Implementation of Infinite Dimensional Interior Point Method for Solving Multi-criteria Linear Quadratic Control Problem

L. Faybusovich [*]
Department of Mathematics,
University of Notre Dame, Notre Dame, IN 46556 USA.

T. Mouktonglang [†]
Department of Mathematics,
University of Notre Dame, Notre Dame, IN 46556 USA.

T. Tsuchiya [‡]
The Institute of Statistical Mathematics,
4-6-7 Minami-Azabu, Minato-Ku, Tokyo, 106-8569, Japan.

11 May 2004

Abstract

We describe an implementation of an infinite-dimensional primal-dual algorithm based on the Nesterov-Todd direction. Several applications to both continuous and discrete-time multi-criteria linear-quadratic control problems and linear-quadratic control problem with quadratic constraints are described. Numerical results show a very fast convergence (typically, within 3-4 iterations) to optimal solutions.

1 Introduction

One of the current trends in the development of interior-point algorithms of optimization is the extension of the domain of their applicability. While at first the linear programming problem was the major target, subsequent developments led to creation of algorithms and software for much broader class of symmetric programming problems. This class includes semidefinite programming and second-order cone programming problems. See e.g. [2] for numerous

engineering applications. Of course, these extensions are not without a price. While the general theory (very similar to the linear programming case) has been developed and cast in a very elegant form with the help of the technique of Euclidean Jordan algebras (see e.g. [1, 3, 4, 5, 6, 13, 14, 16, 18]), the computational cost of implementation grew quite significantly (mostly due to the cost of implementation of the major computational "Newton-like" step). In the present paper we make the next natural step along this road. We consider a class of infinite-dimensional optimization problems and a primal-dual algorithm based on the Nesterov-Todd direction. We implemented this algorithm for a class of control problems: multi-criteria linear-quadratic control problem and linear-quadratic control problem with quadratic constraints (in both continuous and discrete-time settings). A computation of Nesterov-Todd direction is an infinite-dimensional (linear) problem. Thus, the cost of the major computational step is even higher in comparison with previous essentially finite-dimensional schemes. What is crucial here is a careful consideration of the structure of the problem. In particular, in control applications mentioned above "the infinite-dimensional part" of the Newton-like step is reduced to solving the linear-quadratic control problem with extra linear term in the cost function. As we show in this paper, the presence of this extra linear term does not lead to any significant complications and (as in the standard control-theoretic setting) the solution of this problem is essentially reduced to solving one of the standard Riccati equations (differential, difference or algebraic depending on the concrete setting of the problem). The conceptual difference in implementation of the algorithm for continuous-time and discrete-time systems is quite minimal: it is essentially in the organization of data, different schemes for computation of scalar products in various functional spaces, and solving the linear-quadratic problem briefly described above.

A "Newton-like" step for continuous-time problem is computationally expensive operation. Good news here is that it is much cheaper for discrete-time systems. Combining this with an experimental fact (which can be partially explained by t heoretical complexity estimates [10]) that three-four iterations of the algorithm lead to very reasonable approximations of optimal solutions, one can hope t hat on-line versions of the algorithm can be implemented. This, in turn, may lead to entirely new and exciting methodologies for the multi-criteria regulation of control systems.

2   Primal-dual interior-point algorithms

In this section, following [10], we describe a general set up for a class of infinite-dimensional optimization problems. We then proceed with the optimality criterion and describe a primal-dual algorithm based on Nesterov-Todd (NT) direction.

Let $(H, <, >)$ be a Hilbert space, V_1 be a vector space: $V_1 = \mathbb{R} \times H$. Let, further, $V = V_1 \times \cdots \times V_1$ (m-times), $X \subset V$ be a closed vector subspace in V.

