



Dual Newton's Methods for Linear Second-Order Cone Programming

Vitaly Zhadan^{1,2} (✉)

¹ Dorodnicyn Computing Centre, FRC "Computer Science and Control" of RAS,
40, Vavilova str., Moscow 119333, Russia

zhadan@ccas.ru

² Moscow Institute of Physics and Technology (State Research University),
9 Institutskiy per., Dolgoprudny, Moscow Region 141701, Russia

Abstract. The linear second-order cone programming problem is considered. For its solution, two dual Newton's methods are proposed. These methods are constructed with the help of optimality conditions. The nonlinear system of equations, obtained from the optimality conditions and depended only from dual variables, is solved by the Newton method. Under the assumption that there exist strictly complementary solutions of both primal and dual problems the local convergence of the methods with super-linear rate is proved.

Keywords: Linear second-order cone programming · Dual Newton's method · Local convergence · Super-linear rate of convergence

1 Introduction

The second-order cone programming problem (SOCP) is one of the main programs in cone programming. The linear SOCP is a problem in which the linear objective function is minimized on the intersection of a linear manifold with a second-order cone (the Lorentz cone) (see [1]). Many optimization problems, including combinatorial problems, can be reduced to the SOCP programs [1–3].

Today, there are some numerical methods for solving SOCP programs. From these methods, the primal-dual path-following methods are the most known [4, 5]. In [6] the dual barrier-projection methods have been proposed for SOCP programs. These methods are generalizations of the corresponding methods for linear programming [7]. The primal Newton's method for SOCP have been considered in [8]. Both dual barrier-projection methods and the primal Newton's method had been worked out with the help of optimality conditions.

In dual methods the dual variables depending on primal variable are defined. As a result, the system of nonlinear equations with respect to dual variables, including a slack dual variable, is derived. In [6] this derived system of nonlinear equations is solved by the fix point method. The proposed in [6] dual methods are of the affine-scaling type. In present paper unlike to [6] the Newton method

is used for solving the derived system of nonlinear equations. Under assumption that the solutions of primal and dual problems are strictly complementary dual Newton's methods converge locally to these solutions with super-linear rate.

The paper is organized as follows. In Sect. 1, we formulate the SOCP program. Section 2 is principal in the paper. In this section the dual iterative methods for SOCP programs, based on the Newton method, are constructed. In Sect. 3 we show that in the case of non-degenerate problem these dual methods are well-posed. Finally, in Sect. 4, the local convergence of the methods is proved.

In what follows, the identity matrix of order s is denoted by I_s . The symbol 0_s is used for denoting the zero s -dimensional vector, and the symbol 0_{sk} is used for denoting $s \times k$ zero matrix. By $\text{Diag}(x)$ is denoted the diagonal matrix with a vector x at its diagonal. Similarly, a block diagonal matrix with diagonal blocks M_1, \dots, M_k is denoted by $\text{DIAG}(M_1, \dots, M_k)$.

2 The Linear Second-Order Cone Programming Problem

Let $\mathcal{K} \subset \mathbb{R}^n$ denote a closed convex pointed cone with the nonempty interior. This cone induces in \mathbb{R}^n a partial order, that is: $x_1 \succeq_{\mathcal{K}} x_2$, if $x_1 - x_2 \in \mathcal{K}$. The linear cone programming problem is

$$\min \langle c, x \rangle, \quad \mathcal{A}x = b, \quad x \in \mathcal{K}, \quad (1)$$

where \mathcal{A} is a $m \times n$ matrix, and $c = [c^1; \dots; c^n] \in \mathbb{R}^n$, $b = [b^1; \dots; b^m] \in \mathbb{R}^m$. The semicolons between vectors or components of a vector denote that these vectors or components are placed one under another. The angle brackets denotes the usual Euclidean scalar product.

The linear SOCP program is a special case of the problem (1). Let $c_i \in \mathbb{R}^{n_i}$, $1 \leq i \leq r$. Let also matrices A_i have dimensions $m \times n_i$, $1 \leq i \leq r$. Consider the problem

$$\min \sum_{i=1}^r \langle c_i, x_i \rangle, \quad \sum_{i=1}^r A_i x_i = b, \quad x_1 \succeq_{K_2^{n_1}} 0_{n_1}, \dots, x_r \succeq_{K_2^{n_r}} 0_{n_r}. \quad (2)$$

Here $K_2^{n_i}$ is the second order cone (the Lorentz cone) in \mathbb{R}^{n_i} , defined as

$$K_2^{n_i} = \{[x^0; \bar{x}] \in \mathbb{R} \times \mathbb{R}^{n_i-1} : x^0 \geq \|\bar{x}\|\}, \quad 1 \leq i \leq r,$$

where $\|\cdot\|$ is the Euclidean norm. The cone $K_2^{n_i}$ is self-dual, that is $(K_2^{n_i})^* = K_2^{n_i}$.

The following problem is dual to (2)

$$\max \langle b, u \rangle, \quad A_i^T u + y_i = c_i, \quad 1 \leq i \leq r; \quad y_1 \succeq_{K_2^{n_1}} 0_{n_1}, \dots, y_r \succeq_{K_2^{n_r}} 0_{n_r}, \quad (3)$$

in which $u \in \mathbb{R}^m$.

Denote $n = n_1 + \dots + n_r$. If set $c = [c_1; \dots; c_r]$, $x = [x_1; \dots; x_r]$, $y = [y_1; \dots; y_r]$ and $\mathcal{A} = [A_1, \dots, A_r]$, $\mathcal{K} = K_2^{n_1} \times \dots \times K_2^{n_r}$, then the problem (2) can be written in the form of (1). The cone \mathcal{K} is self-dual. We assume that both

problems (2) and (3) have solutions, and that rows of the matrix \mathcal{A} are linear independent. We assume also that $r > 1$.

Let $y(u) = c - \mathcal{A}^T u$. By

$$\mathcal{F}_P = \{x \in \mathcal{K} : \mathcal{A}x = b\}, \quad \mathcal{F}_D = \{[u, y] \in \mathbb{R}^m \times \mathcal{K} : y = y(u)\}$$

we will denote the feasible sets in problems (2) and (3), respectively. By $\mathcal{F}_{D,u}$ we will denote the projection of the set \mathcal{F}_D onto the space \mathbb{R}^m , i.e. the set $\mathcal{F}_{D,u} = \{u \in \mathbb{R}^m : y(u) \in \mathcal{K}\}$.

If x and $[u, y]$ are solutions of problems (2) and (3), then they satisfy to the following system of equalities

$$\langle x, y \rangle = 0, \quad \mathcal{A}x = b, \quad y = c - \mathcal{A}^T u, \quad (4)$$

and to inclusions: $x \in \mathcal{K}$, $y \in \mathcal{K}$. Taking into account these inclusions, the equality $\langle x, y \rangle = 0$ from (4) can be replaced by n other equalities

$$x_i \circ y_i = 0_{n_i}, \quad 1 \leq i \leq r, \quad (5)$$

where the product between vectors $x_i \in \mathbb{R}^{n_i}$ and $y_i \in \mathbb{R}^{n_i}$ is defined by the following way $x_i \circ y_i = [x_i^T y_i; x_i^0 \bar{y} + y_i^0 \bar{x}_i]$. By introducing the matrix

$$\mathbf{Arr}(x_i) = \begin{bmatrix} x_i^0 & \bar{x}_i^T \\ \bar{x}_i & x_i^0 I_{n-1} \end{bmatrix},$$

the product $x_i \circ y_i$ can be represented as $x_i \circ y_i = \mathbf{Arr}(x_i) y_i = \mathbf{Arr}(y_i) x_i$.

Compose the block-diagonal matrix $\mathbf{Arr}(y) = \text{DIAG}[\mathbf{Arr}(y_1), \dots, \mathbf{Arr}(y_r)]$. Then equalities (4) can be rewritten as

$$\mathbf{Arr}(y) x = \mathbf{Arr}(x) y = 0_n, \quad \mathcal{A}x = b, \quad y = c - \mathcal{A}^T u, \quad (6)$$

where, recall, $x \in \mathcal{K}$, $y \in \mathcal{K}$.

3 The Dual Newton's Methods

Consider the iterative dual methods for solving problems (2) and (3). These methods are analogs of the primal method proposed in [8]. In dual methods the dependence $x(u)$ or more general dependence $x(u, y)$ are used to derive from (6) the system of nonlinear equations depending on only dual variables.

