# Assignment 7 Solution

Nayan Man Singh Pradhan

**Excercise 1.** (Training error)

In this task, we consider the training data

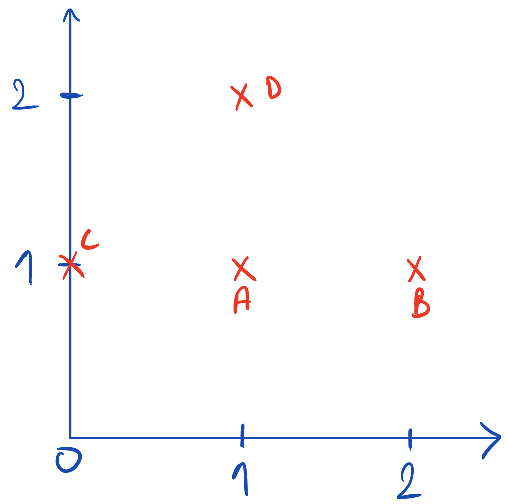
$$\mathcal{T} = \left\{ \left((1,1)^\top, 2\right), \left((2,1)^\top, 5\right), \left((0,1)^\top, 1\right), \left((1,2)^\top, 5\right) \right\}.$$

a) Build a predictor using kNN regression for $k = 3$ and evaluate the training error of the predictor.

b) Build a predictor using linear regression by least squares and evaluate the training error of the linear model.

(4 Points)

a) let $A(1,1)$, $B(2,1)$, $C(0,1)$ and $D(1,2)$




Computing the distance between the points

| | A | B | C | D |
|---|---|---|---|---|
| A | 0 | 1 | 1 | 1 |
| B | 1 | 0 | 2 | 1.414 |
| C | 1 | 2 | 0 | 1.414 |
| D | 1 | 1.414 | 1.414 | 0 |

Using KNN regression (with $k=3$)

For $(1,1)$, the closest 3 neighbors are $(1,1)$, $(2,1)$, $(0,1)$

$$Y(1,1) = \frac{1}{3}\left(2+5+1\right) = \frac{8}{3} = 2.67$$

For $(2,1)$, the closest 3 neighbors are $(2,1)$, $(1,1)$, $(1,2)$

$$Y(2,1) = \frac{1}{3}\left(5+2+5\right) = \frac{12}{3} = 4$$

for $(0,1)$, the closest 3 neighbors are $(0,1)$, $(1,1)$, $(1,2)$

$$Y(0,1) = \frac{1}{3}\left(1+2+5\right) = \frac{8}{3} = 2.67$$

For $(1,2)$, the closest 3 neighbors are $(1,2)$, $(1,1)$, $(2,1)$

$$Y(1,2) = \frac{1}{3}\left(5+2+5\right) = \frac{12}{3} = 4$$

Training Error:

$$TE(f, T) = \frac{1}{N} \sum_{i=1}^{N} L(y_i, f(x_i))$$

$$= \frac{1}{4} \sum_{i=1}^{4} L_2(y_i, f(x_i))$$

$$= \frac{1}{4}\left( (2-2.67)^2 + (5-4)^2 + (1-2.67)^2 + (5-4)^2 \right)$$

$$= 1.309$$

b) $X = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 2 \end{pmatrix}, \quad X^T = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 0 & 1 \\ 1 & 1 & 1 & 2 \end{pmatrix}, \quad y = \begin{pmatrix} 2 \\ 5 \\ 1 \\ 5 \end{pmatrix}$

$$X^T X = \underbrace{\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 0 & 1 \\ 1 & 1 & 1 & 2 \end{pmatrix}}_{3\times4} \underbrace{\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 2 \end{pmatrix}}_{4\times3} = \begin{pmatrix} 4 & 4 & 5 \\ 4 & 6 & 5 \\ 5 & 5 & 7 \end{pmatrix}$$

$$X^T y = \underbrace{\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 0 & 1 \\ 1 & 1 & 1 & 2 \end{pmatrix}}_{3\times4} \cdot \underbrace{\begin{pmatrix} 2 \\ 5 \\ 1 \\ 5 \end{pmatrix}}_{4\times1} = \begin{pmatrix} 13 \\ 17 \\ 18 \end{pmatrix}$$

$$(X^T X)\hat{\beta} = (X^T y)$$

$$\begin{pmatrix} 4 & 4 & 5 \\ 4 & 6 & 5 \\ 5 & 5 & 7 \end{pmatrix} \hat{\beta} = \begin{pmatrix} 13 \\ 17 \\ 18 \end{pmatrix}$$

