Les Cahiers du GERAD ISSN: 0711–2440

A linear fractional pricing problem for solving linear programs

H. Bouarab, G. Desaulniers,
J. Desrosiers

G–2015–82

September 2015

Les textes publiés dans la série des rapports de recherche *Les Cahiers du GERAD* n'engagent que la responsabilité de leurs auteurs.

La publication de ces rapports de recherche est rendue possible grâce au soutien de HEC Montréal, Polytechnique Montréal, Université McGill, Université du Québec à Montréal, ainsi que du Fonds de recherche du Québec – Nature et technologies.

Dépôt légal – Bibliothèque et Archives nationales du Québec, 2015.

The authors are exclusively responsible for the content of their research papers published in the series *Les Cahiers du GERAD*.

The publication of these research reports is made possible thanks to the support of HEC Montréal, Polytechnique Montréal, McGill University, Université du Québec à Montréal, as well as the Fonds de recherche du Québec – Nature et technologies.

Legal deposit – Bibliothèque et Archives nationales du Québec, 2015.

GERAD HEC Montréal
3000, chemin de la Côte-Sainte-Catherine
Montréal (Québec) Canada H3T 2A7

Tél. : 514 340-6053
Téléc. : 514 340-5665
info@gerad.ca
www.gerad.ca

A linear fractional pricing problem for solving linear programs

Hocine Bouarab [a]
Guy Desaulniers [a]
Jacques Desrosiers [b]

[a] *GERAD & Polytechnique Montréal, Montréal (Québec) Canada, H3C 3A7*

[b] *GERAD & HEC Montréal, Montréal (Québec) Canada, H3T 2A7*

hocine.bouarab@gerad.ca
guy.desaulniers@gerad.ca
jacques.desrosiers@hec.ca

September 2015

Les Cahiers du GERAD

G–2015–82

Copyright © 2015 GERAD

Abstract: This paper introduces a new primal algorithm for solving a linear program LP. In this algorithm, a pricing problem, namely a linear fractional program, is solved at each iteration to compute the largest possible minimum reduced cost over all variables. The linear version of this pricing problem obtained from the Charnes-Cooper transformation optimizes the dual vector under the constraint that the primal and dual objective values of LP must be equal. The solution of the pricing problem is used to update the current primal solution or to prove that it is optimal when the minimum reduced cost is null. The main properties of this algorithm are that, at every iteration except the last one, the cost of the primal solution decreases, the sum of the variables in the primal solution increases, and the minimum reduced cost also increases. We show that the minimum reduced cost converges to zero with a super-geometric growth rate.

Key Words: Linear program, linear fractional program, pricing problem, super-geometric growth rate.

Acknowledgments: Guy Desaulniers and Jacques Desrosiers acknowledge the National Science and Engineering Research Council of Canada for its financial support.

1 Introduction

This paper introduces a primal algorithm for solving a linear program LP. Given a feasible solution at the beginning of an iteration, the pricing step solves a *linear fractional program* (LFP) for finding the least possible reduced cost. Using the Charnes-Cooper transformation [2], this pricing problem becomes equivalent to a linear program that has the following two features:

- the dual variables are optimized to achieve such a minimum reduced cost;
- the dual objective of LP is equal to the cost of the current feasible solution.

Several algorithms possess partially these features. For example, in the *primal simplex* (PS) algorithm [see 3], all dual variables are fixed, that is, determined according to the current basic solution, and the second feature is thus present at every iteration, see Section 3.2. The second feature is also shown to be true for the *improved primal simplex* (IPS) algorithm [5, 14, 12], where the dual variables are not fixed but subject to constraints that set to zero the reduced costs of the nondegenerate variables, see Section 3.3. At the other extreme when none of the dual variables are fixed but all optimized, we find the *minimum mean cycle-canceling* (MMCC) algorithm devised for capacitated network flow problems and shown to be strongly polynomial [9]. In that case, the objective of the pricing problem turns out to be a linear fractional function, that is, the cost of the selected cycle divided by the number of arcs in it. However, the optimization of the dual variables for finding the minimum reduced cost does not involve the second feature, that is, the equality between the primal and dual objective values is not required, see Section 3.1.

The paper is organized as follows. Section 2.1 presents the linear fractional pricing problem, elaborates its linear programming primal and dual versions, provides an interpretation of such a pricing step, and ends with an outline of the proposed algorithm. In Section 3, we establish a link between the pricing step of several algorithms from the literature and also derive the new pricing problem using an alternative approach. Section 4 presents various properties of the algorithm. Among others, we show that the minimum reduced cost increases at each iteration and that optimality of LP is reached if and only if the minimum reduced cost is zero. Section 5 analyzes the growth rate of the minimum reduced cost towards zero, which turns out to be super-geometric. This is followed by our conclusions.

2 The pricing step as a linear fractional program

Consider that LP is in standard form

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

where $\mathbf{x}, \mathbf{c} \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^m$, $\mathbf{A} \in \mathbb{R}^{m \times n}$, and $m \leq n$. The vector of dual variables $\boldsymbol{\pi} \in \mathbb{R}^m$ associated with the equality constraints appears within brackets on the right side. The dual formulation of (1) reads as:

$$\begin{aligned} \max \quad & \mathbf{b}^\intercal \boldsymbol{\pi} \\ \text{s.t.} \quad & \mathbf{A}^\intercal \boldsymbol{\pi} \leq \mathbf{c} \quad [\mathbf{x}]. \end{aligned} \tag{2}$$

Notation. Vectors and matrices are written in bold face characters. We denote by $\mathbf{0}$ or $\mathbf{1}$ a vector/matrix with all zero or one entries of appropriate contextual dimensions. A single column of $\mathbf{A}$ is denoted $\mathbf{a}_j,\ j \in \{1, \ldots, n\}$. For a subset $R \subseteq \{1, \ldots, m\}$ of row-indices and a subset $P \subseteq \{1, \ldots, n\}$ of column-indices, we denote by $\mathbf{A}_{RP}$ the sub-matrix of $\mathbf{A}$ containing the rows and columns indexed by R and P, respectively. We further use standard linear programming notation like $\mathbf{A}_B \mathbf{x}_B$, the subset of basic columns of $\mathbf{A}$ indexed by B multiplied by the corresponding vector of basic variables $\mathbf{x}_B$. Finally, let $\bar{\mathbf{c}}^\intercal := \mathbf{c}^\intercal - \boldsymbol{\pi}^\intercal \mathbf{A}$ be the reduced cost vector of the $\mathbf{x}$-variables for a given vector $\boldsymbol{\pi}$ of dual variables.

In the rest of the paper, we assume that vector $\mathbf{b} \neq \mathbf{0}$. Otherwise, the optimal value $z^\star$ is either 0 (with $\mathbf{x}^\star = \mathbf{0}$ as an optimal solution) or the problem is unbounded.

2.1 The pricing problem

Set iteration counter $k = 0$ and assume that a feasible solution $\mathbf{x}^k \neq \mathbf{0}$ of cost $z^k = \mathbf{c}^\intercal \mathbf{x}^k$ is known. At iteration $k+1$, define the pricing problem PP^{k+1} as the following linear fractional program:

$$\begin{aligned} \mu^{k+1} := \quad \min \quad & \frac{\mathbf{c}^\intercal \mathbf{x} - z^k}{\mathbf{1}^\intercal \mathbf{x}} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} = \mathbf{b}, \quad \mathbf{x} \geq \mathbf{0}. \end{aligned} \tag{3}$$

An interpretation of this pricing problem comes next, after its linearization. Using the Charnes-Cooper transformation of a linear fractional program into an equivalent linear program [2], set $\theta = 1/\mathbf{1}^\intercal \mathbf{x} > 0$ and $\mathbf{y} = \theta \mathbf{x}$ to obtain the primal linear programming version of PP^{k+1}:

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

where the domain of θ has been extended from $\theta > 0$ to $\theta \geq 0$. Taking the dual of (4), one finds:

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

2.2 Interpretation

The interpretation of μ^{k+1} in the linear fractional pricing program (3) is done in terms of the alternative primal and dual formulations (4) and (5), respectively. In the primal formulation, define the reduced cost $\bar{c}_\theta = -z^k + \boldsymbol{\pi}^\intercal \mathbf{b}$ for variable θ. Let $(\mathbf{y}^{k+1}, \theta^{k+1}; \ \boldsymbol{\pi}^{k+1}, \mu^{k+1})$ be an optimal primal-dual solution to these formulations at iteration $k+1$, $k \geq 0$. An optimal solution $(\boldsymbol{\pi}^{k+1}, \mu^{k+1})$ to (5) satisfies

$$\mu^{k+1} \leq c_j - (\boldsymbol{\pi}^{k+1})^\intercal \mathbf{a}_j, \qquad \forall j \in \{1, \ldots, n\}, \tag{6}$$

and the equality holds for the largest possible minimum reduced cost value of the $\mathbf{x}$-variables. Observe that $\boldsymbol{\pi}^{k+1}$ is also required to satisfy $\mathbf{b}^\intercal \boldsymbol{\pi}^{k+1} \geq z^k$, that is, $\bar{c}_\theta \geq 0$. For a solution with $\theta^{k+1} > 0$, we have $\mathbf{b}^\intercal \boldsymbol{\pi}^{k+1} = z^k$ by complementary slackness, and consequently, the dual objective in (2) is equal to the cost of the current primal one in (1):

$$\theta^{k+1} > 0 \quad \Rightarrow \quad \mathbf{b}^\intercal \boldsymbol{\pi}^{k+1} = \mathbf{c}^\intercal \mathbf{x}^k. \tag{7}$$

The primal formulation (4) of PP^{k+1} allows us to show that the minimum reduced cost μ^{k+1} is bounded from above by zero. Indeed, given a solution $\mathbf{x}^k \neq \mathbf{0}$ to LP, vector $\mathbf{y} = \theta \mathbf{x}^k$ together with scalar $\theta = 1/\mathbf{1}^\intercal \mathbf{x}^k$ form a feasible solution to (4) with a value $\mathbf{c}^\intercal(\theta\,\mathbf{x}^k) - z^k\theta = 0$, hence $\mu^{k+1} \leq 0$. When $\mu^{k+1} = 0$, the algorithm stops with on hand a primal solution $\mathbf{x}^k$ to (1) of cost z^k and a feasible dual vector $\boldsymbol{\pi}^{k+1}$, as it satisfies $\bar{\mathbf{c}} \geq \mathbf{0}$ in (2), at the same cost. In that case $(\mathbf{x}^k, \boldsymbol{\pi}^{k+1})$ is an optimal primal-dual solution to LP.