Consider the following standard optimization problem:

$$< a, z >_V \to \min \qquad (2.1)$$

$$z \in (b + X) \cap \bar{\Omega} \qquad (2.2)$$

and its dual

$$< b, w >_V \to \min \qquad (2.3)$$

$$b \in (a + X^\perp) \cap \bar{\Omega}^* \qquad (2.4)$$

where our cone Ω_1 is a second order cone(open convex set in V_1):

$$\Omega_1 = \{(s, y) \in \mathbb{R} \times H : s > \|y\|_H\}$$

(sometimes also called the cone of squares) and

$$\Omega = \Omega_1 \times \cdots \times \Omega_1 \ \ (\text{m-times}).$$

It is known that

$$\bar{\Omega}_1 = \{(s, y) \in \mathbb{R} \times H : s \geq \|y\|_H\}.$$

Suppose $z, w \in V$,

$$z = \{(s_1, x_1), \cdots, (s_m, x_m)\}$$

$$w = \{(t_1, y_1), \cdots, (t_m, y_m)\}.$$

Then we define an inner product on V as follows:

$$< z, w >_V = \sum_{i=1}^m (s_i t_i + < x_i, y_i >_H).$$

The cone Ω is self-dual, i.e.

$$\bar{\Omega}^* = \{z \in V :< w, z >\geq 0, \ \forall w \in \bar{\Omega}\} = \bar{\Omega}.$$

Denoted by $X^\perp$ the orthogonal complement of X in V with respect to $<, >_V$. Let, further,

$$\mathcal{F} = [(x + X) \cap \bar{\Omega}] \times [(a + X^\perp) \cap \bar{\Omega}^*].$$

We assume that

$$\text{int}(\mathcal{F}) = [(x + X) \cap \text{int}(\bar{\Omega})] \times [(a + X^\perp) \cap \text{int}(\bar{\Omega}^*)] \neq 0. \qquad (2.5)$$

It is easy to see that if a pair of $\tilde{z}, \tilde{w}$ satisfies (2.2),(2.4) and

$$< \tilde{z}, \tilde{w} >= 0,$$

then $\tilde{z}$ is an optimal solution to (2.1)-(2.2) and $\tilde{w}$ is an optimal solution to (2.3)-(2.4).

Definition 2.1. *Given (z, w) then the duality gap $\mu(z, w)$ is defined as follows:*

$$\mu(z, w) = \frac{<z, w>}{2m}$$

where r is a positive constant.

A typical primal-dual algorithm generates a sequence of pairs primal and dual feasible points $(z^{(k)}, w^{(k)}) \in int(\mathcal{F}), k = 1, 2, \ldots$ such that

$$\mu(z^{(k+1)}, w^{(k+1)}) \leq (1 - \frac{\delta}{r^\omega})\mu(z^{(k)}, w^{(k)}), \tag{2.6}$$

for some positive δ and ω (see e.g. [19]).

The following theorem (see [10]) provides necessary and sufficient conditions for a pair (z^*, w^*) to be an optimal solution to (2.1),(2.2) and (2.3),(2.4).

Theorem 2.2. *Suppose the conditions (2.5) holds. Then both (2.1),(2.2) and (2.3),(2.4) have optimal solutions. Moreover, z^* is an optimal solution to(2.1),(2.2) and w^* is an optimal solution to (2.3),(2.4) if and only if*

$$<z^*, w^*>_V = 0.$$

One of the most important steps in the implementation of primal-dual algorithms is a computation of a descent direction which drives duality gap μ to zero.

One of these directions (the so-called NT-direction) is described below. Let $z_i = (t_i, x_i), w_i = (s_i, y_i) \in V_1$. We define a determinant of z_i as follows:

$$\det(z_i) = t_i^2 - \|x_i\|^2.$$

And the multiplication in V_1 is defined as follows:

$$(t_i, x_i) \circ (s_i, y_i) = (t_i s_i + <x_i, y_i>_H, s_i x_i + t_i y_i)$$

Observe that the above multiplication is commutative but not associative. This defines the structure of a Jordan algebra (see e.g. [10] for more details). Then the inverse z_i^{-1} of z_i is given by the formula

$$z_i^{-1} = \frac{1}{\det(z_i)}(t_i, -x_i).$$

Next we consider a function

$$f(z_i) = -\ln \det(z_i) = -\ln(t_i^2 - \|x_i\|^2).$$

The quadratic representation $P(z_i)$ is defined as follows:

$$P(z_i) = H(z_i)^{-1}$$

where $H(z_i)$ is the Hessian of $f(z_i)$.

