

# EMS702P Statistical Thinking and Applied Machine Learning

**Week 6 – PBL**

---

Yunpeng Zhu

**Q1: (5 min)**

$$\mathbf{A} = \begin{bmatrix} -1 & 0 \\ 2 & 3 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}, \text{ calculate } \mathbf{A}^T(\mathbf{A} + \mathbf{B})^{-1}$$

**Q2: (25 min)**

Consider we have 5 sets of observed data (red points on the coordinate system)

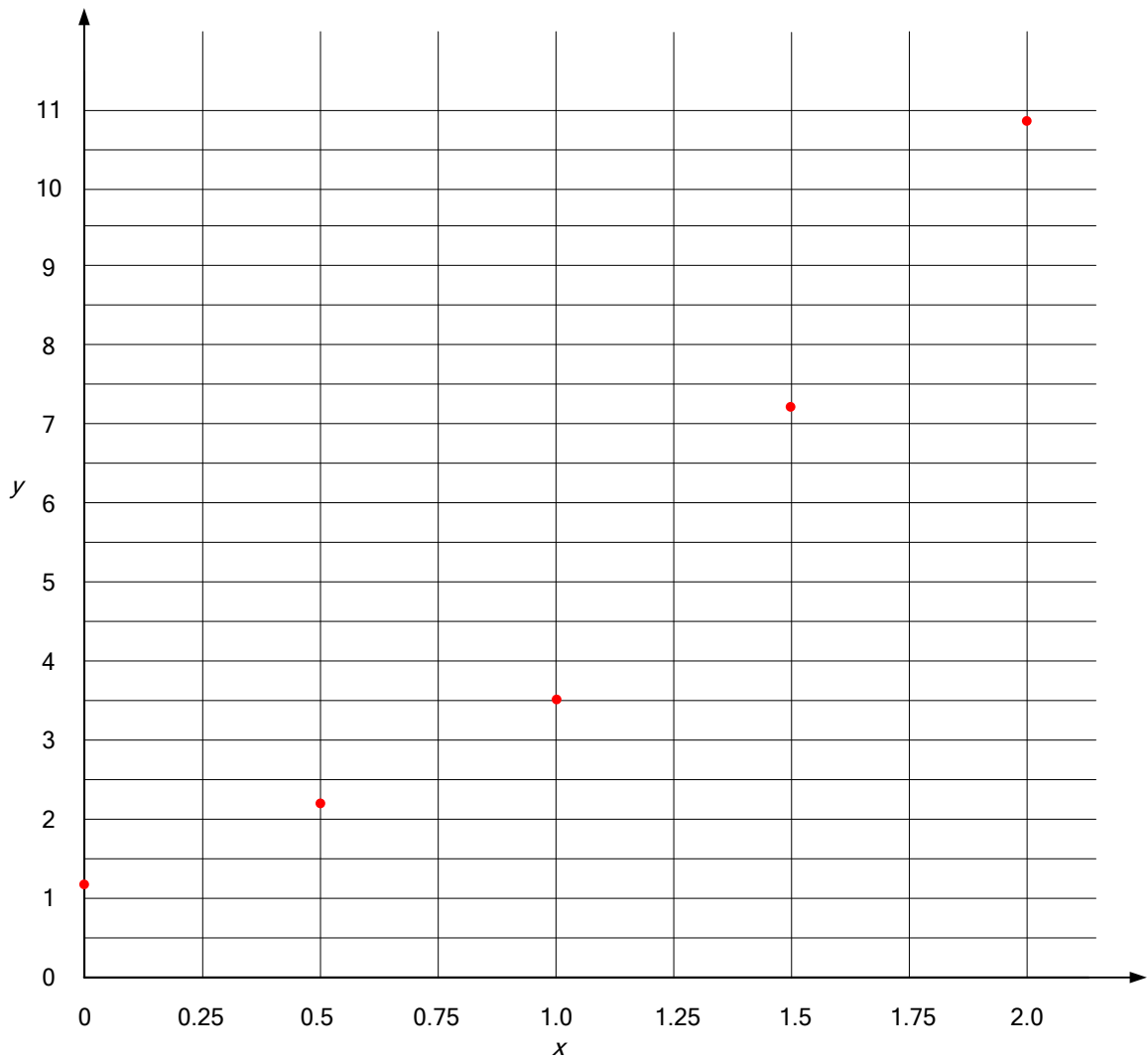
$$(x, y) = [(0.0, 1.2); (0.5, 2.2); (1.0, 3.5); (1.5, 7.3); (2.0, 10.8)]$$

Evaluate the regression model as a second-order polynomial function:

$$y = 1 + a_1x + a_2x^2$$

by

- (1) Formulate the regression model into a matrix form and show the least Squares (LS) representation of  $a_1$  and  $a_2$ ;
- (2) Estimate the values of  $a_1, a_2$  from the observed data by using the LS method;
- (3) Draw the regression model on the coordinate system (Show the coordinate values of the points you used for regression);
- (4) Validate the regression results by calculating the Mean Squared Error (MSE);



**Q3: (15 min)**

Separate the 4 points:  $(x_1, x_2) = (0,1), (1,0), (0,-1), (-1,0)$  by using the K-class discriminant classifiers:

$$y_k = w_1^{(k)} x_1 + w_2^{(k)} x_2, k = 1, 2, 3, 4$$