Practice Midterm.

Problem 1: parameters describe linear regression & logistic regression.

Problem 2. 0~1

Problem 5. $\mathcal{L}_{\text{Square Error}} = \frac{1}{2}\|t-y\|^2$, calcs the residual

Problem 6. $\sqrt{(x_i^a - x_i^b)^2 + (x_j^a - x_j^b)^2 + (x_k^a - x_k^b)^2}$

$= \sqrt{(5-1)^2 + (9-2)^2 + (-3+6)^2} = \sqrt{a}$

Problem 7. $y = w_0x_0 + w_1x_1 + w_2x_2. = w_0 \times 1 + w_1x_1 + w_2x_2$

$= w^T x$

$w = [w_0, w_1, w_2]$

$x = [x_0, x_1, x_2]$

Problem 8. $y = w^T x$ coeff = weight matrix/vector

$(y-t)x$

$(\cancel{-\frac{t}{y}}\ y)\cdot x$

$w \leftarrow w - \cancel{\ }\alpha \frac{\partial \mathcal{J}}{\partial w}$ ✓ $= w - \alpha \frac{1}{N}\left(\frac{\partial \mathcal{L}_i}{\partial w}\right)$ ✓

$w_j \leftarrow w_j - \alpha \frac{\partial \mathcal{J}}{\partial w_j}$ ✓

Slide 26

$$\mathcal{L}_{CrossEntropy} = -t\log y - (1-t)\log(1-y)$$

$$y = \frac{1}{1+e^{(-z)}}, \; z = w^T x$$

$$\frac{\partial \mathcal{L}_{CE}}{\partial w_j} = \frac{\partial \mathcal{L}_{CE}}{\partial y} \cdot \frac{\partial y}{\partial z} \cdot \frac{\partial z}{\partial w_j}$$

$$\frac{\partial \mathcal{L}_{CE}}{\partial y} = \frac{\partial}{\partial y}\left(-t\log y - (1-t)\log(1-y)\right) \quad \vdots \; (\log x)' = \frac{1}{x}$$

$$= -t \cdot \frac{1}{y} + \frac{(1-t)}{1-y} = \frac{-t}{y} + \frac{(1-t)}{(1-y)}$$

$$\frac{\partial y}{\partial z} = \frac{\partial}{\partial z} \frac{1}{1+e^{(-z)}} \qquad g(a) = \frac{u(a)}{v(a)}, \; g' = \frac{u'v - v'u}{v^2}$$

$$= \frac{0 - (-e^{-z}) \times 1}{(1+e^{-z})^2} = \boxed{\frac{e^{-z}}{(1+e^{-z})^2}} \qquad u = 1, \; v = 1 + e^{-z}$$

$$= y(1-y)$$

$$y(1-y) = y - y^2 = \frac{1}{1+e^{-z}} - \frac{1}{(1+e^{-z})^2} = \frac{1+e^{-z}-1}{(1+e^{-z})^2}$$

$$\frac{1}{1+e^{-z}} = \frac{1+e^{-z}}{(1+e^{-z})^2}$$

$$= \boxed{\frac{e^{-z}}{(1+e^{-z})^2}}$$

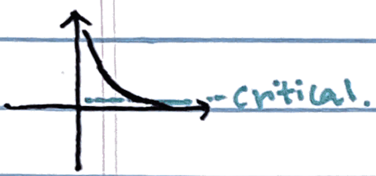
$$\frac{\partial z}{\partial w_j} = \frac{\partial}{\partial w_j} w^T x = \frac{\partial}{\partial w_j} w_0 x_0 + w_1 x_1 + \cdots + \boxed{w_j x_j} + \cdots w_D x_D$$

$$= 0 + 0 + \cdots + x_j + \cdots + 0 = x_j$$

Slide 26

$$\frac{\partial \mathcal{L}_{CE}}{\partial w_j} = \left(\frac{-t}{y} + \frac{(1-t)}{(1-y)}\right) \cdot y(1-y) \cdot x_j = (y-t)x_j$$

$$w_j \leftarrow w_j - \frac{\alpha}{N} \underbrace{\sum_{i=1}^{N} (y^i - t^i) x_{ij}}_{\text{optimize cost } \mathcal{J}\text{, equivalently}}, \text{ gradient descent update}$$

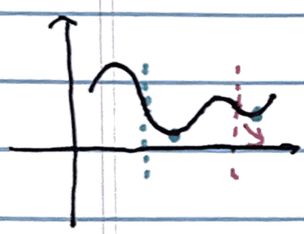


Slide 30

$\{1, 2, 3, 4, \ldots, K\}$ = animals

$y = 1, 2, 3, 4, \ldots$

y = cat (category 1)



instead of having $y=1$, we have $y = [1, \underbrace{0, 0, \ldots, 0}_{K-1}] \in \mathbb{R}^K$

Slide 31

D = features; dealing w/ x

K = classes; dealing w/ y.

$a \cdot b \times b \cdot c = a \cdot c$

$\vec{w} = K \times D$.

$$\underset{\text{bug}}{z_k} = \sum_{j=1}^{D} w_{kj} \cdot x_{kj} + b \text{ ; repeat for } k \in \{1, 2, \ldots K\}$$

Slide 33

$z = Wx + b \Longleftrightarrow z = Wx \text{ ; } x_0 = 1$

$$f\left(\boxed{0}\right) = \begin{bmatrix} 0.5 \\ 0.4 \\ 0.1 \\ 0 \\ 0 \end{bmatrix} \Big\} K.$$

$$\underset{\text{bug}}{y_k} = \text{softmax}(z_1, \ldots z_K) = \frac{e^{z_{k=\text{bug}}}}{\sum_{k'} e^{z_{k'}}}$$

Slide 34

$$\mathcal{L}_{CE}(y, t) = -\sum_{k=1}^{K} t_k \log y_k = -t^T \log(y)$$