Module 10.5: From Regression to Logistic Regression with Sigmoid Function

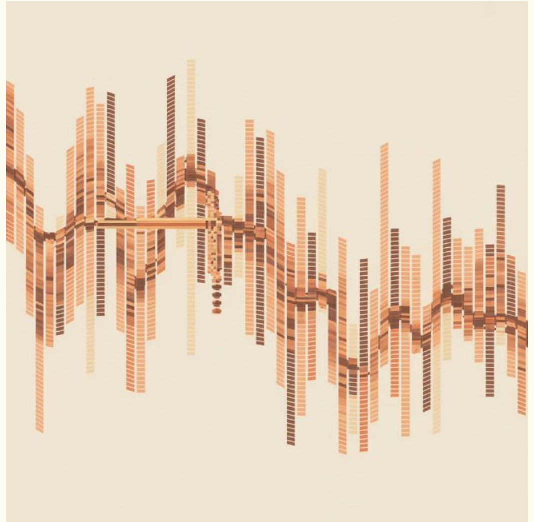
Understanding the Sigmoid Function and Decision Boundary

Let's explore how we move from predicting continuous values to making binary classifications using logistic regression.

Why Linear Regression Fails for Classification

Linear regression produces **unbounded outputs** like -2, 1.7, or 5.1. But classification needs clear categories like Pass/Fail or Yes/No.

When we try to predict binary outcomes, the regression line doesn't respect class boundaries. It shoots past them, creating predictions that don't make sense for classification tasks.



Problem 1

Outputs can be any number

Problem 2

No natural class boundaries

Problem 3

Hard to interpret predictions

We Need Probabilities Instead

Classification tasks require outputs between 0 and 1 that represent the probability of belonging to a class. This gives us a clear, interpretable prediction we can act on.

Diabetes Prediction

Will this patient develop diabetes? We need a probability score between 0% and 100% to make informed medical decisions.

Spam Detection

Is this email spam? A probability helps the system decide whether to filter it or let it through to your inbox.

Loan Approval

Will the applicant repay? Banks use probability scores to assess risk and make lending decisions confidently.

Introducing Logistic Regression

Logistic regression transforms any linear score into a probability using the **sigmoid function**. This elegant transformation ensures our output always stays between 0 and 1.

Here's how it works in practice:

