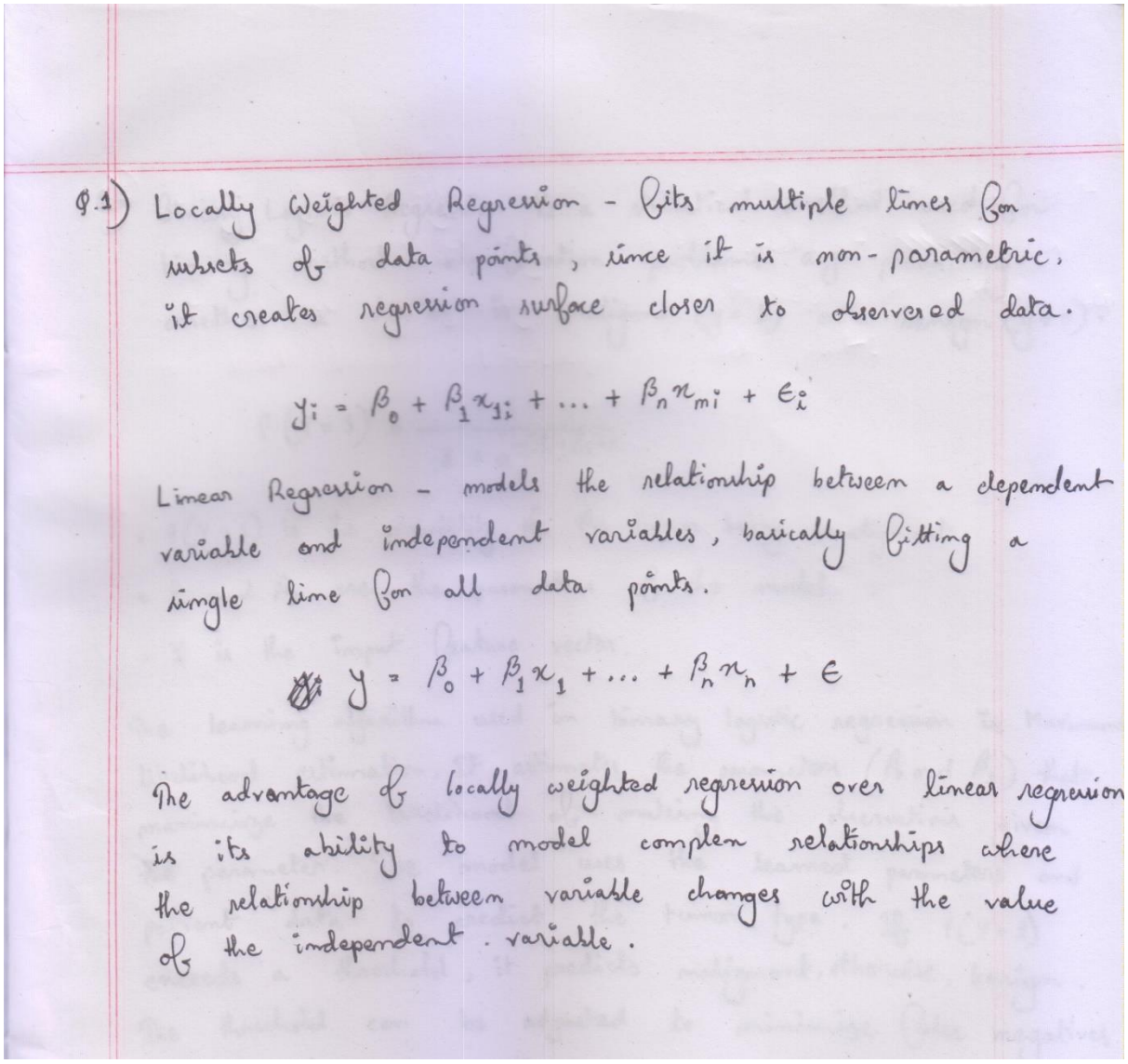
Name: Rezwan Ahmad ID:20-44273-3 Section: C

Final Assignment

Q1) Explain how locally weighted regression differs from linear regression, including their formulas. What is an advantage of locally weighted regression over linear regression? [2 points]



Q.1) Locally Weighted Regression - fits multiple lines for subsets of data points, since it is non-parametric, it creates regression surface closer to observered data.

$$y_i = \beta_0 + \beta_1 x_{1i} + \ldots + \beta_n x_{mi} + \epsilon_i$$

Linear Regression - models the relationship between a dependent variable and independent variables, basically fitting a single line for all data points.

$$y = \beta_0 + \beta_1 x_1 + \ldots + \beta_n x_n + \epsilon$$

The advantage of locally weighted regression over linear regression is its ability to model complex relationships where the relationship between variable changes with the value of the independent variable.

