HW3

1951095 Yiwen Liang

1. (**Math**)

Answer:

First, for the formula $L(\mathbf{h})$, its domain is in the space of $\mathbb{R}^n$, that is, its value range is in a convex set. And the second order partial derivative of this formula exists.

Let the formula $L(\mathbf{h})$ be derived with respect to $\mathbf{h}$ to get :

$$\frac{d\left(L(\mathbf{h})\right)}{d\mathbf{h}} = (\mathbf{J}(\mathbf{x}))^T\mathbf{f}(\mathbf{x}) + (\mathbf{J}(\mathbf{x}))^T\mathbf{J}(\mathbf{x})\mathbf{h} + \mu\mathbf{I}\mathbf{h} \tag{1}$$

Then take the derivative of the above formula again to get the Hessian matrix of $L(\mathbf{h})$ as:

$$\frac{d\left((\mathbf{J}(\mathbf{x}))^T\mathbf{f}(\mathbf{x}) + \mathbf{J}(\mathbf{x}))^T\mathbf{J}(\mathbf{x})\mathbf{h} + \mu\mathbf{I}\mathbf{h}\right)}{d\mathbf{h}} = (\mathbf{J}(\mathbf{x}))^T\mathbf{J}(\mathbf{x}) + \mu\mathbf{I} \tag{2}$$

Let $\mathbf{A} = \mathbf{J}^T\mathbf{J}$, then

$$\forall\mathbf{x} \neq \mathbf{0}, \mathbf{y} = \mathbf{J}\mathbf{x} \tag{3}$$

$$0 \leq \mathbf{y}^T\mathbf{y} = \mathbf{x}^T\mathbf{J}^T\mathbf{J}\mathbf{x} = \mathbf{x}^T\mathbf{A}\mathbf{x}, \tag{4}$$

$\therefore$ $\mathbf{A}$ is positive semi-definite. Let λ be the eigenvalue of $\mathbf{A}$, that is

$$\begin{gathered}\{\lambda_i \geq 0, i = 1, \ldots, n\}, \\ \mathbf{A}\mathbf{v}_i = \lambda_i\mathbf{v}_i, \\ (\mathbf{A} + \mu\mathbf{I})\mathbf{v}_i = (\lambda_i + \mu)\mathbf{v}_i\end{gathered} \tag{5}$$

$$\begin{gathered}\because \mu > 0, \\ \therefore \{\lambda_i + \mu\} > 0, \\ \therefore \mathbf{A} + \mu\mathbf{I} \text{ is positive definite, which is the Hessian matrix of } L(\mathbf{h}).\end{gathered}$$

By the following theorem:

If a function L(h) is differentiable up to at least second order, L is strictly convex if its Hessian matrix is positive definite.

we can get that $L(\mathbf{h})$ is strictly convex.

---

2. **(Math)**

Answer:

According to the definition of gradient:

$$\nabla_\theta J(\theta) = \left[\frac{\partial J(\theta)}{\partial\theta_1}\frac{\partial J(\theta)}{\partial\theta_2}\cdots\frac{\partial J(\theta)}{\partial\theta_m}\right]^T, \tag{1}$$

where $\mathbf{x}$ is the input vector, assuming a length of n; y is ground-truth; θ is the linear regression parameter of $\mathbf{x}$,

$$\theta^T \mathbf{x}_i = [\theta_1 x_{i1}, \theta_2 x_{i2}, \ldots \theta_n x_{in}]^T \tag{2}$$

$h_\theta(x_i)$ is the logistic regression function:

$$h_\theta(\mathbf{x}_i) = \frac{1}{1+\exp(-\theta^T \mathbf{x}_i)} \tag{3}$$

$$\frac{\partial h_\theta(\mathbf{x}_i)}{\partial \theta_j} = -\frac{e^{-\theta^T \mathbf{x}_i}(x_{ij})}{[1+\exp(-\theta^T \mathbf{x}_i)]^2} \tag{4}$$

$$h_\theta(\mathbf{x}_i)[1-h_\theta(\mathbf{x}_i)] =- \frac{e^{-\theta^T \mathbf{x}_i}}{[1+\exp(-\theta^T \mathbf{x}_i)]^2} \tag{5}$$

According to equation (1)(2)(3)(4)(5), we can calculate:

$$\begin{aligned}
\frac{\partial J(\theta)}{\partial \theta_j} &= -\sum_{i=1}^{m}\left[y_i \frac{1}{h_\theta(\mathbf{x}_i)} \frac{\partial h_\theta(\mathbf{x}_i)}{\partial \theta_j} + (1-y_i)\frac{-\frac{\partial h_\theta(\mathbf{x}_i)}{\partial \theta_j}}{1-h_\theta(\mathbf{x}_i)}\right] \\
&= -\sum_{i=1}^{m}\left[\frac{\partial h_\theta(\mathbf{x}_i)}{\partial \theta_j}\left(\frac{y_i}{h_\theta(\mathbf{x}_i)} + \frac{y_i-1}{1-h_\theta(\mathbf{x}_i)}\right)\right] \\
&= -\sum_{i=1}^{m}\left[\frac{\partial h_\theta(\mathbf{x}_i)}{\partial \theta_j}\left(\frac{y_i-h_\theta(\mathbf{x}_i)}{h_\theta(x_i)[1-h_\theta(\mathbf{x}_i)]}\right)\right] \\
&= -\sum_{i=1}^{m} x_{ij}\cdot(y_i - h_\theta(\mathbf{x}_i)) \\
&= \sum_{i=1}^{m} x_{ij}\cdot(h_\theta(\mathbf{x}_i) - y_i)
\end{aligned} \tag{6}$$

For each $\theta_j \in \boldsymbol{\theta}$, the above equation (6) holds, so the gradient of the cost function is:

$$\nabla_\theta J(\boldsymbol{\theta}) = \sum_{i=1}^{m} \mathbf{x}_i (h_\theta(\mathbf{x}_i) - y_i)$$

---

3. (**Programming**)

The model used is YoloV4. Because there is only one cover of "Xiangla" in my collected data set, so it will not be recognized for another variety. However, the model has a high accuracy rate for the recognition of "Luxiang".

supermarket_detect.mp4
xiangla (83%)
luxiang (99%)
luxiang
xiangla (90%)
xiangla (94%)
luxiang (100%)
luxiang (99%)
香辣牛肉面
香辣牛肉面
00:00:17 / 00:00:31
标记
SVIP
0.75x
高清
字幕
NEW