Commun. Optim. Theory 2027 (2027) 8 https://doi.org/10.23952/cot.2027.8

Communications in Optimization Theory
Available online at http://cot.mathres.org

LINEAR PROGRAMMING DECOMPOSITION VIA NONLINEAR RESCALING

ROMAN A. POLYAK

Department of Mathematical Sciences, George Mason University, Fairfax, USA

Dedicated to Professor Vladimir M. Tikhomirov on the occasion of his 90th birthday

Abstract. We introduce and study decomposition methods for linear programming (LP) problems with both linking and block constraints. The methods are based on the nonlinear rescaling (NR) theory. We rescale the linking constraints into an equivalent set of constraints using scalar functions of scalar argument with particular properties. The classical Lagrangian for the equivalent problem with only linking constraints, which we call truncated Lagrangian (TL), is our main tool. Using TL, we remove the linking constraints from the entire set of constraints, which decompose the LP. Each NR decomposition (NRD) step finds approximation for the primal TL maximizer under only blocks constraints. Then the approximation is used for updating Lagrange multipliers, which correspond to the linking constraints. We call it the master problem. To find the primal TL maximizer, we use conditional or projected gradient methods applied to the TL under block constraints only. It decomposes the original LP into LPs for blocks, which are independent, have much smaller sizes and can be solved simultaneously. The critical issue is the feasibility of the linking constraints, which is achieved due to the update of the correspondent Lagrange multipliers. In fact, the master problem finds the best economic outcome for each local branch under given local resource and prices for general resources. Then, the approximation for the primal maximizer is used to correct the prices for general resources. Thus, the NRD is a pricing mechanism for finding the best outcome for an economy that has limited general and local resources. The pricing mechanism leads to such allocation of general resources that the best outcome for branches turns out to be the best outcome for the entire economy.
Keywords. Duality; Linear programming decomposition; Lagrange multipliers; Master problem; Pricing mechanism.
2020 Mathematics Subject Classification. 91B52.

1. Introduction

The LP decomposition has a long and interesting history, starting with remarkable Dantzig-Wolfe (DW) decomposition; see [1]. The DW decomposition is a brilliant combination of the revised simplex method with column generation, on the one hand, and the barycentric representation of the block constraints, on the other hand. However, the DW decomposition is a product of simplex methodology, which translates its combinatorial nature into LP decomposition. Also, the DW decomposition does not leave much space for parallel computation, which

E-mail address: rpolyak@gmu.edu
Received January 14, 2026; Accepted April 30, 2026.

is one of the primary purposes of LP decomposition. We consider an alternative approach to the LP decomposition, which is based on the NR theory (see [7] and the references therein). The NR decomposition (NRD) removes the linking constraints from the entire set of constraints by using TL. Each NRD step alternates finding an approximation for the TL primal maximizer with Lagrange multipliers update for the linking constraints. The approximation for the primal TL maximizer we find by applying conditional or projection gradient methods (see [5]) for TL maximization under block constraints. This is the master problem. The master problem decomposes the original LP into much smaller LPs. Each LPs finds the best economic outcome for a local branch of the economy under the given prices for general resources and given local resources. The LPs for the local branches are independent of each other and can be solved simultaneously. The master problem is solved when none of the local branches can improve its economic outcome. Then approximation for the primal TL maximizer, is used for prices correction. Namely, the primal TL maximizer is user for update Lagrange multipliers of the linking constraints. The formula for the Lagrange multipliers update defines the pricing mechanism, which is in the heart of the NRD. The price correction guarantees the feasibility of the linking constraints without any extra care. For the master problem, one can use any LP method. We use the modified log-barrier (MBF) method, which is fast and produces accurate results (see [4])

In the following section, we describe the primal and dual LP with block diagonal structure. In Section 3, we describe the Lagrangians used in the NRD. In Section 4, we consider the master NRD problem. In Section 5, we describe the NRD, based on the modified log-barrier function (MBF). In Section 6, we discuss the convergence properties of the NRD. In Section 7, we consider a version of the NRD with dynamic scaling parameters and hyperbolic barrier, which for some LP guarantees convergence with up to quadratic rate.

2. Problem Formulation

We consider an economy which has r branches and m_0 general resources, that define the linking constraints.

Let $A_i : \mathbb{R}^{n_i} \to \mathbb{R}^{m_0}$, $1 \le i \le r$ be the matrix with elements a^i_{kl} define the amount of general resource $1 \le k \le m_0$ required to produce one item of product $1 \le l \le n_i$ by branch $1 \le i \le r$. Let $B_i : \mathbb{R}^{n_i} \to \mathbb{R}^{m_i}$ be the matrix with elements b^i_{tl} define the amount of local resource $1 \le t \le m_i$ required to produce an item of product $1 \le l \le n_i$ by branch $1 \le i \le r$. Components of vector $p_i = (p_{i1}, \dots, p_{il}, \dots p_{in_i})$, define prices of units of products $1 \le l \le n_i$ from branch $1 \le i \le r$. Vectors $b_i \in \mathbb{R}^{m_i}$, $1 \le i \le r$ is the given local resources of branch $1 \le i \le r$. Finally, vector $a \in \mathbb{R}^{m_0}_+$ defines the given general resources of the economy and $x = (x_1, \dots, x_i, \dots, x_r) \in \mathbb{R}^N_+$ is its production vector, where $x_i \in \mathbb{R}^{n_i}_+$ is the production vector of the branch $1 \le i \le r$ and $N = \sum_{i=1}^r n_i$.

