On using an improved Benders method for cell suppression [1]

Daniel Baena*, Jordi Castro**, Antonio Frangioni***

* Dept. of Statistics and Operations Research, Universitat Politècnica de Catalunya, danibaena@gmail.com

** Dept. of Statistics and Operations Research, Universitat Politècnica de Catalunya, jordi.castro@upc.edu

*** Dipartimento di Informatica, Universit di Pisa, frangio@di.unipi.it

Abstract. The cell suppression problem (CSP) is one of the most widely applied methods for tabular data protection. Given a set of primary cells to be protected, CSP aims at finding a set of secondary cells to be additionally removed to guarantee that estimates of values of primary cells fall out of a predefined protection interval. From a computational point of view, CSP is very challenging even for tables of moderate size and number of primary cells. Currently, the only effective optimal approach for CSP is Benders decomposition (also known as cutting planes). However, the convergence to the optimal solution is often too slow due to well known instability issues of Benders decomposition. This work discusses a recently developed improved Benders method, which focus on finding new solutions in the neighborhood of "good" points. Some results are reported in the solution of realistic and real-world CSP instances, showing the effectiveness of this approach.

1 Introduction

Benders decomposition [1] is an iterative method that allows to decompose the original MILP in several smaller and theoretically easier to solve subproblems (referred to relaxed master and slaves) and after a finite number of iterations the method provides an optimal solution. Originally, this method was suggested for problems with two types of variables where one of them are considered as "complicating variables". In MILP models complicating variables are the binary/integer ones y. Although Benders decomposition is widely used in many real-world applications, it suffers from well-known instability issues that limit its efficiency. The convergence to the optimum is often too slow due to the fact that the solutions tend to oscillate wildly

[1]This is an brief summary of the paper "D. Baena, J. Castro, A. Frangioni, Stabilized Benders methods for large combinatorial optimization problems" which is currently under review.

among different feasible regions by jumping from a good point, i.e. close to optimality, to a much worse one.

This paper addresses this issue and proposes a stabilized Benders decomposition in order to prevent this behaviour. In particular, we focus on finding new solutions inside trust feasible regions, i.e. neighborhoods of well considered points, where we expect to find better solutions. The main cause for slow convergence is due to the generation of weak Benders cuts as a result of obtaining “bad” points y when we solve the master problem [3]. The idea behind the stabilized Benders decomposition is to search new solutions y as close as possible to properly chosen points, so called stability center points.

For binary MILPs, the stabilization can be done by adding linear constraints that restrict the feasible region of relaxed master problems. This is made possible by using the Hamming distance defined from a stability center point $(\bar{y})$, not necessarily feasible, and a radius $K \geq 1$ [6]. This restricted feasible region of size K is called trust region (TR). Note that K can be either a constant or dynamically updated at each iteration. TR is defined by a well-known local branching constraint which limits the "switching" of binary variables only at most K [4]. These local branching constraints prevent the master problem solution from moving too far from the stability center point $\bar{y}$.

2 Application to data privacy: Cell Suppression Problem

Cell Suppression Problem (CSP) formulates a very large MILP problem whose direct solution is impractical with state-of-the-art MILP solvers for tables of moderate size. Because of that, a Benders decomposition approach was suggested in the past for its solution [5]. In this paper we try to improve the behaviour of the Benders decomposition applied to CSP through the application of stabilization techniques. A series of computational experiments have been designed to empirically validate the efficiency of the proposed stabilized Benders decomposition for CSP. The numerical experiments have been performed on a set of real-world general and synthetic 1H2D tables. Real-world general tables are standard instances used in the literature [2]. Synthetic instances were obtained with a generator of random 1H2D tables. A total of 48 randomly 1H2D instances and 15 real-world tables were considered. The results have been successful and show the stabilized Benders applied to CSP to be an excellent strategy compared to the state-of-the-art classical Benders of [5]. In 92% of the synthetic 1H2D tables, stabilized Benders outperformed classical Benders in terms of both CPU time and gap of the feasible solution found. With stabilized strategy, the average GAP was 0.87% whereas for classical CSP Benders was 2.51%. Moreover, the stabilized Benders was 1.8 faster than classical Benders. For real-world general tables the stabilized approach was not as competitive as for the 1H2D case, probably due to the absence of a hierarchical structure. However, it is worth noting that stabilized Benders can be a promising approach because, when applied to real-world tables, the average gap was lower (2.61%) than for classical Benders (4.57%).

References

[1] J.F Benders (1962). Partitioning procedures for solving mixed-variables programming problems. *Numerische Mathematik*, 4, 238–252.

[2] J. Castro (2012), Recent advances in optimization techniques for statistical tabular data protection, *European Journal of Operational Research*, 21, 257–269.

[3] J.B. Hiriart-Urruty and C. Lemaréchal (1996). Convex Analysis and Minimization Algorithms II. *Springer-Verlag*, Berlin Heidelberg.

[4] M. Fischetti and A. Lodi (2003), Local Branching, *Mathematical Programming*, 98 (1-3), 23–47.

[5] M. Fischetti and J.J. SalazarGonzález (2001), Solving the cell suppression problem on tabular data with linear constraints, *Management Science*, 47, 1008–1026.

[6] T. Santoso, S. Ahmed, M. Goetschalckx and A. Shapiro (2005). A stochastic programming approach for supply chain network design under uncertainty. *European Journal of Operational Research* 167 (1), 96–115.