Discovering Hidden Prerequisites

Anumat Srivastava, Lowell Bander, Paul Jayazeri

MATH 191 Topics in Data Science, UCLA

December 1, 2015

Outline

Introduction

- The objective is to use ranking algorithms to discover hidden prerequisites in the Math and Computer Science curriculums
- We have access to data containing 13,000+ students' GPA's and the order in which they have taken their classes
- We will be using the rank centrality and serial rank algorithms to rank the courses
- By contrasting the different orderings of classes taken for high gpa students vs. low gpa students, we should be able to find courses that should precede other courses

Outline

Serial Rank

- For Serial Rank, we use a matrix C of size n x n of pairwise comparisons (where n is the number of courses offered by the CS/Math departments), defined as follows :

$$C_{ij} = \begin{cases} 1 & \text{if course i is taken before course j (or i=j)} \\ 0 & \text{for i and j are tied or have no available comparison} \\ -1 & \text{if course j is taken before course i} \end{cases} \tag{1}$$

The pairwise similarity matrix is constructed as follows

$$S_{ij}^{match} = \sum_{k=1}^{n} \left(\frac{1 + C_{ik} C_{jk}}{2} \right)$$

Then compute the Laplacian matrix $L_S = diag(S1) - S$ and finally computing and sorting the Fiedler vector of S will give the ranking.

Serial Rank

- For Serial Rank, we shall split the data into sets of high gpa students and low gpa students.
- Then we build similarity matrices for each student in both sets as described in the previous slide
- We combine all the similarity matrices into 1 normalized similarity matrix for each set and look at the pairwise difference of each entry in the two matrices.
- If the difference is large, the course taken by the high gpa students should be a prerequisite to the course taken by the low gpa students!
- We then use these two similarity matrixes to rank courses according to the algorithm described.

Rank Centrality

Outline

Summary of our work

- We hope to discover some hidden prerequisites in the curriculums and improve the experience for future UCLA students by presenting our findings to the respective departments