2.3 The algorithm

This leads us to the proposed algorithm to solve LP based on the *linear fractional pricing* problem PP^{k+1} described by the pair of primal-dual formulations (4)-(5). This algorithm is called Algorithm LFP.

2.3.1 Algorithm LFP

1. Let $\mathbf{x}^0$ be a feasible solution to LP (1) of cost $z^0 = \mathbf{c}^\intercal \mathbf{x}^0$. Set $k = 0$.
2. Solve PP^{k+1} (4) with optimal primal-dual solution $(\mathbf{y}^{k+1}, \theta^{k+1};\ \boldsymbol{\pi}^{k+1}, \mu^{k+1})$.
3. If $\mu^{k+1} < 0$ and $\theta^{k+1} > 0$, update to an improved solution: $\mathbf{x}^{k+1} := \mathbf{y}^{k+1} / \theta^{k+1}$. Set $k := k + 1$ and return to Step 2.
4. If $\mu^{k+1} < 0$ and $\theta^{k+1} = 0$, stop: LP is unbounded.
5. Otherwise $\mu^{k+1} = 0$ and primal-dual pair $(\mathbf{x}^k, \boldsymbol{\pi}^{k+1})$ is optimal for LP.

In Step 1, Algorithm LFP assumes a known feasible solution $\mathbf{x}^0$. This may require a phase 1 step to establish the existence or not of such a solution. Observe however that, because $\mathbf{x}^0$ does not appear in the pricing problem formulation (3), the algorithm can also be initialized with an upper bound $z^0 > z^\star$ on the optimal cost. Iteration counter k is set to zero.

In Step 2, PP^{k+1} computes a primal-dual solution $(\mathbf{y}^{k+1}, \theta^{k+1};\ \boldsymbol{\pi}^{k+1}, \mu^{k+1})$ to (4). If $\mu^{k+1} < 0$ and $\theta^{k+1} > 0$ in Step 3, one updates the primal solution to $\mathbf{x}^{k+1} := \mathbf{y}^{k+1} / \theta^{k+1}$ for which the cost decreases, as computed using the objective function of (4):

$$\mu^{k+1} = \mathbf{c}^\intercal \mathbf{y}^{k+1} - z^k \theta^{k+1} \quad \Leftrightarrow \quad z^{k+1} = \mathbf{c}^\intercal \mathbf{x}^{k+1} = z^k + \frac{\mu^{k+1}}{\theta^{k+1}} < z^k. \tag{8}$$

If $\mu^{k+1} < 0$ and $\theta^{k+1} = 0$ occurs in Step 4, vector $\mathbf{y}^{k+1}$ is a negative cost ray in the cone defined by $\{\mathbf{y} \geq \mathbf{0} \mid \mathbf{A}\mathbf{y} = \mathbf{0}\}$ and the algorithm stops for unboundedness of LP: in fact, $\mathbf{x}^{k+1} := \mathbf{x}^k + \rho \mathbf{y}^{k+1}$ is feasible at a cost $z^k + \rho \mu^{k+1}$ for any value of $\rho > 0$.

Finally, $\mu^{k+1} = 0$ in Step 5 and, as discussed above, we have on hand a primal solution $\mathbf{x}^k$ and a feasible dual vector $\boldsymbol{\pi}^{k+1}$ at the same cost. Therefore the primal-dual pair $(\mathbf{x}^k, \boldsymbol{\pi}^{k+1})$ is optimal for LP.

3 Related works

This section presents three algorithms that indirectly use a linear fractional pricing problem. They are mainly described in the context of network flow problems.

3.1 The minimum mean cycle-canceling algorithm

A linear fractional pricing problem defines the pricing step of the *minimum mean cycle-canceling algorithm* (MMCC) algorithm of Goldberg and Tarjan [9] designed to solve the *capacitated minimum cost flow* ($CMCF$) problem. Let us take a closer look at it.

Consider a directed graph $G = (V, A)$, where V is the set of vertices and A the set of arcs. The vertices are associated with a set $b_i, i \in V$, of supplies and demands defined, respectively, by positive and negative values such that $\sum_{i \in V} b_i = 0$. The arc costs are given by $\mathbf{c} := [c_{ij}]_{(i,j) \in A}$ while lower and upper bounds $\boldsymbol{\ell} := [\ell_{ij}]_{(i,j) \in A}$ and $\mathbf{u} := [u_{ij}]_{(i,j) \in A}$ restrict the flow on each arc. Denoting $\mathbf{x} := [x_{ij}]_{(i,j) \in A}$ the arc flow variables, the $CMCF$ problem is formulated as:

$$\begin{aligned} z^\star_{CMCF} := \quad \min \quad & \sum_{(i,j) \in A} c_{ij} x_{ij} \\ \text{s.t.} \quad & \sum_{j:(i,j) \in A} x_{ij} - \sum_{j:(j,i) \in A} x_{ji} = b_i \quad [\pi_i] \quad \forall i \in V \\ & 0 \leq \ \ell_{ij} \leq x_{ij} \leq u_{ij} \qquad \forall (i,j) \in A, \end{aligned} \tag{9}$$

where $\boldsymbol{\pi} := [\pi_i]_{i \in V}$ is the vector of dual variables. When the right-hand side $\mathbf{b} := [b_i]_{i \in V}$ is the null vector, formulation (9) corresponds to a *circulation* problem, a situation that occurs in the reformulation of $CMCF$ using the residual network described next.

Assume given a feasible solution $\mathbf{x}^k$, $k \geq 0$, to a $CMCF$ problem. The pricing problem at iteration $k+1$ in the MMCC algorithm is defined on an associated residual network, see Ahuja et al. [1]. The residual network depends on $\mathbf{x}^k$ and is denoted $G(\mathbf{x}^k) = (V, A(\mathbf{x}^k))$. Each arc $(i,j) \in A$ is replaced by two arcs representing upwards and downwards possible flow variations:

- arc (i,j) with cost $d_{ij} = c_{ij}$ and residual flow $0 \leq f_{ij} \leq r_{ij}^k := u_{ij} - x_{ij}^k$;
- arc (j,i) with cost $d_{ji} = -c_{ij}$ and residual flow $0 \leq f_{ji} \leq r_{ji}^k := x_{ij}^k - \ell_{ij}$.

Denote $A' := \{(i,j) \cup (j,i) \mid (i,j) \in A\}$ the arc support of any residual network. The current one $G(\mathbf{x}^k)$ consists of only the arcs with positive residual capacities, that is, $A(\mathbf{x}^k) := \{(i,j) \in A' \mid r_{ij}^k > 0\}$. The combination of the current solution $\mathbf{x}^k$ with an optimal residual flow $\mathbf{f} := [f_{ij}]_{(i,j)\in A(\mathbf{x}^k)}$ computed on $G(\mathbf{x}^k)$ is optimal for the original formulation (9). In fact, the residual network with respect to $\mathbf{x}^k$ corresponds to the change of variables $x_{ij} = x_{ij}^k + (f_{ij} - f_{ji}), \forall (i,j) \in A$. Observe that passing flow in both directions along an arc would be counterproductive and could be simplified to the net flow in a single direction. This means that the residual flow must be such that $f_{ij}\, f_{ji} = 0, \forall (i,j) \in A$. The formulation of $CMCF$ on the residual network is as follows:

$$
\begin{aligned}
z^\star_{CMCF} = \mathbf{c}^\intercal \mathbf{x}^k \; + \; \min \quad & \sum_{(i,j)\in A(\mathbf{x}^k)} d_{ij} f_{ij} \\
\text{s.t.} \quad & \sum_{j:(i,j)\in A(\mathbf{x}^k)} f_{ij} - \sum_{j:(j,i)\in A(\mathbf{x}^k)} f_{ji} = 0 \quad [\pi_i] \quad \forall i \in V \\
& 0 \leq f_{ij} \leq r_{ij}^k \qquad \forall (i,j) \in A(\mathbf{x}^k).
\end{aligned}
\tag{10}
$$

In MMCC, finding an improving directed cycle on $G(\mathbf{x}^k)$ at iteration $k+1$ disregards the residual capacities and looks for the *minimum mean cost cycle*, the average being taken on the number of arcs in the selected cycle. Although usually stated in words [9, 13, 1], this is a linear fractional pricing problem and it can be formulated as

$$
\begin{aligned}
\mu^{k+1} := \quad \min \quad & \frac{\sum_{(i,j)\in A(\mathbf{x}^k)} d_{ij} f_{ij}}{\sum_{(i,j)\in A(\mathbf{x}^k)} f_{ij}} \\
\text{s.t.} \quad & \sum_{j:(i,j)\in A(\mathbf{x}^k)} f_{ij} - \sum_{j:(j,i)\in A(\mathbf{x}^k)} f_{ji} = 0 \quad [\pi_i] \quad \forall i \in V \\
& f_{ij} \geq 0 \qquad \forall (i,j) \in A(\mathbf{x}^k).
\end{aligned}
\tag{11}
$$

The domain of (11) is a cone with $\mathbf{0}$ as the unique extreme point. For an optimal solution $\mathbf{f}^{k+1} \neq \mathbf{0}$ of cost $\mu^{k+1} < 0$, observe that vector $\alpha \mathbf{f}^{k+1}$ (where $\alpha > 0$) is also an optimal solution at the same cost. Therefore one can use any flow value on the arcs of the selected cycle, for example, a unit flow, or equivalently, an alternative flow equal to $1/|W|$, where $|W|$ is the size of the cycle in terms of the number of arcs. This corresponds to the Charnes-Cooper transformation presented next, a linearization of the pricing problem for which the flow variables sum to one.

