Practice Problems Week-04: Logistic Regression & Softmax Regression

Instructions: Solve the following problems step-by-step. Show all your work.

1. **Sigmoid Calculation:** Calculate the sigmoid function output for the following values: $z = -2, 0, 1.5$.

2. **Simple Logistic Prediction:** A simple logistic regression model for spam classification is defined as $P(\text{Spam}|x) = \sigma(-2 + 0.5x)$, where x is the number of exclamation marks in an email. Predict the probability that an email with $x = 6$ exclamation marks is spam.

3. **Decision Boundary:** For the model $P(Y = 1|x) = \sigma(3 - 2x)$, find the value of x where the decision boundary lies (i.e., where $P(Y = 1|x) = 0.5$).

4. **Log Loss Calculation (Single Point):** The true label is $y = 1$ and the model predicts $p = 0.85$. Calculate the log loss for this single data point.

5. **Log Loss Calculation (Multiple Points):** Calculate the total log loss for three data points: (True: 1, Pred: 0.9), (True: 0, Pred: 0.2), (True: 1, Pred: 0.6).

6. **Two-Feature Prediction:** A logistic regression model has parameters $\theta_0 = -1, \theta_1 = 0.4, \theta_2 = -0.8$. For a new sample with features $x_1 = 2, x_2 = 3$, calculate the probability $P(Y = 1|\mathbf{x})$.

7. **Cost Function Derivative:** The derivative of the log loss cost $J(\theta)$ with respect to a parameter θ_j has a term $(y - \hat{y})x_j$. For a point where $y = 1$, $\hat{y} = 0.7$, and $x_j = 2$, calculate the value of this term. What does its sign indicate?

8. **Softmax Logits:** For a 3-class problem, the logits for a data point are $[z_1, z_2, z_3] = [1.2, -0.5, 0.8]$. Calculate the final probability distribution using the softmax function.

9. **Softmax Prediction:** Using the probabilities from the previous question, which class would be predicted?

10. **Softmax Constant Shift:** Show that subtracting a constant (e.g., the maximum value) from all logits does not change the softmax output. Use the logits from question 8 and subtract the max value (1.2) before recalculating softmax.

11. **Multi-class Log Loss:** For a single data point belonging to class 2 (indexing starts at 1), the predicted probabilities are $[0.1, 0.7, 0.2]$. Calculate the multi-class log loss for this point.

12. **Parameter Update Interpretation:** A logistic regression model is being trained with a learning rate $\eta = 0.1$. For a data point, the gradient for weight θ_1 is calculated as -2.5. What does this gradient value imply, and what will be the update to θ_1?

13. **Effect of Feature Scaling:** Why is feature scaling (standardization) often recommended for logistic regression, especially when using gradient descent?

14. **Overfitting Prevention:** Name two common techniques used to prevent overfitting in logistic regression models.

15. **Softmax with High Confidence:** A softmax model for a 4-class problem outputs probabilities $[0.94, 0.02, 0.03, 0.01]$. Calculate the multi-class log loss if the true class is the first one.

16. **Softmax Symmetry:** In a binary classification scenario, show that the softmax function reduces to the sigmoid function.

17. **Complex Decision Boundary:** A logistic regression model uses the features x_1 and x_2 and their interaction term x_1x_2. The model is $P(Y = 1|\mathbf{x}) = \sigma(1 + 2x_1 - 3x_2 + 0.5x_1x_2)$. For a point with $x_1 = 1, x_2 = 1$, calculate the probability.

18. **Gradient Calculation:** For a logistic regression model, compute the gradient of the log loss with respect to θ_0 for a single data point where $y = 1$, $\hat{y} = 0.3$, and $x_0 = 1$ (the bias term).

19. **Regularization Effect:** If we add L2 regularization with $\alpha = 0.1$ to a logistic regression model, and the current value of θ_1 is 2.5, what would be the additional term in the gradient for θ_1 due to regularization?

20. **Multi-class Decision:** In a 3-class softmax regression problem, the computed probabilities for an instance are $[0.1, 0.6, 0.3]$. What is the predicted class, and what is the probability of the true class if the instance actually belongs to class 3?