In order to obtain $x(u)$ we multiply the second equality from (6) by the matrix \mathcal{A}^T and sum it with the first equality (6). As a result, we get the equation with respect to x :

$$\Phi(y)x = \mathcal{A}^T b, \quad (7)$$

where by $\Phi(y)$ is denoted the matrix: $\Phi(y) = \mathcal{A}^T \mathcal{A} + \mathbf{Arr}(y)$. The matrix $\Phi(y)$ is symmetric of order n . If $\Phi(y)$ is nonsingular, then, solving the Eq. (7), we obtain

$$x = x(y) = \Phi^{-1}(y) \mathcal{A}^T b.$$

Taking $y = y(u)$, we conclude that in fact the matrix $\Phi(y)$ depends on u .

Substituting the founded $x(u) = x(y(u))$ into the second equation from (6), we get the system of nonlinear equations with respect to u , namely,

$$\left[I_m - \mathcal{A}\Phi^{-1}(y(u))\mathcal{A}^T \right] b = 0_m. \quad (8)$$

The system (8) consists of m equations. The number of unknowns is also equal to m .

Applying the Newton method to solve (8), we obtain the iterative process

$$u_{k+1} = u_k - \mathbf{G}^{-1}(u_k) (\mathcal{A}x_k - b). \quad (9)$$

Here $x_k = x(u_k)$ and $\mathbf{G}(u) = \frac{d}{du} \mathcal{A}x(u) = \mathcal{A}x_u(u)$.

Treating (7) as the identity with respect to u , we obtain after differentiating

$$\mathbf{Arr}_u(y(u))x(u) + \Phi(y(u))x_u(u) = 0_{nm}.$$

If $\Phi(y(u))$ is a nonsingular matrix, then

$$x_u(u) = -\Phi^{-1}(y(u))\mathbf{Arr}_u(y(u))x(u). \quad (10)$$

Since $y(u) = c - \mathcal{A}^T u$, we get $\mathbf{Arr}_u(y(u)) = -\mathbf{Arr}_y(y)\mathcal{A}^T$.

Proposition 1. For any $x \in \mathbb{R}^n$ the equality

$$\mathbf{Arr}_u(y(u))x = -\mathbf{Arr}(x)\mathcal{A}^T \quad (11)$$

holds.

Proof. Let us take the product z_i of any matrix $\mathbf{Arr}(y_i)$ on the vector x_i and differentiate each row of z_i by y_i separately. First of all, the “null” row is the following: $z_i^0 = \sum_{j=0}^{n-1} x_i^j y_i^j$. Therefore,

$$\frac{d}{dy} z_i^0 = [x_i^0; x_i^1; \dots; x_i^{n-1}]. \quad (12)$$

Further, for any consequent j -th row: $z_i^j = y_i^j x_i^0 + y_i^0 x_i^j$. Hence

$$\frac{d}{dy} z_i^j = [x_i^j; 0; \dots; 0; x_i^0; 0; \dots, 0], \quad 1 \leq j \leq n-1. \quad (13)$$

From (12) and (13) we derive that $\mathbf{Arr}_y(y)x = \mathbf{Arr}(x)$. Hence, the equality (11) takes place. \square

According to Proposition 1 and to (10) $\mathbf{G}(u) = \mathcal{A}\Phi^{-1}(u)\mathbf{Arr}(x(u))\mathcal{A}^T$. Thus, the iterative method (9) can be written in the following form

$$u_{k+1} = u_k - \left[\mathcal{A}\Phi^{-1}(y_k)\mathbf{Arr}(x_k)\mathcal{A}^T \right]^{-1} (\mathcal{A}x_k - b), \quad (14)$$

where $x_k = x(u_k)$, $y_k = y(u_k)$.

It is possible to consider the more general with respect to (9) iterative process. In this process both variables u and y are updated at each iteration. In order to construct the method we add to the right side of Eq. (7) the second equality from (6), multiplied by some parameter $\tau > 0$. As a result, we obtain instead of (7) the system of equations

$$\Phi(y)x = \mathcal{A}^T b + \tau (y + \mathcal{A}^T u - c) \quad (15)$$

with the solution $x(u, y) = \Phi^{-1}(y)f(u, y)$, where $f(u, y) = \mathcal{A}^T b + \tau(y + \mathcal{A}^T u - c)$.

Substituting $x(u, y)$ in first and second equalities from (6), we obtain the system of $n + m$ equations

$$\mathcal{A}\Phi^{-1}(y)f(u, y) - b = 0_m, \quad \mathbf{Arr}(y)\Phi^{-1}(y)f(u, y) = 0_n. \quad (16)$$

Denote $w = [u; y]$ and $\Psi(w) = [\Psi^{(1)}(w); \Psi^{(2)}(w)]$, where

$$\Psi^{(1)}(w) = \mathcal{A}\Phi^{-1}(y)f(u, y) - b, \quad \Psi^{(2)}(w) = \mathbf{Arr}(y)\Phi^{-1}(y)f(u, y).$$

Lemma 1. *Let the point $w = [u; y] \in \mathcal{F}_D$ be such that the matrix $\Phi(y)$ is nonsingular. Then the matrix $\Psi_w(w)$ has the form*

$$\Psi_w(w) = \begin{bmatrix} \tau\mathcal{A}\Phi^{-1}\mathcal{A}^T & \mathcal{A}\Phi^{-1}[\tau I_n - \mathbf{Arr}(x(w))] \\ \tau\mathbf{Arr}(y)\Phi^{-1}\mathcal{A}^T & [I_n - \mathbf{Arr}(y)\Phi^{-1}]\mathbf{Arr}(x(w)) + \tau\mathbf{Arr}(y)\Phi^{-1} \end{bmatrix}. \quad (17)$$

where $\Phi^{-1} = \Phi^{-1}(y)$.

Proof. Differentiating $\Psi^{(1)}$, we obtain: $\Psi_u^{(1)}(w) = \mathcal{A}x_u(w)$, $\Psi_y^{(1)}(w) = \mathcal{A}x_y(w)$. Moreover,

$$\Psi_u^{(2)}(w) = \mathbf{Arr}(y)x_u(w), \quad \Psi_y^{(2)}(w) = \mathbf{Arr}(x(w)) + \mathbf{Arr}(y)x_y(w).$$

Because of (15), the function $x(w)$ is satisfied to the identity

$$[\mathcal{A}^T\mathcal{A} + \mathbf{Arr}(y)]x(w) \equiv \mathcal{A}^T b + \tau(y + \mathcal{A}^T u - c). \quad (18)$$

After differentiation (18) by u we obtain

$$[\mathcal{A}^T\mathcal{A} + \mathbf{Arr}(y)]x_u(w) = \tau\mathcal{A}^T. \quad (19)$$

Respectfully, after differentiation (18) by y we derive the equality $\mathcal{A}^T\mathcal{A}x_y(w) + \frac{\partial}{\partial y}\mathbf{Arr}(y)x(w) = \tau I_n$ or

$$\mathcal{A}^T\mathcal{A}x_y(w) + \mathbf{Arr}(x(w)) + \mathbf{Arr}(y)x_y(w) = \tau I_n. \quad (20)$$

Equalities (19) and (20) can be written as

$$\Phi(y)x_u(w) = \tau\mathcal{A}^T, \quad \Phi(y)x_y(w) + \mathbf{Arr}(x(w)) = \tau I_n.$$

If $\Phi(y)$ is a nonsingular matrix, we derive from here that

$$x_u(w) = \tau\Phi^{-1}(y)\mathcal{A}^T, \quad x_y(w) = \Phi^{-1}(y)[\tau I_n - \mathbf{Arr}(x(w))].$$

Thus, the matrix $\Psi_w(w)$ has the form (17). \square

If the matrix $\Psi_w(w)$ is nonsingular for all points w in some neighbourhood of the solution w_* of the problem (3), then it is possible to apply the Newton method for solving the system of nonlinear equations (16). We obtain the dual iterative method

$$w_{k+1} = w_k - \Psi_w^{-1}(w_k)\Psi(w_k). \quad (21)$$

The point w_0 must be taken from some vicinity of the solution w_* .