We consider the following block-diagonal LP:

$$P(x^*) = \sum_{i=1}^r p_i^T x_i^* = \max \left\{ \sum_{i=1}^r p_i^T x_i \,\middle|\, c(x) = -\sum_{i=1}^r A_i x_i + a \ge 0 \right.$$
$$\left. q_i(x_i) = -B_i x_i + b_i \ge 0, x_i \in \mathbb{R}^{n_i}_+, 1 \le i \le r \right\} \quad (2.1)$$

Along with the primal block-diagonal LP (2.1), we consider the dual LP

$$D(y^*) = a^T u^* + b^T v^* = \min \left\{ a^T u + \sum_{i=1}^{r} b_i^T v_i \right|$$

$$\left. d_i(y) = A_i^T u + B_i^T v_i - p_i \geq 0, u \in \mathbb{R}_+^{m_0}, v_i \in \mathbb{R}_+^{m_i}, 1 \leq i \leq r, v = (v_1, \dots, v_r) \in \mathbb{R}_+^{M_0} \right\}, \quad (2.2)$$

where $u \in \mathbb{R}_+^{m_0}$ is the price-vector for general resources, $v_i \in \mathbb{R}_+^{m_i}$ is the price-vector for local resources of the branch $1 \leq i \leq r$ and $M_0 = \sum_{i=1}^{r} m_i$ be the total amount of given local resources.

We assume that the feasible set of LP (2.1) is not empty and bounded, therefore x^* and $y^* = (u^*, v^*)$ exist and

$$\sum_{i=1}^{r} p_i^T x_i^* = a^T u^* + \sum_{i=1}^{r} b_i^T v_i^* \quad (2.3)$$

The primal LP (2.1) finds such production vector $x^* = (x_1^*, \dots, x_r^*)$, which maximizes the total production output under two types of constraints: general, which are typical for the entire economy, and local, which are typical for a particular branch of the economy.

The dual LP (2.2) finds prices u^* for general resources and prices v_i^*, $1 \leq i \leq r$ for local (branch) resources such that the total cost $a^T u + \sum_{i=1}^{r} b_i^T v_i$ of all resources involved in the production reaches its minimum under condition: price vectors $p_i, 1 \leq i \leq r$ are such, that for any product it's price per item P_{il} can not exceed the total cost $\left(A_i^T u + B_i^T v_i\right)_l$ of general and branch resources required to produce this item.

The fundamental difficulty of problem (2.1) is its size. The LP (2.1) has $M = M_0 + m_0$ rows and $N = \sum_{i=1}^{r} n_i$ columns. Application of simplex methods for solving such LP might lead to very large basic matrices.

The application of DW decomposition reduces the size of the basic matrix but might increase the number of steps required, due to the substantial increase of the number of vertices in the barycentric system of coordinate.

The main idea of NR decomposition is to remove the linking constraints from the given set of constraints by using Lagrangian for a problem equivalent to the initial one. It decomposes the problem and neither increase the number of vertices nor the sizes of LP one has to solve at each step. It requires, however, instroduction a pricing mechanism, which guarantees primal feasibility for the linking constraints. The NR theory offers several ways to ensure the linking constraints' feasibility (see [7]). Some of them we considered in the following sections.

3. The Lagrangians

LP (2.1) is associated the classical Lagrangian $L : \mathbb{R}_+^N \times \mathbb{R}_+^{m_0} \times \mathbb{R}_+^{M_0} \Longrightarrow \mathbb{R}$ given by

$$L(x, u, v) = p^T x + u^T c(x) + \sum_{i=1}^{r} v_i^T q_i(x_i). \quad (3.1)$$

For the primal-dual solution $(x^*; y^*) = (x_i^*; u_j^*, v_i^*)$ $1 \leq j \leq m_0$, $1 \leq i \leq r$, we have

$$L(x, y^*) \leq L(x^*, y^*) \leq L(x^*, y), \forall x \in \mathbb{R}_+^N, y \in \mathbb{R}_+^M,$$

i.e., the primal-dual solution (x^*, y^*) is a saddle point of the classical Lagrangian (3.1)

Along with the classical Lagrangian (3.1), we will use Lagrangian for the equivalent to (2.1) problem.

Let Ψ be a set of smooth concave scalar functions of scalar arguments $\psi : (a,b) \to \mathbb{R}$, $-\infty < a < 0 < b < \infty$, with following properties:

1°. $\psi(0) = 0$;

2°. (a) $\psi'(0) = 1$,

(b) $\psi'(\tau) > 0$

3°. $-t_0^{-1} \le \psi''(\tau) < 0, \forall \tau \in [a,b]$

4°. $\psi''(\tau) \le -T_0^{-1}, \tau \in [a,0], 0 < t_0 < T_0 < \infty$

We use two functions from the set Ψ:

$$\begin{aligned} &\text{Log-barrier function:} \qquad && \psi_1(\tau) = \ln(\tau+1), \\ &\text{Hyperbolic barrier:} && \psi_2(\tau) = \frac{\tau}{\tau+1}. \end{aligned} \tag{3.2}$$

In the framework of NR theory, the functions transform constraints of the problem (2.1) into an equivalent set of constraints. The transformation ψ_1 leads to the logarithmic, and ψ_2 leads to the hyperbolic barrier transformations and corresponding NR theory and methods, (see [7] and the references therein).

