Journal of King Saud University – Science xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

Journal of King Saud University – Science

journal homepage: www.sciencedirect.com

A new modified deflected subgradient method

Rachid Belgacem [a,*], Abdessamad Amir [b]

[a] Université Hassiba Benbouali de Chlef, Chlef 02000, Algeria
[b] Universié Abdelhamid Ibn Badis de Mostaganem, Mostaganem 02700, Algeria

ARTICLE INFO

Article history:
Received 19 April 2017
Accepted 13 August 2017
Available online xxxx

MSC:
90C26
90C10
90C27
90C06

Keywords:
Integer linear programming
Subgradient method
Nonsmooth optimization
Travelling salesman problem

ABSTRACT

A new deflected subgradient algorithm is presented for computing a tighter lower bound of the dual problem. These bounds may be useful in nodes evaluation in a Branch and Bound algorithm to find the optimal solution of large-scale integer linear programming problems. The deflected direction search used in the present paper is a convex combination of the Modified Gradient Technique and the Average Direction Strategy. We identify the optimal convex combination parameter allowing the deflected subgradient vector direction to form a more acute angle with the best direction towards an optimal solution. The modified algorithm gives encouraging results for a selected symmetric travelling salesman problem (TSPs) instances taken from TSPLIB library.

1. Introduction

In this paper we consider the following integer linear program:

$$(IP)\begin{cases} & z^* = \min \top x \\ \text{s.t.} & A_1 x \leqslant b_1 \\ & x \in X = \{x \in \mathbb{Z}^n : A_2 x \leqslant b_2\}, \end{cases} \tag{1}$$

where x is an $n \times 1$ vector, $\mathbb{Z}^n$ is the set of integers, c, b_1, b_2, A_1 and A_2 are $n \times 1$, $m \times 1$, $k \times 1$, $m \times n$ and $k \times n$ matrices, respectively. We assume that the problem (IP) is feasible and that X is a bounded and finite set. The problem (IP) is called the "*primal problem*" and z^* is called the "*primal optimal value*". The constraints $A_2 x \leqslant b_2$ are generally called the easy constraints, in the sense that an integer linear program with only these constraints is easy to solve. Lagrangian duality (Bazaraa and Sherali, 1981) is the most computationally useful idea for solving hard integer programs. The Lagrangian dual problem is obtained via Lagrangian relaxation approach (Fisher, 1985), where the constraints $A_1 x \leqslant b_1$, which are called the "*complicated constraints*", are relaxed by introducing a multiplier vector $\lambda \in \mathbb{R}^m_+$, called "*Lagrangian multiplier*". The Lagrangian relaxation problem is formulated as follows:

$$(RP)\begin{cases} w(\lambda) = \min \; c^\top x + \lambda^\top (A_1 x - b_1) \\ \text{s.t.} \quad x \in X, \end{cases} \tag{2}$$

It is easy to prove that $w(\lambda) \leqslant z$ for all $\lambda \geqslant 0$ (weak duality (Bazaraa et al., 2006)). The best choice for λ would be the optimal solution of the following problem, called the dual problem:

$$(D)\begin{cases} w^* = \max \; w(\lambda) \\ \lambda \geqslant 0. \end{cases} \tag{3}$$

With some suitable assumptions, the dual optimal value w^* is equal to z^* (strong duality (Bazaraa et al., 2006)). In general, w^* provides a tighter lower bound of z^*. These bounds may be useful in nodes evaluation in exact methods such as Branch and Bound algorithm to find the optimal solution of (IP). The function $w(\lambda)$ is continuous and concave but non-smooth. The most widely adopted method for solving the dual problem is the subgradient optimization, see for instance Polyak (1967), Shor (1985), Nedic and Bertsekas (2010), Nesterov (2014) and Hu et al. (2015). The

* Corresponding author.
E-mail addresses: belgacemrachid02@yahoo.fr (R. Belgacem), amirabdessamad@yahoo.fr (A. Amir).

Peer review under responsibility of King Saud University.

Production and hosting by Elsevier

http://dx.doi.org/10.1016/j.jksus.2017.08.004
1018-3647/© 2017 The Authors. Production and hosting by Elsevier B.V. on behalf of King Saud University.

pure subgradient optimization method is an iterative procedure that can be used to solve the problem of maximizing (minimizing) a non-smooth concave (convex) function $w(\lambda)$ on a closed convex set Ω. This procedure is summarized in Algorithm 1, and it is used in various fields in science and engineering (Sra et al., 2012).

Algorithm 1 (Based Subgradient Algorithm).

1. Choose an initial point $\lambda_0 \in \Omega$.
2. Construct a sequence of points $\left(\lambda^k\right)_k \subset \Omega$ which eventually converges to an optimal solution using the rule $\lambda^{k+1} = P_\Omega\left(\lambda^k + t_k s^k\right)$, where $P_\Omega(.)$ is a projection operator on the set Ω and t_k is a positive scalar called step length such that
$$t_k = \delta_k \frac{w^* - w^k}{\|s^k\|^2}, \tag{4}$$
where $w^k = w(\lambda^k)$ is the dual function at the current iteration, $\delta_k \in]0, 2[$ and s^k is a subgradient of the function w at λ^k.
3. Replace k by $k+1$ and repeat the process until some stopping criteria.

