ISSN 0005-1179 (print), ISSN 1608-3032 (online), Automation and Remote Control, 2026, Vol. 87, No. 9, pp. 831–845.
© The Author(s), 2026 published by Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, 2026.
Russian Text © The Author(s), 2026, published in Avtomatika i Telemekhanika, 2026, No. 9, pp. 91–109.

LINEAR SYSTEMS

Block Coordinate Descent Method Applied to Stabilization and Time-Optimization Problems for Linear Discrete Systems

K. A. Tsarkov

Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia
e-mail: k6472@mail.ru

Received January 12, 2026
Revised March 13, 2026
Accepted March 23, 2026

Abstract—In the paper, we propose algorithms for solving terminal stabilization and time-optimization problems for discrete-time linear dynamical systems with convex sets of geometric constraints on the control. The algorithms are based on the well-known block coordinate descent method. We present some corollaries to classical results and give new assertions on the convergence of the constructed algorithms to a solution in the problems.

Keywords: terminal stabilization, optimal time, discrete time, linear systems, sequentional global improvement

DOI: 10.7868/S1608303226090039

1. INTRODUCTION

The present study is motivated by the results obtained in [1, 2] on applications of V.F. Krotov's method of sequentional global improvements to problems of terminal stabilization and time-optimization for discrete-time linear systems. An analysis of these results shows that one of the possible implementations of Krotov method in the case under consideration leads to an algorithm widely known as block coordinate descent. With that, a natural desire arises to present an approach to solving the mentioned problems, based on the use of results on the convergence of block coordinate descent. In this regard, we intend to formulate the most significant known results and derive some new meaningful consequences from them. Ultimately, we plan to obtain a transparent justification for the convergence of the algorithms from [1, 2] and propose potentially more efficient numerical methods for solving these problems.

It should be noted that Krotov method, in general, differs significantly from the block coordinate descent method in its design, applicability, and results, although they share many similarities. For example, the formulation of Krotov method involves an unspecified function (the Krotov function), which can be selected from a fairly wide class of admissible functions. Obtaining identical algorithmic implementations for these methods is largely due to the fact that Krotov method is used here with a very specific choice of the Krotov function. The stages of constructing such a function are described in detail in [1, 2].

One of the advantages of the block coordinate descent method is its widespread popularity. There is an extensive list of works devoted to various theoretical and computational aspects of its application to various problems (see, for example, a list of more than two dozen relevant applications of the method in [5]). It is also known that the block coordinate descent can be a very effective computational tool for the problems with explicit block-composition structure. At the same time, it is practically never used to solve terminal stabilization and time-optimization problems for discrete-time linear systems, although it is precisely in the theory of optimal control of discrete systems

that the domain of admissible values of variables is decomposed into a product of a finite number of sets of geometric constraints (blocks) by the very problem statement. In fact, specialized methods are rarely used in studying these problems at all.

Among recent works related to the study of original methods for solving time-optimization problem for a linear discrete system, we should specially mention [6]. In [6], a numerical approach to time-optimization problem is developed, based on support functions of sets of geometric constraints on the control. A distinctive feature is that it presupposes the solution of a series of non-convex and non-smooth optimization problems, while the original problem, generally speaking, decomposes into a sequence of convex problems with a smooth objective function. Nevertheless, almost all of these problems can be replaced by equivalent convex problems. However, a difficulty remains associated with the need to use support functions of sets of geometric constraints in the structure of the objective function to be optimized. Moreover, after determining the optimal time, the problem of constructing an optimal process remains. In [6], it is shown that an optimal process can be constructed using the solution of the last of the mentioned optimization problems. At the same time, such a construction is not always trivial and, in the non-strictly convex case, requires the use of certain additional constructions and algorithms. Note that some of the ideas behind this method for solving time-optimization problem appeared earlier in [7].

The methods in [1, 2], also discussed here, propose solving only low-dimensional convex problems in which the objective is quite simple. Moreover, the construction of an optimal process is carried out simultaneously with the calculation of the optimal time.

2. STATEMENT OF THE PROBLEMS

Consider a linear nonstationary discrete-time system

$$x(k+1) = A(k)x(k) + B(k)u(k), \quad k = 0, 1, \dots, \tag{1}$$

where $x(k) \in \mathbb{R}^n$ is the state of the system, $u(k) \in U(k)$ is the control, $U(k)$ are convex compact sets in $\mathbb{R}^m$ containing the origin, $A(k) \in \mathbb{R}^{n\times n}$ and $B(k) \in \mathbb{R}^{n\times m}$ are arbitrary given matrices. The initial condition for the system is fixed:

$$x(0) = x_0 \in \mathbb{R}^n. \tag{2}$$

First, let us discuss the formulation of terminal stabilization problem for system (1) with initial condition (2). Let us fix some value $N > 0$ of the discrete time at which the controllable process terminates. Consider the problem

$$\|x(N)\| \to \min_{u\in\mathcal{U}_N}, \tag{3}$$

where $\mathcal{U}_N := \{k \mapsto u(k) : \{0, \dots, N-1\} \to \mathbb{R}^m \mid u(k) \in U(k)\}$, and $x(N) \in \mathbb{R}^n$ is determined by the system (1)–(2) for each $u \in \mathcal{U}_N$. Here and below, $\|\cdot\|$ denotes the Euclidean norm in $\mathbb{R}^n$.

In explicit form, the resulting problem (1)–(3) can be rewritten as follows:

$$J_N(u) := \left\| \mathcal{A}_N(0)x_0 + \sum_{k=0}^{N-1} \mathcal{A}_N(k+1)B(k)u(k) \right\| \to \min_{u\in\mathcal{U}_N}, \tag{4}$$

where $\mathcal{A}_N(k) := A(N-1)\dots A(k)$ for $0 \leqslant k \leqslant N-1$, $\mathcal{A}_N(N) := I$, I is the identity matrix of dimensions $n \times n$. Since the elements of the set $\mathcal{U}_N$ are naturally identified with the corresponding vectors in the Euclidean space $\mathbb{R}^{mN}$, $\mathcal{U}_N$ itself can be represented as the Cartesian product $U(0)\dots\times U(N-1)$ of the initially given convex and compact sets $U(k) \subset \mathbb{R}^m$.

Problem (4) has a solution, since for any $N > 0$ the function J_N is continuous on the compact set $\mathcal{U}_N$ in $\mathbb{R}^{mN}$.

Now let us consider time-optimization problem. In this problem, we need to calculate the minimum number of steps $N_{\min}$ required to transfer the system (1) from the given initial state x_0 to the origin, and construct an optimal process $\{\overline{x}(k), \overline{u}(k-1)\}_{k=1}^{N_{\min}}$ satisfying the condition $\overline{x}(N_{\min}) = 0$. The number $N_{\min}$ is usually called the optimal time of system (1) with initial condition (2).

In general, time-optimization problem may have no solution. For example, let $n = m = 1$ and the system (1)–(2) have the form

$$x(k+1) = x(k) + u(k), \quad x(0) = x_0 \in \mathbb{R}, \quad u(k) \in U(k) = [-1/2^k; 1/2^k].$$

For this system, time-optimization problem is solvable precisely when $|x_0| < 2$.