Applying the Charnes-Cooper transformation is done by setting $\theta = 1/\mathbf{1}^\intercal \mathbf{f} > 0$ and $\mathbf{y} := \theta \mathbf{f}$. The linear version of the fractional pricing problem (11) becomes

$$
\begin{aligned}
\mu^{k+1} = \quad \min \quad & \sum_{(i,j)\in A(\mathbf{x}^k)} d_{ij} y_{ij} \\
\text{s.t.} \quad & \sum_{j:(i,j)\in A(\mathbf{x}^k)} y_{ij} - \sum_{j:(j,i)\in A(\mathbf{x}^k)} y_{ji} = 0 \quad [\pi_i] \quad \forall i \in V \\
& \sum_{(i,j)\in A(\mathbf{x}^k)} y_{ij} = 1 \quad [\mu] \\
& y_{ij} \geq 0 \qquad \forall (i,j) \in A(\mathbf{x}^k),
\end{aligned}
\tag{12}
$$

with associated dual vector $\boldsymbol{\pi} = [\pi_i]_{i \in V}$ and dual scalar μ. Taking the dual of (12), one obtains

$$\begin{aligned} \mu^{k+1} = \quad & \max \quad \mu \\ & \text{s.t.} \quad \mu + \pi_i - \pi_j \quad \leq d_{ij} \quad [y_{ij}] \qquad \forall (i,j) \in A(\mathbf{x}^k). \end{aligned} \tag{13}$$

With respect to $\boldsymbol{\pi}$, the reduced cost of $x_{ij}, (i,j) \in A$, in (9) is given by $\bar{c}_{ij} := c_{ij} - \pi_i + \pi_j$. Let the reduced cost $\bar{d}_{ij}$ of $f_{ij}, (i,j) \in A(\mathbf{x}^k)$, be computed in the same way, $\bar{d}_{ij} := d_{ij} - \pi_i + \pi_j$. The primal formulation (12) of the pricing problem and the dual version (13) are intimately related to the following two optimality conditions on the residual network $G(\mathbf{x}^k)$, each one being equivalent to the classical complementary slackness conditions only verified on the original network G, see Ahuja et al. [1, Theorems 9.1, 9.3, and 9.4]:

- *Primal:* $G(\mathbf{x}^k)$ contains no negative cost directed cycle.
- *Dual:* $\exists \boldsymbol{\pi}$ such that $\bar{d}_{ij} \geq 0,\ \forall (i,j) \in A(\mathbf{x}^k)$.
- *Complementary slackness:* $\exists \boldsymbol{\pi}$ such that, for every arc $(i,j) \in A$,

$$x_{ij}^k = \ell_{ij} \ \text{ if } \bar{c}_{ij} > 0; \quad x_{ij}^k = u_{ij} \ \text{ if } \bar{c}_{ij} < 0; \quad \bar{c}_{ij} = 0 \ \text{ if } \ell_{ij} < x_{ij}^k < u_{ij}.$$

Clearly, (13) searches for the minimum reduced cost μ while optimizing the dual variables $\pi_i,\ i \in V$, but there is no reference to the cost $z^k = \mathbf{c}^\intercal \mathbf{x}^k$. This pricing problem can be solved in polynomial time by the dynamic programming approach devised by Karp [10] which runs in $O(|V||A|)$ time. Strongly polynomial convergence properties of MMCC are derived from the nondecreasing behavior of μ, see the seminal paper by Goldberg and Tarjan [9], and the additional papers by Radzik and Goldberg [13] and Gauthier et al. [6] for successive improved complexity results.

We next show that for the MMCC algorithm applied on $CMCF$, the dual objective $(\boldsymbol{\pi}^{k+1})^\intercal \mathbf{b}$ is not equal to the cost $\mathbf{c}^\intercal \mathbf{x}^k$ if $\mu^{k+1} < 0$. In fact, vector $\boldsymbol{\pi}^{k+1}$ retrieved from the solution of (13) is optimal if the arc costs are redefined as $c_{ij} := c_{ij} - \mu^{k+1}, \forall (i,j) \in A$. In that case, the set of constraints

$$(d_{ij} - \mu^{k+1}) - \pi_i + \pi_j \geq 0, \qquad \forall (i,j) \in A(\mathbf{x}^k), \tag{14}$$

satisfies the dual optimality condition for the redefined network flow problem, and as a consequence the primal and dual objective functions are equal, that is,

$$(\boldsymbol{\pi}^{k+1})^\intercal \mathbf{b} = (\mathbf{c} - \mathbf{1}\mu^{k+1})^\intercal \mathbf{x}^k = z^k - \mu^{k+1} \mathbf{1}^\intercal \mathbf{x}^k. \tag{15}$$

The same line of argumentation can be used to prove that applying the MMCC algorithm on the standard linear program LP in (1) also provides expression (15).

3.2 The primal simplex algorithm

Let us remain in the context of a network flow problem. Surprisingly, the pricing step of the *primal simplex* (PS) algorithm [see 3] with the minimum reduced cost rule for the selection of the entering variable in the basis can also be seen as a linear fractional problem, but in a shrunken version. This is also true for the *improved primal simplex* (IPS) algorithm [5, 14, 12]. Indeed, MMCC, PS and IPS are shown, amongst other algorithms, to be special cases of a *vector space decomposition* framework for linear programs [7].

To describe the framework from which these three algorithms can be derived and hence to illustrate the presence of a linear fractional pricing problem in PS and IPS (as the one described above for MMCC), let us analyze these on the residual network $G(\mathbf{x}^k)$ provided in Figure 1. In this figure, variables x_{ij}^k at their lower/upper bounds correspond to arcs with a single direction while the other variables with $\ell_{ij} < x_{ij}^k < u_{ij}$ are illustrated with bidirectional arcs (indeed, two arcs for each variable). Let us denote F the latter subset of arcs whose flow can move upwards and downwards in $G(\mathbf{x}^k)$. Recall that the cost and the reduced cost of a directed cycle are equal because the dual part of the arc reduced cost sums up to zero. Hence, the pricing problem (12) in the MMCC algorithm finds in fact a directed cycle on $G(\mathbf{x}^k)$ with *minimum mean reduced cost* while optimizing all dual variables $\pi_i,\ i \in V$.

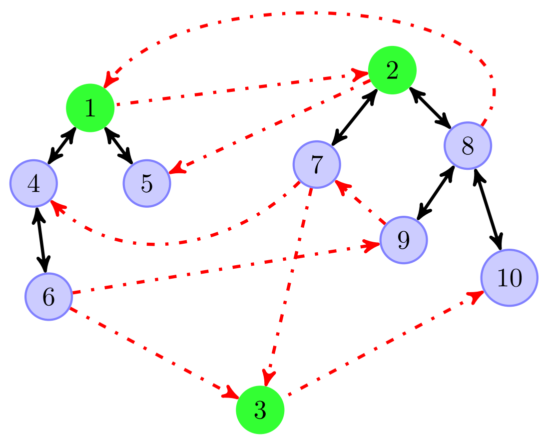


Figure 1: A residual network $G(\mathbf{x}^k)$ in MMCC.

At the other extreme is the PS algorithm where all the dual variables are fixed, that is, at iteration $k+1$, vector $\boldsymbol{\pi}^{k+1}$ is evaluated according to the current basic solution. Let us start the analysis with the notation used in linear programming. Let $\mathbf{x}^k = (\mathbf{x}_B^k, \mathbf{x}_N^k)$ be a basic solution to LP, where $\mathbf{A}_B\mathbf{x}_B^k = \mathbf{b}$ whereas $\mathbf{x}_N^k = \mathbf{0}$. At iteration $k+1$, the reduced cost of the basic variables is set to zero, i.e., $\bar{\mathbf{c}}_B^\intercal = \mathbf{c}_B^\intercal - (\boldsymbol{\pi}^{k+1})^\intercal \mathbf{A}_B = \mathbf{0}$, from which one derives $(\boldsymbol{\pi}^{k+1})^\intercal = \mathbf{c}_B^\intercal \mathbf{A}_B^{-1}$. It also follows that

$$z^k = \mathbf{c}_B^\intercal \mathbf{x}_B^k = (\boldsymbol{\pi}^{k+1})^\intercal \mathbf{A}_B \mathbf{x}_B^k = (\boldsymbol{\pi}^{k+1})^\intercal \mathbf{b}. \tag{16}$$

Therefore, at every iteration of PS, the dual objective $(\boldsymbol{\pi}^{k+1})^\intercal \mathbf{b}$ is equal to the cost of the current solution z^k. The pricing step in PS finds the minimum reduced cost μ^{k+1} over the nonbasic variables in set N, and this can be found by enumeration, or by solving the following linear program

$$\begin{aligned} \mu^{k+1} := \max \quad & \mu \\ \text{s.t.} \quad & \mu \leq \bar{c}_j \quad [y_j] \quad \forall j \in N \end{aligned} \tag{17}$$

or its dual

$$\begin{aligned} \min \quad & \sum_{j \in N} \bar{c}_j y_j \\ \text{s.t.} \quad & \sum_{j \in N} y_j = 1 \quad [\mu] \\ & y_j \geq 0 \qquad j \in N. \end{aligned} \tag{18}$$

Observe once again the presence of the convexity constraint in (18). An extreme point solution to the above identifies a single nonbasic variable $y_j = 1$ for which x_j uniquely impacts some basic variables that altogether form a combination of variables at a cost $\mu^{k+1} = \bar{c}_j$, that is, the reduced cost of any nonbasic variable reflects the unit cost of the associated combination. The step size is determined with the classical ratio-test using the variables of the combination.