4

Remark: Since V is a product of m copies of V_1. We define a quadratic representation $P(z)$ for $z = (z_1, \cdots, z_m) \in V$ as follows:

$$P(z) = (P(z_1), \cdots, P(z_m)).$$

One of the main ingredients in the construction of NT-direction is the so-called scaling point.

Proposition 2.3. *Given $(z_1, z_2) \in \Omega \times \Omega$. Then there exists a unique $z_3 \in \Omega$ such that*

$$P(z_3)z_1 = z_2. \tag{2.7}$$

The following Corollary gives an explicit formula for the scaling point.

Corollary 2.4. *Let $z_1, z_2 \in \Omega$, $z_1 = (s, y)$, $z_2 = (t, x)$. Then consider $z_3 = (r, u)$ with*

$$r = \sqrt{\frac{\mu_1}{\mu_2}} \frac{\mu_1 s + \mu_2 t}{\sqrt{2 + \mu_1 \mu_2 < (s, y), (t, x) >}}$$

$$u = \sqrt{\frac{\mu_1}{\mu_2}} \frac{\mu_2 x - \mu_1 y}{\sqrt{2 + \mu_1 \mu_2 < (s, y), (t, x) >}}$$

$$\mu_i = \frac{1}{\sqrt{\det(z_i)}}, i = 1, 2.$$

Then $z_3 \in \Omega_1$ is a unique solution to (2.7) for $V_1 = \mathbb{R} \times H$.

Proposition 2.5. *Every element $z \in V_1$ admits the following spectral decomposition*

$$z = \lambda_1 e_1 + \lambda_2 e_2$$

where $\lambda_i \in \mathbb{R}$, $e_i \circ e_j = \delta_{ij} e_i$ for $i, j = 1, 2$.

The following proposition gives us an explicit formula for the spectral decomposition of an element $z \in V_1$.

Proposition 2.6. *Let $z = (s, y) \in V_1$, $y \neq 0$. Consider*

$$e_1 = \frac{1}{2}(1, \frac{y}{\|y\|}), \quad e_2 = \frac{1}{2}(1, -\frac{y}{\|y\|})$$

$$\lambda_1 = s + \|y\|, \quad \lambda_2 = s - \|y\|.$$

Then

$$(s, y) = \lambda_1 e_1 + \lambda_2 e_2$$

and $e_i \circ e_j = \delta_{ij} e_i$ for $i, j = 1, 2$.

Proposition 2.7. *Let Ω_1 be the cone of squares of V_1 and $(s, y) \in \Omega_1$. Consider*

$$z = (\frac{\mu}{2}, \frac{y}{\mu}), \mu = \sqrt{s + \|y\|} + \sqrt{s - \|y\|}.$$

Then $z \in \Omega_1$ and $z^2 = (s, y)$. Moreover if

$$z^2 = \lambda_1 e_1 + \lambda_2 e_2$$

is the spectral decomposition of z^2, then

$$z = \sqrt{\lambda_1} e_1 + \sqrt{\lambda_2} e_2.$$

Denote z by $(s, y)^{\frac{1}{2}}$.

Proposition 2.8. *We have*

$$P(z^{\frac{1}{2}})^2 = P(z).$$

Hence,

$$P(z^{\frac{1}{2}}) = P(z)^{\frac{1}{2}}. \tag{2.8}$$

For proofs of the above propositions and corollary, see [10]. We are now in a position to introduce Nesterov-Todd direction (see e.g. [15]). Given $z_1, z_2 \in \Omega$, let $z_3 \in \Omega$ be the scaling point of z_1 and z_2 defined in (2.7). Then the NT-direction $(\xi, \eta) \in X \times X^\perp$ is defined to be the solution to the following system of linear equations:

$$P(z_3)\xi + \eta = \gamma \mu(z_1, z_2) z_1^{-1} - z_2, \tag{2.9}$$

$$\xi \in X, \ \eta \in X^\perp. \tag{2.10}$$

Here $0 < \gamma < 1$ is a real parameter. One can show (see [10]) that (2.9),(2.10) admits a unique solution.