In the paper, we will assume that time-optimization problem for the system (1)–(2) is solvable and that the corresponding finite value $N_{\min}$ exists. Excluding the trivial case $x_0 = 0$, we will also assume that $N_{\min} > 0$.

Since, by hypothesis, we have $0 \in U(k)$ for all $k = 0, 1, \ldots$, the optimal value $\overline{J_N}$ in problem (4) is zero if and only if $N \geqslant N_{\min}$. Accordingly, under the assumption $0 < N_{\min} < \infty$, time-optimization problem can be formulated as follows:

$$N \to \min, \quad N \in \left\{ M > 0 \;\middle|\; \overline{J_M} := \min_{u \in \mathcal{U}_M} J_M(u) = 0 \right\}. \tag{5}$$

3. BLOCK COORDINATE DESCENT

The problems (4) and (5) are related to the minimization of the function J_N on the set $\mathcal{U}_N$ in $\mathbb{R}^{mN}$. The function $J_N : \mathbb{R}^{mN} \to \mathbb{R}_+$ and the set $\mathcal{U}_N \subset \mathbb{R}^{mN}$ have a number of useful properties. In particular, $u \mapsto J_N(u)^2$ is a convex and smooth function on $\mathbb{R}^{mN}$, and the set $\mathcal{U}_N$ is compact and convex. Moreover, $\mathcal{U}_N$ can be decomposed into a Cartesian product of sets $U(0) \times \ldots \times U(N-1)$, where all $U(k)$ are also convex and compact.

Note that the dimension of the set $\mathcal{U}_N$ can be significantly higher than the dimension of each of the sets $U(k)$. Moreover, it is the sets $U(k)$ that are part of the original problem statement, while $\mathcal{U}_N$ is an artificial construction. For these reasons, minimization on sets $U(k)$ can be significantly simpler rather than minimization on the entire set $\mathcal{U}_N$. A suitable minimization algorithm that takes these features into account is the block coordinate descent method, for which the theoretical foundations of convergence were laid in [3]. We present its general formulation below.

3.1. Formulation of the Method

Let $K_1, \ldots, K_\nu$ be nonempty compact sets in finite-dimensional Euclidean spaces $V_1, \ldots, V_\nu$ respectively. Let a continuous function $f : V \to \mathbb{R}$ be defined on the Cartesian product $V := V_1 \times \ldots \times V_\nu$. Consider the problem

$$f(v) = f(v_1, \ldots, v_\nu) \to \min_{v \in K}, \quad K := K_1 \times \ldots \times K_\nu. \tag{6}$$

Fix an arbitrary point $v^{(0)} = (v_1^{(0)}, \ldots, v_\nu^{(0)}) \in K$ and write the following iterative procedure. Suppose that for some $l \geqslant 0$ the point $v^{(l)} \in K$ is given. Then we define a point $v^{(l+1)} \in K$ by the rule

$$v_i^{(l+1)} \in \operatorname{Arg}\min_{\mu \in K_i} f(v_1^{(l+1)}, \ldots, v_{i-1}^{(l+1)}, \mu, v_{i+1}^{(l)}, \ldots, v_\nu^{(l)}), \quad i = 1, \ldots, \nu. \tag{7}$$

In the case $V_1 = \ldots = V_\nu = \mathbb{R}$, rule (7) defines a step of the classical numerical Gauss–Seidel method (or coordinate descent method) for solving the problem of conditional finite-dimensional optimization of a function f depending on ν scalar variables. In the general case, the resulting

algorithm is usually called the block coordinate descent method or the generalized (nonlinear) Gauss-Seidel method [4, 5].

By construction, it is clear that for any choice of $v^{(0)} \in K$,

$$f(v^{(l+1)}) \leqslant f(v^{(l)}), \quad l = 0, 1, 2, \ldots.$$

However, under the assumptions made, there can be no convergence of the sequence of numbers $f(v^{(l)})$ to the minimum value in the problem (6). Moreover, as was shown in [3], even the convexity conditions for all sets K_i and the function f itself are insufficient. Especially, it is impossible to speak of any reasonable convergence of the sequence of elements $v^{(l)}$ to a solution in problem (6). At the same time, a number of known results guarantee this convergence in one form or another under additional assumptions.

3.2. Convergence

The main known result on the convergence of the method (7) is the following.

Theorem 1. *Let all compact sets K_i in problem* (6) *be convex, and let the continuous function f be convex on the set $K = K_1 \times \ldots \times K_\nu$. Assume, moreover, that the following conditions hold:*

(i) *function f is continuously differentiable on the set K;*

(ii) *function $\mu \mapsto f(v_1, \ldots, v_{i-1}, \mu, v_{i+1}, \ldots, v_\nu)$ has a unique minimum on the set K_i for each i and arbitrary values $v_j \in K_j$, $j \neq i$.*

Then the following is true. If the point $\overline{v}$ is a limit point for the sequence $\{v^{(l)}\}$ constructed according to the rule (7) *from an arbitrary initial point $v^{(0)} \in K$, then*

$$f(\overline{v}) = \min_{v \in K} f(v).$$

Remark 1. In the statement of Theorem 1 and throughout the text, a limit point for a sequence $\{v^{(l)}\}$ with elements in a finite-dimensional Euclidean space V is understood to be a limit point of the set $\{v^{(l)}\} \subset V$, i.e., any point $\overline{v} \in V$ for which there exists a subsequence of $\{v^{(l)}\}$ converging to it.

The validity of Theorem 1 follows from [4, Proposition 2.7.1] and [4, Proposition 2.1.2]. Note that in [3], the assertion of Theorem 1 is obtained under somewhat different assumptions. Specifically, condition (i) of Theorem 1 is weakened in [3] to local continuous differentiability of f in a neighborhood of every fixed point of the method (7), and condition (ii) is strengthened to the strict convexity of all functions $\mu \mapsto f(v_1, \ldots, v_{i-1}, \mu, v_{i+1}, \ldots, v_\nu)$.

As will be shown below, condition (ii) of Theorem 1 can be weakened, but apparently cannot be abandoned completely. The importance of condition (i) in its local version is demonstrated by the following simple example from [3].

Example 1. On the set $K = [0; 1]^2 \subset \mathbb{R}^2$, consider the function

$$f(v_1, v_2) = \begin{cases} v_1 - 2v_2, & v_2 \leqslant v_1, \\ v_2 - 2v_1, & v_1 < v_2. \end{cases}$$

In this example, the sets $K_i = [0; 1] \subset \mathbb{R}$ are convex and the function $f(v_1, v_2) = |v_1 - v_2| - \min\{v_1, v_2\}$ is convex on K, with $\operatorname{Arg\,min}_{0 \leqslant \mu \leqslant 1} f(v_1, \mu) = \{v_1\}$ and $\operatorname{Arg\,min}_{0 \leqslant \mu \leqslant 1} f(\mu, v_2) = \{v_2\}$, however, condition (i) of Theorem 1 is not satisfied in any neighborhood of any of the fixed points of the method (7) defined by the equality $v_1 = v_2$, and for any initial point $v^{(0)} = (v_1^{(0)}, v_2^{(0)})$, we have $v^{(l)} = (v_2^{(0)}, v_2^{(0)})$ and $f(v^{(l)}) = -v_2^{(0)}$ for $l \geqslant 1$. Consequently, the assertion of Theorem 1 also fails for $v_2^{(0)} \neq 1 = -\min_{v \in K} f(v)$. The example is complete.