Denote $\mathbf{\Gamma}(y) = \mathcal{A}\Phi^{-1}(y)\mathcal{A}^T$, $\mathbf{K}_1(y) = \mathcal{A}\Phi^{-1}(y)$, $\mathbf{K}_2(y) = \mathbf{Arr}(y)\Phi^{-1}(y)$. Then the matrix (17) can be written as

$$\mathbf{\Psi}_w(w) = \begin{bmatrix} \tau\mathbf{\Gamma}(y) & \mathbf{K}_1(y) [\tau I_n - \mathbf{Arr}(x(w))] \\ \tau\mathbf{Arr}(y)\mathbf{K}_1^T(y) & [I_n - \mathbf{K}_2(y)] \mathbf{Arr}(x(w)) + \tau\mathbf{K}_2(y) \end{bmatrix}.$$

Let the point $w \in \mathcal{F}_D$ be such that the matrix $\mathbf{\Psi}_w(w)$ is nonsingular. In this case the matrix $\mathbf{\Gamma}(y)$ is also nonsingular. Denote $\mathbf{K}_3(y) = \mathbf{K}_1^T(y)\mathbf{\Gamma}^{-1}\mathbf{K}_1(y)$ and

$$\mathbf{\Omega}(w) = [I_n - \mathbf{K}_2(y)] \mathbf{Arr}(x(w)) + \tau\mathbf{K}_2(y) - \mathbf{Arr}(y)\mathbf{K}_3(y) [\tau I_n - \mathbf{Arr}(x(w))].$$

Then, by the Frobenius formula, we obtain

$$\mathbf{\Psi}_w^{-1}(w) = \begin{bmatrix} \mathbf{P}_1 & \mathbf{P}_2 \\ \mathbf{P}_3 & \mathbf{P}_4 \end{bmatrix},$$

where $\mathbf{P}_4 = \mathbf{\Omega}^{-1}$ and

$$\mathbf{P}_1 = \tau^{-1}\mathbf{\Gamma}^{-1} + \tau^{-1}\mathbf{\Gamma}^{-1}\mathbf{K}_1(y) [\tau I_n - \mathbf{Arr}(x(w))] \mathbf{\Omega}^{-1} \mathbf{Arr}(y)\mathbf{K}_1^T(y)\mathbf{\Gamma}^{-1},$$

$$\mathbf{P}_2 = -\tau^{-1}\mathbf{\Gamma}^{-1}\mathbf{K}_1(y) [\tau I_n - \mathbf{Arr}(x(w))] \mathbf{\Omega}^{-1}, \quad \mathbf{P}_3 = -\mathbf{\Omega}^{-1} \mathbf{Arr}(y)\mathbf{K}_1^T(y)\mathbf{\Gamma}^{-1}.$$

At last, denote $\mathbf{W} = \mathbf{\Gamma}^{-1}\mathbf{K}_1(y) [\tau I_n - \mathbf{Arr}(x(w))] \mathbf{\Omega}^{-1}$. It follows from previous formulas that $\mathbf{P}_2 = -\tau^{-1}\mathbf{W}$ and $\mathbf{P}_1 = \tau^{-1} [I_m - \mathbf{W}\mathbf{Arr}(y)\mathbf{K}_1^T(y)] \mathbf{\Gamma}^{-1}$.

Therefore, the formulas (21) for updating the point $[u_k; y_k]$ are following:

$$u_{k+1} = u_k - \tau^{-1} \left[(I_m - \mathbf{W}\mathbf{Arr}(y)\mathbf{K}_1^T(y)) \mathbf{\Gamma}^{-1} (\mathcal{A}x_k - b) + \mathbf{W}\mathbf{Arr}(y_k)x_k \right],$$

$$y_{k+1} = y_k + \mathbf{\Omega}^{-1} \left[\mathbf{Arr}(y)\mathbf{K}_1^T(y)\mathbf{\Gamma}^{-1} (\mathcal{A}x_k - b) - \mathbf{Arr}(y_k)x_k \right],$$

where $x_k = x(w_k) = \Phi^{-1}(y_k)f(w_k)$.

4 Non-degeneracy in the Dual Problem

Let us show that the matrix $\Phi(y)$ is nonsingular, if the point $[u, y] \in \mathcal{F}_D$ is non-degenerate.

Definition 1. [1]. *The point $[u, y] \in \mathcal{F}_D$ is called non-degenerate if $\mathcal{T}_{\mathcal{K}}(y) + \mathcal{R}(\mathcal{A}^T) = \mathbb{R}^n$, where $\mathcal{T}_{\mathcal{K}}(y)$ is the tangent space to the cone \mathcal{K} at the point y and $\mathcal{R}(\mathcal{A}^T)$ is the image of the matrix \mathcal{A}^T .*

Let $[u, y] \in \mathcal{F}_D$, and let the vector $y \in \mathcal{K}$ be partitioned onto three blocks of components: $y = [y_F; y_I; y_N]$. We assume for definiteness that these blocks are consisted from components y_i ordered in the following way: $y_F = [y_1; \dots; y_{r_F}]$, $y_I = [y_{r_F+1}; \dots; y_{r_F+r_I}]$, $y_N = [y_{r_F+r_I+1}; \dots; y_{r_F+r_B+r_N}]$. Recall that $r = r_F + r_I + r_N$.

The partition of the vector y induces the partition of the index set $J^r = [1 : r]$ onto three subsets:

$$J_F^r(y) = [1, \dots, r_F], \quad J_I^r(y) = [r_F + 1, \dots, r_F + r_I], \quad J_N^r(y) = [r_F + r_I + 1, \dots, r].$$

If $i \in J_F^r(y)$, then $y_i \neq 0_{n_i}$ and $y_i \in \partial K_2^{n_i}$, where $\partial K_2^{n_i}$ is the boundary of the cone $K_2^{n_i}$. If $i \in J_I^r(y)$, then $y_i = 0_{n_i}$. At last, in the case, where $i \in J_N^r(y)$, the inclusion $y_i \in \text{int}K_2^{n_i}$ holds. According to the partition of the vector y onto three

blocks of components we partition also the matrix $\mathcal{A} = [\mathcal{A}_F, \mathcal{A}_I, \mathcal{A}_N]$ and the vector $c = [c_F; c_I; c_N]$.

For any nonzero component $y_i \in \mathbb{R}^{n_i}$, $i \in J^r$, the following spectral decomposition

$$y_i = \theta_{i,1} \mathbf{d}_{i,1} + \theta_{i,n_i} \mathbf{d}_{i,n_i} \quad (22)$$

takes place [1]. Here the pair of vectors

$$\mathbf{d}_{i,1} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1; \frac{\bar{y}_i}{\|\bar{y}_i\|} \end{bmatrix}, \quad \mathbf{d}_{i,n_i} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1; -\frac{\bar{y}_i}{\|\bar{y}_i\|} \end{bmatrix},$$

is a Jordan frame. The coefficients $\theta_{i,1}$ and θ_{i,n_i} in (22) are following:

$$\theta_{i,1} = \frac{1}{\sqrt{2}} (y_i^0 + \|\bar{y}_i\|), \quad \theta_{i,n_i} = \frac{1}{\sqrt{2}} (y_i^0 - \|\bar{y}_i\|).$$

Both vectors $\mathbf{d}_{i,1}$ and \mathbf{d}_{i,n_i} are orthogonal each to other and their lengths equal to one.

If $y_i \in K_2^{n_i}$, then $\theta_{i,1} \geq 0$ and $\theta_{i,n_i} \geq 0$. In the case, where $y_i \neq 0_{n_i}$ and $y_i \in \partial K_2^{n_i}$, the equality $y_i^0 = \|\bar{y}_i\|$ holds. Hence, only the first coefficient $\theta_{i,1} = \sqrt{2}y_i^0 = \sqrt{2}\|\bar{y}_i\|$ differs from zero.

Let us assume that $y_i \in K_2^{n_i}$ and $y_i \neq 0_{n_i}$. The matrix $\text{Arr}(y_i)$ is symmetric. Denote by H_i the orthogonal matrix with columns being eigenvectors of $\text{Arr}(y_i)$. The vectors $\mathbf{d}_{i,1}$ and \mathbf{d}_{i,n_i} are among eigenvectors of $\text{Arr}(y_i)$. The matrix H_i can be taken in the following form

$$H_i = [\mathbf{d}_{i,1}, h_{i,2}, \dots, h_{i,n_i-2}, \mathbf{d}_{i,n_i}].$$

The eigenvectors $h_{i,2}, \dots, h_{i,n_i-2}$ are arbitrary vectors from the subspace

$$\mathbb{R}_0^{n_i} = \{z = [z^0; \bar{z}] \in \mathbb{R}^{n_i} : z^0 = 0\}.$$

All these vectors have the unit length and are orthogonal each to others. Moreover, they are orthogonal to the vectors $\mathbf{d}_{i,1}$ and \mathbf{d}_{i,n_i} .