Along with $\psi_i \in \Psi, i = 1,2$, we use their Legendre-Fenchel transform

$$\begin{aligned} \psi_1^*(s) &= \inf_\tau \{s\tau - \psi_1(\tau)\} = \ln s - s + 1, \\ \psi_2^*(s) &= \inf_\tau \{s\tau - \psi_2(\tau)\} = 2\sqrt{s} - s - 1. \end{aligned}$$

From properties 1°. and 2°. for any $k > 0$, it follows

$$\begin{aligned} &\Omega(k) = \left\{x \in \mathbb{R}_+^N : c_i(x) \ge 0, 1 \le i \le m_0\right\} \Leftrightarrow \\ &\Omega(k) = \left\{x \in \mathbb{R}_+^N : k^{-1}\psi(kc_i(x)) \ge 0, 1 \le i \le m_0\right\}. \end{aligned} \tag{3.3}$$

Let us fix $k > 0$ and consider an LP with primal objective function P and constraints (3.3).

For the corresponding Lagrangian $L_k : \mathbb{R}_+^N \times \mathbb{R}_+^{m_0} \to \mathbb{R}$, which is given by the following formula

$$L_k(x,u) = \sum_{i=1}^r p_i^T x_j + k^{-1} \sum_{i=1}^{m_0} u_i \psi(kc_i(x)), \tag{3.4}$$

we call truncated Lagrangian (TL), which is our main tool. Along with $\Omega(k)$, we consider the brunches constraints

$$Q_i = \left\{x_i \in \mathbb{R}_+^{n_i} : q_{it}(x_i) = (b_i - Bx_i)_t \ge 0, 1 \le t \le m_i\right\}, \quad 1 \le i \le r$$

and correspondent equivalent sets

$$Q_i(k) = \left\{x_i \in \mathbb{R}_+^{n_i} : k^{-1}\psi(kq_{it}(x_i)) \ge 0, 1 \le t \le m_i\right\}$$

Let $Q(k) = Q_1(k) \times \cdots \times Q_r(k)$. Then the original block-digonal LP (2.1) can be written as follows

$$P(x^*) = \max \left\{ P(x) = \sum_{i=1}^{r} p_i^T x_i \middle| x \in \Omega(k) \cap Q(k) \right\}. \tag{3.5}$$

The correspondent to (3.5) Lagrangian $\mathscr{L}_k$ is given by the following formula

$$\mathscr{L}_k(x,u,v) = P(x) + k^{-1} \sum_{i=1}^{m_0} u_i \psi(kc_i(x)) + k^{-1} \sum_{i=1}^{r} \sum_{t=1}^{m_i} v_{it} \psi(kq_{it}(x_i)). \tag{3.6}$$

Therefore,

$$\mathscr{L}_k(x,u,v) = L_k(x,u) + k^{-1} \sum_{i=1}^{r} \sum_{t=1}^{m_i} v_{it} \psi(kq_{it}(x_i)).$$

The NR method for solving (3.5) generates the primal-dual sequence $\{x_s; y_s\}_{s=1}^{\infty}$ as follows.

First, we find

$$x_{s+1} : \mathscr{L}_k(x_{s+1}, u_s, v_s) = \max \left\{ \mathscr{L}_k(x, u_s, v_s) \,\middle|\, x \in \mathbb{R}^N_+ \right\}. \tag{3.7}$$

Therefore, for x_{s+1}, we have

$$\nabla_{x_i} \mathscr{L}_k(x_{s+1}, u_s, v_s) = \left(p_i - A_i^T \Psi'(kc(x_{s+1}))u_s - B_i^T \Psi'(kq_i(x_{s+1}))v_{is} \right) \leq 0, 1 \leq i \leq r \tag{3.8}$$

where $\Psi'(kc_i(x_{s+1})) = \operatorname{diag}\left(\psi'(kc_i(x_{s+1}))\right)_{i=1}^{m_0}$, $\Psi'(kq_i(x_{s+1})) = \operatorname{diag}\left(\psi'(kq_{it}(x_{s+1}))\right)_{t=1}^{m_i}$, $1 \leq i \leq r$.

Second, we update the Lagrange multipliers

$$u_{s+1} = \Psi'(kc(x_{s+1}))u_s, \tag{3.9}$$

$$v_{i,s+1} = \Psi'(kq_i(x_{s+1}))v_{i,s}\,, \quad 1 \leq i \leq r \tag{3.10}$$

From (3.8) - (3.10), for any $1 \leq i \leq r$, it follows

$$\nabla_{x_i} \mathscr{L}(x_{s+1}, u_s, v_s) = \left(p_i - A_i^T u_{s+1} - B_i^T v_{i,s+1} \right) = \nabla_x L(x_{s+1}, u_{s+1}, v_{s+1}) \leq 0, 1 \leq i \leq r \tag{3.11}$$

and complementarity condition

$$(x_{i,s+1}, \nabla_{x_i} \mathscr{L}(x_{s+1}, u_{s+1}, v_{s+1})) = 0, \;\; 1 \leq i \leq r \tag{3.12}$$

holds. Thus, at each NR step, the max of the Lagrangian for the equivalent problem in primal space, is coincides with the correspondent max of the classical Lagrangian for the original LP with updated Lagrange multipliers.