In the context of applications of the above results to solving problems in the previous section, we will need the following minor generalization of Theorem 1, which weakens condition (ii).

Corollary 1. *Let all compact sets K_i in problem* (6) *be convex, and let the continuous function f have the form*

$$f(v) = g(\Lambda_1 v_1 + \lambda_1, \ldots, \Lambda_\nu v_\nu + \lambda_\nu),$$

where Λ_i is a linear operator from the finite-dimensional Euclidean space V_i to a finite-dimensional Euclidean space V_i', $\lambda_i \in V_i'$, $i = 1, \ldots, \nu$, $g : V_1' \times \ldots \times V_\nu' \to \mathbb{R}$. Let function g satisfy the conditions of Theorem 1 on the set

$$\Lambda K + \lambda := (\Lambda_1 K_1 + \lambda_1) \times \ldots \times (\Lambda_\nu K_\nu + \lambda_\nu).$$

Then the following is true. If the point $\overline{v}$ is a limit point for the sequence $\{v^{(l)}\}$ constructed according to the rule (7) *from an arbitrary initial point $v^{(0)} \in K$, then*

$$f(\overline{v}) = \min_{v \in K} f(v).$$

Proof. Consider the point $z^{(0)} = \Lambda v^{(0)} + \lambda$, where $\Lambda v + \lambda := (\Lambda_1 v_1 + \lambda_1, \ldots, \Lambda_\nu v_\nu + \lambda_\nu) \in \Lambda K + \lambda$ for $v \in K$. Every sequence $\{v^{(l)}\} \subset K$ constructed according to the rule (7) from the point $v^{(0)} \in K$ generates a sequence $\{z^{(l)}\} = \{\Lambda v^{(l)} + \lambda\} \subset \Lambda K + \lambda$ satisfying the rule

$$z_i^{(l+1)} = \arg \min_{\mu \in \Lambda K_i + \lambda_i} g(z_1^{(l+1)}, \ldots, z_{i-1}^{(l+1)}, \mu, z_{i+1}^{(l)}, \ldots, z_\nu^{(l)}), \quad i = 1, \ldots, \nu, \quad l \geqslant 0.$$

Since the sets $\Lambda_i K_i + \lambda_i$ are nonempty, compact and convex, and the function g on $\Lambda K + \lambda$ satisfies the conditions of Theorem 1, we have

$$g(\overline{z}) = \min_{z \in \Lambda K + \lambda} g(z)$$

for every limit point $\overline{z}$ of the sequence $\{z^{(l)}\}$. Now setting $\overline{z} = \Lambda \overline{v} + \lambda$ and using the continuity of finite-dimensional linear operators Λ_i, we obtain the desired result. Corollary 1 is proved.

We demonstrate the result of Corollary 1 with the following important example.

Example 2. Consider the square of the values of J_N from the problem (4). Let $\nu = N$, $v_i = u(i-1)$, $K_i = U(i-1)$. We obtain the function

$$f(v) = \left\| \mathcal{A}_\nu(0) x_0 + \sum_{i=1}^{\nu} \mathcal{A}_\nu(i) B(i-1) v_i \right\|^2.$$

Here $\mathcal{A}_\nu(i)$ and $B(i)$ are known matrices, x_0 is a known vector. Let $\Lambda_i = \mathcal{A}_\nu(i) B(i-1)$, $\lambda_1 = \mathcal{A}_\nu(0) x_0$, $\lambda_2 = \ldots = \lambda_\nu = 0$. Then the function f and the sets K_i satisfy the conditions of Corollary 1. However, they may not satisfy the conditions of Theorem 1. Indeed, if any of the matrices $\mathcal{A}_\nu(i)$ is singular, then the minimum in the problem $f(v_1, \ldots, v_{i-1}, \mu, v_{i+1}, \ldots, v_\nu) \to \min_{\mu \in K_i}$ may not be unique. At the same time, a function g of the form

$$g(z) = \left\| \sum_{i=1}^{\nu} z_i \right\|^2$$

is continuously differentiable on $\mathbb{R}^{n\nu}$, and each of the functions

$$\mu \mapsto g(z_1, \ldots, z_{i-1}, \mu, z_{i+1}, \ldots, z_\nu) = \left\| \mu + \sum_{j \neq i} z_j \right\|^2$$

is strictly convex on $\Lambda_i K_i + \lambda_i$ for any fixed values $z_j \in \Lambda_j K_j + \lambda_j$, $j \neq i$, where $i = 1, \ldots, \nu$. Therefore, g satisfies the conditions of Theorem 1. Consequently, for $f(v) = g(\Lambda v + \lambda)$, the assertion of Corollary 1 holds. The example is complete.

Note that the smoothness of the function $g(z)$ from Example 2 is not so important. In particular, it can be verified directly that for the function

$$f(v) = \left\| \sum_{i=1}^{\nu} v_i \right\|,$$

where the Euclidean norm can be replaced by any norm in the space $\mathbb{R}^n$, and an arbitrary set of nonempty convex compact sets $K_i \subset \mathbb{R}^n$, the block coordinate descent method for the problem $f(v) \to \min_{v \in K}$, $K = K_1 \times \ldots \times K_\nu$, converges in the same sense as specified in Theorem 1 and Corollary 1, despite the fact that condition (i) of Theorem 1 may not be satisfied even locally in a neighborhood of the solution. We will not dwell on this in detail here.

For convenience, introducing the notation

$$\overline{f} := \min_{v \in K} f(v), \quad \overline{K} := \operatorname{Arg}\min_{v \in K} f(v), \quad \operatorname{dist}(\mu, \overline{K}) := \min_{v \in \overline{K}} \|\mu - v\|,$$

we obtain the following results.

Corollary 2. *Under the conditions of Corollary 1, for any sequence $\{v^{(l)}\}$ constructed according to the rule* (7) *from an arbitrary initial point $v^{(0)} \in K$, the following convergence holds:*

$$f(v^{(l)}) \to \overline{f}, \quad \operatorname{dist}(v^{(l)}, \overline{K}) \to 0, \quad l \to \infty.$$

Proof. Suppose that $f(v^{(l)}) \not\to \overline{f}$. Then there exists a subsequence $\{v^{(l_m)}\}$ with the property $|f(v^{(l_m)}) - \overline{f}| > \delta > 0$ for some $\delta > 0$. Passing to a subsequence again if needed, we have $\{v^{(l_m)}\}$ converges to some $\overline{v} \in K$. Then, by Corollary 1, we obtain $f(\overline{v}) = \overline{f}$. But by the continuity of f on K and the assumption made, we have $|f(\overline{v}) - \overline{f}| \geqslant \delta > 0$ — a contradiction. Similar arguments yield the convergence of $\operatorname{dist}(v^{(l)}, \overline{K}) \to 0$. Corollary 2 is proved.

