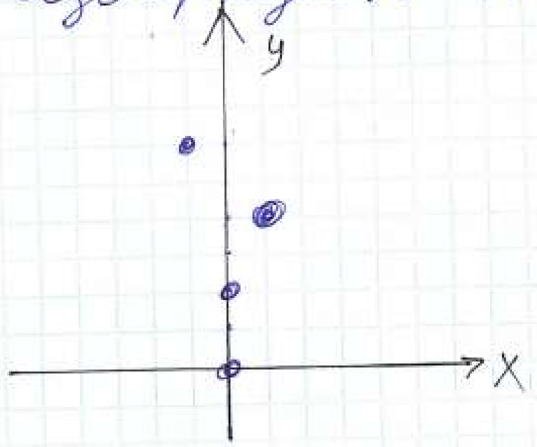
№3

x	1	1	0	0	-1
y	4	4	0	2	6

1) изобразить точки:



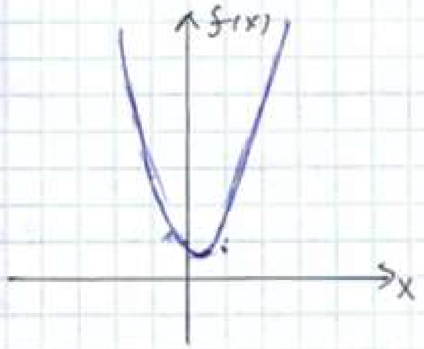
2) методом НК построить модель вида:
$f(x) = \beta_0 + \beta_1 x + \beta_2 x^2$; построить график этой ф-и.

$$X = \begin{array}{c} \\ x=1 \\ x=1 \\ x=0 \\ x=0 \\ x=-1 \end{array} \begin{array}{c} \begin{array}{ccc} \beta_0 & \beta_1 & \beta_2 \end{array} \\ \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & -1 & 1 \end{bmatrix} \end{array}; \quad Y = \begin{bmatrix} 4 \\ 4 \\ 0 \\ 2 \\ 6 \end{bmatrix}$$

$$X^T X \beta = X^T Y \qquad X^T X = \begin{bmatrix} 5 & 1 & 3 \\ 1 & 3 & 1 \\ 3 & 1 & 3 \end{bmatrix}; \quad X^T Y = \begin{bmatrix} 16 \\ 2 \\ 14 \end{bmatrix}$$

$$\left[\begin{array}{ccc|c} 5 & 1 & 3 & 16 \\ 1 & 3 & 1 & 2 \\ 3 & 1 & 3 & 14 \end{array}\right] \longleftrightarrow \left[\begin{array}{ccc|c} 0 & 0 & 1 & 4 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 \end{array}\right]$$

$$f(x) = 1 - x + 4x^2$$

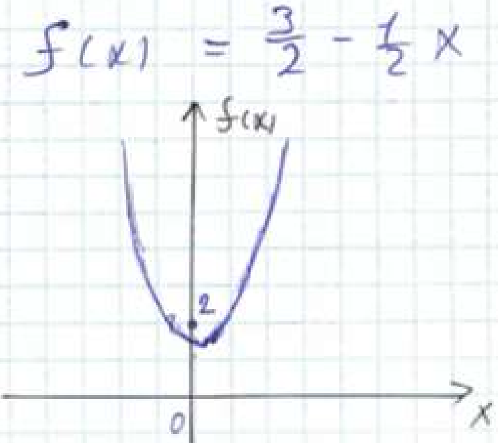


3) построить модель того же вида методом ридж-регрессии с параметром регуляризации $\lambda = 1$; построить график этой функции. $(X^T X + \lambda I)\beta = X^T Y$

$$X^T X = \begin{bmatrix} 5 & 1 & 3 \\ 1 & 3 & 1 \\ 3 & 1 & 3 \end{bmatrix}; \quad X^T Y = \begin{bmatrix} 16 \\ 2 \\ 14 \end{bmatrix}$$

$$\left[\begin{array}{ccc|c} 6 & 1 & 3 & 16 \\ 1 & 4 & 1 & 2 \\ 3 & 1 & 4 & 14 \end{array}\right] \Rightarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & \frac{3}{2} \\ 0 & 1 & 0 & -\frac{1}{2} \\ 0 & 0 & 1 & \frac{5}{2} \end{array}\right]$$

$$f(x) = \frac{3}{2} - \frac{1}{2}x + \frac{5}{2}x^2$$



№40

∃ задача бин классификации.