On the other hand, (3.11) and (3.12) are necessary and sufficient conditions for x_{s+1} to be the primal maximizer for the following problem

$$x_{s+1} : L_k(x_{s+1}, u_s) = \max \left\{ L_k(x, u_s) \,|\, x \in Q(k) \right\}. \tag{3.13}$$

Therefore, (3.7)-(3.10) is equivalent finding the maximizer of TL $L_k(x,u_s)$ following by the Lagrange multipliers update by (3.9) - (3.10).

4. Master Problem

The main computational cost at each step (3.7) - (3.10) is solving problem (3.7), which can be a difficult task due to the size of the problem. Therefore, instead of (3.7), we consider an equivalent constrained optimization problem (3.13).

Application of conditional gradient (CG) or projected gradient (PG) methods (see [5]) for solving (3.13) leads to LP decomposition. In fact, application of the CG method for solving (3.13) leads to solving at each step the following LP

$$(\nabla_x L(x_s,u_s),\hat{x}_{s+1}-x_s) = \max\left\{(\nabla_x L_k(x_s,u_s),x-x_s)\,|x\in Q(k)\right\}. \tag{4.1}$$

Keeping in mind

$$\nabla_x L_k(x_s,u_s) = \left(p_i - A_i^T u_s, 1\le i\le r\right), \tag{4.2}$$

we obtain solution $\hat{x}_{k+1}$ for (4.1) by solving the following LPs for branches

$$\left(p_i - A_i^T u_s,\hat{x}_{i,s+1}-x_{i,s}\right) = \max\left\{\left(p_i - A_i^T u_s, x_i - x_{i,s}\right)|x_i\in Q_i(k)\right\}, 1\le i\le r. \tag{4.3}$$

Both objective functions and feasible sets of LPs (4.3) are independent of each other, therefore the LPs (4.3) for the branches can be solved simultaneously. Usually LPs (4.3) have relatively small size. By solving LPs (4.3), one finds the best outcome for each branch under given prices $u_s=(u_{s1},\ldots,u_{sm_0})$ for global resources and given feasible sets $Q_i(k)$, which defines local resources of each branch $1\le i\le r$. The solutions $\hat{x}_{i,s+1}$ defines directions

$$\zeta_{i,s+1} = \hat{x}_{i,s+1} - x_{i,s},\ \ 1\le i\le r \tag{4.4}$$

in which each branch has to change its production to make its best outcome of the given global and local resources. Then we update production vector $x_{i,s}$ by formula

$$x_{i,s} : x_{i,s} + \tau\zeta_{i,s+1},\ \ 1\le i\le r \tag{4.5}$$

where step-length $0<\tau<1$ is chosen as CG method prescribes (see [5]).

Master problem alternates steps (4.3) and (4.5) till for all $i\le i\le r$ the following bound

$$\left(p_i - A_i^T u_s,\hat{x}_{i,s+1}-x_{i,s}\right)\le\varepsilon,\ \ 1\le i\le r \tag{4.6}$$

holds, where $\varepsilon>0$ is small enough. From (4.6), it follows: none of the branches can meaningfully improve its production outcome under given price-vector u_s for global resources and given polytopes $Q_i(k)$, $1\le i\le r$, which defines local resources. For solving LPs (4.3), the MBF for LP can be used (see [4]). The master problem is completed after price-vector for global resources u_s and local price-vectors $v_{i,s}$ are updated by (3.9) and (3.10).

5. NR Decomposition

We start with optimality criteria for the primal-dual solution of (2.1) and (2.2).

Vectors $\bar{z}=(\bar{x},\bar{y})\in\mathbb{R}^N_+\times\mathbb{R}^M_+$ is the solution of LP (2.1) and (2.2) if it is a primal-dual feasible pair, that is,

$$\pi_1(\bar{z}) = \min\left\{c_i(\bar{x}),1\le i\le m_0, q_{i,t}(\bar{x}_i),1\le i\le r,1\le t\le m_i,d_i(y),1\le i\le r\right\}=0 \tag{5.1}$$

and complementarity conditions

$$\pi_2(\bar{z}) = \max\left\{c_i(\bar{x})\bar{u}_i,1\le i\le m_0;q_{it}(x_i)v_{it},1\le i\le r,1\le t\le m_i,(d_i(\bar{y}))^T\bar{x}_i,1\le i\le r\right\}=0 \tag{5.2}$$

is satisfied. In other words,

$$\pi(\bar{z}) = \max\{\pi_1(\bar{z}), \pi_2(\bar{z})\} = 0, \tag{5.3}$$

is necessary and sufficient condition for $\bar{z} = (\bar{x};\bar{y})$ to be the primal-dual solution for LP (2.1) and (2.2).

The NR decomposition consists of the following steps:

(1) Let $z_o = (x_o; y_o) \in \mathbb{R}^N_+ \times \mathbb{R}^M_+$ be the starting point, $\delta > 0$ be small enough and approximation $\hat{z} = (\hat{x}, \hat{y})$ have been found already.
(2) if $\pi(\hat{z}) \le \delta$ then $z^* := \hat{z}$, otherwise
(3) set $z_0 := \hat{z}$ and solve the master problem;
(4) let $\hat{z} = (\hat{x}; \hat{y})$ be the solution of the master problem then, go to (2).