Corollary 3. *Let the assumptions of Corollary 1 be satisfied. Then a necessary and sufficient condition for the optimality of a point $\overline{v}$ in problem* (6) *is the relation*

$$\overline{v}_i \in \operatorname{Arg}\min_{\mu \in K_i} f(\overline{v}_1, \ldots, \overline{v}_{i-1}, \mu, \overline{v}_{i+1}, \ldots, \overline{v}_\nu), \quad i = 1, \ldots, \nu.$$

Proof. The necessity of the condition is obvious. Assume that it holds. Then, for $v^{(0)} = \overline{v}$, the sequence of points $v^{(l)} \equiv \overline{v}$ satisfies the rule (7). By Corollary 2, this means that $f(\overline{v}) = \overline{f}$, as required. Corollary 3 is proved.

To conclude this subsection, we briefly discuss the convergence of the sequence $\{v^{(l)}\}$, constructed according to the rule (7), to a single point $\overline{v} \in \overline{K}$. The classical result on the presence of such convergence under the conditions of Theorem 1 was obtained in [8] for the case $V_1 = \ldots = V_\nu = \mathbb{R}$ and $K_i = [a_i; b_i]$. It is also noted there that it naturally generalizes to the case where the sets K_i are closed parallelepipeds in $\mathbb{R}^{n_i}$. However, in the general situation, verifying this kind of convergence remains an open question.

The following example shows that in a one-dimensional problem of the form (4), the method (7) converges to a single solution in one iteration.

Example 3. Consider the problem

$$|\lambda_1 v_1 + \ldots + \lambda_\nu v_\nu + \gamma| \to \min, \quad v_i \in [a_i; b_i].$$

Let the numbers $v_i^{(0)} \in [a_i; b_i]$ be given. Then for all i such that $\lambda_i \neq 0$, the numbers $v_i^{(1)}$ are determined according to (7) from the conditions

$$v_i^{(1)} = \arg\min_{\mu \in [a_i; b_i]} |\lambda_i \mu + \sigma_i|, \quad \sigma_i = \lambda_1 v_1^{(1)} + \ldots + \lambda_{i-1} v_{i-1}^{(1)} + \lambda_{i+1} v_{i+1}^{(0)} + \ldots + \lambda_\nu v_\nu^{(0)} + \gamma.$$

For $i = 1, \dots, \nu$ in the case $\lambda_i \neq 0$ we put

$$v_i^{(1)} = \begin{cases} \max\{\lambda_i a_i, \lambda_i b_i\}/\lambda_i, & \sigma_i < -\max\{\lambda_i a_i, \lambda_i b_i\}, \\ \min\{\lambda_i a_i, \lambda_i b_i\}/\lambda_i, & \sigma_i > -\min\{\lambda_i a_i, \lambda_i b_i\}, \\ -\sigma_i/\lambda_i & \text{otherwise.} \end{cases}$$

For the remaining i, we choose $v_i^{(1)} \in [a_i; b_i]$ arbitrarily. It is directly verified that the points $v_i^{(1)}$ constructed step by step in this way satisfy the rule (7) and provide a minimum in the problem. The example is complete.

3.3. Stopping Rules

The usual stopping rule for the classical coordinate descent method is the condition

$$f(v^{(l)}) - f(v^{(l+1)}) < \varepsilon \tag{8}$$

for a pre-selected small value $\varepsilon > 0$. If the convergence of the sequence $\{v^{(l)}\}$ is proven, one can also apply the alternative condition $\|v^{(l+1)} - v^{(l)}\| < \varepsilon$. However, in the general case, such convergence does not hold for the procedure (7), so only condition (8) remains. An undoubted advantage of the condition (8) is the simplicity of its verification. At the same time, inequality (8) in the absence of additional assumptions cannot guarantee that the value $f(v^{(l+1)})$ is close to the unknown optimal value $\overline{f}$.

A more reliable stopping rule can be obtained by exploiting the following property of the class of problems under consideration. Suppose that in problem (6) the sets K_i are convex, and the function f is continuously differentiable and convex on K (i.e., for example, the convergence conditions from Theorem 1 or Corollary 1 are satisfied). Then for any $w \in K$ we have the two-sided estimate

$$f(w) \geqslant \overline{f} \geqslant f(w) + \min_{w+h \in K} \langle f'(w), h \rangle = f(w) - \sum_{i=1}^{\nu} \max_{\mu \in K_i} \left\langle -\frac{\partial f(w)}{\partial v_i}, \mu - w_i \right\rangle.$$

Indeed, the first inequality follows from the definition of the minimum value $\overline{f}$, the second inequality follows from the fact that the epigraph of a proper convex function f is located above any of its supporting and, in particular, tangent hyperplanes, and the last equality follows from the decomposition

$$\langle f'(w), h \rangle = \sum_{i=1}^{\nu} \left\langle \frac{\partial f(w)}{\partial v_i}, h_i \right\rangle.$$

Now we can take as a stopping rule the condition

$$\sum_{i=1}^{\nu} \max_{\mu \in K_i} \left\langle \frac{\partial f(v^{(l)})}{\partial v_i}, v_i^{(l)} - \mu \right\rangle < \varepsilon. \tag{9}$$

The fulfillment of this condition (9) guarantees the validity of the inequality $f(v^{(l)}) < \overline{f} + \varepsilon$ for a predetermined error value $\varepsilon > 0$. Moreover, maximization in (9) is performed over the same sets K_i that appear in (7). Therefore, the stopping rule (9) is consistent with the method (7).

Note also that the equality

$$\max_{v \in K} \langle f'(\overline{v}), \overline{v} - v \rangle = \sum_{i=1}^{\nu} \max_{\mu \in K_i} \left\langle \frac{\partial f(\overline{v})}{\partial v_i}, \overline{v}_i - \mu \right\rangle = 0$$

coincides with the classical necessary and sufficient condition for the optimality of a point $\overline{v} \in K$ in a smooth and convex problem (6) [10, Theorem 7.1.3].

The stopping rules (8) and (9) can be used together. The computationally simpler condition (8) can be checked first, and then, if it holds, we check the guaranteeing condition (9).

Finally, the two-sided estimate presented above allows us to refine the interval of values to which the number $\overline{f}$ belongs at each step when constructing the sequence $v^{(l)}$. Namely, for every $l \in \{0, 1, \dots\}$, holds the inclusion

$$\overline{f} \in \left[f(v^{(l)}) - \sum_{i=1}^{\nu} \max_{\mu \in K_i} \left\langle \frac{\partial f(v^{(l)})}{\partial v_i}, v_i^{(l)} - \mu \right\rangle ; f(v^{(l)}) \right]. \tag{10}$$

If the conditions of Theorem 1 or Corollary 1 are satisfied, the boundaries of the intervals specified here converge, as l increases, to the point $\overline{f}$. Furthermore, this interval estimate can provide an early answer to the question of what values the unknown $\overline{f}$ can and cannot take. The latter is relevant, for example, when considering the previously posed time-optimization problem (5), in which the value $\overline{f} = 0$ plays a key role.

