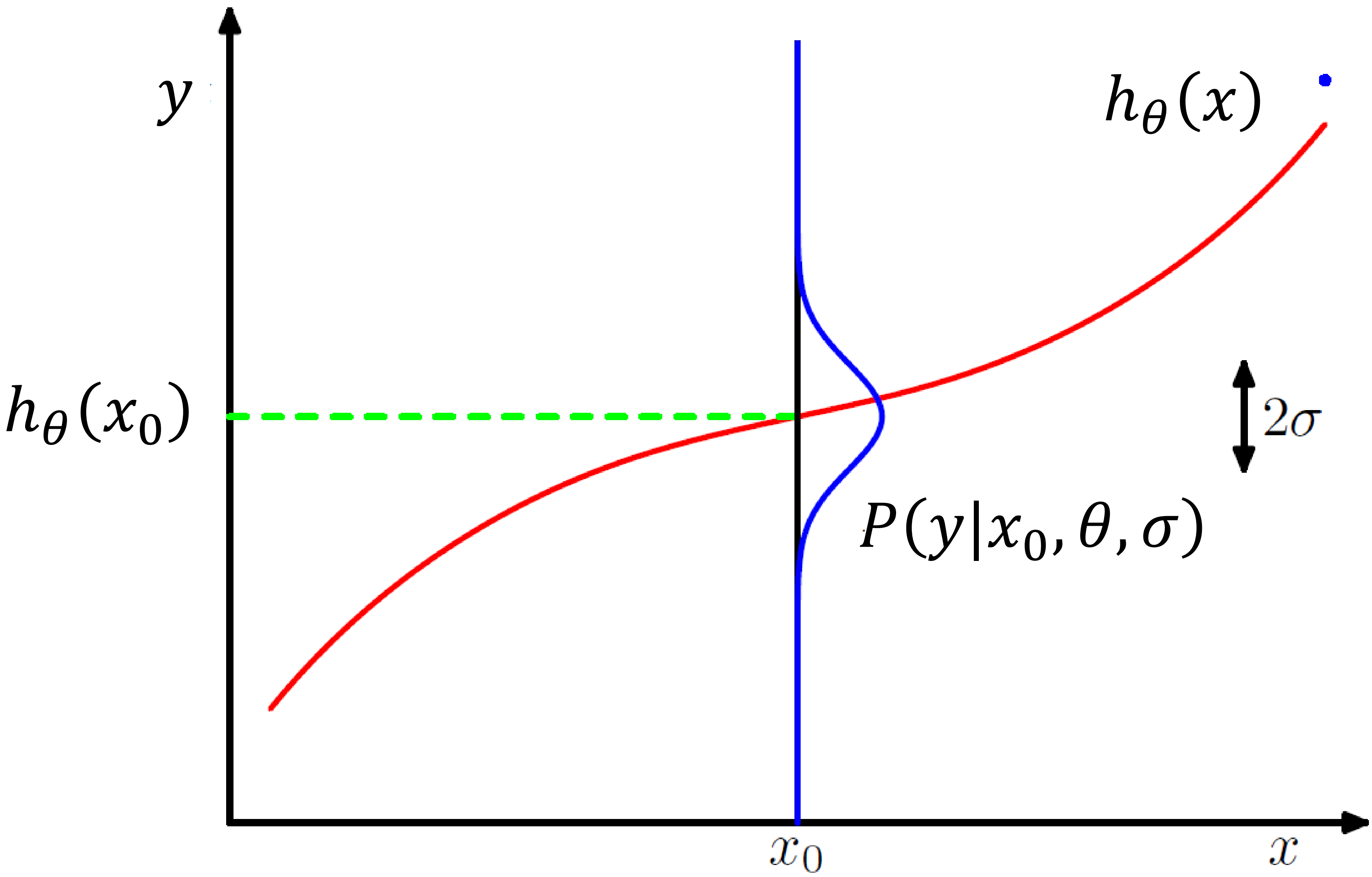
$$y^{(n)} = h_\theta(x^{(n)}) + \varepsilon^{(n)} = \theta^T x^{(n)} + \varepsilon^{(n)}$$