6. Convergence Properties of NR Decomposition

From the boundedness of Y^*, we see the boundedness of $Y_0 = \left\{y = (u, v) \in \mathbb{R}^M_+ : D(y) \le D(y_0)\right\}$ for any dual feasible vector y_0. Along with the dual sequence $\{y_s\}_{s=0}^\infty$ generated by the NR method (3.7) - (3.10), we consider the following dual level sets:

$$Y_s = \left\{y \in \mathbb{R}^M_+ : D(y) \le D(y_s)\right\}$$

and their boundaries:

$$\partial Y_s = \{y \in Y_s : D(y) = D(y_s)\}.$$

The following theorem for the NR method with log-barrier transform $\psi_1(t) = \ln(t+1)$ follows from the correspondent MBF results (see [3, 7, 8]).

Theorem 6.1. *If Y^* is not empty and bounded, then the NR method (3.7) - (3.10) generates such sequence $\{y_s\}_{s=0}^\infty$ that*

(1) $D(y_{s+1}) < D(y_s),\ s > 0$,
(2) $\lim_{s->\infty} D(y_s) = D(y^*)$, $\lim_{s->\infty} P(x_s) = P(x^*)$,
(3) for the Hausdorf distance $d_H(\partial y_s, Y^*)$, $\lim_{s\to\infty} d_H(\partial y_s, Y^*) = 0$,
(4) there exists subsequence $\{s_l\}_{l\in\mathbb{N}}$ such that, for the sequence $\left\{\bar{x}_l = \sum_{s=s_l}^{s_{l+1}} (s_{l+1} - s_l)^{-1} x_s\right\}_{l\in\mathbb{N}}$,

$$\lim_{l\to\infty} \bar{x}_l = x^* \in X^*$$

holds

The efficiency of the NR decomposition depends on both the convergence rate of the NR method and convergence rate of the subroutine used for solving the master problem. The CG method, used for solving the master problem, can be rather slow (see [5]-[7]). Therefore, it is important to speed up the basic NR method. It can be done for some LP by replacing one scaling parameter $k > 0$ for all constraints by a vector $\mathbb{K} \in \mathbb{R}^m_+$ of scaling parameters, one parameter for each linking constraint. At this point, we also replace the log-barrier transformation $\psi_1(t) = \ln(t+1)$ by hyperbolic transformation $\psi_2(t) = \frac{t}{t+1}$ for reasons, which will be clarified later. For truncated Lagrangian, we obtain

$$L_{\mathbb{K}}(x, u) = \sum_{i=1}^{r} p_i^T x_i + \sum_{i=1}^{m_0} k_i^{-1} u_i \psi_2\left(k_i\left(a - \sum_{j=1}^{r} A_j x_j\right)_i\right). \tag{6.1}$$

The Lagrangian

$$\mathscr{L}_{\mathbb{K}}(x,u,v)=L_{\mathbb{K}}(x,u)+\sum_{i=1}^{r}\sum_{t=1}^{m_i}k_{it}^{-1}v_{it}\psi_2\left(k_{it}q_{it}(x_i)\right) \tag{6.2}$$

for the equivalent problem, is still our main tool. The Lagrangian (6.2), however, is fundamentally different from (3.6) because in (6.2) each constraint has its own scaling parameter, which along with Lagrangian multipliers will be updated from step to step. It allows to improve NR convergence rate substantially for a wide class of LPs. Let us consider the new NR scheme. At each step, we find the primal maximizer

$$x_{s+1}:\mathscr{L}_{\mathbb{K}}(x_{s+1},u_s,v_s)=\max\left\{\mathscr{L}_{\mathbb{K}}(x,u_s,v_s)\,\middle|\,x\in\mathbb{R}^N_+\right\} \tag{6.3}$$

and update the Lagrangian multipliers using the primal maximizer, then we update scaling parameters, using new Lagrange multipliers. For the primal maximizer x_{s+1}, we have

$$\nabla_x\mathscr{L}_{\mathbb{K}}(x_{s+1},u_s,v_s)=\\ \left(p_i-A_i^T\Psi_2'(k_s^u c(x_{s+1}))u_s-B_i^T\Psi_2'\left(k_{is}^v q_i(x_{s+1})\right)v_{is},1\le i\le r\right)\le 0, \tag{6.4}$$

where

$$\Psi_2'(k_s^u c(x_{s+1}))=\operatorname{diag}\left(\psi_2'(k_{is}^u c_i(x_{s+1}))\right)_{i=1}^{m_0}$$

and

$$\Psi_2'(k_{is}^v q_i(x_{s+1}))=\operatorname{diag}\left(\psi_2'(k_{ist}^v q_{it}(x_{s+1}))\right)_{t=1}^{m_i}.$$

For the Lagrange multipliers and the scaling parameters, we update by formulas

$$u_{s+1}=\Psi_2'(k_s^u c(x_{s+1}))u_s, \tag{6.5}$$

$$v_{s+1}=\left(v_{i,s+1}=\Psi_2'\left(k_{is}^v q_i(x_{s+1})\right)v_{is}\right),\ \ 1\le i\le r; \tag{6.6}$$

$$k_{i,s+1}=k(u_{i,s+1})^{-1},\ \ 1\le i\le m_0,\ \ \mathbb{K}^u_{s+1}=\operatorname{diag}(k_{i,s+1})_{i=1}^{m_0}, \tag{6.7}$$

$$k_{i,s+1,t}=k(v_{i,s+1,t})^{-1},\ \ 1\le i\le r,1\le t\le m_i,\ \ \mathbb{K}^v_{i,s+1}=\operatorname{diag}(k_{i,s+1,t})_{t=1}^{m_i}. \tag{6.8}$$