Proposition 2.9. *By following the NT-direction, the duality gap decreases.*

Proof Let $(z_1^{(k)}, z_2^{(k)}) \in \Omega \times \Omega$, $(\xi^{(k)}, \eta^{(k)})$ be NT-direction at $(z_1^{(k)}, z_2^{(k)})$, and $t^{(k)} > 0$. Define

$$z_1^{(k+1)} = z_1^{(k)} + t^{(k)} \xi^{(k)},$$
$$z_2^{(k+1)} = z_2^{(k)} + t^{(k)} \eta^{(k)}.$$

Then, consider the following equality:

$$
\begin{aligned}
< z_1^{(k+1)}, z_2^{(k+1)} > &= \ < z_1^{(k)} + t^{(k)} \xi^{(k)}, z_2^{(k)} + t^{(k)} \eta^{(k)} > \\
&= \ < z_1^{(k)}, z_2^{(k)} > + t^{(k)} (< z_1^{(k)}, \eta^{(k)} > + < z_2^{(k)}, \xi^{(k)} >)
\end{aligned}
$$

So it suffices to show that $< z_1^{(k)}, \eta^{(k)} > + < z_2^{(k)}, \xi^{(k)} > < 0$.

$$
\begin{aligned}
< z_1^{(k)}, \eta^{(k)} > + < z_2^{(k)}, \xi^{(k)} > &= \ < z_1^{(k)}, \eta^{(k)} > + < P(z_3) z_1^{(k)}, \xi^{(k)} > \\
&= \ < z_1^{(k)}, \eta^{(k)} + P(z_3) \xi^{(k)} > \\
&= \ < z_1^{(k)}, \gamma \mu(z_1^{(k)}, z_2^{(k)})(z_1^{(k)})^{-1} - z_2^{(k)} > \\
&= \ (\gamma - 1) < z_1^{(k)}, z_2^{(k)} > \\
&< \ 0 \quad (\text{ since: } 0 < \gamma < 1).
\end{aligned}
$$

Here we used (2.9). Hence we have

$$\begin{aligned} <z_1^{(k+1)}, z_2^{(k+1)}> &< <z_1^{(k)}, z_2^{(k)}> \\ \mu(z_1^{(k+1)}, z_2^{(k+1)}) &< \mu(z_1^{(k)}, z_2^{(k)}). \end{aligned}$$

$\square$

We conclude this section by a brief description of a primal-dual algorithm based on NT-direction.

For a fix $\epsilon > 0$, suppose that

$$(z^{(0)}, w^{(0)}) \in \operatorname{int}(\overline{\Omega}) \cap (b + X) \times \operatorname{int}(\overline{\Omega}) \cap (a + X^\perp).$$

Let (ξ_k, η_k) be the descent NT-direction obtained by solving the system of linear equations (2.9),(2.10), $\bar{t} > 0$ be the largest value such that

$$(z^{(k)} + \bar{t}\xi_k, w^{(k)} + \bar{t}\eta_k) \in \operatorname{int}(\overline{\Omega}) \cap (b + X) \times \operatorname{int}(\overline{\Omega}) \cap (a + X^\perp).$$

Then we set $(z^{(k+1)}, w^{(k+1)}) = (z^{(k)} + \bar{t}\xi_k, w^{(k)} + \bar{t}\eta_k)$. We stop the iteration when $\mu(z^{(k)}, w^{(k)}) \leq \epsilon$. In [10], it is shown that by choosing the right stepsize carefully, the algorithm runs in polynomial-time, i.e. it generates a sequence of primal and dual feasible iterates satisfying (2.6) with $\omega = 0.5$ or 1 and with δ being an independent positive constant.

