ME598/494 Homework 3

1. (40 Points, problem from Professor M. Kokkolaras, McGill University) Vapor-liquid equilibria data are correlated using two adjustable parameters A_{12} and A_{21} per binary mixture. For low pressures, the equilibrium relation can be formulated as:

$$\begin{aligned} p &= x_1 \exp\left(A_{12}\left(\frac{A_{21}x_2}{A_{12}x_1 + A_{21}x_2}\right)^2\right) p_1^{sat} \\ &+ x_2 \exp\left(A_{21}\left(\frac{A_{12}x_1}{A_{12}x_1 + A_{21}x_2}\right)^2\right) p_2^{sat}. \end{aligned} \tag{1}$$

   Here the saturation pressures are given by the Antoine equation

$$\log(p^{sat}) = a_1 - \frac{a_2}{T + a_3}, \tag{2}$$

   where $T = 20(°\text{C})$ and $a_{1,2,3}$ for a water - 1,4 dioxane system is given below.

	a_1	a_2	a_3
Water	8.07131	1730.63	233.426
1,4 dioxane	7.43155	1554.679	240.337

   The following table lists the measured data. Recall that in a binary system $x_1 + x_2 = 1$.

x_1	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
p	28.1	34.4	36.7	36.9	36.8	36.7	36.5	35.4	32.9	27.7	17.5

   Estimate A_{12} and A_{21} using data from Table 1: (1) Formulate the least square problem; (2) solve using your own gradient descent or Newton's implementation; (3) solve using Matlab functions "lsqnonlin" or "lsqcurvefit".

   Solution Let $p_i(A_{12}, A_{21})$ be the model for the ith input, and p_i be the ith output, the least square problem is as follow:

$$\min_{A_{12},A_{21}} \quad \sum_{i=1}^{11}(p_i(A_{12}, A_{21}) - p_i)^2. \tag{3}$$

   The gradient can be calculated using chain role, or through finite difference. See code.

2. (30 Points) Download the data homework3data.mat. The data contains a set of topologically optimal brackets. Each row of "X" represents a bracket structure and "y" the corresponding loading force angle. Build a metamodel using "X" as input and "y" as

output, e.g., using a neural network. Instruction: If your Matlab does not come with the neural net toolbox, you can find it from the CAD lab (on GWC 4th floor). The neural net toolbox automatically reports test performance.

Solution Not much to say.

3. (20 Points, Optional for MAE494) Logistic regression is commonly used to approximate systems with binary outputs, e.g., the result of an election, a football game, a purchase on Amazon, etc. The mathematical form of a logistic regression model is as follows:

$$p(y; \mathbf{x}, \theta) = \frac{1}{1 + \exp(-y\theta^T \mathbf{x})}, \tag{4}$$

where y is the output that takes either 1 or -1, $\mathbf{x}$ are the input variables (covariates), θ are the unknown parameters. Given a dataset $D = \{(\mathbf{x}_i, y_i)\}_{i=1}^N$, the likelihood of θ can be written as

$$L(\theta; D) = \prod_{i=1}^{N} \frac{1}{1 + \exp(-y_i\theta^T \mathbf{x}_i)}. \tag{5}$$

Is the maximum likelihood estimate of θ unique? Hint: To obtain the maximum likelihood estimate of θ, one needs to minimize the negative log-likelihood function.

Solution The negative log-likelihood is

$$f(\theta) = -\log L(\theta; D) = \sum_{i=1}^{N} \log(1 + \exp(-y_i\theta^T \mathbf{x}_i)). \tag{6}$$

The gradient is

$$\nabla_\theta f = \sum_{i=1}^{N} \frac{-y_i \mathbf{x}_i}{1 + \exp(y_i\theta^T \mathbf{x}_i)}. \tag{7}$$

The Hessian is

$$\nabla_\theta^2 f = \sum_{i=1}^{N} \frac{\exp(y_i\theta^T \mathbf{x}_i)\mathbf{x}_i\mathbf{x}_i^T}{(1 + \exp(y_i\theta^T \mathbf{x}_i))^2}. \tag{8}$$

Let $\alpha_i = \frac{\exp(y_i\theta^T\mathbf{x}_i)}{(1+\exp(y_i\theta^T\mathbf{x}_i))^2}$, we have $\nabla_\theta^2 f = \sum_{i=1}^N \alpha_i \mathbf{x}_i\mathbf{x}_i^T$. With any non-zero vector $\mathbf{w}$, we have $\mathbf{w}^T\nabla_\theta^2 f\mathbf{w} = \sum_{i=1}^N \alpha_i \mathbf{w}^T\mathbf{x}_i\mathbf{x}_i^T\mathbf{w} = \sum_{i=1}^N \alpha_i (\mathbf{w}^T\mathbf{x}_i)^2$. Since $\alpha_i > 0$ by definition, and $(\mathbf{w}^T\mathbf{x}_i)^2 > 0$ for arbitrary $\mathbf{w}$, we have $\mathbf{w}^T\nabla_\theta^2 f\mathbf{w} > 0$. Thus $\nabla_\theta^2 f$ is positive definite everywhere, and f is strictly convex.

4. (10 Points) Please go through this tutorial on creating a metamodel using a deep convolutional neural network to predict Young's modulus of sandstone structures. There is an instruction on installing Keras and Tensorflow. Please attach your results.

Solution Not much to say. I will give you full credit if you struggle a lot and can't get the code to run...