ZADATAK 1

$$E(w|D) = \frac{1}{2}\sum_{i=1}^{N}\left(y^{(i)} - \phi(x^{(i)})^T w\right)^2$$

istovjetno s minim. log-izglednosti

$$\ln \mathcal{L}(\vec{w}|D) = \ln p(D|\vec{w}) \overset{iid}{=} \ln \prod_{i=1}^{n} p(x^{(i)}, y^{(i)}) = \ln \prod_{i=1}^{n} p(x^{(i)}|y^{(i)})\, p(x^{(i)})$$

$$= \ln \prod_{i=1}^{n} p(x^{(i)}|y^{(i)}) + \underbrace{\cancel{\ln \prod_{i=1}^{n} p(x^{(i)})}}_{\text{konst.}} =$$

$$= \ln \mathcal{N}(h(x^{(i)}|\vec{w}), \sigma^2) =$$

$$= \ln \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{ - \frac{(y^{(i)} - h(x^{(i)}|\vec{w}))^2}{2\sigma^2} \right\}$$

$$= \underbrace{\cancel{n \cdot \ln \frac{1}{\sqrt{2\pi}\sigma}}}_{\text{konst.}} - \sum_{i=1}^{n} \frac{(y^{(i)} - h(x^{(i)}|\vec{w}))^2}{2\underbrace{\cancel{\sigma}^2}_{\text{konst.}}} =$$

$$= -\frac{1}{2}\sum_{i=1}^{n}\left(y^{(i)} - h(x^{(i)}|\vec{w})\right)^2$$

$\Rightarrow$ minimizacija $E(w|D)$ je istovjetna min. neg. log. izglednosti

ZADATAK 2

a) $h(x|w) = \sum_{j=0}^{m} w_j \phi_j(x) = \vec{w}^T \phi(x)$

DIZAJN MARICA

$$\Phi = \begin{pmatrix} 1 & \phi_1(x^{(1)}) & \dots & \phi_m(x^{(1)}) \\ \vdots & \vdots & & \vdots \\ 1 & \phi_1(x^{(N)}) & & \phi_m(x^{(N)}) \end{pmatrix} = \begin{pmatrix} \phi(x^{(1)})^T \\ \vdots \\ \phi(x^{(N)})^T \end{pmatrix}$$

vektor izlaznih vrijednosti

$$y = \begin{pmatrix} y^{(1)} \\ \vdots \\ y^{(N)} \end{pmatrix}$$

$\rightarrow$ $\Phi \vec{w} = y$

$$\vec{w} = \Phi^{-1} y$$

$$E(\vec{w}|D) = \frac{1}{2}\sum_{i=1}^{N}\left(\vec{w}^T \phi(x^{(i)}) - y^{(i)}\right)^2$$

$$= \frac{1}{2}(\Phi\vec{w} - \vec{y})^T(\Phi\vec{w} - \vec{y})$$

$$= \frac{1}{2}\left(\vec{w}^T \Phi^T \Phi \vec{w} - \vec{w}^T \Phi^T \vec{y} - \vec{y}^T \Phi \vec{w} + \vec{y}^T \vec{y}\right)$$

$$= \frac{1}{2}\left(\vec{w}^T \Phi^T \Phi \vec{w} - 2 y^T \Phi \vec{w} + \vec{y}^T \vec{y}\right) \quad / \frac{\partial E}{\partial \vec{w}}$$

$$\frac{1}{2}\left(\vec{w}^T \left(\Phi^T \Phi + (\Phi^T \Phi)^T\right) - 2 y^T \Phi\right) = 0$$

$$\vec{w}^T \Phi^T \Phi - \vec{y}^T \Phi = 0 \quad /^T$$

$$\Phi^T \Phi \vec{w} - \Phi^T \vec{y} = 0 \quad / \cdot (\Phi^T \Phi)^{-1}$$

$$\vec{w} = (\Phi^T \Phi)^{-1} \Phi^T \vec{y}$$

$$\vec{w} = \Phi^+ \vec{y}$$

↳ pseudoinverz

b) L2 - regularizirani model

$$E'(\vec{w}|D) = \frac{1}{2} \sum_{i=1}^{N} (\vec{w}^T \phi(x^{(i)}) - y^{(i)})^2 + \frac{\lambda}{2} \|\vec{w}\|^2$$

$$= \frac{1}{2}(\Phi \vec{w} - \vec{y})^T (\Phi \vec{w} - \vec{y}) + \frac{\lambda}{2} \vec{w}^T \vec{w}$$