1

Step 1: Linear Score

Calculate a score (like 1.2) from input features

2

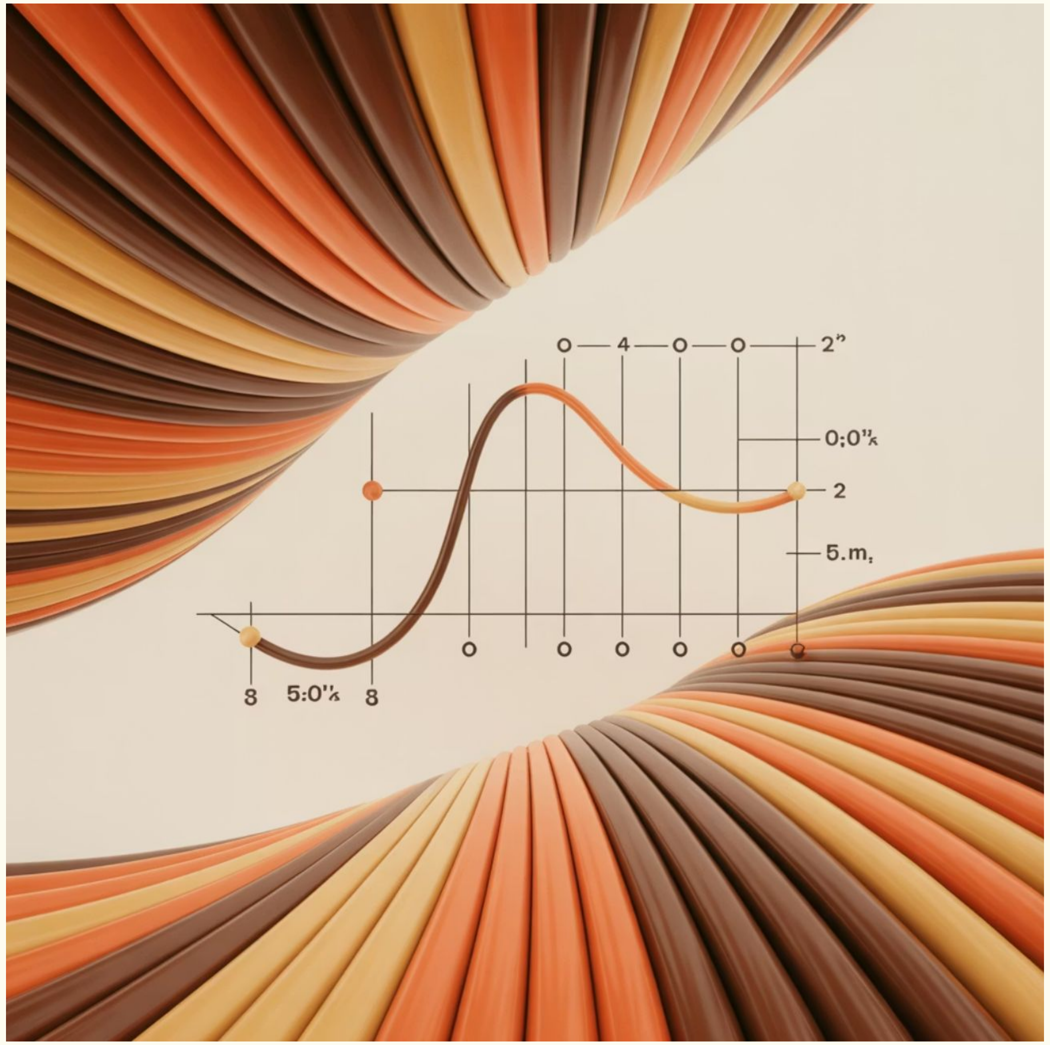
Step 2: Apply Sigmoid

Transform score to probability (0.77)

3

Step 3: Interpret

"77% chance of passing"



The Sigmoid Function Visualized

The sigmoid function creates a smooth **S-shaped curve** that maps any input to a probability between 0 and 1. This is the mathematical heart of logistic regression.

Left Region (Near 0)

Strong prediction for Class 0

Example: Very low probability of passing

Middle Region (Around 0.5)

