1.4.

① $\nabla_\theta^2 J_\theta = H = \frac{1}{m}(X^T S X + \lambda I)$, show also H is positive def.

$\Rightarrow \nabla_\theta J_\theta = \frac{1}{m}\left(\underline{\underline{1_m^T}}\,(\mathrm{diag}(h_\theta - y))\,\underline{\underline{X}}\right)^T$ ... $(1 \times d$ shape$)$ $\quad d.$

x_i's row's y.

sum over column

$$\underbrace{[1\,1 \cdots 1]}_{1\times m}\ \underbrace{\begin{bmatrix} h_\theta(x_1) - y(x_1) & & \\ & h_\theta(x_2) - y(x_2) & \\ & & \ddots \\ & & & h_\theta(x_m) - y(x_m) \end{bmatrix}}_{m\times m}\ \underbrace{\begin{bmatrix} \\ \\ \\ \end{bmatrix}}_{\substack{X \\ m\times d}}$$

$\overset{m}{\underbrace{[1\,1\,1\,1\,1]}_{m.}}$ $\quad X.$

$$\begin{array}{cc} & m \\ m & \begin{bmatrix} h_\theta(x_1) - y(1) & \cdots & 0 \\ \vdots & h_\theta(2) - y_2 & \\ 0 & & \ddots \end{bmatrix} \end{array}$$

$\Rightarrow H(\theta) = \nabla_\theta\left(\frac{1}{m}\sum_{i=1}^{m}(h_\theta(x^{(i)}) - y^{(i)})X_i\right)$

$= \frac{1}{m}\sum_{i=1}^{m}\left(h_\theta(x^{(i)})(1 - h_\theta(x^{(i)}))\,x^{(i)}\right)x^{(i)T}$

$= \frac{1}{m}X^T \cdot \mathrm{diag}(h_1(1-h_1), h_2(1-h_2) \ldots h_m(1-h_m))\,X$

$= \frac{1}{m}X^T S X.$

②

To show H being positive def, assume for any nonzero $\vec{v}$

$\vec{v}^T X^T S X \vec{v} = (\vec{v}^T X^T S^{\frac{1}{2}})(S^{\frac{1}{2}} X \vec{v})$ $\quad$ ($\because S$ diagonal)

$= A^T A$, where $A = S^{\frac{1}{2}} X \vec{v}$ > 0

#

1.5.

$\theta \leftarrow \theta - H(\theta)^{-1}\nabla J(\theta)$