Assignment 3 (Due: Jun. 12, 2022)

1. (**Math**) Nonlinear least-squares. Suppose that $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), ..., f_m(\mathbf{x})): \mathbb{R}^n \to \mathbb{R}^m$, $\mathbf{x} \in \mathbb{R}^n, \mathbf{f} \in \mathbb{R}^m$ and some $f_i(\mathbf{x}): \mathbb{R}^n \to \mathbb{R}$ is a (are) non-linear function(s). Then, the problem,

$$\mathbf{x}^* = \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{f}(\mathbf{x})\|_2^2 = \arg\min_{\mathbf{x}} \frac{1}{2}(\mathbf{f}(\mathbf{x}))^T \mathbf{f}(\mathbf{x})$$

is a nonlinear least-squares problem. In our lecture, we mentioned that Levenberg-Marquardt algorithm is a typical method to solve this problem. In L-M algorithm, for each updating step, at the current **x**, a local approximation model is constructed as,

$$L(\mathbf{h}) = \frac{1}{2}(\mathbf{f}(\mathbf{x}+\mathbf{h}))^T \mathbf{f}(\mathbf{x}+\mathbf{h}) + \frac{1}{2}\mu \mathbf{h}^T\mathbf{h}$$
$$= \frac{1}{2}(\mathbf{f}(\mathbf{x}))^T \mathbf{f}(\mathbf{x}) + \mathbf{h}^T (\mathbf{J}(\mathbf{x}))^T \mathbf{f}(\mathbf{x}) + \frac{1}{2}\mathbf{h}^T (\mathbf{J}(\mathbf{x}))^T \mathbf{J}(\mathbf{x})\mathbf{h} + \frac{1}{2}\mu \mathbf{h}^T\mathbf{h}$$

where **J**(**x**) is **f**(**x**)'s Jacobian matrix, and $\mu > 0$ is the damped coefficient. Please prove that $L(\mathbf{h})$ is a strictly convex function. (Hint: If a function $L(\mathbf{h})$ is differentiable up to at least second order, L is strictly convex if its Hessian matrix is positive definite.)

2. (**Math**) In our lecture, we mentioned that for logistic regression, the cost function is,

$$J(\boldsymbol{\theta}) = -\sum_{i=1}^{m} y_i \log\left(h_{\boldsymbol{\theta}}(\boldsymbol{x}_i)\right) + (1-y_i)\log\left(1-h_{\boldsymbol{\theta}}(\boldsymbol{x}_i)\right)$$

Please verify that the gradient of this cost function is

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

3. (**Programming**) In intelligent retail, one task is to investigate the proportion of each commodity occupying shelves. In this assignment, suppose that you are provided a surveillance video of a shelf and you need to recognize and locate two

specific kinds of products, “康师傅香辣牛肉面” and “康师父卤香牛肉面” in real time. You are recommended to use YoloV4 (an object detection approach) for this task. The project page for YoloV4 is here https://github.com/AlexeyAB/darknet.

The test video is given on the course website. You only need to send the video with detection labels (similar as the figure given above) to TA.