Using gaussian elimination to solve,

$$\left(\begin{array}{ccc|c} 4 & 4 & 5 & 13 \\ 4 & 6 & 5 & 17 \\ 5 & 5 & 7 & 18 \end{array}\right)$$

$$\left(\begin{array}{ccc|c} 4 & 4 & 5 & 13 \\ 0 & 2 & 0 & 4 \\ 0 & 0 & 3/4 & 7/4 \end{array}\right)$$

$$\frac{3}{4}\beta_3 = \frac{7}{4} \Rightarrow \beta_3 = \frac{7}{3} \,//$$

$$2\beta_2 = 4 \Rightarrow \beta_2 = 2 \,//$$

$$4\beta_1 + 4\beta_2 + 5 \times \beta_3 = 13 \Rightarrow 4\beta_1 + 8 + \frac{35}{3} = 13 \Rightarrow \beta_1 = -\frac{5}{3} \,//$$

$$\therefore f(\beta) = -\frac{5}{3} + 2\beta_1 + \frac{7}{3}\beta_2$$

$$f(1,1) = -\frac{5}{3} + 2 + \frac{7}{3} = 2.67$$

$$f(2,1) = -\frac{5}{3} + 2\times 2 + \frac{7}{3} = 4.67$$

$$f(0,1) = -\frac{5}{3} + 0 + \frac{7}{3} = 0.67$$

$$f(1,2) = -\frac{5}{3} + 2\times 1 + \frac{7}{3}\times 2 = 5$$

Training Error:

$$TE(f,T) = \frac{1}{N}\sum_{i=1}^{N} L(y_i, f(x_i))$$

$$= \frac{1}{4}\sum_{i=1}^{4} L_2(y_i, f(x_i))$$

$$= \frac{1}{4}\left((2-2.67)^2 + (5-4.67)^2 + (1-0.67)^2 + (5-5)^2\right)$$

$$= 0.167 \,//$$

**Excercise 2.** (Cross validation)

In this task, we consider the data set

$$\mathcal{T} = \{(-2,4),(2,4),(1,1),(-1,1),(0,0),(3,9)\}\,.$$

a) Evaluate the (expected) generalization error of the kNN regressor with $k = 2$ by $\mathcal{K}$-fold cross validation with $\mathcal{K} = 3$. (Do a deterministic, i.e. not randomized, splitting of the given data following the ordering of the samples.

b) Evaluate the (expected) generalization error of the linear model by leave-one-out cross validation. (It is fine to use the help of a computer to fit the individual linear models.)

(4 Points)