Eigenvalues $y_i^0 + \|\bar{y}_i\|$ and $y_i^0 - \|\bar{y}_i\|$ correspond to the eigenvectors $\mathbf{d}_{i,1}$ and \mathbf{d}_{i,n_i} , respectively. The eigenvalue y_i^0 has the multiplicity $n_i - 2$ and corresponds to eigenvectors $h_{i,2}, \dots, h_{i,n_i-2}$. Denoting by Σ_i the diagonal matrix

$$\Sigma_i = \text{Diag} \left(\sqrt{2}\theta_{i,1}, y_i^0, \dots, y_i^0, \sqrt{2}\theta_{i,n_i} \right),$$

we have $\text{Arr}(y_i) = H_i \Sigma_i H_i^T$.

If $i \in J_I^r(y)$, then $y_i = 0_{n_i}$. In this case the identity matrix I_{n_i} can be taken as the orthogonal matrix H_i . It is evident that $\Sigma_i = 0_{n_i n_i}$ for this $\text{Arr}(y_i)$.

Introduce into consideration the block-diagonal matrices

$$\mathbf{H}_F = \text{DIAG} [H_1, \dots, H_{r_F}], \quad \mathbf{H}_I = \text{DIAG} [H_{r_F+1}, \dots, H_{r_F+r_I}],$$

The matrices \mathbf{H}_F and \mathbf{H}_I are orthogonal. In the same way we combine the diagonal matrices Σ_i :

$$\Sigma_F = \text{DIAG} [\Sigma_1, \dots, \Sigma_{r_F}], \quad \Sigma_I = \text{DIAG} [\Sigma_{r_F+1}, \dots, \Sigma_{r_F+r_I}],$$

Let $\mathcal{A}_F^{\mathbf{H}} = \mathcal{A}_F \mathbf{H}_F$, and let $\bar{\mathcal{A}}_F^{\mathbf{H}}$ be the matrix $\mathcal{A}_F^{\mathbf{H}}$, from which all columns are removed, except the columns being the first columns of matrices $A_i H_i$, $i \in J_F^r(y)$. The matrix $\bar{\mathcal{A}}_F^{\mathbf{H}}$ has the dimension $m \times r_F$. Denote $\mathcal{A}_{FI}^{\mathbf{H}} = [\bar{\mathcal{A}}_F^{\mathbf{H}}, \mathcal{A}_I^{\mathbf{H}}]$, where $\mathcal{A}_I^{\mathbf{H}} = \mathcal{A}_I \mathbf{H}_I$. The following criterion of the non-degeneracy of the point $[u, y]$ is valid [1].

Proposition 2. *The point $[u, y] \in \mathcal{F}_D$ is non-degenerate if and only if columns of the matrix $\mathcal{A}_{F_I}^H$ are linear independent.*

It follows from Proposition 2, that at a non-degenerate point $[u, y]$ the inequality $r_F + n_I \leq m$ takes place, where $n_I = \sum_{i \in J_I^r(y)} n_i$.

Proposition 3. [6] *Let the point $[u, y] \in \mathcal{F}_D$ be non-degenerate. Then the matrix $\Phi(y)$ is nonsingular.*

We call the dual problem (3) *non-degenerate*, if all points $[u, y] \in \mathcal{F}_D$ are non-degenerate. Below we suppose that the problem (3) *is non-degenerate*.

5 Local Convergence of the Dual Methods

Let x_* and $[u_*, y_*]$ be the solutions of problems (2) and (3), respectively. We assume without loss of generality that for the vector x_* the following partition onto three blocks of components $x_* = [x_{*,F}, x_{*,I}, x_{*,N}]$ holds. Moreover, the number of component in blocks $x_{*,F}$, $x_{*,I}$ and $x_{*,N}$ is equal to r_F , r_I and r_N , respectively. Each component $x_{*,i}$ from the block $x_{*,F}$ belongs to the boundary of the cone $K_2^{n_i}$. Each component $x_{*,i}$ from the block $x_{*,I}$ is an interior point of $K_2^{n_i}$. All $x_{*,i}$ from the block $x_{*,N}$ are zero vectors.

Besides, let for the vector y_* the decomposition onto block of components $y_* = [y_{*,F}, y_{*,I}, y_{*,N}]$ take place. Moreover, the number of components in blocks is equal to \bar{r}_F , \bar{r}_I and \bar{r}_N , respectively. But unlike to x_* , components $y_{*,i}$ from the block $y_{*,I}$ are zero vectors. On the contrary, $y_{*,i}$ from the block $y_{*,N}$ is an interior point of the cone $K_2^{n_i}$.

According to (5) the following *complementary condition* $x_{*,i} \circ y_{*,i} = 0$, $1 \leq i \leq r$, holds. The strict complementary condition means that additionally $x_{*,i} + y_{*,i} \in \text{int } K_2^{n_i}$. In this case $\bar{r}_F = r_F$, $\bar{r}_I = r_I$ and $\bar{r}_N = r_N$. Furthermore, the matrices $\text{Arr}(x_{*,i})$ and $\text{Arr}(y_{*,i})$ commute between themselves. The following decompositions

$$\text{Arr}(x_{*,i}) = H_i \Lambda_i H_i^T, \quad \text{Arr}(y_{*,i}) = H_i \Sigma_i H_i^T, \quad (23)$$

take place. Here H_i is an orthogonal matrix, and Λ_i and Σ_i are diagonal matrices with eigenvalues of $\text{Arr}(x_{*,i})$ and $\text{Arr}(y_{*,i})$ at their diagonals, respectively. Below we set $r_{FI} = r_F + r_I$ and $J_F^r = [1 : r_F]$, $J_I^r = [r_F + 1 : r_{FI}]$, $J_N^r = [r_{FI} + 1 : r]$, $J_{F_I}^r = J_F^r \cup J_I^r$.

Similar to (22) for $x_{*,i}$ the spectral decomposition

$$x_{*,i} = \eta_{i,1} \mathbf{e}_{i,1} + \eta_{i,n_i} \mathbf{e}_{i,n_i} \quad (24)$$

holds, where

$$\mathbf{e}_{i,1} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1; \frac{\bar{x}_{*,i}}{\|\bar{x}_{*,i}\|} \end{bmatrix}, \quad \mathbf{e}_{i,n_i} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1; -\frac{\bar{x}_{*,i}}{\|\bar{x}_{*,i}\|} \end{bmatrix}$$

are frame vectors. The coefficients $\eta_{i,1}$ and η_{i,n_i} in (24) are following:

$$\eta_{i,1} = \frac{1}{\sqrt{2}} (x_{*,i}^0 + \|\bar{x}_{*,i}\|), \quad \eta_{i,n_i} = \frac{1}{\sqrt{2}} (x_{*,i}^0 - \|\bar{x}_{*,i}\|).$$

Both $\mathbf{e}_{i,1}$ and \mathbf{e}_{i,n_i} are unit vectors and orthogonal each to other.

The orthogonal matrix H_i in (23) has the form

$$H_i = [\mathbf{e}_{i,1}, h_{i,2}, \dots, h_{i,n_i-2}, \mathbf{e}_{i,n_i}], \quad (25)$$

where $h_{i,2}, \dots, h_{i,n_i-2}$ are unit vectors from the subspace $\mathbb{R}_0^{n_i}$. The matrix $\Lambda_i = \text{Diag}(\sqrt{2}\eta_{i,1}, x_{*,i}^0, \dots, x_{*,i}^0, \sqrt{2}\eta_{i,n_i})$ is diagonal with eigenvalues of $\text{Arr}(x_{*,i})$ on its diagonal, $i \in J_{F_I}^r$. Remark, that for $i \in J_{F_I}^r$ the last eigenvalue is zero, that is $\Lambda_i = \text{Diag}(2x_{*,i}^0, x_{*,i}^0, \dots, x_{*,i}^0, 0)$.

At solutions x_* and y_* according to the complementary condition the vector \mathbf{e}_{i,n_i} must be collinear to the vector $\mathbf{d}_{i,1}$ from the spectral decomposition (22) for $y_{*,i}$. Hence, the orthogonal matrix (25) can be used also in the spectral decomposition of the matrix $\text{Arr}(y_{*,i})$, $i \in J_{F_I}^r$, i.e. $\text{Arr}(y_{*,i}) = H_i \Sigma_i H_i^T$, where Σ_i is a diagonal matrix with the vector of eigenvalues of the matrix $\text{Arr}(y_{*,i})$ at its diagonal. For $i \in J_{F_I}^r$ the matrix Σ_i has the form $\Sigma_i = \text{Diag}(0, y_{*,i}^0, \dots, y_{*,i}^0, 2y_{*,i}^0)$. The matrix Λ_i is zero for $i \in J_N^r$, and, vice versa, the matrix Σ_i is zero, when $i \in J_I^r$.