gde $\varepsilon$ predstavlja šum

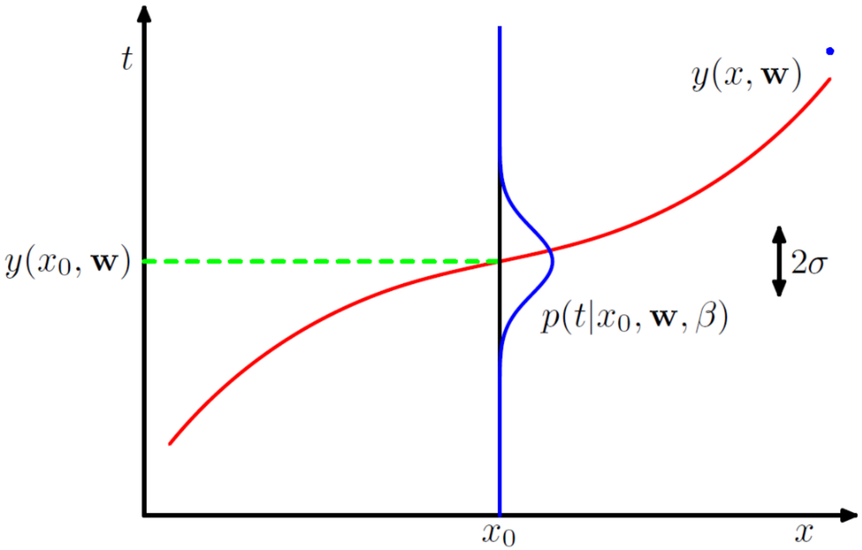
- Pretpostavka 2: $\varepsilon^{(n)}$ su IID po Gausovoj raspodeli sa srednjom vrednošću 0 i nekom varijansom σ$^2$:

$$\varepsilon^{(i)} \sim \mathcal{N}(0, \sigma^2)$$

# Fitovanje polinomijalne krive



# Fitovanje polinomijalne krive




Iskoristićemo trening podatke da odredimo $\theta$

$$P\left(\varepsilon^{(i)}\right) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{\left(\varepsilon^{i}\right)^2}{2\sigma^2}\right\}$$

$$P\left(y^{(i)} | x^{(i)}; \theta\right) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{\left(y^{(i)} - \theta^T x^{(i)}\right)^2}{2\sigma^2}\right\}$$

# Fitovanje polinomijalne krive

- Metod maksimalne verodostojnosti: odabrati $\theta$ tako da su uočeni podaci najverovatniji – $\theta$ koje maksimizuje $L(\theta)$

$$L(\theta) = \prod_{i=1}^{N} p\left(y^{(i)} | x^{(i)}; \theta\right) =$$

$$= \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\left(y^{(i)} - \theta^T x^{(i)}\right)^2}{2\sigma^2}\right)$$

# Fitovanje polinomijalne krive

$$L(\theta) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\left(y^{(i)} - \theta^T x^{(i)}\right)^2}{2\sigma^2}\right)$$

$$\ln L(\theta) = -\frac{1}{2\sigma^2} \sum_{i=1}^{N} (y^{(i)} - \theta^T x^{(i)})^2 - \frac{N}{2} \ln \sigma^2 - \frac{N}{2} \ln(2\pi)$$

$$\frac{\partial L(\theta, \sigma)}{\partial \theta}$$

# $\partial L(\theta, \sigma)/\partial\theta$

$$\ln L(\theta) = -\frac{1}{2\sigma^2}\sum_{i=1}^{N}(y^{(i)} - \theta^T x^{(i)})^2 - \frac{N}{2}\ln\sigma^2 - \frac{N}{2}\ln(2\pi)$$

Zamenićemo sa $\frac{1}{2}$ (skaliranje ne utiče na poziciju minimuma)

$=0$ (ne zavisi od $\theta$)

- Umesto maksimuma logaritma, tražićemo minimum negativnog logaritma:

$$\frac{1}{2}\sum_{i=1}^{N}\left(y^{(i)} - \theta^T x^{(i)}\right)^2$$

Metod najmanjih kvadrata

# Fitovanje polinomijalne krive

- Pod uvedenim probabilističkim pretpostavkama o podacima, metod najmanjih kvadrata odgovara pronalaženju ocene maksimalne verodostojnosti $\theta$

- Primetite da $\theta_{ML}$ nije zavisilo od $\sigma$