Operations Research Letters 45 (2017) 503–507

Contents lists available at ScienceDirect

Operations Research Letters

journal homepage: www.elsevier.com/locate/orl

Linear fractional approximations for master problems in column generation

Hocine Bouarab [a], Guy Desaulniers [a], Jacques Desrosiers [b,*], Jean Bertrand Gauthier [b]

[a] Polytechnique Montréal & GERAD, Montréal (Québec), Canada H3C 3A7,
[b] HEC Montréal & GERAD, Montréal (Québec), Canada H3T 2A7,

ARTICLE INFO

Article history:
Received 29 January 2017
Received in revised form 28 July 2017
Accepted 1 August 2017
Available online 10 August 2017

Keywords:
Column generation
Master problem
Linear fractional program
Super-geometric growth rate

ABSTRACT

In the context of large-scale linear programs solved by a column generation algorithm, we present a primal algorithm for handling the master problem. Successive approximations of the latter are created to converge to optimality. The main properties are that, for every approximation except the last one, the cost of the solution decreases whereas the sum of the variable values increases. Moreover, the minimum reduced cost of the variables also increases and converges to zero with a super-geometric growth rate.

1. Introduction

We consider a linear program in standard form called the *master problem* (*MP*):

$$z^\star := \min \ \mathbf{c}^\intercal\boldsymbol{\lambda} \quad \text{s.t.} \ \ \mathbf{A}\boldsymbol{\lambda} = \mathbf{b} \ \ [\boldsymbol{\pi}], \quad \boldsymbol{\lambda} \geq \mathbf{0}, \tag{1}$$

where $\boldsymbol{\lambda}, \mathbf{c} \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^m$, $\mathbf{A} \in \mathbb{R}^{m\times n}$. The vector of dual variables $\boldsymbol{\pi} \in \mathbb{R}^m$ associated with the equality constraints appears within brackets on the right side. When the number of variables is very large, the latter is solved by a column generation algorithm. Furthermore, we assume that $\boldsymbol{\lambda} = \mathbf{0}$ is infeasible for (1), the sum of the λ-variables $s := \mathbf{1}^\intercal\boldsymbol{\lambda} \leq s_{max}$ is bounded from above, and $\mathbf{c} \neq \mathbf{0}$, three natural assumptions in the formulation of the *MP* for many applications. The dual of (1) reads as:

$$z^\star = \max \ \mathbf{b}^\intercal\boldsymbol{\pi} \quad \text{s.t.} \ \ \mathbf{A}^\intercal\boldsymbol{\pi} \leq \mathbf{c} \ \ [\boldsymbol{\lambda}]\,. \tag{2}$$

Notation. Vectors and matrices are written in bold face characters by respectively using lower and upper case notations. In particular, the matrix $\mathbf{A} := [\mathbf{a}_j]_{j\in\{1,\ldots,n\}}$ contains n column vectors. We also use standard linear programming notation like $\mathbf{A}_B\boldsymbol{\lambda}_B$, the subset of basic columns of $\mathbf{A}$ indexed by B multiplied by the corresponding vector of basic variables $\boldsymbol{\lambda}_B$. Furthermore, we denote by $\mathbf{0}$ or $\mathbf{1}$ a vector/matrix with all zero or one entries of appropriate contextual dimensions. The index set of nonbasic columns N is used analogously. Finally, let $\bar{\mathbf{c}}^\intercal := \mathbf{c}^\intercal - \boldsymbol{\pi}^\intercal\mathbf{A}$ be the reduced cost vector of the λ-variables for any vector $\boldsymbol{\pi}$ of dual variables.

This paper examines the properties of an approximation scheme for solving the *MP*. Given a feasible primal solution with a finite cost, an approximation problem is expressed in three different ways: a linear fractional program and the two equivalent primal and dual linear programming versions. The first program is essentially used to demonstrate the convergence properties. The primal version is the one that would be solved by a column generation algorithm whereas the dual version brings insightful interpretation with the following two features: the dual variables are optimized for maximizing the minimum reduced cost of the λ-variables and the dual objective value is forced to be equal to the cost of the current primal solution. In fact, the motivation behind these formulations is that solving the dual one is a way to verify if the optimality conditions provided by the *Strong Duality Theorem* [2] are satisfied.

Several algorithms possess these features partially. For instance, in the *primal simplex (PS)* algorithm [see [2]], all the dual variables are fixed at every iteration, that is, determined according to the current basic solution and the second feature is present by construction, see Section 2. The second feature is also shown to be true for the *improved primal simplex* algorithm [3], where the dual variables are optimized subject to constraints that only set to zero the reduced costs of the positive-valued variables (rather than doing so for all the basic variables as in *PS*). At the other extreme, when all of the dual variables are optimized for maximizing the minimum reduced cost, we find the *minimum mean*

* Corresponding author.
E-mail addresses: hocine.bouarab@gerad.ca (H. Bouarab), guy.desaulniers@gerad.ca (G. Desaulniers), jacques.desrosiers@hec.ca (J. Desrosiers), jean-bertrand.gauthier@hec.ca (J.B. Gauthier).

http://dx.doi.org/10.1016/j.orl.2017.08.004
0167-6377/

cycle-canceling (MMCC) algorithm devised for solving network flow problems and shown to be strongly polynomial [5]. However, *MMCC* does not require the equality between the primal and dual objective values at every iteration.