In addition, let the orthogonal matrix H_i for $i \in J_N^r$ be defined by the matrix $\text{Arr}(y_{*,i})$, that is $\text{Arr}(y_{*,i}) = H_i \Sigma_i H_i^T$. Then $H_i = [\mathbf{d}_{i,1}, h_{i,2}, \dots, h_{i,n_i-2}, \mathbf{d}_{i,n_i}]$ and $\Sigma_i = \text{Diag}(y_{*,i}^0 + \|\bar{y}_{*,i}\|, y_{*,i}^0, \dots, y_{*,i}^0, y_{*,i}^0 - \|\bar{y}_{*,i}\|)$. Moreover, Λ_i is a zero matrix for $i \in J_N^r$.

Let $\Lambda = \text{DIAG}(\Lambda_F, \Lambda_I, \Lambda_N)$, $\Sigma = \text{DIAG}(\Sigma_F, \Sigma_I, \Sigma_N)$, where

$$\Lambda_F = \text{DIAG}(\Lambda_1, \dots, \Lambda_{r_F}), \quad \Sigma_F = \text{DIAG}(\Sigma_1, \dots, \Sigma_{r_F})$$

and

$$\begin{aligned} \Lambda_I &= \text{DIAG}(\Lambda_{r_F+1}, \dots, \Lambda_{r_{F_I}}), & \Sigma_I &= \text{DIAG}(\Sigma_{r_F+1}, \dots, \Sigma_{r_{F_I}}), \\ \Lambda_N &= \text{DIAG}(\Lambda_{r_{F_I}+1}, \dots, \Lambda_r), & \Sigma_N &= \text{DIAG}(\Sigma_{r_{F_I}+1}, \dots, \Sigma_r). \end{aligned}$$

Set also $\mathbf{H} = \text{DIAG}(H_1, \dots, H_r)$ and denote: $\mathcal{A}^{\mathbf{H}_F} = \mathcal{A}\mathbf{H}_F$, $\mathcal{A}^{\mathbf{H}_I} = \mathcal{A}\mathbf{H}_I$, $\mathcal{A}^{\mathbf{H}_N} = \mathcal{A}\mathbf{H}_N$, $\mathcal{A}^{\mathbf{H}} = \mathcal{A}\mathbf{H}$. For $\mathcal{A}^{\mathbf{H}}$ the decomposition $\mathcal{A}^{\mathbf{H}} = [\mathcal{A}^{\mathbf{H}_F}, \mathcal{A}^{\mathbf{H}_I}, \mathcal{A}^{\mathbf{H}_N}]$ is valid. With the introduced notations the matrix $\mathbf{G}(u_*)$ can be submitted in the form

$$\mathbf{G}(u_*) = \mathcal{A}^{\mathbf{H}} \Phi^{-\mathbf{H}}(y_*) \Lambda(x_*) (\mathcal{A}^{\mathbf{H}})^T, \quad \Phi^{-\mathbf{H}}(y_*) = \left(\Phi^{\mathbf{H}}(y_*) \right)^{-1}, \quad (26)$$

where $y_* = y(u_*)$ and $\Phi^{\mathbf{H}}(y_*) = \mathbf{H}^T \Phi(y_*) \mathbf{H}$.

We have by aforesaid

$$\Phi^{\mathbf{H}}(y_*) = \begin{bmatrix} (\mathcal{A}_F^{\mathbf{H}})^T \mathcal{A}_F^{\mathbf{H}} + \Sigma_F & (\mathcal{A}_F^{\mathbf{H}})^T \mathcal{A}_I^{\mathbf{H}} & (\mathcal{A}_F^{\mathbf{H}})^T \mathcal{A}_N^{\mathbf{H}} \\ (\mathcal{A}_I^{\mathbf{H}})^T \mathcal{A}_F^{\mathbf{H}} & (\mathcal{A}_I^{\mathbf{H}})^T \mathcal{A}_I^{\mathbf{H}} & (\mathcal{A}_I^{\mathbf{H}})^T \mathcal{A}_N^{\mathbf{H}} \\ (\mathcal{A}_N^{\mathbf{H}})^T \mathcal{A}_F^{\mathbf{H}} & (\mathcal{A}_N^{\mathbf{H}})^T \mathcal{A}_I^{\mathbf{H}} & (\mathcal{A}_N^{\mathbf{H}})^T \mathcal{A}_N^{\mathbf{H}} + \Sigma_N \end{bmatrix}.$$

All diagonal entrees of the matrix Σ_N are strictly positive. The diagonal matrix Σ_F is such that there are r_F zero entrees at its diagonal. All these zero entrees are first diagonal elements of the matrices Σ_i , $i \in J_{F_I}^r$.

Compute $\Phi^{-\mathbf{H}}(y_*)$. For this purpose we firstly rearrange rows and columns of the matrix. Suppose that first columns of the matrices $\mathcal{A}_i^{\mathbf{H}}$, $i \in J_{F_I}^r$, are removed from $\mathcal{A}_i^{\mathbf{H}}$, and the separate sub-matrix $\tilde{\mathcal{A}}_F^{\mathbf{H}}$ is composed from these first columns. The dimension of $\tilde{\mathcal{A}}_F^{\mathbf{H}}$ is $m \times r_F$. Denote by $\tilde{\mathcal{A}}_F^{\mathbf{H}}$ the sub-matrix of the matrix $\mathcal{A}_F^{\mathbf{H}}$ composed from the rest columns of $\mathcal{A}_F^{\mathbf{H}}$. Add the sub-matrix $\tilde{\mathcal{A}}_F^{\mathbf{H}}$ to the matrix \mathcal{A}_I , putting it before \mathcal{A}_I . The resulting $m \times (r_F + n_I)$ matrix denote by $\mathcal{A}_{F_I}^{\mathbf{H}}$. Moreover, denote by $\tilde{\Sigma}_F$ the diagonal sub-matrix of the matrix Σ_F , from which first diagonal entrees of the matrices Σ_i , $i \in J_{F_I}^r$, are eliminated. Let Π be a permutation matrix, realizing the mentioned

changes of rows and columns of $\Phi^{\mathbf{H}}(y_*)$. Then the matrix $\Phi^{\mathbf{H}}(y_*)$ can be written in the form

$$\Phi^{\mathbf{H}}(y_*) = \Pi \begin{bmatrix} (\hat{\mathcal{A}}_F^{\mathbf{H}})^T \hat{\mathcal{A}}_F + \hat{\Sigma}_F & (\hat{\mathcal{A}}_F^{\mathbf{H}})^T \mathcal{A}_{FI}^{\mathbf{H}} & (\hat{\mathcal{A}}_F^{\mathbf{H}})^T \mathcal{A}_N^{\mathbf{H}} \\ (\mathcal{A}_{FI}^{\mathbf{H}})^T \hat{\mathcal{A}}_F & (\mathcal{A}_{FI}^{\mathbf{H}})^T \mathcal{A}_{FI}^{\mathbf{H}} & (\mathcal{A}_{FI}^{\mathbf{H}})^T \mathcal{A}_N^{\mathbf{H}} \\ (\mathcal{A}_N^{\mathbf{H}})^T \hat{\mathcal{A}}_F & (\mathcal{A}_N^{\mathbf{H}})^T \mathcal{A}_{FI}^{\mathbf{H}} & (\mathcal{A}_N^{\mathbf{H}})^T \mathcal{A}_N^{\mathbf{H}} + \Sigma_N \end{bmatrix} \Pi^T. \quad (27)$$

Partition the matrix (27) onto four blocks:

$$\Phi^{\mathbf{H}}(y_*) = \Pi \begin{bmatrix} \mathcal{W}_{11} & \mathcal{W}_{12} \\ \mathcal{W}_{12}^T & \mathcal{W}_{22} \end{bmatrix} \Pi^T,$$

where

$$\mathcal{W}_{11} = \begin{bmatrix} (\hat{\mathcal{A}}_F^{\mathbf{H}})^T \hat{\mathcal{A}}_F + \hat{\Sigma}_F & (\hat{\mathcal{A}}_F^{\mathbf{H}})^T \mathcal{A}_{FI}^{\mathbf{H}} \\ (\mathcal{A}_{FI}^{\mathbf{H}})^T \hat{\mathcal{A}}_F & (\mathcal{A}_{FI}^{\mathbf{H}})^T \mathcal{A}_{FI}^{\mathbf{H}} \end{bmatrix}, \quad \mathcal{W}_{12} = \begin{bmatrix} (\hat{\mathcal{A}}_F^{\mathbf{H}})^T \mathcal{A}_N^{\mathbf{H}} \\ (\mathcal{A}_{FI}^{\mathbf{H}})^T \mathcal{A}_N^{\mathbf{H}} \end{bmatrix}$$

and $\mathcal{W}_{22} = \Sigma_N + (\mathcal{A}_N^{\mathbf{H}})^T \mathcal{A}_N^{\mathbf{H}}$.