$$= \frac{1}{2}\left(\vec{w}^T \Phi^T \Phi \vec{w} - \vec{w}^T \Phi^T \vec{y} - \vec{y}^T \Phi \vec{w} + \vec{y}^T \vec{y} + \lambda \vec{w}^T \vec{w}\right)$$

$$= \frac{1}{2}\left(\vec{w}^T \Phi^T \Phi \vec{w} - 2 \vec{y}^T \Phi^T \vec{w} + \vec{y}^T \vec{y} + \lambda \vec{w}^T \vec{w}\right)$$

$$\frac{\partial E}{\partial \vec{w}} = \Phi^T \Phi \vec{w} - \Phi^T \vec{y} + \lambda \vec{w} = 0$$

$$(\Phi^T \Phi + I\lambda)\vec{w} - \Phi^T \vec{y} = 0$$

$$(\Phi^T \Phi + I\lambda)\vec{w} = \Phi^T \vec{y}$$

$$\vec{w} = (\Phi^T \Phi + \lambda I)^{-1} \Phi^T \vec{y}$$

c) $D = \{(x^{(i)}, y^{(i)})\}_{i=1}^{4} = \{(0,4), (1,1), (2,2), (4,5)\}$ $\quad \mathcal{X} = \mathbb{R}$

$h(x) = w_0 + w_1 x + w_2 x^2$ uz $\lambda = 10$

$$\Phi = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 4 & 16 \end{bmatrix} \qquad y = \begin{bmatrix} 4 \\ 1 \\ 2 \\ 5 \end{bmatrix}$$

$$\vec{w} = \left(\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 4 & 16 \end{bmatrix}^T \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 4 & 6 \end{bmatrix} + 10 \begin{bmatrix} 1 & 0 & \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 4 & 6 \end{bmatrix}^T \begin{bmatrix} 4 \\ 1 \\ 2 \\ 5 \end{bmatrix}$$

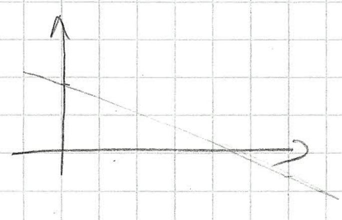
d) velik broj značajki, jer o značajkama ovise dim. matrice koju trebamo invertirati

3. $f(x) = \sin(\pi x)$

$D = \{(0.25, 0.707), (0.5, 1), (1, 0), (1.5, -1), (2, 0)\}$

$\phi(x) = (1, x)$

a) $\Phi_a = \begin{bmatrix} 1 & 0.25 \\ 1 & 0.5 \\ 1 & 1 \\ 1 & 1.5 \\ 1 & 2 \end{bmatrix}$ $\vec{y}_a = \begin{bmatrix} 0.707 \\ 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}$



$f_1(x) = 0.943 - 0.764x$

$\vec{w} = (\Phi^T \Phi)^{-1} \Phi^T \vec{y}$

$= \begin{bmatrix} 5 & 5.25 \\ 5.25 & 7.5625 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0.25 & 0.5 & 1 & 1.5 & 2 \end{bmatrix} \begin{bmatrix} 0.707 \\ 1 \\ 0 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.943292 \\ -0.763708 \end{bmatrix}$

b) $\phi(x) = (1 \quad x \quad x^2)$

$\Phi_b = \begin{bmatrix} 1 & 0.25 & 0.0625 \\ 1 & 0.5 & 0.25 \\ 1 & 1 & 1 \\ 1 & 1.5 & 2.25 \\ 1 & 2 & 4 \end{bmatrix}$ $\vec{y}_b = \vec{y}_a$

$\vec{w} = [1.75354 \quad -2.94026 \quad 0.97529]^T$

$f_2(x) = 1.753 - 2.94x + 0.975x^2$

c) $\phi(x) = (1, x, x^2, x^3, x^4)$

L2-regularizacija

$\lambda = 1$

$\Phi_c = \begin{bmatrix} 1 & 0.25 & 0.0625 & 1/64 & \frac{1}{256} \\ 1 & 0.5 & 0.25 & 0.125 & 0.0625 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1.5 & 2.25 & 3.375 & 5.0625 \\ 1 & 2 & 4 & 8 & 16 \end{bmatrix}$ $\vec{y}_c = \vec{y}_a$

