www.elsevier.com/locate/caor

Discrete optimization methods to fit piecewise affine models to data points

E. Amaldi, S. Coniglio, L. Taccari

PII: S0305-0548(16)30102-2
DOI: http://dx.doi.org/10.1016/j.cor.2016.05.001
Reference: CAOR3995

To appear in: *Computers and Operation Research*

Received date: 4 March 2015
Revised date: 26 April 2016
Accepted date: 1 May 2016

Cite this article as: E. Amaldi, S. Coniglio and L. Taccari, Discrete optimizatio methods to fit piecewise affine models to data points, *Computers and Operatio Research,* http://dx.doi.org/10.1016/j.cor.2016.05.001

Discrete optimization methods to fit piecewise affine models to data points

E. Amaldi[a], S. Coniglio[b,1], L. Taccari[a]

[a] *Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano, Italy*

[b] *Department of Mathematical Sciences, University of Southampton, University Road, Southampton, SO17 1BJ, UK*

Abstract

Fitting piecewise affine models to data points is a pervasive task in many scientific disciplines. In this work, we address the *k-Piecewise Affine Model Fitting with Piecewise Linear Separability problem* (k-PAMF-PLS) where, given a set of m points $\{\mathbf{a}_1, \dots, \mathbf{a}_m\} \subset \mathbb{R}^n$ and the corresponding observations $\{b_1, \dots, b_m\} \subset \mathbb{R}$, we have to partition the domain $\mathbb{R}^n$ into k piecewise linearly (or affinely) separable subdomains and to determine an affine submodel (function) for each of them so as to minimize the total linear fitting error w.r.t. the observations b_i.

To solve k-PAMF-PLS to optimality, we propose a mixed-integer linear programming (MILP) formulation where symmetries are broken by separating shifted column inequalities. For medium-to-large scale instances, we develop a four-step heuristic involving, among others, a point reassignment step based on the identification of critical points and a domain partition step based on multi-category linear classification. Differently from traditional approaches proposed in the literature for similar fitting problems, in both our exact and heuristic methods the domain partitioning and submodel fitting aspects are taken into account simultaneously.

Computational experiments on real-world and structured randomly generated instances show that, with our MILP formulation with symmetry breaking constraints, we can solve to proven optimality many small-size instances. Our four-step heuristic turns out to provide close-to-optimal solutions for small-size instances, while allowing to tackle instances of much larger size. The experiments also show that the combined impact of the main features of our heuristic

Email addresses: edoardo.amaldi@polimi.it (E. Amaldi), s.coniglio@soton.ac.uk (S. Coniglio), leonardo.taccari@polimi.it (L. Taccari)

[1] The work of S. Coniglio was carried out, for a large part, while he was with Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano and with Lehrstuhl II für Mathematik, RWTH Aachen University, supported. While with the latter, he was supported by the German Federal Ministry of Education and Research (BMBF), grant 05M13PAA, and Federal Ministry for Economic Affairs and Energy (BMWi), grant 03ET7528B.