If the non-degeneracy condition holds at the point $[u_*, y_*]$, then according to Proposition 3 the matrix $\Phi^{\mathbf{H}}(y_*)$ is positive definite. Therefore, the diagonal blocks \mathcal{W}_{11} and \mathcal{W}_{22} are also positive definite matrices.

Using the Frobenius formula, we obtain

$$\Phi^{-\mathbf{H}}(y_*) = \Pi \begin{bmatrix} \mathcal{V}_{11} & \mathcal{V}_{12} \\ \mathcal{V}_{12}^T & \mathcal{V}_{22} \end{bmatrix} \Pi^T,$$

where

$$\mathcal{V}_{11} = \mathcal{W}_{11}^{-1} + \mathcal{W}_{11}^{-1} \mathcal{W}_{12} \mathcal{Z}^{-1} \mathcal{W}_{12}^T \mathcal{W}_{11}^{-1}, \quad \mathcal{V}_{12} = -\mathcal{W}_{11}^{-1} \mathcal{W}_{12} \mathcal{Z}^{-1}, \quad \mathcal{V}_{22} = \mathcal{Z}^{-1} \quad (28)$$

and $\mathcal{Z} = \mathcal{W}_{22} - \mathcal{W}_{12}^T \mathcal{W}_{11}^{-1} \mathcal{W}_{12}$.

Firstly, compute the matrix \mathcal{W}_{11}^{-1} . According to Proposition 2 the matrix $(\mathcal{A}_{FI}^{\mathbf{H}})^T \mathcal{A}_{FI}^{\mathbf{H}}$ at the non-degenerate point $[u_*, y_*]$ is nonsingular. Denote

$$\mathcal{Y} = (\hat{\mathcal{A}}_F^{\mathbf{H}})^T \hat{\mathcal{A}}_F + \hat{\Sigma}_F - (\hat{\mathcal{A}}_F^{\mathbf{H}})^T \mathcal{A}_{FI}^{\mathbf{H}} \left[(\mathcal{A}_{FI}^{\mathbf{H}})^T \mathcal{A}_{FI}^{\mathbf{H}} \right]^{-1} (\mathcal{A}_{FI}^{\mathbf{H}})^T \hat{\mathcal{A}}_F. \quad (29)$$

Denote also $\mathcal{P} = \mathcal{A}_{FI}^{\mathbf{H}} \left[(\mathcal{A}_{FI}^{\mathbf{H}})^T \mathcal{A}_{FI}^{\mathbf{H}} \right]^{-1} (\mathcal{A}_{FI}^{\mathbf{H}})^T$. The matrix \mathcal{P} is an orthogonal projector onto the linear sub-space \mathcal{L} , generated by columns of the matrix $\mathcal{A}_{FI}^{\mathbf{H}}$. The matrix $\mathcal{P}_{\perp} = I - \mathcal{P}$ projects onto the orthogonal complement \mathcal{L}^{\perp} to the sub-space \mathcal{L} . By (29)

$$\mathcal{Y} = \hat{\Sigma}_F + (\hat{\mathcal{A}}_F^{\mathbf{H}})^T \mathcal{P}_{\perp} \hat{\mathcal{A}}_F. \quad (30)$$

Let $\mathcal{E} = \left[(\mathcal{A}_{FI}^{\mathbf{H}})^T \mathcal{A}_{FI}^{\mathbf{H}} \right]^{-1}$, $\mathcal{S} = \hat{\mathcal{A}}_F \mathcal{Y}^{-1} (\hat{\mathcal{A}}_F^{\mathbf{H}})^T$. With the help of the Frobenius formula, we obtain

$$\mathcal{W}_{11}^{-1} = \begin{bmatrix} \mathcal{Y}^{-1} & -\mathcal{Y}^{-1} (\hat{\mathcal{A}}_F^{\mathbf{H}})^T \mathcal{A}_{FI}^{\mathbf{H}} \mathcal{E} \\ -\mathcal{E} (\mathcal{A}_{FI}^{\mathbf{H}})^T \hat{\mathcal{A}}_F \mathcal{Y}^{-1} & \mathcal{E} + \mathcal{E} (\mathcal{A}_{FI}^{\mathbf{H}})^T \mathcal{S} \mathcal{A}_{FI}^{\mathbf{H}} \mathcal{E} \end{bmatrix}, \quad (31)$$

The matrix \mathcal{P}_\perp is idempotent, that is $\mathcal{P}_\perp = \mathcal{P}_\perp^2$. Using the Sherman-Morrison-Woodbury formula, we derive from (30)

$$\mathcal{Y}^{-1} = \hat{\Sigma}_F^{-1} - \hat{\Sigma}_F^{-1} \left(\hat{\mathcal{A}}_F^{\mathbf{H}} \right)^T \mathcal{P}_\perp \left[I_m + \mathcal{P}_\perp \hat{\mathcal{A}}_F^{\mathbf{H}} \hat{\Sigma}_F^{-1} \left(\hat{\mathcal{A}}_F^{\mathbf{H}} \right)^T \mathcal{P}_\perp \right]^{-1} \mathcal{P}_\perp \hat{\mathcal{A}}_F^{\mathbf{H}} \hat{\Sigma}_F^{-1}. \quad (32)$$

Introduce the additional notation $\hat{\mathcal{A}}_{FI}^{\mathbf{H}} = \left[\hat{\mathcal{A}}_F^{\mathbf{H}}, \mathcal{A}_{FI}^{\mathbf{H}} \right]$. Then the matrix \mathcal{Z} can be written in the form

$$\mathcal{Z} = \Sigma_N + \left(\mathcal{A}_N^{\mathbf{H}} \right)^T \left[I_m - \hat{\mathcal{A}}_{FI}^{\mathbf{H}} \mathcal{W}_{11}^{-1} \left(\hat{\mathcal{A}}_{FI}^{\mathbf{H}} \right)^T \right] \mathcal{A}_N^{\mathbf{H}}. \quad (33)$$

It can be seen from (33) that the matrix \mathcal{Z} is a Schur complement of the positive definite matrix \mathcal{W}_{11} at (27). Therefore, \mathcal{Z} is a positive definite matrix too.

Proposition 4. *Let $\hat{\mathcal{S}} = \mathcal{P} + \mathcal{P}_\perp \mathcal{S} \mathcal{P}_\perp$. Then $\hat{\mathcal{A}}_{FI}^{\mathbf{H}} \mathcal{W}_{11}^{-1} \left(\hat{\mathcal{A}}_{FI}^{\mathbf{H}} \right)^T = \hat{\mathcal{S}}$.*

Proof. This equality can be obtained by direct calculations. \square

Corollary 1. *According to (33) $\mathcal{Z} = \Sigma_N + \left(\mathcal{A}_N^{\mathbf{H}} \right)^T \left(I_m - \hat{\mathcal{S}} \right) \mathcal{A}_N^{\mathbf{H}}$. Since Σ_N is a positive definite diagonal matrix, we obtain by the Sherman-Morrison-Woodbury formula*

$$\begin{aligned} \mathcal{Z}^{-1} &= \Sigma_N^{-1} - \Sigma_N^{-1} \left(\mathcal{A}_N^{\mathbf{H}} \right)^T \left(I - \hat{\mathcal{S}} \right)^{1/2} \\ &\cdot \left[I + \left(I - \hat{\mathcal{S}} \right)^{1/2} \mathcal{A}_N^{\mathbf{H}} \Sigma_N^{-1} \left(\mathcal{A}_N^{\mathbf{H}} \right)^T \left(I - \hat{\mathcal{S}} \right)^{1/2} \right]^{-1} \left(I - \hat{\mathcal{S}} \right)^{1/2} \mathcal{A}_N^{\mathbf{H}} \Sigma_N^{-1} \end{aligned} \quad (34)$$

Using the matrices (31) and (34), it is possible by (28) compute the matrix $\Phi^{-\mathbf{H}}(y_*)$.