3 Abstract Formulation of the Problem and Implementation

Let H be a Hilbert space. Consider the following optimization problem:

$$t \to \min, \tag{3.1}$$

$$\|W_i(y - y_i)\|_H \leq t, \ \ i = 1, .., m, \tag{3.2}$$

$$y \in c + Z. \tag{3.3}$$

Here $\| \ \|_H$ is the norm induced by the scalar product $< | >_H$, $y_i \in H$ are fixed ; Z is a closed vector subspace in H, $W_i : H \to H$ is a bounded linear operator for $i = 1, \cdots, m$. The problem (3.1)-(3.3) is an example of an infinite-dimensional second-order cone programming problem which has been analyzed in detail in [10].

We have implemented a version of a primal-dual algorithm based on Nestreov-Todd direction described in [10].

The problem (3.1)-(3.3) can be easily rewritten in the conic form: Let

$$\Omega_1 = \{(s, y) \in R \times H : s \geq \|y\|_H\}; \ \ V_1 = \mathbb{R} \times H;$$

$$\Omega = \Omega_1 \times \Omega_1 \times ... \times \Omega_1 \ (\text{m-times}) \ ; V = V_1 \times ... \times V_1 \ (\text{m-times}).$$

Let, further, $\Lambda : V_1 \to V$ be a linear operator such that:

$$\Lambda(s,y) = \{(s, W_1 y), ..., (s, W_m y)\},$$

$a = ((1,0),(0,0),...,(0,0)) \in V$, $b = ((0, W_1(c - y_1)), ..., (0, W_m(c - y_m)))$, and $X = \Lambda(\mathbb{R} \times Z)$. The scalar product $<,>_V$ in V is defined as follows:

$$< (t_1, x_1), ..., (t_m, x_m), (s_1, y_1), ..., (s_m, y_m) >_V = \sum_{i=1}^{m} \{ t_i s_i + < x_i, y_i >_H \} b.$$

With these notations we can rewrite the problem (3.1)-(3.3) in the conic form (2.1),(2.2).

Proposition 3.1. *We have*

$$X^\perp = \{(r_1, u_1), ..., (r_m, u_m) \in V \; ; r_1 + ... + r_m = 0, \sum_{i=1}^{m} W_i^* u_i \in Z^\perp \} \quad (3.4)$$

where $Z^\perp$ is the orthogonal complement of Z and W_i^* is the adjoint of W_i for each i.

A conic dual to (2.1),(2.2) will have the following form ([10]):

$$\sum_{i=1}^{m} < W_i(c - y_i), u_i >_H \to \min, \quad (3.5)$$

$$\sum_{i=1}^{m} W_i^* u_i \in Z^\perp, \quad \|u_i\| \leq r_i, \quad i = 1, ..., m, \quad (3.6)$$

$$\sum_{i=1}^{m} r_i = 1. \quad (3.7)$$

Here W_i^* is the adjoint of W_i.

3.1 Calculation of NT-direction

Given (z_1, z_2) a pair of feasible solutions to the problem (2.1),(2.2), and (2.3),(2.4). To obtain NT-direction for each iteration, we need to solve the following equation for (ξ, η),

$$P(z_3)\xi - \Delta \in X^\perp, \quad \xi \in X \quad (3.8)$$

where z_3 is the scaling point of z_1 and z_2 ,and $\Delta = \gamma \mu(z_1, z_2) z_1^{-1} - z_2$. (Compare with (2.9), (2.10).)

The equation (3.8) is equivalent to :

$$\frac{< P(z_3)\xi, \xi >}{2} - < \xi, \Delta > \to \min,$$

$$\xi \in X.$$

8

Let

$$z_3 = ((t_1, x_2), (t_2, x_2), \cdots, (t_m, x_m)) \text{ and } \Delta = ((r_1, u_1), \cdots, (r_m, u_m)).$$

For $(\mu, \zeta) \in \mathbb{R} \times Z, \ \xi = \Lambda(\mu, \zeta)$,

$$\rho(\mu, \zeta) = \frac{< P(z_3)\xi, \xi >}{2} - < \xi, \Delta >$$

$$= \frac{1}{2} \sum_{i=1}^{m} (t_i^2 - \|x_i\|^2) \|W_i \zeta\|^2 + \sum_{i=1}^{m} < x_i, W_i \zeta >^2 - \sum_{i=1}^{m} < u_i, W_i \zeta > + \frac{\nu_1 \mu^2}{2} + \nu_2 \mu,$$

where

$$\nu_1 = \sum_{i=1}^{m} (t_i^2 + \|x_i\|^2), \ \ \nu_2 = 2 \sum_{i=1}^{m} t_i < x_i, W_i \zeta > - \sum_{i=1}^{m} r_i.$$