We now turn directly to solving the problems (4) and (5).

4. PROBLEM SOLVING

4.1. Terminal Stabilization Problem

We apply the block coordinate descent method to solve problem (4). As noted in Example 2, problem (4) satisfies the conditions of Corollary 1, which, thanks to Corollary 2, guarantees the convergence of the objective values obtained by method (7) to minimum, as well as the convergence of the control sequence constructed by it to the set of solutions. As the stopping rules we can use (8) and (9). It remains to specify the formulas (7)–(9).

In problem (4) we have $\nu = N$, $K_i = U(i-1)$, $V_i = \mathbb{R}^m$, $f = J_N^2$.

Formula (7) takes the form

$$u^{(l+1)}(k) \in \operatorname{Arg}\min_{\mu \in U(k)} \Bigg\| \mathcal{A}_N(0)x_0 + \sum_{i=0}^{k-1} \mathcal{A}_N(i+1)B(i)u^{(l+1)}(i) + \mathcal{A}_N(k+1)B(k)\mu + \sum_{i=k+1}^{N-1} \mathcal{A}_N(i+1)B(i)u^{(l)}(i) \Bigg\|, \quad k = 0, \dots, N-1.$$

By the definition of the matrix $\mathcal{A}_N(k)$, we have

$$\mathcal{A}_N(0)x_0 + \sum_{i=0}^{k-1} \mathcal{A}_N(i+1)B(i)u^{(l+1)}(i) = \mathcal{A}_N(k) \left(\mathcal{A}_k(0)x_0 + \sum_{i=0}^{k-1} \mathcal{A}_k(i+1)B(i)u^{(l+1)}(i) \right).$$

Hence, using the values $x^{(l)}(k)$ and $\xi^{(l)}(k)$, calculated by the formulas

$$x^{(l)}(k+1) = A(k)x^{(l)}(k) + B(k)u^{(l)}(k), \quad x^{(l)}(0) = x_0,$$

$$\xi^{(l)}(k) = \xi^{(l)}(k+1) - \mathcal{A}_N(k+1)B(k)u^{(l)}(k), \quad \xi^{(l)}(N) = 0,$$

we find that method (7) in problem (4) is written as

$$u^{(l+1)}(k) \in \operatorname{Arg}\min_{\mu \in U(k)} \left\| \mathcal{A}_N(k+1)B(k)\mu + \mathcal{A}_N(k)x^{(l+1)}(k) - \xi^{(l)}(k+1) \right\| \tag{11}$$

for $k = 0, \dots, N-1$. Here $x^{(l+1)}(k)$ are determined by the values $u^{(l+1)}(i)$, $i < k$, and $\xi^{(l)}(k+1)$ are determined by the values $u^{(l)}(i)$, $i > k$, so the formula (11) correctly implements method (7).

Formula (8) takes the form

$$\|x^{(l)}(N)\|^2 - \|x^{(l+1)}(N)\|^2 < \varepsilon. \tag{12}$$

Finally, we write formula (9) using the classical dual variable $\psi(k)$ in problem (1)–(3). Let $\psi^{(l)}(k)$ satisfy the equalities

$$\psi^{(l)}(k) = A(k)^{\mathrm{T}}\psi^{(l)}(k+1), \quad \psi^{(l)}(N) = -2x^{(l)}(N).$$

Then for $\mu \in \mathbb{R}^m$ we have

$$\begin{aligned}\left\langle \frac{\partial f(v^{(l)})}{\partial v_{k+1}}, \mu \right\rangle &= \left\langle \frac{\partial J(u^{(l)})^2}{\partial u(k)}, \mu \right\rangle \\ &= 2\left\langle \mathcal{A}_N(0)x_0 + \sum_{i=0}^{N-1} \mathcal{A}_N(i+1)B(i)u^{(l)}(i), \mathcal{A}_N(k+1)B(k)\mu \right\rangle \\ &= \langle 2x^{(l)}(N), \mathcal{A}_N(k+1)B(k)\mu\rangle = -\langle \mathcal{A}_N(k+1)^{\mathrm{T}}\psi^{(l)}(N), B(k)\mu\rangle = -\langle \psi^{(l)}(k), B(k)\mu\rangle.\end{aligned}$$

Therefore, the condition (9) takes the form

$$\sum_{k=0}^{N-1} \max_{\mu\in U(k)} \left\langle \psi^{(l)}(k), B(k)(\mu - u^{(l)}(k)) \right\rangle < \varepsilon. \tag{13}$$

Let us consider two ways of constructing the algorithm. The first is to use method (7) with the stopping rules (12) and (13) without any additional transformations.

Algorithm 1.

0. Set the permissible computational error $\varepsilon > 0$.
1. Put $l = 0$, define the initial approximation by the equalities $u^{(0)}(k) = 0 \in \mathbb{R}^m$, $k = 0, \dots, N-1$, calculate the matrices $\mathcal{A}_N(k) = A(N-1)\dots A(k)$, $k = 0, \dots, N-1$, put the matrix $\mathcal{A}_N(N)$ to be the identity matrix of size $n \times n$.
2. Calculate the values $\xi^{(l)}(k)$, $k = 1, \dots, N$, by the formulas

$$\xi^{(l)}(k) = \xi^{(l)}(k+1) - \mathcal{A}_N(k+1)B(k)u^{(l)}(k), \quad \xi^{(l)}(N) = 0.$$

3. For each $k = 0, \dots, N-1$, sequentially find and fix some solution $u^{(l+1)}(k)$ to the extremum problem

$$\|\mathcal{A}_N(k+1)B(k)\mu + \mathcal{A}_N(k)x^{(l+1)}(k) - \xi^{(l)}(k+1)\| \to \min_{\mu\in U(k)},$$

where the values $x^{(l+1)}(k)$ are calculated step by step using the formulas

$$x^{(l+1)}(k+1) = A(k)x^{(l+1)}(k) + B(k)u^{(l+1)}(k), \quad x^{(l+1)}(0) = x_0.$$

4. Check the first stopping condition

$$\|x^{(l)}(N)\|^2 - \|x^{(l+1)}(N)\|^2 < \varepsilon,$$

if true, go to step 5; otherwise, increment l by one and go to step 2.
5. Calculate the values $\psi^{(l+1)}(k)$, $k = 1, \dots, N$, by the formulas

$$\psi^{(l+1)}(k) = A(k)^{\mathrm{T}}\psi^{(l+1)}(k+1), \quad \psi^{(l+1)}(N) = -2x^{(l+1)}(N).$$

6. Calculate the value $D_N^{(l+1)}$ by the formula

$$D_N^{(l+1)} = \sum_{k=0}^{N-1} \max_{\mu \in U(k)} \langle \psi^{(l+1)}(k+1), B(k)(\mu - u^{(l+1)}(k)) \rangle.$$

7. Check the second stopping condition

$$D_N^{(l+1)} < \varepsilon,$$

if true, terminate the calculations with the answer $u_\varepsilon = u^{(l+1)}$; otherwise, increase l by one and go to step 2.