The correspondence on a network flow problem is as follows. In the pricing problem, the basic variables, whose associated arcs form a spanning tree reaching every node of the network, are hidden from the residual network $G(\mathbf{x}^k)$. One only sees a single vertex (the root node of the spanning tree) together with the nonbasic arcs in $N = A \setminus B$, see network $B(\mathbf{x}^k)$ on Figure 2. The pricing problem finds on $B(\mathbf{x}^k)$ a loop of *minimum mean reduced cost*, the average being therefore taken on that single entering arc. Network $B(\mathbf{x}^k)$ is next expanded to identify on $G(\mathbf{x}^k)$ the associated directed cycle on which as much flow as possible is sent to compute the next solution $\mathbf{x}^{k+1}$. The step size is zero if the associated directed cycle comprises a basic arc in a direction opposite to that of the selected entering arc; otherwise the step size is equal to the minimum residual capacity on the arcs of the directed cycle.

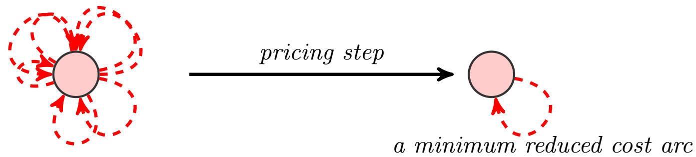
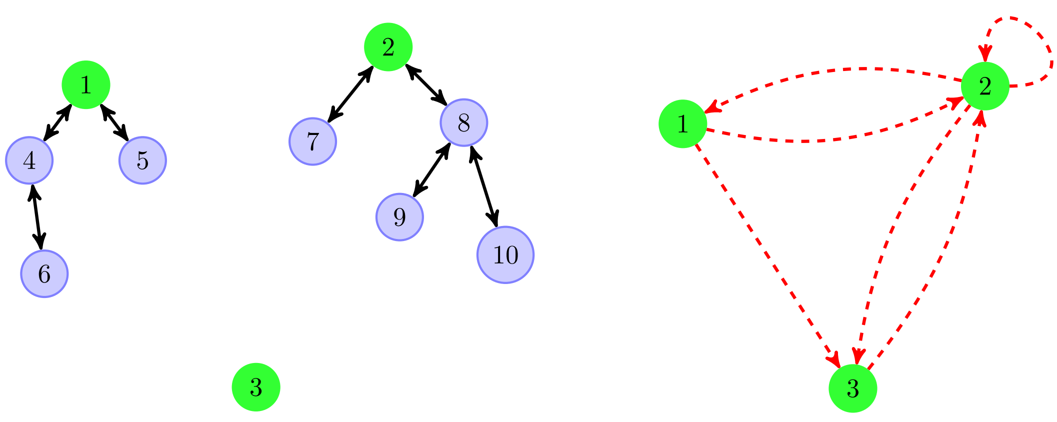


Figure 2: Shrunk network $B(\mathbf{x}^k)$ in the *primal simplex* algorithm.

3.3 The improved primal simplex algorithms

Degeneracy in the PS algorithm for network flow problems comes from the opposite orientation (compared to that of the entering arc) of at least one degenerate basic arc of the selected directed cycle in $G(\mathbf{x}^k)$. In fact, the orientation of the basic arcs is hidden while solving the pricing problem. IPS resolves this difficulty by hiding only the arcs of subset $F \subseteq B$, where F is nonempty because $\mathbf{x}^k \neq \mathbf{0}$.

The IPS algorithm first identifies the forest (a collection of trees) formed by the arcs of F before setting the reduced costs of these arcs to zero (as in PS), see Figure 3(a) derived from $G(\mathbf{x}^k)$ in Figure 1. Network $F(\mathbf{x}^k)$ in Figure 3(b) comprises one root node per tree of the forest, that is, only three vertices. Removing the dominated multiarcs of $A \setminus F$ based on their reduced costs, the pricing problem of IPS finds on $F(\mathbf{x}^k)$ a directed cycle of *minimum mean reduced cost*, the average being taken on the number of selected arcs. Expanding network $F(\mathbf{x}^k)$ to identify on $G(\mathbf{x}^k)$ the associated directed cycle always results in a cycle with positive residual capacities on all arcs, hence a nondegenerate pivot at every iteration.



(a) Forest induced by the arcs of F at $\mathbf{x}^k$

(b) $F(\mathbf{x}^k)$, without the dominated arcs

Figure 3: Contracted network in the *improved primal simplex* algorithm.

What remains to show is that, at iteration $k+1$ of the IPS algorithm, the dual objective $(\boldsymbol{\pi}^{k+1})^\intercal \mathbf{b}$ is equal to the cost of the current solution z^k. To do so, we rewrite formulation LP in (1) by using a partition $\{F, L\}$ of the variables and a partition $\{R, S\}$ of the constraints. First consider a basic solution $\mathbf{x}^k = (\mathbf{x}_F^k, \mathbf{x}_L^k)$, where $\mathbf{x}_j^k > 0, \forall j \in F$, and $\mathbf{x}_j^k = 0, \forall j \in L$ (the variables at the zero lower bound). Because all the positive variables are basic ($F \subseteq B$), columns of $\mathbf{A}_F$ are linearly independent and one can extract from these a nonsingular matrix $\mathbf{A}_{RF}$ of dimension $|F| \times |F|$. These rows define set R while the remaining ones are in set

$S := \{1, \ldots, m\} \setminus R$. Given $\mathbf{x}^k$, the linear system of LP reads as follows:

$$\begin{array}{rcccl} z^k = & \mathbf{c}_F^\intercal \mathbf{x}_F^k & + & \mathbf{c}_L^\intercal \mathbf{x}_L^k & \\ & \mathbf{A}_{RF}\mathbf{x}_F^k & + & \mathbf{A}_{RL}\mathbf{x}_L^k & = \mathbf{b}_R \quad [\boldsymbol{\pi}_R] \\ & \mathbf{A}_{SF}\mathbf{x}_F^k & + & \mathbf{A}_{SL}\mathbf{x}_L^k & = \mathbf{b}_S \quad [\boldsymbol{\pi}_S], \end{array} \tag{19}$$

where we have $z^k = \mathbf{c}_F^\intercal \mathbf{x}_F^k$, $\mathbf{x}_F^k = \mathbf{A}_{RF}^{-1}\mathbf{b}_R$, and also $\mathbf{A}_{SF}\mathbf{x}_F^k = \mathbf{b}_S$ because $\mathbf{x}_L^k = \mathbf{0}$. In the IPS algorithm, the reduced costs of the positive variables are fixed to zero while the pricing problem optimizes that of the variables in L (basic and nonbasic null variables), that is, $\bar{\mathbf{c}}_F^\intercal = \mathbf{c}_F^\intercal - (\boldsymbol{\pi}_R^{k+1})^\intercal \mathbf{A}_{RF} - (\boldsymbol{\pi}_S^{k+1})^\intercal \mathbf{A}_{SF} = \mathbf{0}$, or equivalently,

$$(\boldsymbol{\pi}_R^{k+1})^\intercal \quad = \quad \mathbf{c}_F^\intercal \mathbf{A}_{RF}^{-1} - (\boldsymbol{\pi}_S^{k+1})^\intercal \mathbf{A}_{SF}\mathbf{A}_{RF}^{-1}. \tag{20}$$

Therefore, the dual objective value at iteration $k+1$ becomes

$$\begin{aligned} (\boldsymbol{\pi}^{k+1})^\intercal \mathbf{b} &= (\boldsymbol{\pi}_R^{k+1})^\intercal \mathbf{b}_R && + \quad (\boldsymbol{\pi}_S^{k+1})^\intercal \mathbf{b}_S \\ &= \mathbf{c}_F^\intercal \mathbf{A}_{RF}^{-1}\mathbf{b}_R - (\boldsymbol{\pi}_S^{k+1})^\intercal \mathbf{A}_{SF}\mathbf{A}_{RF}^{-1}\mathbf{b}_R && + \quad (\boldsymbol{\pi}_S^{k+1})^\intercal \mathbf{b}_S \\ &= \mathbf{c}_F^\intercal \mathbf{x}_F^k - (\boldsymbol{\pi}_S^{k+1})^\intercal \mathbf{A}_{SF}\mathbf{x}_F^k && + \quad (\boldsymbol{\pi}_S^{k+1})^\intercal \mathbf{A}_{SF}\mathbf{x}_F^k \\ &= \mathbf{c}_F^\intercal \mathbf{x}_F^k \\ &= z^k. \end{aligned} \tag{21}$$

3.4 An alternative derivation of the pricing problem

Primal and dual formulations (4)–(5) of the pricing problem can also be derived using the MMCC strategy which, as described above, indirectly uses the linear fractional formulation (11) for solving $CMCF$. This is done on a reformulation of LP at every iteration.