Using a similar scheme, it has been observed that finding a near optimal integer solution to set partitioning models with the integral simplex (using decomposition) algorithm [7] occurs after solving a very small number of approximation problems [6]. To explain this behavior, we prove the following two properties for the approximation scheme of the *MP*: If the current solution is non-optimal, we proceed to the next approximation problem which provides a new λ-solution where the objective function strictly decreases while the sum of the variable values strictly increases. Moreover, the so-called *optimality parameter*, that is, the largest minimum reduced cost, also strictly increases and converges to zero with a super-geometric growth rate.

The paper is organized as follows. Section 2 presents the approximation problem, the fundamental properties, and the resulting algorithmic framework. We finally study two convergence growth rates in Section 3 as well as a few special cases.

2. The approximation scheme

This section presents the approximation problem along with some insights regarding its solutions. Fundamental properties surrounding the optimality parameter are stated. Then the algorithm comes naturally.

2.1. Approximation of the master problem

The scheme consists of solving the *MP* through successive approximations. Let AMP^k denote the *approximation problem* solved at iteration $k \geq 0$. We assume that a feasible solution to the *MP* (1) is known at $k \geq 1$, say $\boldsymbol{\lambda}^{k-1}$ of cost $z^{k-1} = \mathbf{c}^\intercal \boldsymbol{\lambda}^{k-1}$. We define AMP^k as

$$\mu^k := \min \frac{\mathbf{c}^\intercal \boldsymbol{\lambda} - z^{k-1}}{\mathbf{1}^\intercal \boldsymbol{\lambda}} \quad \text{s.t. } \mathbf{A}\boldsymbol{\lambda} = \mathbf{b}, \ \boldsymbol{\lambda} \geq \mathbf{0}, \tag{3}$$

for which an optimal solution with an objective value μ^k is obtained from a vector $\boldsymbol{\lambda}^k$. The scalar μ^k measures the difference between two successive λ-solution costs in *MP* (1) divided by the sum of the variable values in the new one.

Using the Charnes–Cooper transformation of a linear fractional program [1], set $\theta := 1/\mathbf{1}^\intercal \boldsymbol{\lambda}$ and $\mathbf{y} = \theta \boldsymbol{\lambda}$ to obtain an equivalent linear programming version of AMP^k:

$$\begin{array}{lrrll} \mu^k = \min & \mathbf{c}^\intercal \mathbf{y} & - z^{k-1}\theta & & \\ \text{s.t.} & \mathbf{A}\mathbf{y} & - \mathbf{b}\,\theta & = \mathbf{0} & [\boldsymbol{\pi}] \\ & \mathbf{1}^\intercal \mathbf{y} & & = 1 & [\mu] \\ & \mathbf{y} \geq \mathbf{0}, & \theta \geq 0, & & \end{array} \tag{4}$$

where the domain of θ has been extended from $\theta > 0$ to $\theta \geq 0$ to satisfy linear programming requirements. Since we assume that the *MP* (1) has a finite optimal solution $\lambda^\star$ with $s^\star := \mathbf{1}^\intercal \boldsymbol{\lambda}^\star \leq s_{max}$, an optimal solution to (4) with $\mu < 0$ and $\theta = 0$ is not feasible, otherwise it would imply an extreme ray of negative cost leading to unboundedness for the *MP*. Similarly to the *MP* (1), the problem (4) can be solved by a column generation algorithm. In that case, θ can be seen as a static variable as opposed to the y-variables that are dynamically generated as needed. Taking the dual of (4), we find:

$$\begin{array}{lrll} \mu^k = \max & \mu & & \\ \text{s.t. } \mathbf{1}\mu & + \mathbf{A}^\intercal \boldsymbol{\pi} \leq \mathbf{c} & & [\mathbf{y}] \\ & - \mathbf{b}^\intercal \boldsymbol{\pi} \leq -z^{k-1}. & & [\theta] \end{array} \tag{5}$$

From an optimal solution $\mathbf{y}^k$ and θ^k to AMP^k, the decision vector relevant to the *MP* (1) can be recovered as $\boldsymbol{\lambda}^k = \mathbf{y}^k/\theta^k$ with a cost z^k. Notice that the scalar $s^k = \mathbf{1}^\intercal \boldsymbol{\lambda}^k = 1/\theta^k$ is also obtained as a by-product of solving AMP^k. The process is initialized at iteration $k = 0$ by solving:

$$\mu^0 := \min \frac{\mathbf{c}^\intercal \boldsymbol{\lambda} - UB}{\mathbf{1}^\intercal \boldsymbol{\lambda}} \quad \text{s.t. } \mathbf{A}\boldsymbol{\lambda} = \mathbf{b}, \ \boldsymbol{\lambda} \geq \mathbf{0}, \tag{6}$$

where the z-parameter in (3) is replaced by *UB*, an upper bound on $z^\star$.