The convergence of Algorithm 1 follows from Theorem 1 and its Corollaries 1–3. Steps 0–4 were used in [1] for finding an approximate solution to time-optimization problem (5) for $m = n$, $A(k) \equiv A$, $\det A \neq 0$, $B(k) \equiv I$ and a known two-sided estimate on optimal time.

The second way of constructing the algorithm involves changing the variables in the problem, which simplifies the iterations from (7). However, to find the desired solution based on the results of applying this method, it is necessary to perform the inverse transformation, which is not uniquely defined and leads to a separate subproblem.

The main idea here is to pass from system (1)–(2) to a linear system of the simplest form

$$y(k+1) = y(k) + \theta(k), \quad y(0) = y_0$$

through the degenerate change of variables $\theta(k) = \mathcal{A}_N(k+1)B(k)u(k)$, $y_0 = \mathcal{A}_N(0)x_0$. In this case, $\theta(k) \in \Theta(k) := \mathcal{A}_N(k+1)B(k)U(k) \subset \mathbb{R}^n$ and $y(k) = \mathcal{A}_N(k)x(k)$. In particular, $y(N) = x(N)$. From the latter it follows that if $\overline{\theta}(k) \in \Theta(k)$ minimize the value $\|y(N)\|$, then any element $\overline{u} \in \mathcal{U}_N$ satisfying the system of equations

$$\mathcal{A}_N(k+1)B(k)\overline{u}(k) = \overline{\theta}(k), \quad k = 0, \ldots, N-1, \tag{14}$$

is a solution to problem (4). If $m = n$ and for some k the matrix $\mathcal{A}_N(k+1)B(k)$ is invertible, then $\overline{u}(k)$ is uniquely determined. In the general case, we are faced with a system of equations (14) with constraints $\overline{u}(k) \in U(k)$. By construction, this problem is solvable. The values $\overline{\theta}(k)$ can be found using algorithm 1, since the sets $\Theta(k)$ are convex, compact, and contain the origin in $\mathbb{R}^n$. The difference is that $\overline{\theta}(k)$ are uniquely determined in this case.

Algorithm 2.

0. Set the permissible computational error $\varepsilon > 0$.
1. Define the initial approximation by the equalities $\theta^{(0)}(k) = 0 \in \mathbb{R}^n$, $k = 0, \ldots, N-1$, put $l = 0$ and, given the sets $U(k)$, determine the sets $\Theta(k) = \mathcal{A}_N(k+1)B(k)U(k)$, where $\mathcal{A}_N(k) = A(N-1)\ldots A(k)$, $k = 0, \ldots, N-1$, $\mathcal{A}_N(N)$ is the identity $n \times n$-matrix.
2. Calculate the values $\eta^{(l)}(k)$, $k = 1, \ldots, N$, by the formulas

$$\eta^{(l)}(k) = \eta^{(l)}(k+1) - \theta^{(l)}(k), \quad \eta^{(l)}(N) = 0.$$

3. For each $k = 0, \ldots, N-1$, sequentially find and fix the unique solution $\theta^{(l+1)}(k)$ to the extremum problem

$$\|\mu + y^{(l+1)}(k) - \eta^{(l)}(k+1)\| \to \min_{\mu \in \Theta(k)},$$

where the values $y^{(l+1)}(k)$ are calculated step by step using the formulas

$$y^{(l+1)}(k+1) = y^{(l+1)}(k) + \theta^{(l+1)}(k), \quad k = 0, \ldots, N-1, \quad y^{(l+1)}(0) = \mathcal{A}_N(0)x_0.$$

4. Check the first stopping condition

$$\|y^{(l)}(N)\|^2 - \|y^{(l+1)}(N)\|^2 < \varepsilon,$$

if true, go to step 5; otherwise, increment l by one and go to step 2.

5. Calculate the value $D_N^{(l+1)}$ by the formula

$$D_N^{(l+1)} = 2\sum_{k=0}^{N-1} \max_{\mu\in\Theta(k)} \langle y^{(l+1)}(N), \theta^{(l+1)}(k) - \mu\rangle.$$

6. Check the second stopping condition

$$D_N^{(l+1)} < \varepsilon,$$

if true, terminate the calculations and determine the answer u_ε by solving the system

$$\mathcal{A}_N(k+1)B(k)u_\varepsilon(k) = \theta^{(l+1)}(k), \quad u_\varepsilon(k) \in U(k), \quad k = 0, \dots, N-1;$$

otherwise, increase l by one and go to step 2.

Although the inverse change at step 6 of Algorithm 2 is a separate problem, it is actually implicitly present in Algorithm 1 as well, and must be solved at each iteration. It is natural to move this kind of nontrivial transformation outside the iterative process. At the same time, applying Algorithm 2 requires constructing transformed sets $\Theta(k)$, which in itself can be challenging.

4.2. Time-Optimization Problem

As the main algorithm for solving time-optimization problem (5), we will consider the sequential application of method (7) to construct control improvements in problem (4) for $N = 1, 2, \dots$ and the complementary calculation of the interval estimate, ensuring an effective transition to the next value of N. Let us just clarify that the estimate (10) in accordance with the above specification of the formulas (7)–(9) now takes the form

$$\overline{J}_N^2 \in \left[J_N(u^{(l)})^2 - \sum_{k=0}^{N-1} \max_{u\in U(k)} \langle \psi^{(l)}(k+1), B(k)(u - u^{(l)}(k))\rangle; J_N(u^{(l)})^2 \right].$$

Algorithm 3.

0. Set the permissible computational error $\varepsilon > 0$, put $N = 1$.
1. Put $l = 0$, define/supplement the initial approximation with the equality $u^{(0)}(N-1) = 0 \in \mathbb{R}^n$ and calculate the matrices $\mathcal{A}_N(k) = A(N-1)\dots A(k)$, $k = 0, \dots, N-1$, put the matrix $\mathcal{A}_N(N)$ to be the identity matrix of size $n \times n$.
2. Calculate the values $\xi^{(l)}(k)$, $k = 1, \dots, N$, by the formulas

$$\xi^{(l)}(k) = \xi^{(l)}(k+1) - \mathcal{A}_N(k+1)B(k)u^{(l)}(k), \quad \xi^{(l)}(N) = 0.$$

3. For each $k = 0, \dots, N-1$, sequentially find and fix some solution $u^{(l+1)}(k)$ to the extremum problem

$$\|\mathcal{A}_N(k+1)B(k)\mu + \mathcal{A}_N(k)x^{(l+1)}(k) - \xi^{(l)}(k+1)\| \to \min_{\mu\in U(k)},$$

where the values $x^{(l+1)}(k)$ are calculated step by step using the formulas

$$x^{(l+1)}(k+1) = A(k)x^{(l+1)}(k) + B(k)u^{(l+1)}(k), \quad x^{(l+1)}(0) = x_0.$$

4. Check the external stopping condition

$$\|x^{(l+1)}(N)\|^2 < \varepsilon,$$

if true, terminate the calculations with the answer $N_\varepsilon = N$, $u_\varepsilon = u^{(l+1)}$.