Hence, if we denote by

$$\phi(\zeta) = \min\{\rho(\mu, \zeta) : \mu \in \mathbb{R}\}$$

$$\phi(\zeta) = \frac{< \zeta, M\zeta >}{2} + \frac{1}{2} \sum_{i=1}^{m+1} \varepsilon_i < v_i, \zeta >^2 + < v_0, \zeta > - \frac{(\sum_{i=1}^{m} r_i)^2}{2\nu_1}, \quad (3.9)$$

where

$$M = \sum_{i=1}^{m} (t_i^2 - \|x_i\|^2) \begin{bmatrix} Q_i & 0 \\ 0 & R_i \end{bmatrix} \qquad (3.10)$$

$$v_0 = \sqrt{\frac{1}{\nu_1}} (\sum_{i=1}^{m} r_i) v_{m+1} - \sum_{i=1}^{m} W_i u_i. \qquad (3.11)$$

$$v_i = \sqrt{2} W_i^* x_i, \ \ i = 1, 2, ..., m, \ \ v_{m+1} = \frac{2}{\sqrt{\nu_1}} \sum_{i=1}^{m} t_i W_i^* x_i, \qquad (3.12)$$

$$\varepsilon_i = 1, i = 1, ..., m, \ \ \varepsilon_{m+1} = -1, \qquad (3.13)$$

As a result, to obtain NT-direction, it suffices to solve the following optimization problem:

$$\phi(\zeta) \rightarrow \min,$$

$$\zeta \in Z.$$

which is equivalent to solving:

$$\frac{< \zeta, M\zeta >}{2} + \frac{1}{2} \sum_{i=1}^{m+1} \varepsilon_i < v_i, \zeta >^2 + < v_0, \zeta > \ \rightarrow \min, \qquad (3.14)$$

$$\zeta \in Z.$$

The following theorem further simplifies the problem (3.14).

9

Theorem 3.2. *Let ζ_0 be the optimal solution to the problem:*

$$\frac{\langle \zeta, M\zeta \rangle}{2} + \langle v_0, \zeta \rangle \to min, \tag{3.15}$$

$$\zeta \in Z, \tag{3.16}$$

and $\zeta_i, i = 1, 2, \cdots, m+1$, be optimal solutions to the problems:

$$\frac{\langle \zeta, M\zeta \rangle}{2} + \langle V_i, \zeta \rangle \to min, \tag{3.17}$$

$$\zeta \in Z, \tag{3.18}$$

where $V_i = \varepsilon_i v_i$.
Let $S = (s_{ij})$, $s_{ij} = \langle v_i, \zeta_j \rangle$, $i, j = 1, 2, \cdots, m+1$.
Then the set of optimal solutions to the problems (3.14) is in one-to-one correspondence with the set of the solutions of the system of linear equations

$$(I - S) \begin{bmatrix} \delta_1 \\ \vdots \\ \delta_{m+1} \end{bmatrix} = \begin{bmatrix} \langle v_1, \zeta_0 \rangle \\ \vdots \\ \langle v_{m+1}, \zeta_0 \rangle \end{bmatrix} \tag{3.19}$$

More precisely, if $(\delta_1, ..., \delta_{m+1})$ is a solution to (3.19), then

$$\zeta(\delta) = \zeta_0 + \sum_{i=1}^{m+1} \delta_i \zeta_i \tag{3.20}$$

is the optimal solution to the problem (3.14).