Interpretation. First observe that $\mu^k \leq 0, \forall k \geq 0$. Indeed, obviously $\mu^0 \leq 0$ by construction of (6) whereas for $k \geq 1$, $\boldsymbol{\lambda}^{k-1}$ is feasible with $\mathbf{c}^\intercal \boldsymbol{\lambda}^{k-1} - z^{k-1} = 0$ such that we also have $\mu^k \leq 0$ in (3). An optimal solution $(\boldsymbol{\pi}^k, \mu^k)$ to (5) satisfies $\mu^k \leq c_j - \mathbf{a}_j^\intercal \boldsymbol{\pi}^k$, $\forall j \in \{1, \ldots, n\}$, hence the objective function *maximizes the smallest reduced cost of the λ-variables* and μ^k is equal to the smallest one. When solving the *MP* (1), the primal simplex algorithm does essentially the same except that, because $\boldsymbol{\pi}^\intercal = \mathbf{c}_B{}^\intercal \mathbf{A}_B{}^{-1}$ is computed according to the basic variables $\boldsymbol{\lambda}_B$ rather than being optimized, μ^k is simply determined by inspection: $\mu^k = \max \mu$ s.t. $\mu \leq \bar{c}_j, \forall j \in \{1, \ldots, n\}$. Moreover, observe that (5) imposes $\mathbf{b}^\intercal \boldsymbol{\pi} \geq z^{k-1}$, that is, the reduced cost $\bar{c}_\theta \geq 0$. Because $\theta^k > 0$, we have by complementary slackness $\mathbf{b}^\intercal \boldsymbol{\pi}^k = z^{k-1}$, an equality satisfied by construction in the primal simplex algorithm for any basic solution, where $\boldsymbol{\pi}^\intercal \mathbf{b} = \mathbf{c}_B{}^\intercal \mathbf{A}_B{}^{-1} \mathbf{b} = \mathbf{c}_B{}^\intercal \boldsymbol{\lambda}_B$.

Note finally that the equality $\mathbf{b}^\intercal \boldsymbol{\pi} = z^\star$ is only satisfied by an optimal feasible vector $\boldsymbol{\pi}^\star$. Indeed, rather than verifying if the classical conditions of the primal simplex algorithm are satisfied ($\bar{\mathbf{c}}_B = \mathbf{0}$ and $\bar{\mathbf{c}}_N \geq \mathbf{0}$, these being only sufficient optimality conditions), we formulate in (5) a problem to verify if the necessary and sufficient optimality conditions provided by the *Strong Duality Theorem* [2] can be fulfilled, that is, if there exists a feasible dual solution $\boldsymbol{\pi}$ with cost $\mathbf{b}^\intercal \boldsymbol{\pi}$ equal to that of the current primal solution. In the affirmative, we have an optimality certificate with $\mu = 0$, otherwise we provide a new λ-solution.

2.2. The optimality parameter

In this section, we derive two fundamental properties regarding the outcome of solving AMP^k, $k \geq 0$. The first characterizes an optimality condition for the *MP* while the second establishes the behavior of z, μ, and s across two successive approximations. An algorithm is then drafted upon these properties.

Proposition 1. *For $k \geq 0$, if $\mu^k = 0$, then $\boldsymbol{\lambda}^k$ is an optimal solution for the MP (1).*

Proof. If $k = 0$ and $\mu^0 = 0$, the vector $\boldsymbol{\lambda}^0$ is an optimal solution to the *MP* because its cost z^0 reaches the upper bound *UB* on $z^\star$. Otherwise $k \geq 1$ and $\mu^k = 0 \Leftrightarrow \mathbf{c}^\intercal \boldsymbol{\lambda}^k = z^k = z^{k-1}$. The pair $(\boldsymbol{\lambda}^{k-1}, \boldsymbol{\pi}^{k-1})$ then satisfies the strong duality theorem for linear programs and therefore provides feasible primal and dual solutions for the *MP* (1). The same is true for $\boldsymbol{\lambda}^k$ and observe that whether it is different from $\boldsymbol{\lambda}^{k-1}$ or not is irrelevant although it provides an alternative optimal solution in the affirmative. □

In light of Proposition 1, the value of μ grants a simple way to ascertain optimality, thus the term *optimality parameter* coined by Goldberg and Tarjan [5] is still convenient. In fact, Proposition 2 shows that if $\boldsymbol{\lambda}^k$ is non-optimal, the cost z^{k+1} strictly decreases while both μ^{k+1} and s^{k+1} strictly increase.

Proposition 2. *For $k \geq 0$, let $\boldsymbol{\lambda}^k$ be a non-optimal solution for the MP (1). Then*

$$z^k > z^{k+1}, \ \ \mu^k < \mu^{k+1}, \ \ \text{and } s^k < s^{k+1}. \tag{7}$$