0036542159 Bruno Marković

VOL

2.2 Ulazni prostor $X = \{0,1\}^3$ ima 8 točaka jer je to trodimenzionalni prostor koji je binaran.
Pomoću parametara $\theta_{11}, \theta_{12} \ldots$ određujemo intervale unutar kojih x_1 x_2 x_3 moraju biti da bi vratio 1

Fora je što su x_i binarni što znači da svaki interval odlučuje oće li prihvatiti

a) $x_i = 0$
b) $x_i = 1$
c) $x_i = 0$ ili $x_i = 1$

Tako dobijamo da je ukupni broj $|H| = 3 \cdot 3 \cdot 3 = 27$
Naravno ne smijemo zaboravit na poseban slučaj kada se ništa ne prihvaća i tako dobivamo $|H| = 27 + 1 = \boxed{28}$

Točan odgovor je D $\boxed{28}$

VOZ

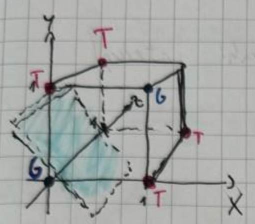
Strojno učenje

DZ

0036542159 Bruno Marković

2.8

$D = \{((0,0,0),0), ((1,1,0),0), ((1,0,0),1), ((1,0,1),1), ((0,1,0),1), ((0,1,1),1)\}$



Razmatramo LINEARNI MODEL

LINEARNI MODEL je pre jednostavan

Iz slike vidimo da ne postoji ravnina koja može točno klasificirati sve primjere, također ne postoji ravnina koja može krivo klasificirati sve primjere. Postoji ravnina koja može imati samo jednu grešku odnosno samo jedno točno i to specifično lažno pozitivan primjer kod opcije s greškom.

Takva ravnina je nacrtana i obojana plavim markerom (napomena ona prolazi ispod točke (T(0,1,0))

Znamo da je kazna za lažno pozitivan primjer 0.5 a lažno negativan 1

Najbolji slučaj: sve iznad ravnine točno a sve ispod krivo => biti će samo jedna greška, to lažno pozitivan primjer => Empirijska greška =>

$$E(h|D) = \frac{1}{6}(0+0+0+0+0+0{,}5) = \frac{1}{6} \cdot \frac{1}{2} = \frac{1}{12}$$

Najgori slučaj: Obrnemo da je iznad ravnine krivo a ispod točno => biti će 1 lažno pozitivan, 4 lažno negativna, i 1 točno određen

$$E(h|D) = \frac{1}{6}(1+1+1+1+0+0.5) = \frac{1}{6} \cdot 4.5 = \frac{4.5}{6} = \boxed{\frac{3}{4}} = 0.75$$

Rješenje je c $\boxed{\frac{1}{12} \le E(h|D) \le \frac{3}{4}}$

P 03

00365 42159 Bruno Markovi

2.3

$x = 0 \; 2 \; 3 \; 5$

$y = 0 \; 0 \; 2 \; 2$

$\bar{x} = \frac{0+2+3+5}{4} = 2.5 \qquad N=4$

$\bar{y} = \frac{0+0+2+2}{4} = 1$

$w_0 = \bar{y} - w_1 \bar{x}$

$w_1 = \frac{\sum_i^N x^i y^i - N \bar{x} \bar{y}}{\sum_i^N (x^i)^2 - N \bar{x}^2}$

$N \cdot \bar{x} \cdot \bar{y} = 4 \cdot 2.5 \cdot 1 = 10$

$\sum_i^N x^i y^i = 0 \cdot 0 + 0 \cdot 2 + 3 \cdot 2 + 5 \cdot 2 = 16$

$\sum_i^N (x^i)^2 = 0^2 + 2^2 + 3^2 + 5^2 = 38$

$N \cdot \bar{x}^2 = 4 \cdot 2.5^2 = 25$

$w_1 = \frac{16-10}{38-25} = \frac{6}{13}$

$w_0 = \bar{y} - w_1 \bar{x}$

$w_0 = 1 - \frac{6}{13} \cdot 2.5 = \boxed{-\frac{2}{13}}$

$h(x_i) = -\frac{2}{13} + \frac{6}{13} x$

$L(y^i, h(x^i)) = (y^i - h(x^i))^2 \Rightarrow L_1 = \frac{4}{169}$

$L_2 = \left(0 - \frac{10}{13}\right)^2 = \frac{100}{169}$

$L_3 = \left(2 - \left(-\frac{2}{13} + \frac{6}{13} \cdot 3\right)\right)^2 = \frac{100}{169}$

$L_4 = \frac{4}{169}$

$\boxed{L_1 = L_4 < L_2 = L_3 \quad \text{D}}$

-15

0.4615

P04 (0036542159 Bruno Machović)

