

①

$$\text{empirical risk} = \frac{1}{n} \sum_i \ell(\hat{y}_i, y_i) = \frac{3}{6} = \frac{1}{2}$$

②

$$C = \begin{matrix} & \begin{matrix} A & B & C & D & E \end{matrix} \\ \begin{matrix} A \\ B \\ C \\ D \\ E \end{matrix} & \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \end{matrix}$$

③ $\ell_2 = 3 \times 1^2 = 3$

④ Yes. Mapping A...F sequentially to 1...5 is better, because every adjacent letter grade should have numerical distance of 1.

⑤ $\frac{1}{2}$. They should be close.

⑥ No.

$$f'(x) = \begin{cases} x \geq 18.6 & , A \\ 18.6 > x \geq 16.5 & , B \\ 16.5 > x \geq 14.8 & , C \\ 14.8 > x \geq 1 & , D \\ 1 > x \geq 0 & , F \end{cases}$$