At iteration $k+1$, $k \geq 0$, let $\mathbf{x}^k$ be a feasible solution of cost z^k. Then let $\lambda \in [0,1]$ be an auxiliary variable of cost z^k and rewrite LP as the following linear program LP^{k+1}:

$$\begin{array}{rlcccl} z^\star = & \min & \mathbf{c}^\intercal \mathbf{x} & + & z^k \lambda & \\ & \text{s.t.} & \mathbf{A}\mathbf{x} & + & \mathbf{b}\lambda & = \mathbf{b} \quad [\boldsymbol{\pi}] \\ & & \mathbf{0} \leq \mathbf{x} & & & \\ & & & & 0 \leq \lambda & \leq 1. \end{array} \tag{22}$$

For all $k \geq 0$, vector $(\hat{\mathbf{x}}, \hat{\lambda}) = (\mathbf{0}, 1)$ is feasible for LP^{k+1} at a cost z^k. We define the residual problem RP^{k+1} at $(\hat{\mathbf{x}}, \hat{\lambda})$ by using the following change of variables with residual vector $\mathbf{y} \in \mathbb{R}_+^n$ and residual variable $\theta \in [0,1]$:

$$\begin{pmatrix} \mathbf{x} \\ \lambda \end{pmatrix} = \begin{pmatrix} \hat{\mathbf{x}} \\ \hat{\lambda} \end{pmatrix} + \begin{pmatrix} \mathbf{y} \\ -\theta \end{pmatrix} = \begin{pmatrix} \mathbf{y} \\ 1-\theta \end{pmatrix}, \qquad \mathbf{y} \geq \mathbf{0},\ 0 \leq \theta \leq 1. \tag{23}$$

Substituting for $\mathbf{x}$ and λ in (22), one obtains RP^{k+1}, equivalent to LP^{k+1}:

$$\begin{array}{rcclcccl} z^\star = & z^k & + & \min & \mathbf{c}^\intercal \mathbf{y} & - & z^k \theta & \\ & & & \text{s.t.} & \mathbf{A}\mathbf{y} & - & \mathbf{b}\theta & = \mathbf{0} \quad [\boldsymbol{\pi}] \\ & & & & \mathbf{0} \leq \mathbf{y} & & & \\ & & & & & & 0 \leq \theta & \leq 1. \end{array} \tag{24}$$

A feasible solution to RP^{k+1} is $(\mathbf{y}, \theta) = (\mathbf{0}, 0)$ and it is optimal if and only if the reduced cost of all variables is greater than or equal to zero. This characterization of optimality on the residual problem is a well known result for network flow problems [see 1] and it has been extended to linear programs [7]. Let $\bar{\mathbf{c}}^\intercal = \mathbf{c}^\intercal - \boldsymbol{\pi}^\intercal \mathbf{A}$ be the reduced cost vector for the $\mathbf{y}$-variables and $\bar{c}_\theta = -z^k + \boldsymbol{\pi}^\intercal \mathbf{b}$ the reduced cost of variable θ. The proposed pricing problem computes the smallest reduced cost μ^{k+1}, where dual vector $\boldsymbol{\pi}$ is optimized but is also required to satisfy $\bar{c}_\theta \geq 0$:

$$\begin{array}{rllcrll} \mu^{k+1} := & \max & \mu & & & & \\ & \text{s.t.} & \mathbf{1}\mu & + & \mathbf{A}^\intercal\boldsymbol{\pi} & \leq \mathbf{c} & [\mathbf{y}] \\ & & & & -\mathbf{b}^\intercal\boldsymbol{\pi} & \leq -z^k & [\theta]. \end{array} \tag{25}$$

This is formulation (5) and its dual is (4), expressed in terms of the residual vector $\mathbf{y} \geq \mathbf{0}$ and the residual variable $\theta \geq 0$.

- If $\mu^{k+1} < 0$ and $\theta^{k+1} > 0$ in Step 3 of Algorithm LFP, one updates the solution with improving direction $(\mathbf{y}^{k+1}, \theta^{k+1})$ on $(\hat{\mathbf{x}}, \hat{\lambda})$, as expressed in (23). Given $(\hat{\mathbf{x}}, \hat{\lambda}) = (\mathbf{0}, 1)$, the maximum step size ρ must satisfy the nonnegativity constraints:

$$\begin{pmatrix} \hat{\mathbf{x}} \\ \hat{\lambda} \end{pmatrix} + \rho \begin{pmatrix} \mathbf{y}^{k+1} \\ -\theta^{k+1} \end{pmatrix} \geq \begin{pmatrix} \mathbf{0} \\ 0 \end{pmatrix}. \tag{26}$$

  Step size $\rho^{k+1} = 1/\theta^{k+1}$ and a new primal solution to LP is $\mathbf{x}^{k+1} := \mathbf{y}^{k+1}/\theta^{k+1}$ at an improved cost

$$z^{k+1} = z^k + \rho^{k+1}\mu^{k+1} = \mathbf{c}^\intercal\mathbf{y}^{k+1}/\theta^{k+1} = \mathbf{c}^\intercal\mathbf{x}^{k+1} < z^k. \tag{27}$$

- If $\mu^{k+1} < 0$ and $\theta^{k+1} = 0$ occurs in Step 4, step size is infinite and the algorithm stops for unboundedness of LP.
- Otherwise $\mu^{k+1} = 0$ in Step 5 and $(\mathbf{y}, \theta) = (\mathbf{0}, 0)$ is optimal for RP^{k+1}, that is, we have on hand a primal solution $\mathbf{x}^k$ and a feasible dual vector $\boldsymbol{\pi}^{k+1}$ with the same cost. Therefore the primal-dual pair $(\mathbf{x}^k, \boldsymbol{\pi}^{k+1})$ is optimal for LP.

4 Basic properties

In this section, we examine the behavior of three quantities throughout the iterative process: the minimum reduced cost μ, the objective function value z, and the sum $s = \mathbf{1}^\intercal\mathbf{x} = 1/\theta$ of the $\mathbf{x}$-variables.

The discussion on the algorithm at the end of Section 2.1 leads us to two propositions. Proposition 1 refers to successive values of the objective function, the proof of which being already given through Steps 3–5 and relation (8), whereas Proposition 2 characterizes an optimality condition for the linear program LP.

Proposition 1 *For $k \geq 0$, if $\mu^{k+1} < 0$, then $z^{k+1} < z^k$ else $z^{k+1} = z^k$.*

Proposition 2 *For $k \geq 0$, a feasible solution $\mathbf{x}^k$ is optimal for LP if and only if $\mu^{k+1} = 0$.*

Proof. Recall $(\mathbf{y}^{k+1}, \theta^{k+1};\ \boldsymbol{\pi}^{k+1}, \mu^{k+1})$ is a primal-dual solution to PP^{k+1}. [$\Leftarrow$] Suppose $\mu^{k+1} = 0$. Then $(\mathbf{x}^k; \boldsymbol{\pi}^{k+1})$ is a pair of primal and dual feasible solutions whose costs are equal, that is, $\mathbf{b}^\intercal\boldsymbol{\pi}^{k+1} = z^k = \mathbf{c}^\intercal\mathbf{x}^k$. Therefore, it also satisfies the complementary slackness optimality conditions for linear programming, see Schrijver [15]. [$\Rightarrow$] The second part of the proof is by contradiction. Suppose $\mu^{k+1} < 0$. Then solution $\mathbf{x}^k$ can be improved: For $\theta^{k+1} > 0$, $\mathbf{x}^{k+1} := \mathbf{y}^{k+1}/\theta^{k+1}$ at a cost $z^{k+1} < z^k$ by expression (27). For $\theta^{k+1} = 0$, LP is unbounded and again $\mathbf{x}^k$ is nonoptimal. □

We take a step back to analyze the two cases that may occur when $\mu^{k+1} = 0$ and the algorithm stops on an optimal solution. PP^{k+1} determines $\boldsymbol{\pi}^{k+1}$ as an optimal dual vector but primal solutions depend on the value of θ^{k+1}.

- If $\theta^{k+1} = 0$, $\mathbf{y}^{k+1}$ is a zero-cost ($\mu^{k+1} = \mathbf{c}^\intercal\mathbf{y} = 0$) ray in the cone $\{\mathbf{y} \geq \mathbf{0} \mid \mathbf{A}\mathbf{y} = \mathbf{0}\}$ in (4) and $\mathbf{x}^{k+1} = \mathbf{x}^k + \rho\,\mathbf{y}^{k+1},\ \rho > 0$, provides alternative optimal solutions to $\mathbf{x}^k$.
- Otherwise $\theta^{k+1} > 0$ and $\mathbf{x}^{k+1} = \mathbf{y}^{k+1}/\theta^{k+1}$ is either $\mathbf{x}^k$ or another optimal solution.

Assumption. For the rest of the paper, we assume that no unboundedness situation occurs, that is, the sum of the $\mathbf{x}$-variables is bounded from above by a constant, $0 < s \leq s_{max}$ in any feasible solution.

We know that the algorithms LFP, MMCC, and IPS converge to an optimal solution and that, in these algorithms, the objective function value z decreases at every iteration ($z^k > z^{k+1},\ k \geq 0$). This is not the case for PS for which degenerate pivots may occur. It is also known that the minimum reduced cost μ oscillates in PS and IPS throughout the iterations [7] and may even not converge to zero in the case of PS (possibility of cycling). It is however nondecreasing for MMCC ($\mu^k \leq \mu^{k+1},\ k \geq 1$), see Goldberg and Tarjan [9] for network flow problems and Gauthier et al. [7] for linear programs. Moreover, in these same two references, one finds that μ increases after a number of iterations at most equal to the number of $\mathbf{x}$-variables in the original formulation, that is, after at most the number of arcs in a network flow problem ($\mu^k < \mu^{k+|A|}, k \geq 1$) or the number of variables in LP ($\mu^k < \mu^{k+n},\ k \geq 1$). We show in Proposition 3 a much stronger property for LFP: μ increases at every iteration.

Proposition 3 *For $k \geq 1$, let $\mathbf{x}^k$ be a nonoptimal solution to LP. Then $\mu^k < \mu^{k+1}$.*

Proof. Let $(\mathbf{y}^k, \theta^k;\ \mu^k)$ and $(\mathbf{y}^{k+1}, \theta^{k+1};\ \mu^{k+1})$ be retrieved from optimal solutions to PP^k and PP^{k+1}, respectively. Then

$$\mu^k \ = \ \mathbf{c}^\intercal \mathbf{y}^k - z^{k-1}\theta^k \ \leq \ \mathbf{c}^\intercal \mathbf{y}^{k+1} - z^{k-1}\theta^{k+1} \ < \ \mathbf{c}^\intercal \mathbf{y}^{k+1} - z^{k}\theta^{k+1} \ = \ \mu^{k+1}. \tag{28}$$

The first inequality is valid at iteration $k \geq 1$ because $(\mathbf{y}^k, \theta^k)$ is optimal for PP^k. The strict inequality follows from $z^k < z^{k-1}$ in Proposition 1. Consequently, $\mu^k < \mu^{k+1}$. □

Having proved that μ is increasing from one iteration to the next in LFP (except at the last iteration when optimality is proven), we further characterize an *improvement factor* in terms of the sum s of the $\mathbf{x}$-variables.