1.3 a)

$$X = \begin{bmatrix} 1 & 0.25 \\ 1 & 0.5 \\ 1 & 1 \\ 1 & 1.5 \\ 1 & 2 \end{bmatrix} \quad X^T = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0.25 & 0.5 & 1 & 1.5 & 2 \end{bmatrix} \quad y = \begin{bmatrix} 0.707 \\ 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}$$

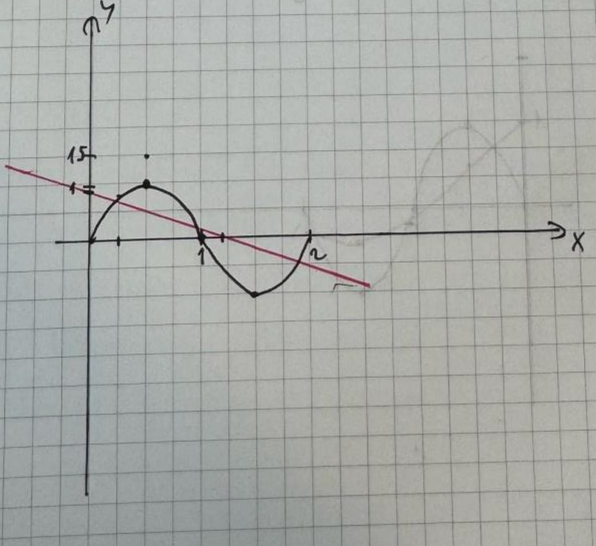
$$w = (X^T \cdot X)^{-1} X^T y$$

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

$$w = \begin{bmatrix} \frac{121}{164} & -\frac{21}{41} \\ -\frac{21}{41} & \frac{20}{41} \end{bmatrix} \cdot X^T \cdot y$$

$$w = \begin{bmatrix} \frac{25}{41} & \frac{79}{164} & \frac{37}{164} & -\frac{5}{164} & -\frac{47}{164} \\ -\frac{16}{41} & -\frac{11}{41} & -\frac{1}{41} & \frac{9}{41} & \frac{19}{41} \end{bmatrix} \cdot \begin{bmatrix} 0.707 \\ 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}$$

$$w = \begin{bmatrix} 0.9433 \\ -0.7637 \end{bmatrix}$$



0036542159 Bruno Marković

IPOV

1.3 b)

$$\phi(x) = (1 \;\; x \;\; x^2)$$

$$\phi = \begin{bmatrix} 1 & 0.25 & 0.25^2 \\ 1 & 0.5 & 0.5^2 \\ 1 & 1 & 1 \\ 1 & 1.5 & 1.5^2 \\ 1 & 2 & 4 \end{bmatrix} \qquad \phi^T = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0.25 & 0.5 & 1 & 1.5 & 2 \\ 0.0625 & 0.25 & 1 & 2.25 & 4 \end{pmatrix}$$

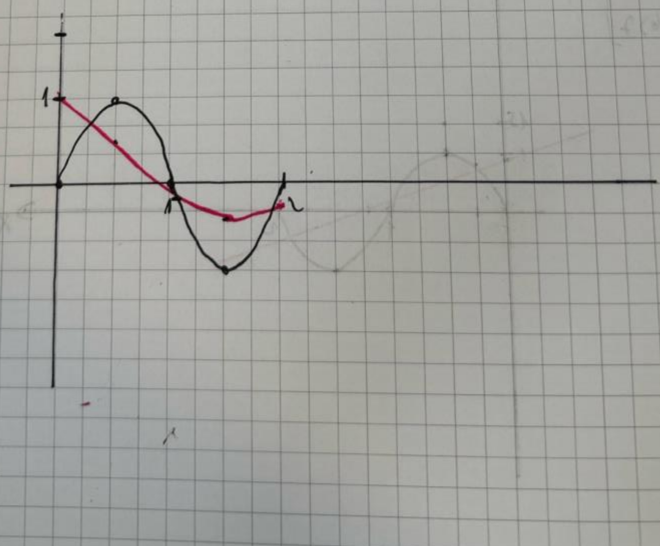
$$\vec{w} = (\phi \cdot \phi^T)^{-1} \cdot \phi^T \cdot y$$

$$\vec{w} = \begin{bmatrix} 5.00 & 5.25 & 7.5625 \\ 5.25 & 7.5625 & 12.5156 \\ 7.56 & 12.5 & 22.13 \end{bmatrix}^{-1} \cdot \phi^T \cdot y$$

$$\vec{w} = \begin{bmatrix} 2.17 & -4.37 & 1.73 \\ -4.37 & 10.85 & -4.64 \\ 1.73 & -4.64 & 2.08 \end{bmatrix} \cdot \phi^T \cdot y$$