Доказать, что если известно сколько в выборке представителей каждого из двух классов, то по двум показателям из списка TPR, TNR, PPV, NPV. определяются остальные два.

$$(1)\quad \begin{cases} TPR = \dfrac{TP}{\underbrace{TP+FN}_{N_1}} \\ TNR = \dfrac{TN}{\underbrace{TN+FP}_{N_2}} \\ PPV = \dfrac{TP}{TP+FP} \\ NPV = \dfrac{TN}{TN+FN} \end{cases}$$

$$(2)\quad N = \underbrace{TP+FN}_{N_1} + \underbrace{FP+TN}_{N_2}$$

	+	−
+	TP	FP
−	FN	TN

∃ известно TPR, N_1, TNR, N_2

из (1) $TP = TPR \cdot N_1$, $TN = TNR \cdot N_2$

$FN = N_1 - TP$, $FP = N_2 - TN$

Теперь нам известны все TP, TN, FN, FP

отсюда находим PPV, NPV из (1)

Остальные комбинации доказываются аналогичным способом

№41

$\exists$ $\triangle$ задача бинарной классификации

Верно ли, что

1) если у двух классификаторов на одной и той же выборке совпадают PPV и совпадают TPR, то будут совпадать TNR и NPV

$$\begin{cases} PPV_1 = PPV_2 \\ TPR_1 = TPR_2 \end{cases} \overset{?}{\Rightarrow} \begin{cases} TNR_1 = TNR_2 \\ NPV_1 = NPV_2 \end{cases}$$

$$\Downarrow$$

$$\begin{cases} \dfrac{TP_1}{TP_1 + FP_1} = \dfrac{TP_2}{TP_2 + FP_2} \\ \dfrac{TP_1}{TP_1 + FN_1} = \dfrac{TP_2}{TP_2 + FN_2} \end{cases} \quad (*)$$

из условий задачи известно, что

(1) $TP_1 + FN_1 = TP_2 + FN_2$

(1) $\rightarrow$ (*) $\Rightarrow TP_1 = TP_2$ (2)

$\cancel{TP_1} + FP = \cancel{TP_2} + FP_2 \Rightarrow FP_1 = FP_2$

$TN_1 + \cancel{FP_1} = TN_2 + \cancel{FP_2} \Rightarrow TN_1 = TN_2$

$\Downarrow$

$TP_1 = TP_2,\ TN_1 = TN_2,\ FP_1 = FP_2,\ FN_1 = FN_2$

$\Rightarrow TNR_1 = TNR_2$

$NPV_1 = NPV_2$

Ответ Верно

2) если у двух классификаторов на одной и той же выборке совпадают TNR и совпадают NPV, то будут совпадать PPV и TPR

$$\begin{cases} TNR_1 = TNR_2 \\ NPV_1 = NPV_2 \end{cases} \overset{?}{\Rightarrow} \begin{cases} PPV_1 = PPV_2 \\ TPR_1 = TPR_2 \end{cases}$$

$$\begin{cases} \dfrac{TN_1}{TN_1 + FP_1} = \dfrac{TN_2}{TN_2 + FP_2} \\ \dfrac{TN_1}{TN_1 + FN_1} = \dfrac{TN_2}{TN_2 + FN_2} \end{cases} \quad (*)$$

(1) $TN_1 + FN_1 = TN_2 + FN_2$ $\quad (1) \to (*)$

$\Rightarrow TN_1 = TN_2$ (2) $\quad (2) \to (1) \Rightarrow FN_1 = FN_2$

$TP_1 + \cancel{FN_1} = TP_2 + \cancel{FN_2} \Rightarrow TP_1 = TP_2$

$FP_1 + \cancel{TN_1} = FP_2 + \cancel{TN_2}$

$\Downarrow$

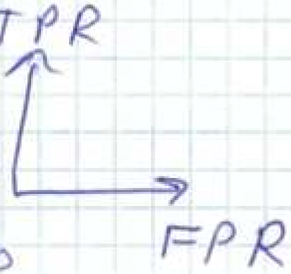
$$\begin{cases} TN_1 = TN_2 \\ TP_1 = TP_2 \\ FP_1 = FP_2 \\ FN_1 = FN_2 \end{cases} \Rightarrow \begin{cases} PPV_1 = PPV_2 \\ NPV_1 = NPV_2 \end{cases}$$

Ответ Верно

3) совпадение ROC-кривых (--//--) влечет совпадение Precision-Recall кривых и наоборот

$$P = \frac{TP}{TP+FP}$$

$$ROC = \frac{TPR}{FPR}$$



$$R = \frac{TP}{TP+FN}$$

$$TPR = \frac{TP}{TP+FN}$$

$$FPR = \frac{FP}{\{FP+TN\}}$$

из ROC

$$\begin{cases} TPR_1 = TPR_2 \\ FPR_1 = FPR_2 \end{cases} \overset{?}{\Longleftrightarrow} \begin{cases} TPR_1 = TPR_2 \\ PPV_1 = PPV_2 \end{cases}$$