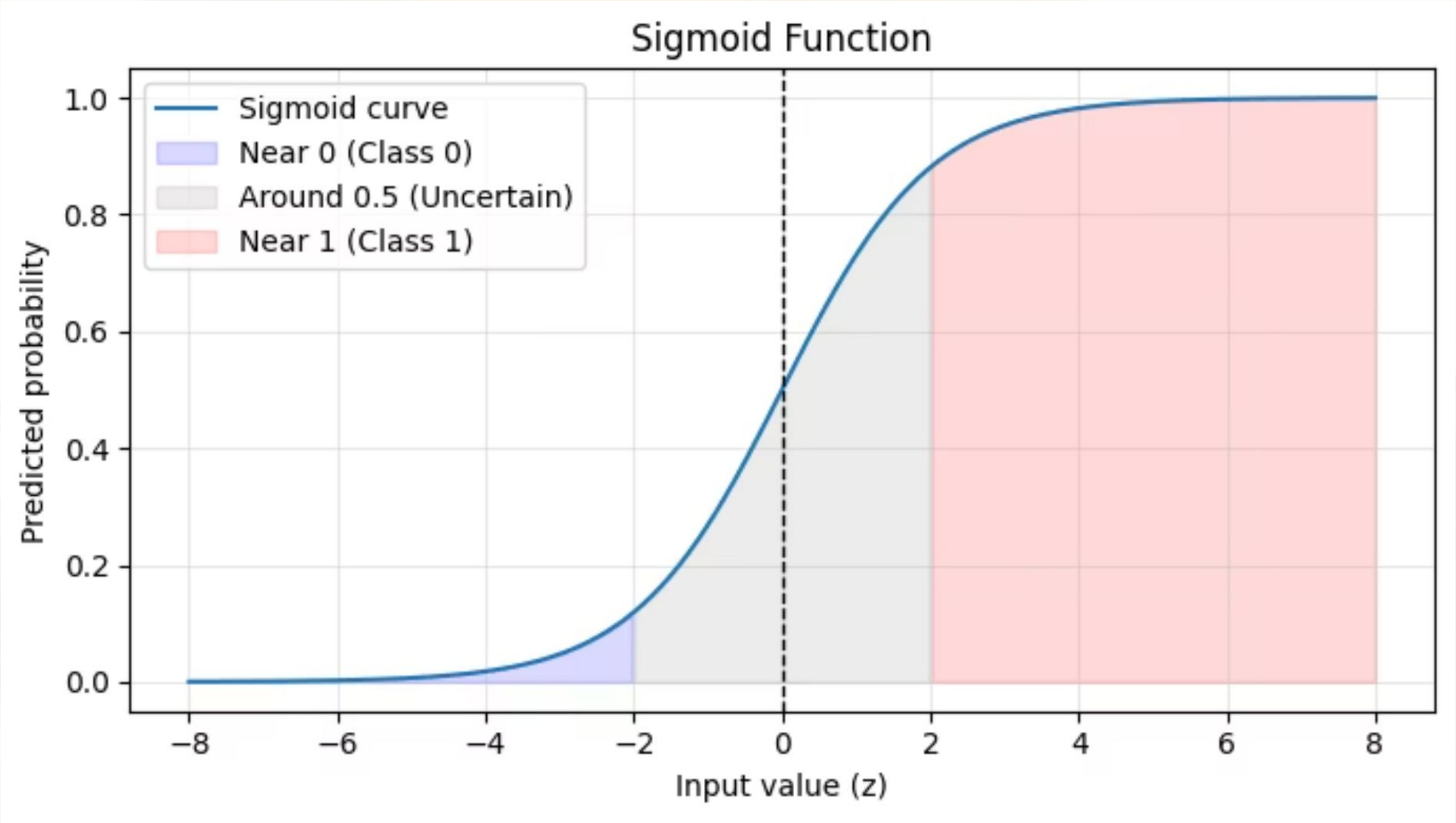
Uncertain prediction

Example: 50-50 chance, needs more information

Right Region (Near 1)

Strong prediction for Class 1

Example: Very high probability of passing



Real Example: Hours Studied

Let's walk through a concrete example with actual numbers to see how logistic regression works in practice.

1

Input Feature

Student studied for **8 hours**

2

Linear Score

Our model calculates: **1.3**

3

Sigmoid Output

Transformed to: **0.785**

4

Interpretation

78.5% chance the student will pass

With nearly 79% probability of passing, we'd classify this student as likely to succeed on the exam.

The Decision Boundary

We need a **threshold** to convert probabilities into final class predictions. The standard choice is 0.5, which creates a decision boundary.

If $p \geq 0.5$

Predict **Class 1**

Example: Pass the exam

If $p < 0.5$

Predict **Class 0**

Example: Fail the exam

Real example: A probability of 0.62 means we predict "Pass" because $0.62 \geq 0.5$. The model is confident this student will succeed!

Visualizing the Decision Boundary

In two dimensions, the decision boundary appears as a **straight line** that separates our two classes. This line represents where the probability equals exactly 0.5.

Blue Cluster (Class 0)

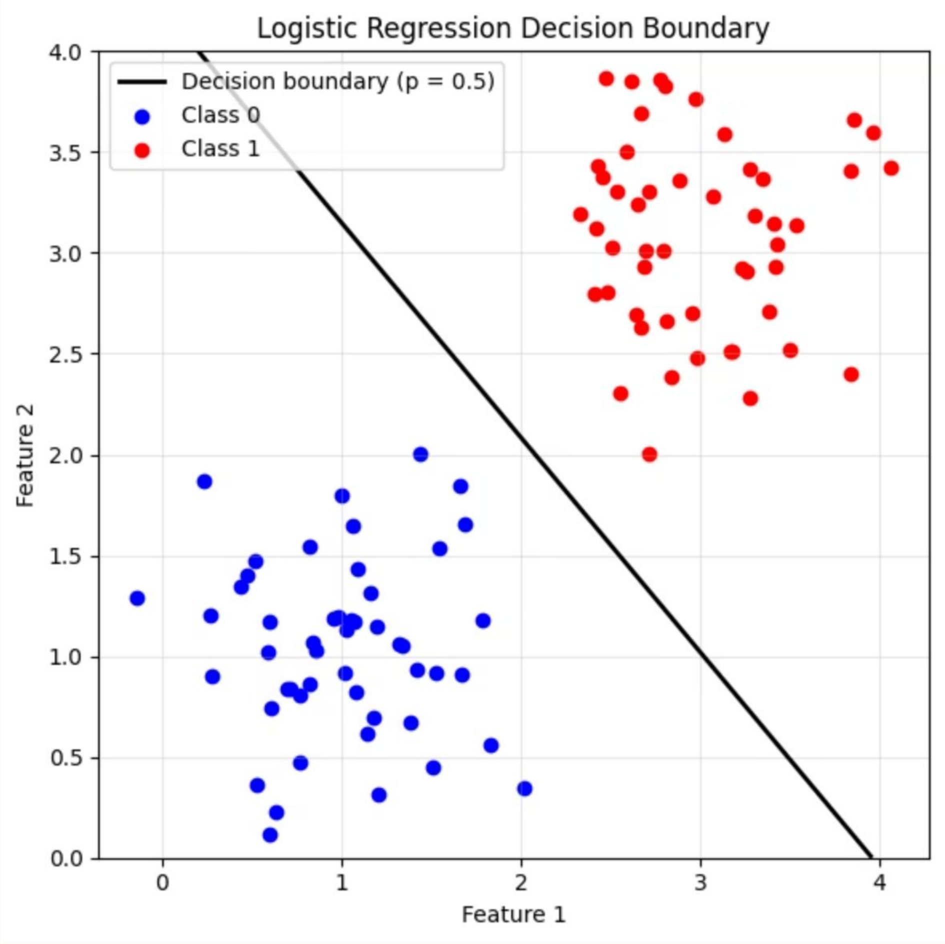
Points on this side have probability < 0.5