$$\vec{w} = \begin{bmatrix} 1.179 & 0.62 & -0.47 & -0.49 & 0.34 \\ -1.94 & -0.11 & 1.84 & 1.66 & -1.24 \\ 0.697 & -0.073 & -0.03 & -0.555 & 0.76 \end{bmatrix} \cdot y$$

$$\vec{w} = \begin{bmatrix} 1.75376 \\ -2.94 \\ 0.975539 \end{bmatrix}$$



0036542159 Bruno Marković

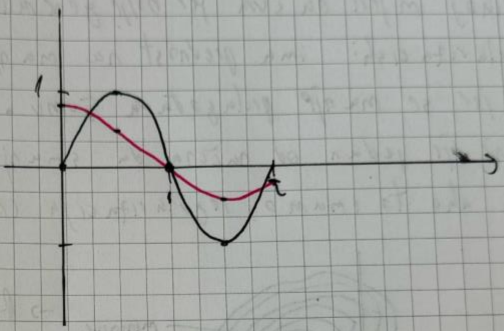
1.3) c)

$$\phi = \begin{bmatrix} 1 & 0.25 & 0.25^2 & 0.25^3 & 0.25^4 \\ 1 & 0.5 & 0.5^2 & 0.5^3 & 0.5^4 \\ 1 & 1 & 1^2 & 1^3 & 1^4 \\ 1 & 1.5 & 1.5^2 & 1.5^3 & 1.5^4 \\ 1 & 2 & 2^2 & 2^3 & 2^4 \end{bmatrix}$$

$$\lambda I = 1 I = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

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

$$\vec{w} = \begin{bmatrix} 0.837662 \\ -0.285 \\ -0.4235 \\ -0.3336 \\ 0.243849 \end{bmatrix}$$



d) Najmanju kvadratnu pogrešku ima C pa zaključujemo da je on najbolji

0036562159 Bruno Marković

V04

1.5 a) Citiram skriptu:

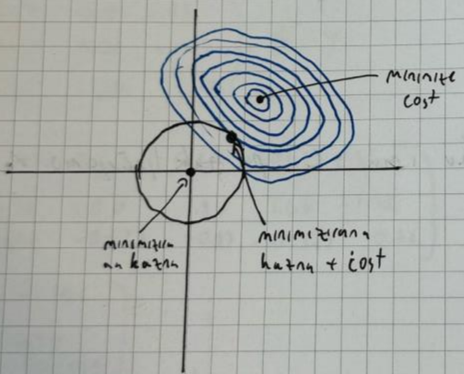
"regularizacija je postupak kojim se implicitno sprječava odabir presloženih modela pri postupku optimizacije parametara modela"

Pretpostavka je da jednostavniji modeli bolje generaliziraju

b) Regularizirani je manje sklon prenaučenosti i ne prilagođava se utolikoj mjeri na šum jer bolje generalizira

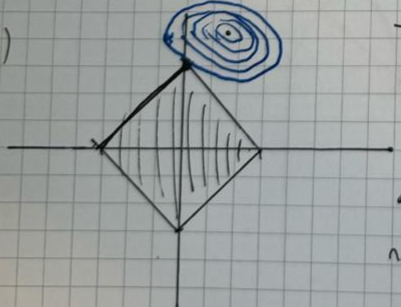
Regularizirani ima prednost na manjim skupom primjera za učenje jer se manje prilagođava šumu. Povećat broj primjera za učenje je jedan od načina da smanjimo prenaučenost tako da ako to imamo regularizacija ima manji koristi.

c)



→ Regularizacija utječe na parametre w_i i w_i^* tako što balansira između minimizacije fje pogreške i minimizacije veličine parametara

d)



→ Izokonture L1 regresije su hiper oktaedri a L2 regresije hipersfere. Izokonture L1 nisu oble nego imaju vrhove, zbog toga će se lakše dogoditi da se minimizator regularizirane fje pogreške nađe na nekoj od koordinatnih osi. Što uzrokuje rijetke modele.

0036542159 Bruno Marković

P04

2.3 $X_1, X_2, X_3, X_4, X_5, X_6, X_7$

$X_1 - X_4 \rightarrow$ ocjene

$X_5 \rightarrow$ prosjek ocjena $\Rightarrow$ linearno ovisan pa ga brišemo

$X_6 \rightarrow$ mat matura

$X_7 \rightarrow$ fiz. matura

Kombinacije $X_i = \binom{6}{1} = 6$ $X_i X_j = \binom{6}{2} = \frac{6 \cdot 5}{2} = 15$

$X_i^2 = \binom{6}{1} = 6$ $X_i X_j X_7 = \binom{6}{3} = \frac{6 \cdot 5 \cdot 4}{3 \cdot 2} = 20$

$\Sigma = 47 = m$

Rješenje $N \geq m + 1$

$\boxed{N \geq 48}$

C