($\Rightarrow$)

если возмем ROC как ф-ю

$$TPR_1(FPR) = TPR_2(FPR)$$

$$TPR_1 = TPR_2 \Rightarrow TP_1 = TP_2$$

$$P_1 = \frac{TP}{TP+FP_1}, \quad P_2 = \frac{TP}{TP+FP_2}$$

если $FP_2 \neq FP_1$, то это неверно

($\Leftarrow$)

$$P_1(R) = P_2(R)$$

из P и $R \Rightarrow TP_1 = TP_2 \quad FP_1 = FP_2$

$$\frac{TP}{TP+FN_1} = \frac{TP}{TP+FN_2}$$

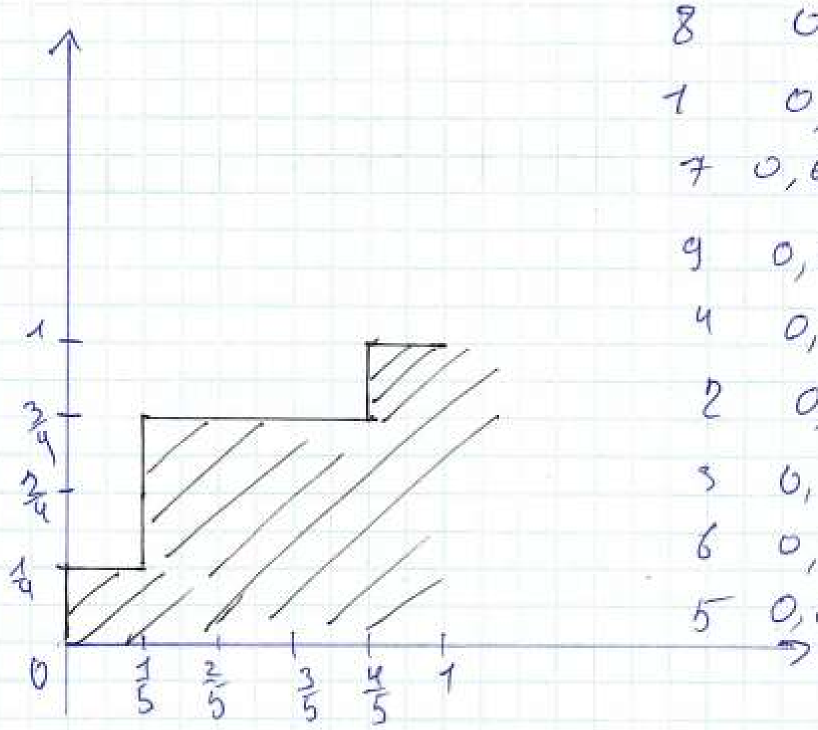
$$\frac{FP}{FP+TN_1} = \frac{FP}{FP+TN_2}$$

№42

∃ в задаче классификации на 2 класса $\{0, 1\}$ некоторый классификатор (байесовский) определяет сред оцен $g(x)$ апостериорной вероятности принадлежности объекта к классу 1

i	1	2	3	4	5	6	7	8	9
$y^{(i)}$	0	0	0	0	0	1	1	1	1
$g(x^{(i)})$	0,75	0,15	0,11	0,23	0,09	0,10	0,66	0,82	0,50

Построить ROC $m = 4$ $n = 5$



8	0,82	1
1	0,75	0
7	0,66	1
9	0,5	1
4	0,23	0
2	0,15	0
3	0,11	0
6	0,1	1
5	0,09	0

AUC - пл под ROC кривой

$$AUC = \frac{1}{4}\cdot\frac{1}{5} + \frac{3}{4}\cdot\frac{3}{5} + 1\cdot\frac{1}{5} = \frac{7}{10}$$

$f(x) = I(g(x) \geq 0{,}5)$ нат рас

	+	−	
TP	3	1	FP
FN	1	4	TN

$FPR = \frac{FP}{FP+TN} = \frac{1}{5}$ $\quad PPV = \frac{TP}{TP+FP} = \frac{3}{4}$

$FNP = \frac{FN}{TP+FN} = \frac{1}{4}$ $\quad TPR = \frac{TP}{TP+FN} = \frac{3}{4}$

$TNR = \frac{TN}{TN+FP} = \frac{4}{5}$

$accuracy = \frac{7}{9}$

$error = 1 - accuracy = \frac{FP+FN}{TP+TN+FP+FN} =$

$= \frac{2}{9}$

$$F_1 = \frac{2}{\frac{1}{P} + \frac{1}{R}} = \frac{2}{\frac{4}{3} + \frac{4}{3}} = \frac{3}{4}$$