From (6.4) - (6.8), it follows

$$\begin{aligned}\nabla_x\mathscr{L}_{\mathbb{K}}(x_{s+1},u_s,v_s)&=\left(p_i-A_i^T u_{s+1}-B^T v_{i,s+1},\ 1\le i\le r\right)\\ &=\nabla_x L(x_{s+1},u_{s+1},v_{s+1})\le 0\end{aligned} \tag{6.9}$$

and complementarity condition

$$(x_{i,s+1},\nabla_{x_i}L(x_{s+1},u_{s+1},v_{s+1}))=0,\ \ 1\le i\le r \tag{6.10}$$

is satisfied. From (6.9) and (6.10), it follows

$$0=(\nabla_x\mathscr{L}_{\mathbb{K}}(x_{s+1},u_s,v_s),x_{s+1})=\sum_{i+1}^{r}p_i^T x_{i,s+1}+u_{s+1}^T\left(\sum_{i+1}^{r}-A_i x_{i,s+1}+a\right)-a^T u_{s+1}\\ +\sum_{i=1}^{r}v_{i,s+1}^T(-B_i x_{i,s+1}+b_i)-\sum_{i=1}^{r}b_i^T v_{i,s+1}.$$

Thus,

$$\max\left\{\mathscr{L}_{\mathbb{K}}(x,u_s,v_s)|x\in\mathbb{R}^N_+\right\}=\mathscr{L}_{\mathbb{K}}(x_{s+1},u_s,v_s)=L(x_{s+1},u_{s+1},v_{s+1})=a^T u_{s+1}^T+\sum_{i=1}^{r}b_i^T v_{i,s+1}. \tag{6.11}$$

In other words, the max of Lagrangian for the equivalent problem in primal space coinsides with max of the classical Lagrangian for the original LP with updated Lagrange multipliers. This is critical for establishing equivalence of the NR method (6.3) - (6.8) and proximal point method for the dual problem with second order φ-divergence entropy-like distance based on the hyperbolic kernel $\varphi_2(s) = -\psi_2^*(s)$, where $\psi_2^*(s) = \inf_+ \left\{ s\tau - \frac{\tau}{\tau+1} \right\} = 2\sqrt{s} - s - 1$. Let p and $q \in \mathbb{R}^n$, to simplify notations we introduce componentwise division and multiplication of p and q as $\frac{p}{q} = \left(\frac{p_1}{q_1}, \cdots, \frac{p_n}{q_n} \right)$ and $pq = (p_1q_1, \ldots, p_nq_n)$.

Theorem 6.2. *The nonlinear rescaling method* (6.3)-(6.8) *with dynamic scaling parameter update for the primal LP* (2.1) *is equivalent to the proximal point method*

$$D(y_{s+1}) = \min \left\{ D(y) + k^{-1} d_1(u, u_s) + k^{-1} d_2(v, v_s) \,\middle|\, A_i^T u_i + B_i^T v_i - p_i \geq 0, i = 1, \ldots, r \right\} \tag{6.12}$$

for the dual problem (2.2)*, where*

$$d_1(u, u_s) = \sum_{i=1}^{m_o} u_{is}^2 \varphi_2 \left(\frac{u_i}{u_{is}} \right)$$

and

$$d_2(v, v_s) = \sum_{i=1}^{r} \sum_{t=1}^{m_i} v_{its}^2 \varphi_2 \left(\frac{v_{it}}{v_{its}} \right)$$

are second order φ-divergence entropy-like distance functions with kernel $\varphi_2 = -\psi_2^*$.

Proof. From (6.5), it follows

$$\frac{u_{s+1}}{u_s} = \Psi_2 \left(\mathbb{K}_s^u c(x_{s+1}) \right) e_0 = \operatorname{diag} \left(\psi_2' \left(k_{is}^u c_i(x_{s+1}) \right) \right)_{i=1}^{m_0} e_0,$$

where $e_{m_0} = (1, \ldots, 1)^T \in \mathbb{R}_+^{m_0}$. From property $2°$.(b) of tranforantion ψ_2, it follows the existence of the inverse ψ_2^{-1}. From Legendre-Fenchel (LF) identity $\psi'^{-1}_2 = \psi_2^{*\prime}$ (see for example [2]) and (6.5), it follows

$$\mathbb{K}_s^u c(x_{s+1}) = \Psi_2^{-1} \left(\frac{u_{s+1}}{u_s} \right) e_{m_0} = -\operatorname{diag} \left(\varphi_2' \left(\frac{u_{i,s+1}}{u_{is}} \right) \right)_{i=1}^{m_0} e_{m_0} = -\Phi_2' \left(\frac{u_{s+1}}{u_s} \right) e_{m_0}.$$

Therefore,

$$c(x_{s+1}) = -\left(\mathbb{K}_s^u \right)^{-1} \Phi_2' \left(\frac{u_{s+1}}{u_s} \right) e_{m_0}. \tag{6.13}$$

It follows from (6.6) that

$$\frac{v_{i,s+1}}{v_{i,s}} = \Psi_2' \left(\mathbb{K}_{is}^v q_i(x_{s+1}) \right) e_{m_i}, \quad 1 \leq i \leq r, \tag{6.14}$$

where $e_{m_i} = (1, \ldots, 1)^T \in \mathbb{R}_+^{m_i}$. Again, keeping in mind property ($2°$.)(b) of transformation ψ_2 and Legendre–Fenchel identity, we obtain

$$q_i(x_{s+1}) = \left(\mathbb{K}_{is}^v \right)^{-1} \cdot \Psi_2'^{-1} \left(\frac{v_{i,s+1}}{v_{is}} \right) e_{m_i} = -\left(\mathbb{K}_{is}^v \right)^{-1} \Phi_2' \left(\frac{v_{i,s+1}}{v_{is}} \right) e_{m_i}, \quad 1 \leq i \leq r. \tag{6.15}$$