5. Calculate the values $\psi^{(l+1)}(k)$, $k = 1, \dots, N$, by the formulas

$$\psi^{(l+1)}(k) = A(k)^{\mathrm{T}}\psi^{(l+1)}(k+1), \quad \psi^{(l+1)}(N) = -2x^{(l+1)}(N).$$

6. Calculate the estimate $E_N^{(l+1)}$ by the formula

$$E_N^{(l+1)} = \|x^{(l+1)}(N)\|^2 - \sum_{k=0}^{N-1} \max_{\mu \in U(k)} \langle \psi^{(l+1)}(k+1), B(k)(\mu - u^{(l+1)}(k)) \rangle.$$

7. Check the internal stopping condition

$$E_N^{(l+1)} > 0,$$

if true, put the found $u^{(l+1)}$ as the new initial approximation, i.e. put $u^{(0)}(k) = u^{(l+1)}(k)$, $k = 0, \dots, N-1$, increment N by one and go to step 1; otherwise increment l by one and go to step 2.

Algorithm 3 is identical to the algorithm for solving time-optimization problem proposed in [2] (to simplify the calculations, in [2] the case $m = n$ and $B(k) \equiv I$ was considered). From the construction, it is clear that the convergence of this algorithm established in [2, Theorem 5], as well as the related results on optimality and improvement (see [2, Theorems 1–4]), follow in the general case from the well-known Theorem 1 and its Corollaries 1–3.

Note that the change of variables discussed in the previous subsection is less relevant here. This is because, for every N, except for large values close to $N_{\min}$, due to the estimate-based stopping conditions, only a small number of transformed problems are solved, and switching to new variables does not provide any significant speedup.

As an alternative approach to solving time-optimization problem, we consider the following algorithm based on the results of [6, 7]. We will apply the numerical method from [6] to calculate the optimal time $N_{\min}$ and then solve the remaining terminal stabilization problem as in the previous subsection. For completeness, we present here the formulation and brief justification of the method for calculating $N_{\min}$ from [6, 7].

For an arbitrary compact set K in $\mathbb{R}^n$, consider its support function

$$s_K(p) := \max_{\mu \in K} \langle p, \mu \rangle.$$

First of all, we consider the support functions of the sets $U(k)$. Since $U(k)$ are convex and contain the origin, $s_{U(k)}$ take finite non-negative values, are convex, continuous, and positively homogeneous [11, Corollary 13.2.2]. Since all sets $U(k)$ are predetermined, the functions $s_{U(k)}$ can, with certain reservations, be considered as known.

Consider the reachability set X_N of the system (1)–(2) at time $N \geqslant 0$. Using the matrices $\mathcal{A}_N(k)$ and the Minkowski set addition operation, it can be represented as

$$X_N = \mathcal{A}_N(0)x_0 + \sum_{k=0}^{N-1} \mathcal{A}_N(k+1)B(k)U(k).$$

The set X_N is convex and compact, $\mathcal{A}_N(0)x_0 \in X_N$. Since $0 \in U(k)$, then

$$0 \in X_N \quad \Leftrightarrow \quad N \geqslant N_{\min}.$$

The support function of the set X_N has the form

$$s_{X_N}(z) = \max_{x\in X_N}\langle z,x\rangle = \max_{u\in\mathcal{U}_N}\left\langle z, \mathcal{A}_N(0)x_0 + \sum_{k=0}^{N-1}\mathcal{A}_N(k+1)B(k)u(k)\right\rangle$$
$$= \langle z,\mathcal{A}_N(0)x_0\rangle + \sum_{k=0}^{N-1}\max_{\mu\in U(k)}\langle z,\mathcal{A}_N(k+1)B(k)\mu\rangle$$
$$= \langle z,\mathcal{A}_N(0)x_0\rangle + \sum_{k=0}^{N-1} s_{U(k)}\left(B(k)^{\mathrm{T}}\mathcal{A}_N(k+1)^{\mathrm{T}}z\right).$$

The function $s_{X_N}:\mathbb{R}^n\to\mathbb{R}$ is convex, continuous and positively homogeneous, and

$$s_{X_N}(z)\geqslant 0\ \ \forall z\in\mathbb{B}^n \quad\Leftrightarrow\quad 0\in X_N \quad\Leftrightarrow\quad N\geqslant N_{\min},$$

where $\mathbb{B}^n$ denotes the closed unit ball in $\mathbb{R}^n$. Therefore, the problem of finding the optimal time for the system (1)–(2) can be reformulated in terms of known functions $s_{U(k)}$ and given point x_0:

$$N\to\min,\quad \min_{z\in\mathbb{B}^n}\left(\langle z,\mathcal{A}_N(0)x_0\rangle + \sum_{k=0}^{N-1} s_{U(k)}\left(B(k)^{\mathrm{T}}\mathcal{A}_N(k+1)^{\mathrm{T}}z\right)\right) = 0. \tag{15}$$

Thus, the question of calculating the optimal time for system (1) with initial condition (2) is reduced to solving a sequence of convex finite-dimensional optimization problems of the same dimension n. Since all functions $s_{U(k)}$ take non-negative values, are continuous, and are equal to zero at the origin, the minimization can be performed on a smaller convex and compact set $\mathbb{B}^n_- := \{z\in\mathbb{B}^n : \langle z,\mathcal{A}_N(0)x_0\rangle\leqslant 0\}$. Moreover, due to the positive homogeneity of the functions $s_{U(k)}$, the solution to problem (15) will not change if $\mathbb{B}^n$ is replaced by the unit sphere $\mathbb{S}^n$ or the compact set $\mathbb{S}^n_- := \{z\in\mathbb{S}^n : \langle z,\mathcal{A}_N(0)x_0\rangle\leqslant 0\}$. The latter circumstance is used in [6] to construct an optimal process. Here, for these purposes, we will apply the block coordinate descent method.

So, another way to solve time-optimization problem is to directly determine the value $N_{\min}$ by solving problem (15), and then find an optimal control using method (7) for $N = N_{\min}$. Note that with this approach, the above-mentioned change of variables can also be effective, i.e., we are talking about the possibility of using both Algorithms 1 and 2. Moreover, there is no need to use the stopping rules (12) and (13). Instead, the calculations are terminated when the required proximity of $\|x(N)\|$ to zero is reached. At the same time, it should be noted that this approach can lead to significant difficulties in those frequent cases when the support functions $s_{U(k)}$ cannot be constructed explicitly and the optimization problems in (15) become minimax.

As a result, we obtain the following two variations of the algorithm.

Algorithm 4.