For a proof see [8, 10].
Remark: The procedure of reduction of (3.14), (3.15), (3.17) is a version of Sherman-Morrison-Woodbury formula.
Solving (3.14) is therefore equivalent to solve $m+2$ problems of the following type

$$\frac{< M\zeta, \zeta >}{2} + < v, \zeta > \to \min, \ \zeta \in Z \tag{3.21}$$

and a system of $(m+1) \times (m+1)$ linear algebraic equations.
Let ζ be an optimal solution to (3.14). Since $\frac{d\rho}{d\mu} = 0$. Hence $\mu = -\frac{\nu_2}{\nu_1}$. Therefore, we obtain

$$\xi = \Lambda(\mu, \zeta) = ((\mu, W_1\zeta), ..., (\mu, W_m\zeta))$$

which is the primal descent direction. A dual descent direction is simply

$$\eta = P(z)\xi - \Delta.$$

10

3.2 Computation of the step size

The Jordan algebraic techniques significantly simplifies the computation of the step-size. Observe that

$$z + t\xi = P(z^{\frac{1}{2}})(e + tP(z^{-\frac{1}{2}})\xi).$$

Hence, $z + t\xi \in \Omega_1$, is equivalent to $e + tP(z^{-\frac{1}{2}})\xi \in \Omega_1$.
If

$$tP(z^{-\frac{1}{2}})\xi = \mu_1 f_1 + \mu_2 f_2$$

is the spectral decomposition, then $z + t\xi \in \overline{\Omega}_1$ is equivalent to

$$1 + t\mu_1 \geq 0, \quad 1 + t\mu_2 \geq 0. \tag{3.22}$$

Since the primal-dual cone is the product of several copies of Ω_1, we obtain step size $\bar{t}$ as the maximal value of t for which all inequalities (3.22) are satisfied. We use $0.9\bar{t}$ for the actual step size in out algorithm.

The stopping rule we used is quite standard: the duality gap $\mu < \epsilon$ for a given $\epsilon > 0$.

Recall (see [10]) that Nesterov-Todd direction (ξ, η) is determined by the following conditions

$$\xi + P(z)^{-1}\eta = \gamma\mu(z_1, z_2)z_2^{-1} - z_1,$$

$$\xi \in X, \quad \eta \in X^{\perp},$$

where $(z_1, z_2) \in \Omega \times \Omega$ and $z_3 \in \Omega$ is the scaling point uniquely determined by the equation:

$$P(z_3)z_1 = z_2.$$

As a result of numical experiments, we arrived at $\gamma = 0.5$ as a value which minimizes the number of iterations.

4 Examples of concrete problems

We now consider several concrete examples of the problem (3.1)-(3.3).

4.1 Multi-criteria linear quadratic control problem

Denote by $L_2^n[0, T]$ the Hilbert space of square integrable functions $f : [0, T] \to \mathbb{R}^n$, $T > 0$. Let $(x, u) \in H = L_2^n[0, T] \times L_2^l[0, T]$. Consider the following optimization problem:

$$J(x, u) = \max_{i \in [1,m]} \int_0^T (x - x_i)^T Q_i(x - x_i) + (u - u_i)^T R_i(u - u_i)dt \to \min \tag{4.1}$$

$$\dot{x}(t) = Ax(t) + Bu(t), \quad x(0) = x_0. \tag{4.2}$$

11

Here $(x_i, u_i) \in H$, $i = 1, 2, ..., m$; $A, B, Q_i, R_i, i = 1, 2, ..., m$ are given matrices of appropriate dimensions. We will assume that Q_i, R_i are symmetric positive definite matrices. The problem (4.1),(4.2) can be easily rewritten in the second order cone programming problem of the form (3.1)- (3.3)

Here $\| \ \|_H$ is the norm induced by the scalar product $< \ | \ >_H$, $y_i = (x_i, u_i)$, $i = 1, 2, ..., m$; Z is a closed vector subspace in H described as follows:

$$Z = \{(x, u); \ \dot{x}(t) = Ax(t) + Bu(t), \ x(0) = 0\}; \tag{4.3}$$

$$W_i = \begin{bmatrix} L_{Q_i}^T & 0 \\ 0 & L_{R_i}^T \end{bmatrix}, \ i = 1, ..., m,$$

where L_{Q_i}, L_{R_i} are lower-triangular matrices obtained by Cholesky factorization of Q_i and R_i respectively.

