Problem Set 5

CS 6375

Due: 4/24/2020 by 11:59pm

Note: all answers should be accompanied by explanations for full credit. Late homeworks will not be accepted.

Problem 1: Poisson Maximum Likelihood Estimation (20pts)

Consider a nonnegative, integer-valued random variable X that is distributed according to a Poisson distribution $X \sim \frac{\lambda^x e^{-\lambda}}{x!}$ for some real-valued parameter $\lambda > 0$.

1. Given data samples $x^{(1)}, \ldots, x^{(m)}$, what is the maximum likelihood estimate for λ?
2. Suppose now that you introduce a prior probability distribution, $\lambda \sim \frac{1}{5} \max\{-\lambda/10 + 1, 0\}$. What is the MAP estimate under this prior probability distribution?
3. Why might you not prefer a prior probability distribution like the above for this estimation task? What might be a better prior?

Problem 2: Log-Normal Maximum Likelihood Estimation (20pts)

Consider a positive, real-valued random variable X that is distributed according to a log-normal distribution: $X \sim \frac{1}{x\sigma\sqrt{2\pi}} \exp(-\frac{(\ln x - \mu)^2}{2\sigma^2})$ for real-valued parameters μ and σ with $\sigma > 0$. Suppose that you are given M data observations $x^{(1)}, \ldots, x^{(M)}$.

1. For a given μ and σ, what is the log-likelihood of the data observations?
2. Find the maximum likelihood estimators for μ and σ. Note that the log-likelihood is not a concave function of μ and σ.
3. Are the maximum likelihood estimators for μ and σ unbiased?
4. Why might a Gaussian distribution not be a good choice for a prior distribution over μ?

Problem 3: Logistic Regression (30pts)

For this problem, consider the Sonar data sets from Homework 4.

1. Fit a logistic regression classifier to training data set. What is the accuracy on the test set? Explain why in standard logistic regression, without any type of regularization, the weights may not converge (even though the predicted label for each data point effectively does) if the input data is linearly separable.

2. Fit a logistic regression classifier with an ℓ_2 penalty on the weights to this data set using the validation set to select a good choice of the regularization constant. Report your selected constant, the learned weights and bias, and the accuracy on the test set.

3. Fit a logistic regression classifier with an ℓ_1 penalty on the weights to this data set using the validation set to select a good choice of the regularization constant. Report your selected constant, the learned weights and bias, and the accuracy on the test set.

4. Does ℓ_1 or ℓ_2 tend to produce sparser weight vectors?

Problem 4: Gaussian Naïve Bayes (30pts)

For this problem, consider the Sonar data set from Homework 4. Suppose that you want to fit a Gaussian NB model to this data. That is, assume that the probability distribution is of the form

$$p(x_1, \ldots, x_n, y) = p(y) \prod_{i=1}^{n} p(x_i|y)$$

where $p(x_i|y)$ is a distinctly parameterized normal distribution, i.e., $p(x_i|y)$ is parameterized by $\mu_{i,y}$ and $\sigma_{i,y}$.

1. Given a data set with m continuous features, what is the log-likelihood of the Gaussian NB model? Compute the MLE for each of the model parameters.

2. Fit a Gaussian NB model to the training data. What is the accuracy of your trained model on the test set?

3. What kind of prior might make sense for this model? Explain.

4. Do you think the NB assumption is reasonable here?