0. Set the permissible computational error $\varepsilon > 0$, put $N = 1$.
1. Calculate matrices $\mathcal{A}_N(k) = A(N-1)\dots A(k)$, $k = 0,\dots,N-1$, put the matrix $\mathcal{A}_N(N)$ to be the identity matrix of size $n\times n$.
2. Calculate the value

$$\overline{S}_N = \min_{z\in\mathbb{B}^n}\left(\langle z,\mathcal{A}_N(0)x_0\rangle + \sum_{k=0}^{N-1} s_{U(k)}\left(B(k)^{\mathrm{T}}\mathcal{A}_N(k+1)^{\mathrm{T}}z\right)\right).$$

3. Check the condition $\overline{S}_N < 0$. If true, increase N by one and go to step 1. Otherwise, go to step 4.

4. Set the initial approximation to $u^{(0)}(k) = 0 \in \mathbb{R}^m$, $k = 0, \dots, N-1$, set $l = 0$.
5. Calculate the values $\xi^{(l)}(k)$, $k = 1, \dots, N$, by the formulas

$$\xi^{(l)}(k) = \xi^{(l)}(k+1) - \mathcal{A}_N(k+1)B(k)u^{(l)}(k), \quad \xi^{(l)}(N) = 0.$$

6. For each $k = 0, \dots, N-1$, sequentially find and fix some solution $u^{(l+1)}(k)$ to the extremum problem

$$\|\mathcal{A}_N(k+1)B(k)\mu + \mathcal{A}_N(k)x^{(l+1)}(k) - \xi^{(l)}(k+1)\| \to \min_{\mu \in U(k)},$$

where the values $x^{(l+1)}(k)$ are calculated step by step using the formulas

$$x^{(l+1)}(k+1) = A(k)x^{(l+1)}(k) + B(k)u^{(l+1)}(k), \quad x^{(l+1)}(0) = x_0.$$

7. Check stopping condition

$$\|x^{(l+1)}(N)\|^2 < \varepsilon,$$

if true, terminate the calculations with the answer $N_\varepsilon = N$, $u_\varepsilon = u^{(l+1)}$; otherwise, increase l by one and go to step 2.

Algorithm 5.

0. Set the permissible computational error $\varepsilon > 0$, put $N = 1$.
1. Calculate matrices $\mathcal{A}_N(k) = A(N-1)\dots A(k)$, $k = 0, \dots, N-1$, put the matrix $\mathcal{A}_N(N)$ to be the identity matrix of size $n \times n$.
2. Calculate the value

$$\overline{S}_N = \min_{z \in \mathbb{B}^n} \left(\langle z, \mathcal{A}_N(0)x_0 \rangle + \sum_{k=0}^{N-1} s_{U(k)} \left(B(k)^{\mathrm{T}} \mathcal{A}_N(k+1)^{\mathrm{T}} z \right) \right).$$

3. Check the condition $\overline{S}_N < 0$. If true, increase N by one and go to step 1. Otherwise, go to step 4.
4. Define the initial approximation with the equalities $\theta^{(0)}(k) = 0 \in \mathbb{R}^n$, $k = 0, \dots, N-1$, put $l = 0$ and, given the sets $U(k)$, determine the sets $\Theta(k) = \mathcal{A}_N(k+1)B(k)U(k)$.
5. Calculate the values $\eta^{(l)}(k)$, $k = 1, \dots, N$, by the formulas

$$\eta^{(l)}(k) = \eta^{(l)}(k+1) - \theta^{(l)}(k), \quad \eta^{(l)}(N) = 0.$$

6. For each $k = 0, \dots, N-1$, sequentially find and fix the unique solution $\theta^{(l+1)}(k)$ to the extremum problem

$$\|\mu + y^{(l+1)}(k) - \eta^{(l)}(k+1)\| \to \min_{\mu \in \Theta(k)},$$

where the values $y^{(l+1)}(k)$ are calculated step by step using the formulas

$$y^{(l+1)}(k+1) = y^{(l+1)}(k) + \theta^{(l+1)}(k), \quad k = 0, \dots, N-1, \quad y^{(l+1)}(0) = \mathcal{A}_N(0)x_0.$$

7. Check stopping condition

$$\|y^{(l+1)}(N)\|^2 < \varepsilon,$$

if true, terminate the calculations with the answer $N_\varepsilon = N$ and u_ε found from solving the system

$$\mathcal{A}_N(k+1)B(k)u_\varepsilon(k) = \theta^{(l+1)}(k), \quad u_\varepsilon(k) \in U(k), \quad k = 0, \dots, N-1;$$

otherwise, increase l by one and go to step 2.

Note that, by construction, applying Algorithms 4 and 5 yields the exact equality $N_\varepsilon = N_{\min}$ for any value of $\varepsilon > 0$.

5. CONCLUSION

The use of the algorithms constructed in this paper appears to be relevant in problems with non-strictly convex sets of geometric constraints, with constraints described by mixed systems of inequalities, and in problems with singular matrices. In other situations, the most efficient way should be using classical finite-dimensional numerical optimization algorithms, as well as the above-mentioned method based on the support function framework from [6].

REFERENCES

1. Ibragimov, D.N. and Tsarkov, K.A., On an Approach to Solving the Time-Optimization Problem for Linear Discrete-Time Systems Based on Krotov Method, *Autom. Remote Control*, 2024, vol. 85, no. 11, pp. 1053–1078.
2. Ibragimov, D.N. and Tsarkov, K.A., Numerical Method for Solving the Time-Optimization Problem for Linear Non-Stationary Discrete-Time Systems of General Form, *Autom. Remote Control*, 2025, vol. 86, no. 11, pp. 989–1014.
3. Warga, J., Minimizing Certain Convex Functions, *J. Soc. Indust. Appl. Math.*, 1963, vol. 11, no. 3, pp. 588–593.
4. Bertsekas, D.P., *Nonlinear Programming*, Athena Scientific, 1999.
5. Peng, L. and Vidal, R., Block Coordinate Descent on Smooth Manifolds: Convergence Theory and Twenty-One Examples, *arxiv.org*, 2023 (arXiv:2305.14744).
6. Bortakovskii, A.S., Applying Support Functions to Solve a Problem of Fast-Acting Discrete Linear Systems with a Convex Set of Controls, *J. Comp. Sys. Sci. Int.*, 2025, vol. 64, no. 2, pp. 203–220.
7. Ibragimov, D.N., On the Optimal Speed Problem for the Class of Linear Autonomous Infinite-Dimensional Discrete-Time Systems with Bounded Control and Degenerate Operator, *Autom. Remote Control*, 2019, vol. 80, no. 3, pp. 393–412.
8. Luo, Z. and Tseng, P., On the Convergence of the Coordinate Descent Method for Convex Differentiable Minimization, *J. Optim. Theory Appl.*, 1992, vol. 72, no. 1, pp. 7–35.
9. Panteleev, A.V. and Letova T.A., *Metody optimizatsii v primerakh i zadachakh* (Optimization Methods in Examples and Problems), St. Petersburg: Lan, 2022.
10. Polyak, B.T., *Vvedenie v optimizatsiyu* (Introduction to Optimization), Moscow: Nauka, 1983.
11. Rockafellar, R.T., *Convex Analysis*, Princeton University Press, 1970.

This paper was recommended for publication by A.I. Kibzun, *a member of the Editorial Board*