The model predicts these belong to the negative class

Red Cluster (Class 1)

Points on this side have probability ≥ 0.5

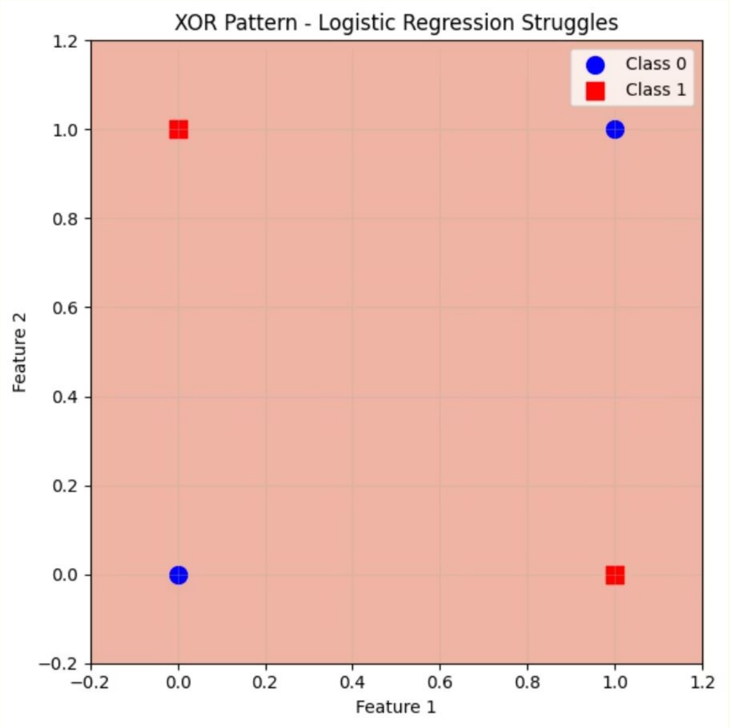
The model predicts these belong to the positive class



When Logistic Regression Struggles

Logistic regression creates **linear decision boundaries**. This works great when classes can be separated by a straight line, but fails when the relationship is non-linear.

The classic example is the **XOR pattern**, where classes are arranged in a checkerboard pattern. No single straight line can separate them!



1 Limitation

Can only draw straight boundaries

2 Solution Needed

More complex patterns require neural networks or other advanced methods

Linear vs. Logistic Regression

Let's compare these two fundamental techniques side by side to understand their key differences and when to use each one.

Aspect	Linear Regression	Logistic Regression
Output Type	Any real number (-∞ to +∞)	Probability (0 to 1)
Best For	Predicting continuous values	Classification tasks
Decision Boundary	No boundary concept	Clear threshold at 0.5
Interpretation	Direct numerical prediction	Probability of class membership
Example Use	Predict house price	Predict spam/not spam

Key Takeaways

1

Classification Needs Special Tools

Linear regression fails for classification because it produces unbounded outputs instead of probabilities

2

Sigmoid Makes It Work

The sigmoid function elegantly transforms any score into a probability between 0 and 1

3

Boundaries Define Classes

The decision boundary at probability 0.5 determines which class we predict

4

Foundation for Deep Learning

Logistic regression principles power neural networks and advanced AI systems