Below, we will need in the definition of non-degeneracy of a point $x \in \mathcal{F}_P$ in the primal problem (2).

Definition 2. [1]. *The point $x \in \mathcal{F}_P$ is called non-degenerate, if $\mathcal{T}_K(x) + \mathcal{N}(A) = \mathbb{R}^n$, where $\mathcal{T}_K(x)$ is a tangent space to the cone \mathcal{K} at the point x , and $\mathcal{N}(A)$ is a null-space of the matrix A .*

Denote by H_i^L , $i \in J_F^r$, the left $n_i \times (n_i - 1)$ sub-matrix of the matrix H_i . In other words, H_i^L is the matrix H_i , from which the last column \mathbf{e}_{i, n_i} is removed. Denote also by $A_i^{\mathbf{H}L} = A_i H_i^L$. Compose from $A_i^{\mathbf{H}L}$, $i \in J_F^r$, the matrix $\mathcal{A}_F^{\mathbf{H}L} = \left[A_1^{\mathbf{H}L}, \dots, A_{r_F}^{\mathbf{H}L} \right]$ with the dimension $m \times (n_F - r_F)$, where $n_F = \sum_{i \in J_F^r} n_i$. Introduce additionally the matrix $\mathcal{A}_{FI}^{\mathbf{H}L} = \left[\mathcal{A}_F^{\mathbf{H}L}, \mathcal{A}_I^{\mathbf{H}} \right]$. The following criterion of non-degeneracy of the point $x \in \mathcal{F}_P$ is valid.

Proposition 5. [1]. *The point $x = [x_F; x_I; x_N]$ is non-degenerate if and only if rows of the matrix $\mathcal{A}_{FI}^{\mathbf{H}L}$ are linear independent.*

Lemma 2. *Let $x_* \in \mathcal{F}_P$ and $[u_*, y_*] \in \mathcal{F}_D$ be non-degenerate solutions of problems (2) and (3), respectively. Let also the solutions x_* and y_* be strictly complementary. Then the matrix $\mathbf{G}(u_*) = \mathcal{A} \Phi^{-1}(u_*) \mathbf{A} \mathbf{r} \mathbf{r}(x(u_*)) \mathbf{A}^T$ is nonsingular.*

Proof. Since $\mathcal{A}x_* = b$ and $\mathbf{A}\mathbf{r}\mathbf{r}(y_*)x_* = 0_n$, we have $[\mathbf{A}\mathbf{r}\mathbf{r}(y_*) + \mathcal{A}^T\mathcal{A}]x_* = \mathcal{A}^Tb$. Hence, $x(u_*) = x_*$. Moreover, the expression (26) for $\mathbf{G}(u_*)$ takes place.

Let show, that the homogeneous system of linear equations

$$\mathbf{G}(u_*)z = 0_m \quad (35)$$

has only zero solution $z = 0_m$. It follows from here that the matrix $\mathbf{G}(u_*)$ is nonsingular.

Denote $\hat{\mathcal{A}}^{\mathbf{H}} = \mathcal{A}^{\mathbf{H}}\mathbf{\Pi}$. By (26) $\mathbf{G}(u_*) = \hat{\mathcal{A}}^{\mathbf{H}}\hat{\Phi}^{-\mathbf{H}}(y_*)\hat{\Lambda}(x_*)\left(\hat{\mathcal{A}}^{\mathbf{H}}\right)^T$, where $\hat{\mathcal{A}}^{\mathbf{H}} = \left[\hat{\mathcal{A}}_{FI}^{\mathbf{H}}, \hat{\mathcal{A}}_N^{\mathbf{H}}\right]$ and $\hat{\Phi}^{-\mathbf{H}}(y_*) = \mathbf{\Pi}^T\hat{\Phi}^{-\mathbf{H}}(y_*)\mathbf{\Pi}$. The block-diagonal matrix $\hat{\Lambda} = \hat{\Lambda}(x_*)$ is obtained from the matrix $\mathbf{\Lambda}$ by rearrangement of rows and columns with the help of the permutation matrix $\mathbf{\Pi}$, that is $\hat{\Lambda} = \mathbf{\Pi}\mathbf{\Lambda}\mathbf{\Pi}^T$. This matrix can be written also in the form $\hat{\Lambda} = \text{DIAG}\left[\hat{\Lambda}_{FI}, \mathbf{\Lambda}_N\right]$.

The right lower block $\mathbf{\Lambda}_N$ is a zero matrix, therefore

$$\mathbf{G}(u_*) = \hat{\mathcal{A}}_{FI}^{\mathbf{H}}\mathcal{V}_{11}\hat{\Lambda}_{FI}\left(\hat{\mathcal{A}}_{FI}^{\mathbf{H}}\right)^T + \mathcal{A}_N^{\mathbf{H}}\mathcal{V}_{12}^T\hat{\Lambda}_{FI}\left(\hat{\mathcal{A}}_{FI}^{\mathbf{H}}\right)^T.$$

Substituting \mathcal{V}_{11} and \mathcal{V}_{12} , we derive that $\mathbf{G}(u_*)$ is the matrix of the following form

$$\begin{aligned} \mathbf{G}(u_*) &= \hat{\mathcal{A}}_{FI}^{\mathbf{H}}\mathcal{W}_{11}^{-1}\hat{\Lambda}_{FI}\left(\hat{\mathcal{A}}_{FI}^{\mathbf{H}}\right)^T \\ &+ \left[\hat{\mathcal{A}}_{FI}^{\mathbf{H}}\mathcal{W}_{11}^{-1}\left(\hat{\mathcal{A}}_{FI}^{\mathbf{H}}\right)^T - I_m\right]\mathcal{A}_N^{\mathbf{H}}\mathcal{Z}^{-1}\left(\mathcal{A}_N^{\mathbf{H}}\right)^T\hat{\mathcal{A}}_{FI}^{\mathbf{H}}\mathcal{W}_{11}^{-1}\hat{\Lambda}_{FI}\left(\hat{\mathcal{A}}_{FI}^{\mathbf{H}}\right)^T. \end{aligned}$$

Denote $\mathcal{U} = \mathcal{A}_N^{\mathbf{H}}\mathcal{Z}^{-1}\left(\mathcal{A}_N^{\mathbf{H}}\right)^T$. Then, using Proposition 4, we come to conclusion that

$$\mathbf{G}(u_*) = \left[I_m - \left(I_m - \hat{\mathcal{S}}\right)\mathcal{U}\right]\hat{\mathcal{A}}_{FI}^{\mathbf{H}}\mathcal{W}_{11}^{-1}\hat{\Lambda}_{FI}\left(\hat{\mathcal{A}}_{FI}^{\mathbf{H}}\right)^T.$$

We multiply the equality (35) from the left on the matrix $\left(\mathcal{A}_{FI}^{\mathbf{H}}\right)^T$. Since $I_m - \hat{\mathcal{S}} = \mathcal{P}_{\perp} - \mathcal{P}_{\perp}\mathcal{S}\mathcal{P}_{\perp}$, we derive that $\left(\mathcal{A}_{FI}^{\mathbf{H}}\right)^T\left(I_m - \hat{\mathcal{S}}\right) = 0$. Thus, we have

$$\left(\mathcal{A}_{FI}^{\mathbf{H}}\right)^T\hat{\mathcal{A}}_{FI}^{\mathbf{H}}\mathcal{W}_{11}^{-1}\hat{\Lambda}_{FI}\left(\hat{\mathcal{A}}_{FI}^{\mathbf{H}}\right)^T z = 0_{l_1}, \quad (36)$$

where $l_1 = r_F + n_I$.

Assume that $z \neq 0_m$ and consider separately two possibilities.

1) $\left(\mathcal{A}_{FI}^{\mathbf{H}}\right)^T z \neq 0_{l_1}$. In this case, taking into account the expression (31) for the matrix \mathcal{W}_{11}^{-1} , we get

$$\left(\mathcal{A}_{FI}^{\mathbf{H}}\right)^T\hat{\mathcal{A}}_{FI}^{\mathbf{H}}\mathcal{W}_{11}^{-1} = \left(\mathcal{A}_{FI}^{\mathbf{H}}\right)^T\left[\mathcal{P}_{\perp}\hat{\mathcal{A}}_F^{\mathbf{H}}\mathcal{Y}^{-1}, (I_m - \mathcal{P}_{\perp}\mathcal{S})\mathcal{A}_{FI}^{\mathbf{H}}\mathcal{E}\right] = [0_{l_1l_2}, I_{l_1}],$$

where $l_2 = n_F - r_F$. Hence, the equation (36) is reduced to the following one: $\Lambda_{FI}\left(\mathcal{A}_{FI}^{\mathbf{H}}\right)^T z = 0_{l_1}$, where Λ_{FI} is a right lower diagonal $l_1 \times l_1$ sub-matrix of the matrix $\hat{\Lambda}_{FI}$. Because all diagonal entrees of the matrix Λ_{FI} are positive numbers, this equality does not fulfilled, when $z \neq 0_m$.