$\vec{w} = (\Phi^T \Phi + \lambda I)^{-1} \Phi^T \vec{y}$

$\vec{w} = \begin{bmatrix} 0.0894 \\ -0.3480 \\ -0.3199 \\ 0.2136 \\ 0.5143 \end{bmatrix}$

d) Drugi model ? greška je dovoljno mala, a model jednostavniji

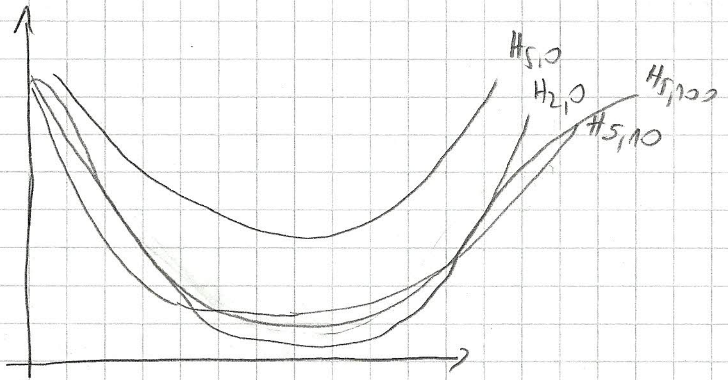
ZADATAK 4

$H_{d,\lambda}$ je model polinomijalne reg. stupnja d s L2-reg. f. λ

$H_{2,0}$ $H_{5,0}$ $H_{100,0}$ $H_{5,1000}$ $X = \mathbb{R}$

- podaci $d=3$

a) SKICA REG. F. $h(x)$



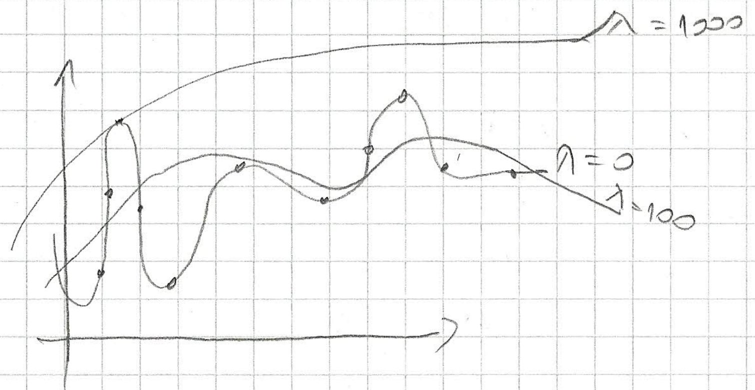
b) emp. pog. i pog. gen.

ZADATAK 5

a) Svrha je ograničiti vrijednost parametara (tj. složenost modela)

b) Prednost -> manje parametara

c)



ZADATAK 6

$\phi: \mathbb{R}^n \to \mathbb{R}^m$

$n > m$, L2 - regularizacija

1) nebitne značajke će imati male parametre odnosno parametre pritegnute na nulu
2) velike razlike između parametara

Neregularizirani model bio bi prenaučen, tj. previše bi se prilagodio primjerima

ZADATAK 7

a) $p(D|\vec{w}) \cdot p(\vec{w}) = \prod_{i=1}^{N} \mathcal{N}(h(x^{(i)}|\vec{w}), \beta^{-1}) \cdot \mathcal{N}(0, \alpha^{-1}I)$

b) $p(\vec{w}) = \mathcal{N}(0, \alpha^{-1}I) = \prod_{i=1}^{m} \frac{1}{\sqrt{2\pi} \cdot |(\alpha^{-1}I)|^{1/2}} \cdot \exp\left(-\frac{\vec{w_i}^2}{2 \cdot \alpha^{-1} I}\right)$

$= \left(\frac{\alpha}{2\pi}\right)^{m/2} \cdot \exp\left(-\left(\frac{\vec{w}^T \vec{w} \cdot \alpha}{2}\right)\right)$

c) $\underset{\vec{w}}{\operatorname{argmax}} (p(D|\vec{w}) \cdot p(\vec{w}))$

$= \ln(p(D|\vec{w})) + \ln p(\vec{w})$

$= \cancel{\ln\left(\frac{\beta}{2\pi}\right)^{N/2}}^{\text{konst.}} - \sum_{i=1}^{N} \frac{(y^{(i)} - h(x^{(i)}|\vec{w}), \beta)^2}{2} + \cancel{\ln\left(\frac{\alpha}{2\pi}\right)^{m/2}}^{\text{konst.}}$

$- \frac{\vec{w}^T \vec{w}\, \alpha}{2} = \beta \sum_{i=1}^{N} \frac{(y^{(i)} - h(x^{(i)}|\vec{w}))^2}{2} - \frac{\alpha}{2} \vec{w}^T \vec{w}$

$= \frac{\beta}{2} \sum_{i=1}^{N} (y^{(i)} - h(x^{(i)}|\vec{w}))^2 - \frac{\alpha}{2} \vec{w}^T \vec{w}$