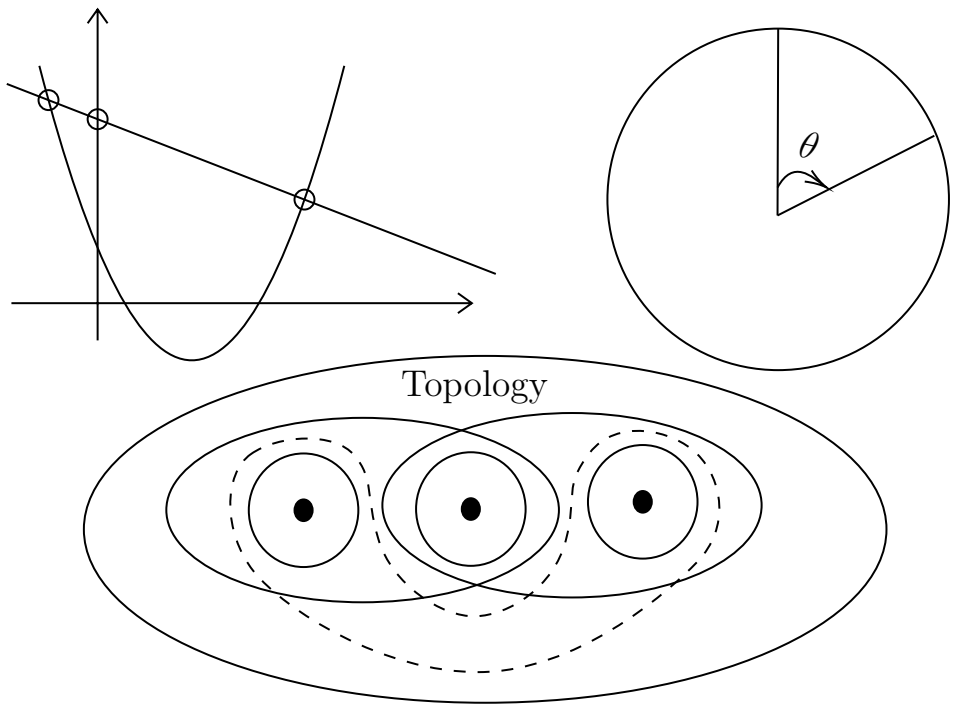
Introduction to Math for DS Group Mini-project

Analysis of factors affecting Premier League match results

IMDS Group 24

Zehao Qian, Chloe Mendez, Mohammad Jamshaid Iqbal



Introduction to Math for DS Group Mini-project

IMDS Group 24
Zehao Qian, Chloe Mendez, Mohammad Jamshaid Iqbal

December 6, 2023

Contents

1 Introduction

2 Model Assumptions

2.1 Performance Assumptions (Correlation Analysis)

- The number of wins positively correlates with the final league standing.
- Teams with a higher goal difference (GF - GA) tend to achieve higher league positions.
- Drawn matches have a minimal impact on final league standings.
- Teams with a higher number of goals scored (GF) are more likely to finish in the top positions.
- The defensive performance, measured by goals against (GA), influences the team's final standing.
- The number of points earned directly correlates with the team's final position in the league.

2.2 Consistency Hypothesis (Principal Component Analysis)

- Consistency in performance, as measured by a balanced distribution of wins, draws, and losses, is associated with a higher league position.

PCA can help identify patterns and relationships among these variables, which can contribute to understanding the consistency in team performance.

2.3 Historical Performance Hypothesis (Entropy Weighting)

- Teams with a consistent performance history over the years are likely to maintain their competitive positions.

3 Data

4 Methods

4.1 Data Feature Extraction with Fourier Transformation

4.2 Correlation Analysis

4.3 Principal Component Analysis

4.4 Comparison: Entropy Weight Method

5 Conclusions