Proposition 4 *For $k \geq 1$, let $\mathbf{x}^k$ be a feasible (optimal or not) solution to LP. Then*

$$\mu^{k+1} \geq \mu^k (1 - \frac{s^k}{s^{k+1}}), \qquad \text{where} \ \ 0 < \frac{s^k}{s^{k+1}} \leq 1. \tag{29}$$

Proof. Let $(\mathbf{y}^k, \theta^k;\ \mu^k)$ and $(\mathbf{y}^{k+1}, \theta^{k+1};\ \mu^{k+1})$ be retrieved from optimal solutions to PP^k and PP^{k+1}, respectively. Because $\theta^k, \theta^{k+1} > 0$, we can derive the finite solutions $\mathbf{x}^k = \mathbf{y}^k/\theta^k$ and $\mathbf{x}^{k+1} = \mathbf{y}^{k+1}/\theta^{k+1}$ for which

$$z^k = \mathbf{c}^\intercal \mathbf{x}^k = \frac{\mathbf{c}^\intercal \mathbf{y}^k}{\theta^k} \quad \text{and} \quad z^{k+1} = \mathbf{c}^\intercal \mathbf{x}^{k+1} = \frac{\mathbf{c}^\intercal \mathbf{y}^{k+1}}{\theta^{k+1}}. \tag{30}$$

Recall from (8) that $z^k - z^{k-1} = \mu^k/\theta^k,\ k \geq 1$. Optimality of $(\mathbf{y}^k, \theta^k)$ for PP^k induces the first inequality in the following sequence:

$$\begin{aligned} \mu^k = &\quad \mathbf{c}^\intercal \mathbf{y}^k - z^{k-1}\theta^k \\ \leq &\quad \mathbf{c}^\intercal \mathbf{y}^{k+1} - z^{k-1}\theta^{k+1} \\ = &\quad \mathbf{c}^\intercal \mathbf{y}^{k+1} - z^{k}\theta^{k+1} + z^{k}\theta^{k+1} - z^{k-1}\theta^{k+1} \\ = &\quad \mu^{k+1} + (z^k - z^{k-1})\theta^{k+1} \\ = &\quad \mu^{k+1} + (\mu^k/\theta^k)\ \theta^{k+1}. \end{aligned} \tag{31}$$

From the above, one obtains the left part of (29):

$$\mu^{k+1} \geq \mu^k (1 - \frac{\theta^{k+1}}{\theta^k}) = \mu^k (1 - \frac{s^k}{s^{k+1}}). \tag{32}$$

To complete the proof, first observe that ratio $s^k/s^{k+1} = \theta^{k+1}/\theta^k$ is positive. Secondly, multiply by $\frac{-1}{\theta^k\theta^{k+1}}$ the first inequality in (31):

$$\frac{z^{k-1}\theta^k - \mathbf{c}^\intercal \mathbf{y}^k}{\theta^k\,\theta^{k+1}} \quad\geq\quad \frac{z^{k-1}\theta^{k+1} - \mathbf{c}^\intercal \mathbf{y}^{k+1}}{\theta^k\,\theta^{k+1}} \quad\Leftrightarrow\quad \frac{z^{k-1} - z^k}{\theta^{k+1}} \quad\geq\quad \frac{z^{k-1} - z^{k+1}}{\theta^k}. \tag{33}$$

Therefore, given that $z^{k-1} > z^k \geq z^{k+1}$ for $k \geq 1$, then

$$0 < \frac{s^k}{s^{k+1}} = \frac{\theta^{k+1}}{\theta^k} \leq \frac{z^{k-1} - z^k}{z^{k-1} - z^{k+1}} \leq 1, \tag{34}$$

where, by Proposition 1, the equality to 1 can only occur when optimality is reached, that is, when $z^k = z^{k+1}$. $\square$

Summarizing the results of Propositions 1–4, we have the following corollary for a nonoptimal solution and for the stopping rule upon optimality:

Corollary 1 *For $k \geq 1$, let $\mathbf{x}^k$ be a feasible solution to LP.*

a) If $\mathbf{x}^k$ is nonoptimal, then in the next iteration, the objective function value z decreases while both the minimum reduced cost μ and the sum s of the $\mathbf{x}$-variables increase, that is,

$$z^k > z^{k+1}, \quad \mu^k < \mu^{k+1}, \quad s^k < s^{k+1}. \tag{35}$$

b) Certifying that $\mathbf{x}^k$ is optimal can be done in various equivalent ways. The stopping rule in Algorithm LFP can be given as follows:

$$z^{k+1} - z^k = 0, \quad \mu^{k+1} = 0, \quad s^{k+1} - s^k = 0. \tag{36}$$

Results from various algorithms are reported in Table 1.

Table 1: Properties of various algorithms given a nonoptimal solution $\mathbf{x}^k$.

	LFP	**MMCC**	**IPS**	**PS**
z^{k+1}	$< z^k$	$< z^k$	$< z^k$	$\leq z^k$
μ^{k+1}	$> \mu^k$	$\geq \mu^k,\ \mu^{k+n} > \mu^k$	*oscillations*	*oscillations*
s^{k+1}	$> s^k$	*oscillations*	*oscillations*	*oscillations*
$(\boldsymbol{\pi}^{k+1})^\intercal \mathbf{b}$	$= z^k$	$= z^k - \mu^{k+1}s^k$	$= z^k$	$= z^k$

A numerical example

We present an example involving eight iterations. Data in Table 2 can be interpreted as the successive extreme point solutions $\mathbf{x}_p$, $p = 1, \ldots, 7$, selected during the iterative process until optimality is proven in the last iteration. These extreme points must be seen as a subset of a much larger set $\mathcal{P}$ of basic solutions. The seven solutions are indexed by increasing sum ($s_p = \mathbf{1}^\intercal \mathbf{x}_p$) and decreasing cost ($\mathbf{c}^\intercal \mathbf{x}_p$) according to Corollary 1. Additional information is also provided for $\theta_p = 1/s_p$ and $\mathbf{c}^\intercal \mathbf{y}_p = \mathbf{c}^\intercal \mathbf{x}_p \theta_p$.

The algorithm is initialized at $k = 0$ with an upper bound $z^0 = 800$ whereas an optimal solution is given by $\mathbf{x}_7$ with a cost of $z_7 = 37$ and a sum $s_7 = 160$. At iteration $k+1$, the value

$$\mu^{k+1} = \min_{p\in\mathcal{P}} \frac{\mathbf{c}^\intercal \mathbf{x}_p - z^k}{s_p} \tag{37}$$

is indicated in boldface character in Table 2, where we also report the ratio $(\mathbf{c}^\intercal \mathbf{x}_p - z^k)/s_p$ for the seven extreme points $\mathbf{x}_p$, $p = 1, \ldots, 7$. Hence, $\mathbf{x}^1 = \mathbf{x}_1$ is selected at the first iteration ($\mu^1 = -30$), $\mathbf{x}^2 = \mathbf{x}_2$ is selected at the second iteration ($\mu^2 = -14,23$), and so on. Figure 4 plots the minimum reduced cost $\mu(s)$ and the objective function value $z(s)$ in terms of the sum of the variables, that is, the sets of coordinates (s^{k+1}, μ^{k+1}) and (s^{k+1}, z^{k+1}), for $k = 0, \ldots, 7$. Observe that $\mu(s)$ is increasing by Proposition 3 and that $z(s)$ is decreasing by Proposition 1.

Table 2: Data and solution process for an example involving eight iterations.

$p \in \{1,\dots,7\}$	z_0	$\mathbf{x}_1$	$\mathbf{x}_2$	$\mathbf{x}_3$	$\mathbf{x}_4$	$\mathbf{x}_5$	$\mathbf{x}_6$	$\mathbf{x}_7$
s_p		10	26	30	46	80	140	160
$\mathbf{c}^\intercal\mathbf{x}_p$	800	500	130	80	60	46	38	37
θ_p		0.1	0.0385	0.0333	0.0217	0.0125	0.00714	0.00625
$\mathbf{c}^\intercal\mathbf{y}_p$		50	5	2.6667	1.3043	0.575	0.27143	0.23125

		Computation of $\frac{\mathbf{c}^\intercal\mathbf{x}_p - z^k}{s_p}$, $p \in \{1,\dots,7\}$						
k	μ^{k+1}	$\mathbf{x}_1$	$\mathbf{x}_2$	$\mathbf{x}_3$	$\mathbf{x}_4$	$\mathbf{x}_5$	$\mathbf{x}_6$	$\mathbf{x}_7$
0	-30.0000	**-30.0**	-25.77	-24.0000	-16.0870	-9.4250	-5.44286	-4.76875
1	-14.2308	0.0	**-14.23**	-14.0000	-9.5652	-5.6750	-3.30000	-2.89375
2	-1.6667	37.0	0.00	**-1.6667**	-1.5217	-1.0500	-0.65714	-0.58125
3	-0.4348	42.0	1.92	0.0000	**-0.4348**	-0.4250	-0.30000	-0.26875
4	-0.1750	44.0	2.69	0.6667	0.0000	**-0.1750**	-0.15714	-0.14375
5	-0.0571	45.4	3.23	1.1333	0.3043	0.0000	**-0.05714**	-0.05625
6	-0.0063	46.2	3.54	1.4000	0.4783	0.1000	0.00000	**-0.00625**
7	0.0000	46.3	3.58	1.4333	0.5000	0.1125	0.00714	0

