

MA477: Data Science
Lesson 12 Board Sheet — 06 February 2026
United States Military Academy, West Point
Instructor: MAJ Patrick Kuiper

1 Logistic Regression Lesson Objectives

- Understand why linear regression is not appropriate for a qualitative response.
- Interpret the coefficients of a multiple logistic regression model and obtain predictions for a set of inputs.
- Use functions in the sklearn module to assess logistic regression models.

Discussion Questions

We will organize this lecture around the following five questions:

1. Describe some of the issues with using a linear regression type model for classification problems - what portion of logistic regression is linear?
2. What function is applied to the linear transformation of the observed features, and how does this function operate?
3. What probability distribution do we use to model do we use to define a error function with Logistic Regression?
4. What is the loss function used for Logistic Regression?
5. How do we minimize this loss?
6. What do we do when we what to consider multiple classes?

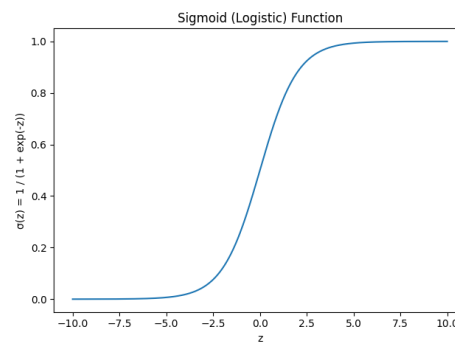


Figure 1

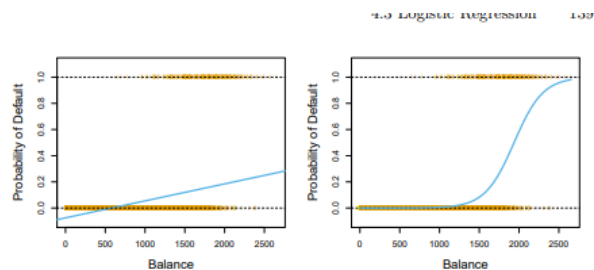


FIGURE 4.2. Classification using the `Default` data. Left: Estimated probability of default using linear regression. Some estimated probabilities are negative! The orange ticks indicate the 0/1 values coded for `default` (No or Yes). Right: Predicted probabilities of default using logistic regression. All probabilities lie between 0 and 1.

Figure 2