Advanced Studies In Mathematics Exercise

Hwijae Son

November 14, 2024

1. Let

$$\mathcal{N}(x|\mu,\Sigma) = \frac{1}{\sqrt{(2\pi)^d \det(\Sigma)}} \exp(-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)),$$

be a p.d.f of a $d-$dimensional Gaussian distribution. Let $x_1, ..., x_N \sim \mathcal{N}(x|\mu,\Sigma)$ be random samples. Show that the Maximum Likelihood Estimators(MLE) of the mean $\mu \in \mathbb{R}^d$ and covariance $\Sigma \in \mathbb{R}^{d\times d}$ given are given by :

$$\hat{\mu}_{MLE} = \frac{1}{N}\sum_{i=1}^{N} x_i,$$
$$\hat{\Sigma}_{MLE} = \frac{1}{N}\sum_{i=1}^{N}(x_i - \hat{\mu}_{MLE})(x_i - \hat{\mu}_{MLE})^T.$$

2. Show that the sigmoid function $\sigma(x) = \frac{1}{1+e^{-x}}$ satisfies $\sigma'(x) = \sigma(1-\sigma)$, $\sigma(-a) = 1-\sigma(a)$, and find the inverse function of $\sigma(x)$.

3. Compute the gradient $\nabla_w \mathcal{J}$ and Hessian $\nabla_w^2 \mathcal{J}$ of the binary cross entropy loss function for the logistic regression:

$$\mathcal{J} = -\sum_{i=1}^{N} y_i \log(\sigma(w^T x_i)) + (1-y_i)\log(1-\sigma(w^T x_i)),$$

where $w, x_i \in \mathbb{R}^d, y_i \in \mathbb{R}$. Show that $\nabla_w^2 J$ is positive semi-definite.

4. (Python) Download the iris dataset from scikit-learn. Write Python programs that performs the following.
(a) Gaussian naïve Bayes classification with the conditional independence assumption:

$$p(x_1, \ldots, x_d|C_k) = \Pi_{i=1}^{d} p(x_i|C_k).$$

(b) Gaussian naïve Bayes classification without the conditional independence assumption.
(c) Multi-class logistic regression via gradient descent.
(d) Split the dataset into train/test sets consisting of 80%/20% of samples. Train each classifier above on the train set and compute the classification accuracy on the test set.