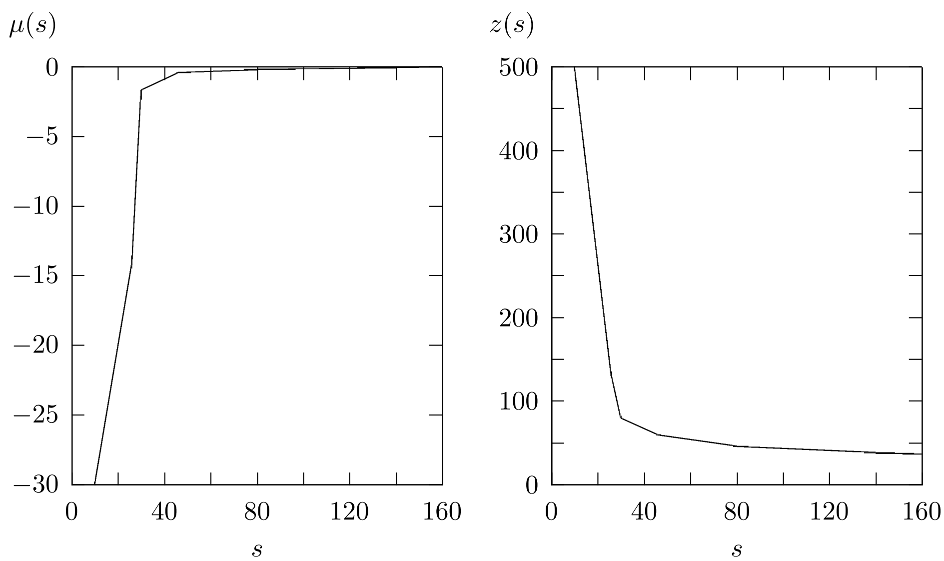


Figure 4: Functions $\mu(s)$ and $z(s)$ for the numerical example.

5 Convergence of the minimum reduced cost

Algorithm LFP optimizes a sequence of pricing problems that all have the same domain $\mathcal{D} = \{\mathbf{y} \geq \mathbf{0},\ \theta \geq 0 \mid \mathbf{A}\mathbf{y} - \mathbf{b}\theta = \mathbf{0},\ \mathbf{1}^\intercal\mathbf{y} = 1\}$ in (4) at every iteration. The iterative process is finite as Algorithm LFP only examines the extreme points of $\mathcal{D}$ which differ from one iteration to the next by a greater value of μ (Proposition 3) until optimality is reached at $\mu = 0$ (Proposition 2). Moreover, for $k \geq 1$, $\mu^{k+1} \geq \mu^k(1 - s^k/s^{k+1})$ by Proposition 4, where $s^k < s^{k+1}$ for a nonoptimal solution $\mathbf{x}^k$. Therefore, recursively applying the previous relation, one obtains

$$\mu^{k+1} \geq \mu^1(1 - \frac{s^1}{s^2})(1 - \frac{s^2}{s^3})\dots(1 - \frac{s^k}{s^{k+1}}), \qquad k \geq 1. \tag{38}$$

5.1 Special cases

We describe in the next paragraphs some special cases for which the number of iterations or the behavior of μ in Algorithm LFP is easily derived.

Algorithm LFP finds an optimal solution in a single iteration if the sum of the variables is imposed as a constraint in LP or is implicitly constant as in a network assignment problem or a balanced transportation problem. This comes out from relation (29) in Proposition 4 or from (38) where $s^2 = s^1$, yielding $\mu^2 = 0$ and confirming the optimality of $\mathbf{x}^1$ by Corollary 1. This indicates that, in these two special cases, PP^1 is in fact equivalent to LP and therefore as easy or as hard to solve. Moreover, we have the following property if the objective function of LP explicitly minimizes the sum of the variables as in the cutting stock application, see Gilmore and Gomory [8].

Proposition 5 *If the objective of LP minimizes $\mathbf{1}^\intercal\mathbf{x}$ (the sum of the variables), then Algorithm LFP determines an optimal solution in the first iteration.*

Proof. If $\mathbf{c}^\intercal\mathbf{x} = \mathbf{1}^\intercal\mathbf{x}$, then $\mathbf{c}^\intercal\mathbf{y} = 1$ in (4) and the objective function of PP^1 minimizes $1 - z^0\theta$, where $z^0 > 0$ is an upper bound on the sum of the variables in any optimal solution. This objective is equivalent to maximizing θ, that is, maximizing $1/\mathbf{1}^\intercal\mathbf{x}$, or finding a solution with the smallest sum of variables at the very first iteration. □

Many vehicle routing and crew scheduling applications can be formulated as huge set partitioning models solved by branch-and-price, i.e., the linear relaxation is solved by column generation, see Desaulniers et al. [4] and Lübbecke and Desrosiers [11]. Using the Algorithm LFP framework within such a solution strategy, the role of the master problem is only to find the step size on the direction found by the pricing problem which is itself solved by column generation. It is expected that finding a *near optimal solution* to the LP relaxation of such large-scale set partitioning models would occur after a very small number of iterations of the master problem. This can be explained as follows.

Let z^0 be a relatively large upper bound on the optimal value $z^\star$. Then PP^1 is solved such that $\mathbf{c}^\intercal\mathbf{y} - z^0\theta$ is minimized, hence assigning a large weight on θ at the first iteration. This favors finding a large value for θ^1 or, equivalently, a relatively small sum of the variables, that is, a value $s^1 = 1/\theta^1$ more or less close to the optimal number of vehicles or crews when the vehicles or crews incur large fixed costs in the objective function. The next iteration confirms the near optimality of $\mathbf{x}^1$. Otherwise it finds an improved solution $\mathbf{x}^2 = \mathbf{y}^2/\theta^2$ with $s^2 > s^1$ closer to the optimal sum of the variables.

5.2 Geometric and super-geometric growth rates

Since the sum of the $\mathbf{x}$-variables is increasing, let us use the least sum s^1 and the largest one s^{k+1} to replace the k ratios within (38) by $r = s^1/s^{k+1}$. By doing so, one obtains

$$\mu^{k+1} \geq \mu^1(1 - \frac{s^1}{s^2})(1 - \frac{s^2}{s^3}) \dots (1 - \frac{s^k}{s^{k+1}}) \geq \mu^1(1-r)^k, \qquad k \geq 1. \tag{39}$$

Since $\mu^1 < 0$, this shows a geometric growth rate on μ. Observe that ratio $r = s^1/s^{k+1}$ can be replaced by $r_{min} = s^1/s_{max}$, where s_{max} is an upper bound on $s^\star$, the optimal sum of the variables.

Proposition 6 establishes a much faster rate, indeed a super-geometric growth rate on μ. It is based on the strict concavity of

$$f(s_2, \dots, s_k) = \quad (1 - \frac{s^1}{s^2})(1 - \frac{s^2}{s^3}) \dots (1 - \frac{s^k}{s^{k+1}}), \qquad s^1 < s^2 < \dots s^k < s^{k+1}, \tag{40}$$

a function involving $k-1$ variables. Strict concavity can easily be verified for the following one-, two-, and three-dimensional functions:

- $f(u) = (1 - \frac{a}{u})(1 - \frac{u}{b}), \quad 0 < a \leq u \leq b,$
  reaches its maximum $(1 - (\frac{a}{b})^{\frac{1}{2}})^2$ at $u = \sqrt{ab}$.
- $f(u, v) = (1 - \frac{a}{u})(1 - \frac{u}{v})(1 - \frac{v}{b}), \quad 0 < a \leq u \leq v \leq b,$
  reaches its maximum $(1 - (\frac{a}{b})^{\frac{1}{3}})^3$ at $(u, v) = (\sqrt[3]{a^2 b}, \sqrt[3]{ab^2})$.
- $f(u, v, w) = (1 - \frac{a}{u})(1 - \frac{u}{v})(1 - \frac{v}{w})(1 - \frac{w}{b}), \quad 0 < a \leq u \leq v \leq w \leq b,$
  reaches its maximum $(1 - (\frac{a}{b})^{\frac{1}{4}})^4$ at $(u, v, w) = (\sqrt[4]{a^3 b}, \sqrt[4]{a^2 b^2}, \sqrt[4]{ab^3})$.

The super-geometric growth rate on μ is based on the following lemma.

Lemma 1 *Let $\mathbf{v} = (v_1, v_2, \ldots, v_k)$. Define function*

$$f(\mathbf{v}) = (1 - \frac{a}{v_1})(1 - \frac{v_1}{v_2}) \ldots (1 - \frac{v_{k-1}}{v_k})(1 - \frac{v_k}{b}), \qquad 0 < a \leq v_1 \leq v_2 \cdots \leq v_k \leq b. \tag{41}$$

Then for $r = a/b$ and $k \geq 1$,

$$\max \ f(\mathbf{v}) = (1 - r^{\frac{1}{k+1}})^{k+1}. \tag{42}$$

Proof. The proof is done by induction on the value of k. The result is obviously valid for $k = 1$ using first and second derivatives of $f(v) = (1 - \frac{a}{v})(1 - \frac{v}{b}), \ 0 < a \leq v \leq b$.

$$\begin{aligned}
f'(v) = &\ \frac{a}{v^2} - \frac{1}{b} && \Rightarrow v^\star = \sqrt{ab} \\
f''(v) = &\ \frac{-2a}{v^3} < 0 && \Rightarrow f(v) \text{ is strictly concave} \\
f(v^\star) = &\ (1 - \frac{a}{\sqrt{ab}})(1 - \frac{\sqrt{ab}}{b}) = \left(1 - \sqrt{\frac{a}{b}}\right)^2
\end{aligned} \tag{43}$$

Assume now that relation (42) is true for up to $k - 1$. Therefore

$$\begin{aligned}
\max_{\mathbf{v}} f(\mathbf{v}) \ &= \ \max_{v_k} \left\{ \ (1 - \frac{v_k}{b}) \max_{v_1, \ldots, v_{k-1}} f(v_1, \ldots, v_{k-1}) \right\} \\
&= \ \max_{v_k} \ (1 - \frac{v_k}{b})(1 - (\frac{a}{v_k})^{\frac{1}{k}})^k
\end{aligned} \tag{44}$$

