

**Operations Research and Engineering Management
Seminar Series**

Research Seminar

**Subspace Clustering with Missing Data
via Integer Programming**



Dr. Jeff Linderoth
Harvey D. Spangler Professor
Department of Industrial and Systems Engineering
University of Madison-Wisconsin
Friday, March 25, 2022
11:00 a.m. – 12:15 p.m.
<https://smu.zoom.us/j/97906267193>

Abstract

In the Subspace Clustering with Missing Data (SCMD) problem, we are given a collection of n data points, arranged into columns of a matrix X , and each of the data points is observed only on a subset of its coordinates. The data points are assumed to be concentrated near a union of low-dimensional subspaces. The goal of SCMD is to recover the missing elements of the data matrix X . State-of-the-art algorithms for SCMD can fail on instances with a large amount of

missing data or if the data matrix X is nearly full rank. We propose a novel integer programming-based approach for SCMD. The approach is based on dynamically determining a set of candidate subspaces and optimally assigning points to selected subspaces. The problem structure is identical to the classical facility-location problem, with subspaces playing the role of facilities and data points that of customers. We propose a column-generation approach for identifying candidate subspaces combined with a Benders decomposition approach for solving the linear programming relaxation of the formulation. An empirical study demonstrates that the proposed approach can achieve better clustering accuracy than state-of-the-art methods when the data is high rank, the percentage of missing data is high, or there are a small number of data points in each subspace.

This is joint work with Jim Luedtke, Daniel Pimentel-Alarcon and Akhilesh Soni, all of UW-Madison.

Biography

Jeff Linderoth is the Harvey D. Spangler Professor of the department of Industrial and Systems Engineering at the University of Wisconsin-Madison. Prof. Linderoth holds a courtesy appointment in the Computer Sciences department and as a Discovery Fellow in the Optimization Theme at the Wisconsin Institutes of Discovery. Dr. Linderoth received his Ph.D. degree from the Georgia Institute of Technology in 1998. He was previously employed in the Mathematics and Computer Science Division at Argonne National Laboratory, with the optimization-based financial products firm of Axioma, and as an Assistant Professor at Lehigh University. His awards include an Early Career Award from the Department of Energy, the SIAM Activity Group on Optimization Prize, and the INFORMS Computing Society (ICS) Prize. He 2016, he was elected to membership as an INFORMS Fellow.