Q2) Given you want to apply a model to predict whether a patient has malignant or benign tumour, where model output y = 1 means malignant and y = 0 means benign. Explain how the binary logistic regression model is used to train on patient data and then predict tumour of a new patient. Include formulas and learning algorithm used in your answer. [2 points]

Q.2) Binary Logistic Regression is a statistical method used for binary method classification problems e.g predicting whether a tumor is maligant ($y = 1$) on benign ($y = 0$).

$$P(Y = 1) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X)}}$$

- $P(Y = 1)$ is the probability of the tumor being malignant.
- β_0 and β_1 are the parameters of the model.
- X is the input feature vector.

The learning algorithm used in binary logistic regression is Maximum likelihood estimation. It estimates the parameters (β_0 and β_1) that maximize the likelihood of making the observations given the parameter. The model uses the learned parameters and patient data to predict the tumor type. If $P(Y = 1)$ exceeds a threshold, it predicts malignant, otherwise, benign. The threshold can be adjusted to minimise false negatives in medical scenarios.

Q3.a) Given the output, y (n), of 3 training items of softmax regression are represented by the following one-hot vectors where y ϵ {1,2,3}: y1 = [1 0 0], y2 = [0 1 0] and y 3 = [0 0 1]. Write the expanded form of the softmax cost function J(w) for these 3 items, and the softmax output function f(x;w). [2 points]

b) What is the relationship between softmax and binary logistic regression? [1 point]

Q.3 (a) Softmax Output Function $f(x;w)$ normalizes a vector of K real numbers into a probability distribution

$$f(x_i;w) = \frac{e^{w^T x_i}}{\sum_{j=1}^{K} e^{w^T x_j}}$$

Softmax Cost Function $J(w)$ is the cross entropy loss used in softmax regression

$$J(w) = -\sum_{i=1}^{N}\sum_{j=1}^{K} y_{ij} \log[f(x_i;w)_j]$$

For the given one-hot vectors $y1 = [1\ 0\ 0]$, $y2 = [0\ 1\ 0]$ and $y3 = [0\ 0\ 1]$, the expanded form of the softmax cost function would be

$$J(w) = -[y_{11}\log(f(x_1;w)_1) + y_{21}\log(f(x_2;w)_2) + y_{31}\log(f(x_3;w)_3)]$$

(b) Softmax regression is a generalization of binary logistic regression for multi-class classification. Binary logistic regression models the probability of a binary outcome, while softmax regression models the probability for multiple outcomes. When softmax regression is applied to a binary classification problem, it reduces to logistic regression.

Q4) What is the penalty term of ridge/L2 regularization and how does it reduce overfitting? [1 point]

Q.4) The penalty term of Ridge or L2 regularization is:

$$\lambda \sum_{i=1}^{n} w_i^2$$

It reduces overfitting by discouraging the model from learning complex patterns, which might be noise in the training data. This is achieved by keeping the model parameters small, thus encouraging the model to learn simpler, more generalizable patterns. The λ ~~patter~~ parameter controls the strength of the regularization. A larger λ means more regularization and simpler models.

Q5.a) Write the pseudocode/steps of applying Policy iteration to solve an MDP, including the equations. [1 point]

b) What is the advantage of using an exploration-based policy like ϵ-greedy, to solve an MDP? [1 point]

Q.5) (a) Steps for applying Policy Iteration to solve MDP are,

(1) Initialization of random policy π

(2) Evaluating the policy for each state (s) in the MDP to calculate the state value function $V(s)$ under policy using the Bellman equation for deterministic policies:

$$V(s) = \sum_{s',r} p(s',r \mid s,\pi(s))[r + \gamma V(s')]$$

Repeat this step until $V(s)$ converges.

(3) Policy improvement updates the policy based on the current value function:

$$\pi(s) = \underset{a}{\operatorname{argmax}} \sum_{s',r} p(s',r \mid s,a)[r + \gamma V(s')]$$

(4) Policy stability check ensuring that if the policy does not change during the improvement step, stop and return the optimal policy and value function. Otherwise go back to the policy evaluation step.

(b) The ϵ-greedy policy in solving an MDP balances exploration and exploitation. It promotes early exploration for information gain and ensures continued exploration with ϵ probability of choosing random actions. This leads to better long term decisions due to a more comprehensive understanding of the environment.

Q6.a) What makes Q-learning an off-policy algorithm? [1 point]

b) What is the difference between on-policy and off-policy algorithms? [1 point]

Q.6 (a) Q-learning is an off-policy algorithm as it learns the optimal policy's value regardless of the agent's actions. It uses a greedy policy based on current Q-values for updates, but does not strictly follow this policy during learning. This means the policy used to select actions can be different from the policy that is evaluated and improved, which is the definition of an off-policy algorithm.

(b) On-policy algorithms, like ~~SAS~~ SARSA, learn and estimate the value of the current policy ~~used~~ in use. Off-policy algorithms, such as Q-learning, evaluate or improve a different policy and can learn the optimal policy irrespective of the agent's actions.