From (6.9)-(6.10), it follows

$$\partial D(y_{s+1}) \ni \begin{pmatrix} c(x_{x+1}) \\ q_i(x_{s+1}), \quad 1 \leq i \leq r, \end{pmatrix} \tag{6.16}$$

where $\partial D(y_{s+1})$ - subdefferntial of the dual objective function at y_{s+1}. Then, from (6.13)-(6.16), it follows

$$\partial D(y_{s+1})+\begin{bmatrix} (\mathbb{K}_s^u)^{-1}\Phi_2'\left(\frac{u_{s+1}}{u_s}\right)e_{m_0} \\ (\mathbb{K}_{is}^v)^{-1}\Phi_2'\left(\frac{v_{i,s+1}}{v_{is}}\right)e_{m_i},\ 1\le i\le r\end{bmatrix}=0. \tag{6.17}$$

System (6.17) is the optimality condition for vector $y_{s+1}=(u_{s+1},v_{s+1})\in\mathbb{R}^M_{++}$ to be the solution of the following problem

$$y_{s+1}=\operatorname{argmin}\Big\{D(y)+k^{-1}d_1(u,u_s)+k^{-1}d_2(v,v_s)=$$
$$=D(y)+k^{-1}\sum_{i=1}^{m_0}u_{is}\varphi_2\left(\frac{u_i}{u_{is}}\right)+k^{-1}\sum_{i=1}^{r}\sum_{t=1}^{m_i}v_{its}^2\varphi_2\left(\frac{v_{it}}{v_{ist}}\right)\Big|$$
$$A_i^Tu+B_i^Tv_i-p_i\ge 0,\ 1\le i\le r\Big\}. \tag{6.18}$$

□

Definition. Kernel φ is well defined if $\varphi(0)<\infty$.

In particular, the logarithmic MBF kernel $\varphi_1(s)=-\ln s+s-1$ is not well defined, while the hyperbolic MBF kernel φ_2 is due to $\varphi_2(0)=1$. This property of the dual kernel is important for the main result, established in the following Theorem.

Theorem 6.3. *If the dual LP(2.2) has a unique solution and the correspondent kernel φ is well defined, then there exists $s_0>0$ such that for the dual sequence $\{y_s=(u_s,v_s)\}_{s\in\mathbb{N}}$, generated by NR method (6.4)-(6.8), the following bound*

$$D(y_{s+1})-D(y^*)\le ck^{-1}\left[D(y_s)-D(y^*)\right]^2,\ \ s\ge s_0 \tag{6.19}$$

holds, and constant $c>0$ is independent on $k>0$.

Proof. For any $y_{s+1}=(u_{s+1},v_{i,s+1},1\le i\le r)\in\mathbb{R}^M_+$, it follows from (6.12) that $D(y_{s+1})\le D(y^*)$. Thus

$$D(y_{s+1})+k^{-1}\sum_{y=1}^{m_0}u_{is}^2\varphi_2\left(\frac{u_{i,s+1}}{u_{i,s}}\right)+k^{-1}\sum_{i=1}^{r}\sum_{t=1}^{m_i}v_{ist}^2\varphi_2\left(\frac{v_{i,s+1,t}}{v_{ist}}\right)$$
$$\le D(y^*)+k^{-1}\sum_{k=1}^{m_0}u_{is}^2\varphi_2\left[\frac{u_{is}^*}{u_{is}}\right]+k^{-1}\sum_{i=1}^{r}\sum_{t=1}^{m_i}v_{ist}^2\varphi_2\left(\frac{v_{it}^*}{v_{ist}}\right). \tag{6.20}$$

From

$$k^{-1}\sum_{i=1}^{m_0}u_{is}^2\varphi_2\left(\frac{u_{i,s+1}}{u_{is}}\right)+k^{-1}\sum_{i=1}^{r}\sum_{t=1}^{m_0}v_{its}^2\varphi_2\left(\frac{v_{i,s+1,t}}{v_{ist}}\right)\ge 0$$

and (6.20) follows

$$D(y_{s+1})-D(y^*)\le k^{-1}\sum_{i=1}^{r}u_{is}^2\varphi_2\left(\frac{u_i^*}{u_{is}}\right)+k^{-1}\sum_{i=1}^{r}\sum_{t=1}^{m_i}v_{ist}^2\varphi_2\left(\frac{v_{it}^*}{v_{ist}}\right). \tag{6.21}$$