Let $v := v_k$. Then $g(v) = (1 - \frac{v}{b})(1 - (\frac{a}{v})^{\frac{1}{k}})^k$ is a single variable function defined on $[a, b]$. Define $h(v) = \ln g(v)$, a transformation by the monotonically increasing logarithmic function. Setting the first derivative to zero,

$$g'(v) = (1 - (\frac{a}{v})^{\frac{1}{k}})^{k-1}(\frac{-1}{b} + \frac{1}{v}(\frac{a}{v})^{\frac{1}{k}}) = 0, \tag{45}$$

or equivalently,

$$h'(v) = \frac{-1}{b - v} + \frac{(\frac{a}{v})^{\frac{1}{k}}}{1 - (\frac{a}{v})^{\frac{1}{k}}} = 0, \tag{46}$$

one obtains $(v^\star)^{k+1} = ab^k$. The reader can verify that the second derivative $h''(v)$ is negative for $0 < a < v < b$. Hence $h(v)$ and $g(v)$ are strictly concave functions. Evaluating $g(v)$ at $v^\star = \sqrt[k+1]{ab^k}$ gives the requested maximum $(1 - (\frac{a}{b})^{\frac{1}{k+1}})^{k+1}$ in (42) for $f(\mathbf{v})$ and this completes the proof of the lemma. □

Proposition 6 *For $k \geq 1$, let $\mathbf{x}^k$ be a nonoptimal solution. Then*

$$\mu^{k+1} \geq \mu^1 (1 - r^{\frac{1}{k}})^k, \qquad \textit{where ratio } r = s^1/s^{k+1}. \tag{47}$$

Proof. Given $s^k < s^{k+1}, \ k \geq 1$, let $(1 - \frac{s^1}{s^2})(1 - \frac{s^2}{s^3}) \ldots (1 - \frac{s^k}{s^{k+1}})$ in expression (38) be written similarly to function $f(\cdot)$ in Lemma 1, that is,

$$f(s_2, \ldots, s_k) = \ \ (1 - \frac{a}{s^2})(1 - \frac{s^2}{s^3}) \ldots (1 - \frac{s^k}{b}), \quad a = s^1 < z^2 < \ldots s^k < s^{k+1} = b. \tag{48}$$

Applying Lemma 1 to $\mu^1 f(s^2, s^3, \ldots, s^k)$ where $\mu^1 < 0$, one obtains

$$\mu^{k+1} \geq \mu^1 f(s^2, s^3, \ldots, s^k) \geq \mu^1 (1 - r^{\frac{1}{k}})^k, \qquad \text{where } r = a/b = s^1/s^{k+1}. \tag{49}$$

This establishes a super-geometric growth rate on μ. $\square$

Here again, because $s^{k+1} \leq s_{max}$, one can instead use ratio $r_{min} = s^1/s_{max}$ in (47). Figure 5 illustrates the super-geometric growth rate of function

$$-(1 - r^{\frac{1}{k}})^k, \qquad 0 < r < 1, \tag{50}$$

for three different values of r up to ten iterations. Ratio r can be interpreted as follows. Let $s^\star$ be the sum of the variables in an optimal solution $\mathbf{x}^\star$. Therefore, $r = {s^1}/{s^\star} = 1/10$ means that $s^\star$ is 10 times larger than s^1 whereas $s^\star$ is only 25% larger than s^1 for $r = 1/1.25$. Note that a unit ratio means that $s^1 = s^\star$, the pricing problem is equivalent to LP, and PP^1 finds an optimal solution in a single iteration.

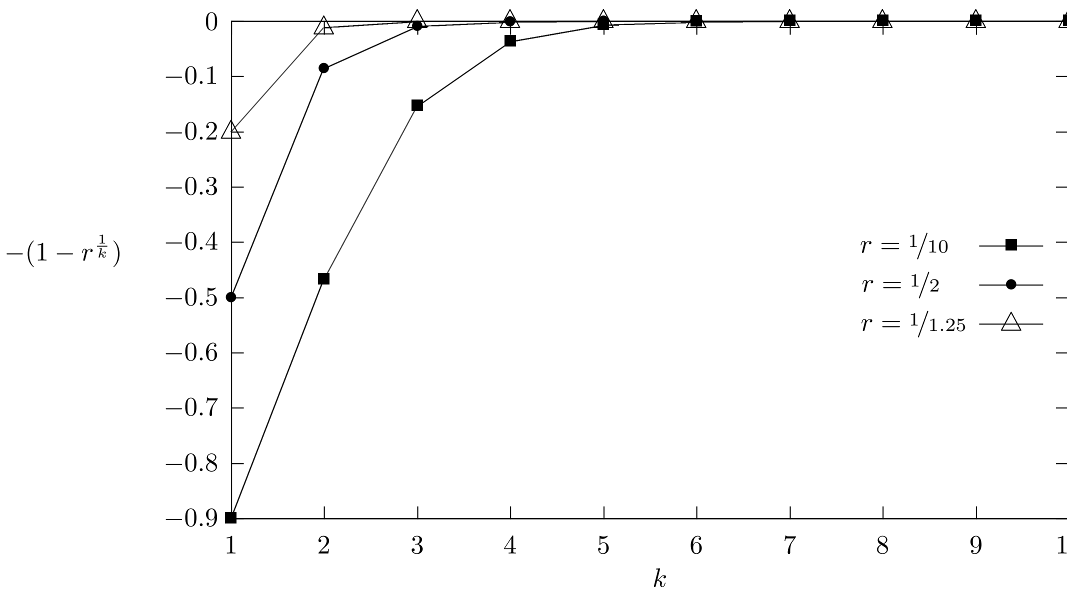


Figure 5: Function $-(1 - r^{\frac{1}{k}})$ for $r \in \{1/10, 1/2, 1/1.25\}$ and $1 \leq k \leq 10$.

6 Conclusion

In this paper, we introduced a primal algorithm, called Algorithm LFP, for solving a linear program LP based on a linear fractional pricing problem. At every iteration of this algorithm, the dual variables are optimized to find the largest possible minimum reduced cost value under the constraint that the primal and dual objective values of LP are equal. The resulting sequence of minimum reduced costs converges to zero and is shown to have a super-geometric growth rate.

Future research will first concentrate on performing computational experiments with Algorithm LFP. In particular, combining it with column generation should yield an efficient algorithm for solving vehicle routing and crew scheduling problems involving large fixed costs for each resource used. Then, various acceleration strategies can be developed to improve the practical performance of Algorithm LFP. For example, the linear fractional pricing problem can be solved heuristically at each iteration as long as it provides a solution with a negative value, this ensuring a decrease of the objective value. In this case, some properties of the algorithm such as the convergence rate and the increasing sum of the variables are not preserved but it may result in significant computational time reductions especially if a near-optimal solution is sought.

References

[1] R.K. Ahuja, T.L. Magnanti, and J.B. Orlin. Network Flows: Theory, Algorithms, and Applications. Prentice Hall, Upper Saddle River, NJ, USA, 1993.

[2] A. Charnes and W.W. Cooper. Programming with linear fractional functionals. Naval Research Logistics Quartely, 9(3-4):181–186, 1962. doi: 10.1002/nav.3800090303.

[3] G.B. Dantzig. Linear programming and extensions. Princeton University Press, Princeton, NJ, USA, 1963.

[4] G. Desaulniers, J. Desrosiers, I. Ioachim, M.M. Solomon, F. Soumis, and D. Villeneuve. A unified framework for deterministic time constrained vehicle routing and crew scheduling problems. In T.G. Crainic and G. Laporte, editors, Fleet Management and Logistics, pages 57–93. Springer, New York, NY, USA, 1998. doi: 10.1007/978-1-4615-5755-5_3.

[5] I. Elhallaoui, A. Metrane, G. Desaulniers, and F. Soumis. An Improved Primal Simplex algorithm for degenerate linear programs. INFORMS Journal on Computing, 23:569–577, 2011. doi: 10.1287/ijoc.1100.0425.

[6] J.B. Gauthier, J. Desrosiers, and M.E. Lübbecke. About the minimum mean cycle-canceling algorithm. Discrete Applied Mathematics, 2014. doi: 10.1016/j.dam.2014.07.005.

[7] J.B. Gauthier, J. Desrosiers, and M.E. Lübbecke. Vector space decomposition for linear programs. Les Cahiers du GERAD G-2015-26, HEC Montréal, Montreal, QC, Canada, March 2015.

[8] P.C. Gilmore and R.E. Gomory. A linear programming approach to the cutting-stock problem. Operations Research, 9(6):849–859, 1961. doi: 10.1287/opre.9.6.849.

[9] A.V. Goldberg and R.E. Tarjan. Finding minimum-cost circulations by canceling negative cycles. Journal of the ACM, 36(4):873–886, 1989. doi: 10.1145/76359.76368.

[10] R.M. Karp. A characterization of the minimum cycle mean in a digraph. Discrete Mathematics, 23(3):309–311, 1978. doi: 10.1016/0012-365X(78)90011-0.

[11] M.E. Lübbecke and J. Desrosiers. Selected topics in column generation. Operations Research, 53(6):1007–1023, 2005. doi: 10.1287/opre.1050.0234.

[12] A. Metrane, F. Soumis, and I. Elhallaoui. Column generation decomposition with the degenerate constraints in the subproblem. European Journal of Operational Research, 207(1):37–44, 2010. doi: 10.1016/j.ejor.2010.05.002.

[13] T. Radzik and A.V. Goldberg. Tight bounds on the number of minimum-mean cycle cancellations and related results. Algorithmica, 11(3):226–242, 1994. doi: 10.1007/BF01240734.

[14] V. Raymond, F. Soumis, and D. Orban. A new version of the Improved Primal Simplex for degenerate linear programs. Computers & Operations Research, 37(1):91–98, 2010. doi: 10.1016/j.cor.2009.03.020.

[15] A. Schrijver. Theory of Linear and Integer Programming. John Wiley & Sons, Inc., Chichester, West Sussex, England, 1986.