a) $T = \{(-2,4), (2,4), (1,1), (-1,1), (0,0), (3,9)\}$

$K = 3$

$T_1 = \{(-2,4), (2,4)\}$

$T_2 = \{(1,1), (-1,1)\}$

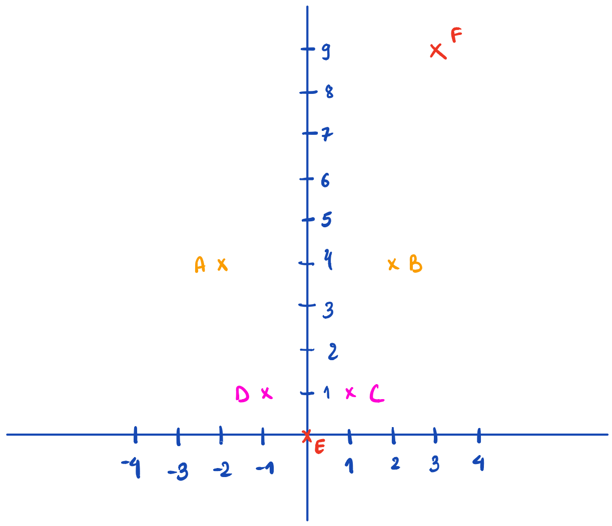
$T_3 = \{(0,0), (3,9)\}$

Let,

$T_1 = A(-2,4), B = (2,4)$

$T_2 = C(1,1), D = (-1,1)$

$T_3 = E(0,0), F(3,9)$




(i) Using $T_1$ and $T_2$ as training and $T_3$ as cross validation set

For E, the 2 nearest neighbors are: C and D

$= \frac{1}{2} \times (1+1) = 1$

for F, the 2 nearest neighbors are B and A

$= \frac{1}{2} \times (4+4) = 4$

$\text{Error}_1 = \frac{1}{2} \times \left[(4-9)^2 + (1-0)^2\right] = \frac{13}{2} = 6.5$

(ii) Using $T_2$ and $T_3$ as training and $T_1$ as cross validation set
for B, the two nearest neighbors are C and D.

$$= \frac{1}{2} \times (1+1) = 1$$

for A, the two nearest neighbors are D and E

$$= \frac{1}{2} \times (1+0) = 0.5$$

$$\text{Error}_2 = \frac{1}{2} \times \left[(1-4)^2 + (0.5-4)^2\right] = 21.25 = 10.625$$

(iii) Using $T_1$ and $T_3$ as training and $T_2$ as cross-validation set
for C, the two nearest neighbors are E and B

$$= \frac{1}{2} \times (4+0) = 2$$

for D, the two nearest neighbors are A and E

$$= \frac{1}{2} \times (4+0) = 2$$

$$\text{Error} = \frac{1}{2} \times \left[(2-1)^2 + (1-2)^2\right] = \frac{2}{2} = 1$$

$$\text{Expected generalization error} = \frac{1}{3} \times (6.5 + 10.625 + 1) = 6.042$$

b) $T = \{(-2,4), (2,4), (1,1), (-1,1), (0,0), (3,9)\}$

let, $T_1(-2,4)$, $T_2(2,4)$, $T_3(1,1)$, $T_4(-1,1)$, $T_5(0,0)$, $T_6(3,9)$

(i) Consider $T_1, T_2, T_3, T_4, T_5$ as training and $T_6$ as cross-validation set.

$$X = \begin{pmatrix} 1 & -2 \\ 1 & 2 \\ 1 & 1 \\ 1 & -1 \\ 1 & 0 \end{pmatrix}, \quad X^T = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ -2 & 2 & 1 & -1 & 0 \end{pmatrix}, \quad y = \begin{pmatrix} 4 \\ 4 \\ 1 \\ 1 \\ 0 \end{pmatrix}$$

$$X^T X = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ -2 & 2 & 1 & -1 & 0 \end{pmatrix}_{2\times5} \begin{pmatrix} 1 & -2 \\ 1 & 2 \\ 1 & 1 \\ 1 & -1 \\ 1 & 0 \end{pmatrix}_{5\times2} = \begin{pmatrix} 5 & 0 \\ 0 & 10 \end{pmatrix}$$

$$X^T y = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ -2 & 2 & 1 & -1 & 0 \end{pmatrix}_{2\times5} \begin{pmatrix} 4 \\ 4 \\ 1 \\ 1 \\ 0 \end{pmatrix}_{5\times1} = \begin{pmatrix} 10 \\ 0 \end{pmatrix}_{2\times1}$$

$$X^T X \beta = X^T y$$

$$\begin{pmatrix} 5 & 0 \\ 0 & 10 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} 10 \\ 0 \end{pmatrix}$$

$5\beta_1 + 0 = 10$ $\qquad$ $0\beta_1 + 10\beta_2 = 0$

$\beta_1 = 2,$ $\qquad$ $\beta_2 = 0$

$Y_p = 2 \cdot + 0$

$= 2$

Testing $T_6(3, 9)$

$Y_p = 2$

$Y_R = 9$