Assume $u^*_{is} > 0,\ i = 1,\ldots,r_0;\ u^*_i = 0,\ i = r_0+1,\ldots,m_0$ and $v^*_{it} > 0,\ t = 1,\ldots,t_{i_0};\ v^*_{it} = 0, t = t_{i_0}+1,\ldots,m_i,\ 1 \le i \le r$. Then keeping in mind $\varphi_2(1) = \varphi_2'(1) = 0$, we obtain

$$\begin{aligned}
&k^{-1}\sum_{i=1}^{r_0} u_{is}^2\varphi_2\left(\frac{u_i^*}{u_{is}}\right) + k^{-1}\sum_{i=1}^{r}\sum_{t=1}^{t_{i_0}} v_{its}^2\varphi_2\left(\frac{v_{it}^*}{v_{its}}\right)\\
&= k^{-1}\left[\sum_{i=1}^{r_0} u_{is}^2\left(\varphi_2\left(\frac{u_i^*}{u_{is}}\right) - \varphi_2\left(\frac{u_{is}}{u_{is}}\right)\right) + \varphi(0)\sum_{i=r_0+1}^{m_0}(u_{is}-u_i^*)^2\right]\\
&\quad + k^{-1}\left[\sum_{i=1}^{r}\sum_{t=1}^{t_{i_0}} v_{ist}^2\left(\varphi_2\left(\frac{v_{it}^*}{v_{ist}}\right) - \varphi_2\left(\frac{v_{ist}}{v_{ist}}\right)\right) + \varphi(0)\sum_{i=1}^{r}\sum_{t=t_{i_0}+1}^{m_i}(v_{ist}-v_{it}^*)^2\right] \qquad (6.22)\\
&= k^{-1}\left[\sum_{i=1}^{r_0}\varphi''(\cdot)(u_{is}-u_{is}^*)^2 + \varphi(0)\sum_{i=r_0+1}^{r}(u_{is}^*-u_i^*)^2\right]\\
&\quad + k^{-1}\left[\sum_{i=1}^{r_i}\sum_{t=1}^{t_{i_0}}\varphi''(\cdot)(v_{ist}-v_{ist}^*)^2 + \varphi(0)\sum_{i=1}^{r}\sum_{t=t_{i_0}+1}^{m_i}(v_{ist}-v_{it}^*)^2\right].
\end{aligned}$$

From properties (3°.) and (4°.) of the transformation ψ_2 for the correspondent kernel φ_2 follows

$$t_0 \le \varphi_2''(\cdot) \le T_0 \qquad (6.23)$$

Combining (6.21) and (6.22) and keeping in mind (6.23), we obtain

$$\begin{aligned}
D(y_{s+1}) - D(y^*) \le k^{-1}\Bigg[&T_0\sum_{i=1}^{r_0}(u_{is}-u_i^*)^2 + \varphi(0)\sum_{i=r_0+1}^{m_0}(u_{is}-u_i^*)^2\\
&+ k^{-1}\sum_{i=1}^{r_i}\sum_{t=1}^{t_{i_0}} T_0(v_{ist}-v_{it}^*)^2 + \varphi(0)\sum_{i=r_i+1}^{r}\sum_{t=t_{i_0}+1}^{m_i}(v_{ist-v_{it}^*})^2\Bigg].
\end{aligned} \qquad (6.24)$$

From (6.23)-(6.24) for $\varphi_0 = \max\{T_0, \varphi(0)\}$, it follows

$$\Delta_{s+1} = D(y_{s+1}) - D(y^*) \le k^{-1}\varphi_0||y_s - y^*||^2. \qquad (6.25)$$

If y^* is a unique, then we see from that [7] the existence $\alpha > 0$ that

$$D(y) - D(y^*) = \alpha||y - y^*|| \qquad (6.26)$$

holds for all dual feasible y From (6.25) and (6.26) for $y = y_s$, we obtain $\Delta_{s+1} \le k^{-1}\varphi_0\alpha^{-2}\Delta_s^2 = ck^{-1}\Delta_s^2$, where $c = \varphi_0\alpha^{-2}$ is independent of $k > 0$. □

7. Concluding Remarks

Removing the linking constraints by using the truncated Lagrangian is the basic idea behind the NR decomposition. To guarantee feasibility for the linking constraints, we used the pricing mechanism, which is a product of the NR theory. The pricing mechanism allows each local branch finding the best economic outcome at each step. Moreover, each branch finds its best economic outcome independent of other branches. Due to the pricing mechanism, by optimizing the local production outputs, the NR decomposition leads to the optimal solution for the entire economy. Such an outcome became possible due to the optimal allocation of general resources and optimal use of local resources. The efficiency of the NR decomposition depends on the NR convergence properties and the ability to solve linear LP for branches fast and accurately. The numerical results obtained so far show that the efficiency of NR decomposition very much depends on our ability to solve the master problem. Therefore, along with the CG method, we are planning to use the fast projected gradient method for solving the master problem; see [6, 7].

References

[1] G. Dantzig, P. Wolfe, The decomposition principle for linear programs, Oper. Res. 8 (1960), 101-111.
[2] A. Ioffe, Tichomirov V. Theory of extremal problems, Elsevier North-Holland, 1979.
[3] D. Jensen, R. Polyak, The convergence of MBF for convex programming, IBM J. Res. Dev. 38 (1994), 301-326.
[4] D. Jensen, R. Polyak R. Schneur, Experience with MBF methods for linear programming, Research Report, Department of Mathematical Sciences, vol 1358 IBM T.J. Watson Research Center, NY, pp. 1-35, 1993.
[5] E. Levitin, B. Polyak, Minimization methods in the presence of constraints, Zh. Vychisl. Mat. Mat. Fiz. 6 (1966), 787-823.
[6] Y. Nesterov, Lectures on Convex Optimization, Springer, 2018.
[7] R. Polyak, Introduction to Continuous Optimization, Springer, 2021.
[8] M. Powell, Some convergence properties of the method of logarithmic barriers for linear programming, SIAM J. Optim. 5 (1995), 699-739.