In the context of Lagrangian relaxation, computing the subgradient direction s^k and the projection $P_\Omega\left(\lambda^k + t_k s^k\right)$ $(\Omega = \mathbb{R}^m_+)$ is a relatively easy problem. Since the subgradient s^k is not necessarily a descent direction, the step-length rule (4) differs from those given in the area of descent methods. In fact, this choice assures the decreasing of the subsequence $\left(\left\|\lambda^k - \lambda^*\right\|\right)_k$ as well as the convergence of $\left(\lambda^k\right)_k$ to λ^*. However, it is impossible to know in advance the value of w^* for most problems. To this end, the most effective way is to use the variable target value methods developed in Kim et al. (1990), Fumero (2001) and Sherali et al. (2000).

Another challenge in subgradient optimization is the choice of direction search that affects the computational performance of the algorithm. It is known that choosing the subgradient direction s^k, leads to the zigzagging phenomenon that might cause slow the procedure to crawl towards optimality (Bazaraa et al., 2006). To overcome this situation, in the spirit of conjugate gradient method (Nocedal and Wright, 2006; Fletcher and Reeves, 1964), we can adopt a direction search that deflects the subgradient pure direction. Accordingly, the direction search d^k at λ^k is computed as:

$$d^k = s^k + \Psi_k d^{k-1}, \tag{5}$$

where $\Psi_k \geqslant 0$ is a deflection parameter, s^k is a subgradient of the function w at λ^k and d^{k-1} is the previous direction ($d^0 = 0$). Then, the new iteration is computed as:

$$\lambda^{k+1} = P_\Omega\left(\lambda^k + t_k d^k\right). \tag{6}$$

Some promising deflection algorithms of this type are the Modified Gradient Technique (MGT) (Camerini et al., 1975) and the Average Direction Strategy (ADS) (Sherali and Ulular, 1989). The MGT method was found to be superior to the pure subgradient method when used in concert with a specially designed step-length selection rule. The deflection parameter Ψ_k^{MGT} is computed according to:

$$\Psi_k^{MGT} = \begin{cases} -\eta_k \frac{s^k d^{k-1}}{\|d^{k-1}\|^2} & \text{if } s^k d^{k-1} < 0, \\ 0 & \text{otherwise}, \end{cases} \tag{7}$$

where $0 < \eta_k \leqslant 2$. With this choice of the deflection parameter, the direction becomes:

$$d^k_{MGT} = s^k + \Psi_k^{MTG} d^{k-1}. \tag{8}$$

The ADS strategy recommends to make the deflection at each iteration point by choosing the direction search which simply bisects the angle between the current subgradient s^k and the previous direction search d^{k-1}. To get this direction, the deflection parameter is computed according to:

$$\Psi_k^{ADS} = \frac{\|s^k\|}{\|d^{k-1}\|}. \tag{9}$$

With this choice of the deflection parameter, the direction becomes:

$$d^k_{ADS} = s^k + \Psi_k^{ADS} d^{k-1}. \tag{10}$$

Nowadays, the deflected subgradient method remains an important tool for nonsmooth optimization problems, especially for linear integer programming, due to its simple formulation and low storage requirement. In this paper, we present a new deflected direction search as a convex combination of the direction d^k_{MGT} (8) and the direction d^k_{ADS} (10). Our main result is the identification of the convex combination parameter which forces the algorithm to have a better deflection search than those given in the pure subgradient, MGT and ADS. For a numerical comparison of our approach and the two concurrent techniques MGT and ADS, we opted for the Travelling Salesman Problem (TSP) where its importance comes from the richness of its application and the fact that it is a typical of other problems of combinatorial optimization (Diaby and Karwan, 2016; El-Sherbeny, 2010).

The remainder of the paper is organized as follows: in Section 2, we describe our deflected subgradient method with convergence analysis. The computational tests, conducted on the Lagrangian relaxation of TSP of different sizes are described in Section 3. In Section 4 we conclude the paper.

2. A new modified deflected subgradient method

In this section, we present a new modified deflected subgradient method (*NMDS*) which determines the direction search as follows:

$$d^k = (1 - \alpha_k) d^k_{MGT} + \alpha_k d^k_{ADS}, \quad \alpha_k \in (0, 1). \tag{11}$$

We then obtain the following deflection parameter:

$$\Psi_k = \begin{cases} \frac{-\eta_k(1-\alpha_k) s^k d^{k-1} + \alpha_k \|s^k\| \|d^{k-1}\|}{\|d^{k-1}\|^2} & \text{if } s^k d^{k-1} < 0, \\ 0 & \text{otherwise}, \end{cases} \tag{12}$$

hence $d^k = s^k + \Psi_k d^{k-1}$.

Algorithm 2 (The Deflected Subgradient Algorithm).

1. (Initialization): Choose a starting point $\lambda^0 \in \Omega = \mathbb{R}^m_+$, let $d^0 = 0$ and $k = 0$.
2. Determine a subgradient $s^k \in \partial w(\lambda^k)$ and compute
$$d^k = s^k + \Psi_k d^{k-1},$$
$$\lambda^{k+1} = P_\Omega(\lambda^k + t_k d^k),$$
where Ψ_k is given by relation (12) and t_k will be specified later.
3. Replace k by $k+1$ if a stopping condition is not yet met and return to step 2.

Consider the deflected subgradient method algorithm given in Algorithm 2. The following proposition extends important properties of the subgradient vector s^k and the deflected subgradient