2) $\left(\mathcal{A}_{FI}^{\mathbf{H}}\right)^T z = 0_{l_1}$. Under this assumption $z \in \mathcal{L}^{\perp}$, therefore,

$$\hat{\mathcal{A}}_{FI}^{\mathbf{H}}\mathcal{W}_{11}^{-1}\hat{\Lambda}_{FI}\left(\hat{\mathcal{A}}_{FI}^{\mathbf{H}}\right)^T z = \mathcal{P}_{\perp}\hat{\mathcal{A}}_F^{\mathbf{H}}\mathcal{Y}^{-1}\hat{\Lambda}_F\left(\hat{\mathcal{A}}_F^{\mathbf{H}}\right)^T\mathcal{P}_{\perp}z. \quad (37)$$

By (32) $\mathcal{P}_\perp \hat{\mathcal{A}}_F^H \mathcal{Y}^{-1} = \left(I_m + \hat{\mathcal{C}}\right)^{-1} \mathcal{P}_\perp \hat{\mathcal{A}}_F^H \hat{\Sigma}_F^{-1}$, where $\hat{\mathcal{C}} = \mathcal{P}_\perp \hat{\mathcal{A}}_F^H \hat{\Sigma}_F^{-1} (\hat{\mathcal{A}}_F^H)^T \mathcal{P}_\perp$. It follows from here and (37) that

$$\hat{\mathcal{A}}_{FI}^H \mathcal{W}_{I1}^{-1} \hat{\Lambda}_F \left(\hat{\mathcal{A}}_{FI}^H\right)^T z = \left(I_m + \hat{\mathcal{C}}\right)^{-1} \mathcal{P}_\perp \hat{\mathcal{A}}_F^H \hat{\Sigma}_F^{-1} \hat{\Lambda}_F \left(\hat{\mathcal{A}}_F^H\right)^T \mathcal{P}_\perp z = 0.$$

But the matrix $\left(I_m + \hat{\mathcal{C}}\right)^{-1}$ is positive definite. Therefore, this equality is possible only when $\mathcal{P}_\perp \hat{\mathcal{A}}_F^H \hat{\Sigma}_F^{-1} \hat{\Lambda}_F \left(\hat{\mathcal{A}}_F^H\right)^T \mathcal{P}_\perp z = 0$.

All diagonal entrees of the diagonal matrix $\Sigma_F^{-1} \hat{\Lambda}_F$ are positive except of r_F diagonal entrees equal to zero. All these zero entrees correspond to the last diagonal entrees of matrices Λ_i in the decomposition $\text{Arr}(x_{*,i}) = H_i \Lambda_i H_i^T$, $i \in J_F^r$. Denote by \mathcal{A}_F^{HL} the sub-matrix of the matrix \mathcal{A}_F^H , from which the last columns of the matrices \mathcal{A}_i^H are removed.

Let $p = \hat{\Sigma}_F^{-1} \hat{\Lambda}_F \left(\hat{\mathcal{A}}_F^H\right)^T \mathcal{P}_\perp z$. The vector p is non-zero. Really, otherwise, because of $\left(\mathcal{A}_{FI}^H\right)^T z = 0$, the rows of the matrix $[\mathcal{A}^{HL}, \mathcal{A}_{FI}^H]$ are linear dependent. This contradicts to non-degeneracy of the point x_* .

By the same reason the equality $\mathcal{P}_\perp \hat{\mathcal{A}}_F^H p = 0$ is also impossible, since in the opposite case we have contradiction with Proposition 5. \square

Lemma 3. *Let assumptions of Lemma 2 hold. Then the matrix $\Psi_w(w_*)$ is nonsingular, where $w_* = [u_*; y_*]$ is the solution of problem (3).*

Proof. Multiplying the right column of the matrix $\Psi_w(w_*)$ from the right by \mathcal{A}^T and subtracting this column from the left column, we obtain the matrix

$$\left[\begin{array}{cc} \mathcal{A}\Phi^{-1} \mathbf{Arr}(x_*) \mathcal{A}^T & \mathcal{A}\Phi^{-1} [\tau I_n - \mathbf{Arr}(x(w))] \\ [\mathbf{Arr}(y_*) \Phi^{-1} - I_n] \mathbf{Arr}(x_*) \mathcal{A}^T & [I_n - \mathbf{Arr}(y_*) \Phi^{-1}] \mathbf{Arr}(x_*) + \tau \mathbf{Arr}(y) \Phi^{-1} \end{array} \right], \quad (38)$$

where $x_* = x(w_*)$ is the solution of the primal problem (2), and $\Phi^{-1} = \Phi^{-1}(y_*)$.

Further, we multiply the first row of the matrix (38) from the left by the matrix \mathcal{A}^T and sum it with the second row. Since $[\mathcal{A}^T \mathcal{A} + \mathbf{Arr}(y_*)] \Phi^{-1}(y_*) = I_n$, we amount to the matrix

$$\left[\begin{array}{cc} \mathcal{A}\Phi^{-1}(y_*) \mathbf{Arr}(x_*) \mathcal{A}^T & \mathcal{A}\Phi^{-1}(y) [\tau I_n - \mathbf{Arr}(x(w))] \\ 0_{nm} & \tau I_n \end{array} \right]. \quad (39)$$

By Lemma 2 the left upper sub-matrix $\mathcal{A}\Phi^{-1}(y_*) \mathbf{Arr}(x_*) \mathcal{A}^T$ is non-singular. Under $\tau > 0$ the right lower sub-matrix of the matrix (39) is also non-singular. Therefore, the matrix (38) is non-singular too. \square

Remark 1. If the point $[u_*, y_*] \in \mathcal{F}_D$ is non-degenerate, then due to continuity the points $[u, y]$ in some vicinity of $[u_*, y_*]$ are also non-degenerate. Thus, the algorithmic mappings in methods (14) and (21) are completely defined in some domain containing points u_* and w_* , respectively.

Theorem 1. *Let all conditions of Lemma 2 be valid. Then the iterative methods (14) and (21) converge locally to the solutions u_* and w_* with super-linear rate.*

Proof. The proof follows from well-known results concerning the Newton method and from Lemmas 2, 3. \square

6 Conclusion

We have proposed two variants of the dual Mewton's method for solving linear second order cone programming problems. Both variants of the method converge locally with the super-linear rate. From theoretical point of view dual methods are preferable in compare with primal methods, when the number of equalities in the primal problem is not large.

References

1. Alizadeh, F., Goldfarb, D.: Second-order cone programming. *Math. Program. Ser. B.* **95**, 3–51 (2003)
2. Lobo, M.S., Vandenberghe, L., Boyd, S., Lebret, H.: Applications of second order cone programming. *Linear Algebra Appl.* **284**, 193–228 (1998)
3. Anjos, M.F., Lasserre, J.B. (eds.): *Handbook of Semidefinite, Cone and Polynomial Optimization: Theory, Algorithms, Software and Applications*, p. 915. Springer, New York (2011). <https://doi.org/10.1007/978-1-4614-0769-0>
4. Nesterov, Y.E., Todd, M.J.: Primal-dual interior-point methods for self-scaled cones. *SIAM. J. Optim.* **8**, 324–364 (1998)
5. Monteiro, R.D.C., Tsuchiya, T.: Polynomial convergence of primal-dual algorithms for second-order cone program based on the MZ-family of directions. *Math. Program.* **88**(1), 61–83 (2000)
6. Zhadan, V.: Dual multiplicative-barrier methods for linear second-order cone programming. In: Jaćimović, M., Khachay, M., Malkova, V., Posypkin, M. (eds.) *OPTIMA 2019. CCIS*, vol. 1145, pp. 295–310. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-38603-0_22
7. Evtushenko, Y.G., Zhadan, V.G.: Dual barrier-projection and barrier-newton methods for linear programming. *Comp. Maths. Math. Phys.* **36**(7), 847–859 (1996)
8. Zhadan, V.G.: Primal Newton method for the linear cone programming problem. *Comput. Mathe. Mathe. Physics.* **58**(2), 207–214 (2018)