Then by following the scheme described in the previous section, we have to show that the condition (2.5) is satisfied, and then compute NT-direction.

Constructing feasible solutions

Primal problem: After some numerical experiments, we end up with the following construction.

Let K_0 be a unique stabilizing solution to the algebraic Riccati equation:

$$KBB^T K - A^T K - KA - I = 0.$$

Here stabilizing means that the close-loop matrix $A - BB^T K_0$ is stable. Let x be the solution to the following system of linear differential equations:

$$\dot{x} = (A - BB^T K_0)x; \ x(0) = x_0, \ u = -B^T K_0 x.$$

It is quite obvious that $y = (x, u) \in c + Z$. and hence if we choose

$$s = \max_{i \in [1, m]} \|W_i(y - y_i)\| + \delta$$

for some $\delta > 0$, the initial point for primal problem is $\Lambda(s, y)$.

Dual problem: We simply take $u_i = 0, r_i = \frac{1}{m}, i = 1, 2, ..., m$.

Computation of NT-direction

For $T < \infty$, the problem (3.21) is simply classical LQ problem with a linear term on a finite interval (see e.g. [8, 11]). Hence on each iteration, we need to solve a differential Riccati equation, system of linear differential equations and a system of $(m + 1) \times (m + 1)$ linear algebraic equations (see [10]).

For the case when $T = \infty$, the problem (3.21) is LQ problem with a linear term on a semi-infinite interval. The complete solution to such a problem is described in the following section.

4.1.1 Solution to LQ problem with a linear term on a semi-infinite interval

Let $(x, u) \in H = L_2^n[0, \infty) \times L_2^m[0, \infty)$, $\Sigma = \Sigma^T$. Consider the following LQ-control problem with a linear term on a semi-infinite

$$J(x, u) = \frac{1}{2} \int_0^\infty < \left[\begin{array}{c} x \\ u \end{array} \right], \Sigma \left[\begin{array}{c} x \\ u \end{array} \right] > dt + \int_0^\infty < \left[\begin{array}{c} x \\ u \end{array} \right], \left[\begin{array}{c} y \\ v \end{array} \right] > dt \to \min \tag{4.4}$$

$$\dot{x} = Ax + Bu, \quad x(0) = x_0 \tag{4.5}$$

By using block partition, denote

$$\Sigma = \left[\begin{array}{cc} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12}^T & \Sigma_{22} \end{array} \right].$$

For proofs of the theorems in this section, see [9].

Theorem 4.1. *Let A be an antistable n by n matrix (i.e. real parts of all eigenvalues of A are positive). Consider the following system of linear differential equations:*

$$\dot{x} = Ax + f, \tag{4.6}$$

where $f \in L_2^n[0, \infty)$. There exists a unique solution $\mathcal{L}(f)$ of (4.6) such that $\mathcal{L}(f) \in L_2^n[0, \infty)$. Moreover the map $f \to \mathcal{L}(f)$ is linear and bounded. Explicitly:

$$\mathcal{L}(f)(t) = - \int_0^{+\infty} e^{-A\tau} f(\tau + t) d\tau \tag{4.7}$$

Theorem 4.2. *The following conditions are equivalent:*

i) Σ_{22} is a positive definite (symmetric) matrix and the following Riccati equation has a stabilizing solution.

$$KLK + K\tilde{A} + \tilde{A}^T K - Q = 0 \tag{4.8}$$

where

$$\tilde{A} = A - B\Sigma_{22}^{-1}\Sigma_{21}, \quad Q = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}, \quad L = B\Sigma_{22}^{-1}B^T.$$

ii) The pair (A, B) is stabilizable and there exists $\epsilon > 0$ such that

$$\Gamma(x, u) = \int_0^{+\infty} < \left[\begin{array}{c} x \\ u \end{array} \right], \Sigma \left[\begin{array}{c} x \\ u \end{array} \right] > dt \geq \epsilon \int_0^{+\infty} (\|x\|^2 + \|u\|^2) dt$$

for all $(x, u) \in Z$.

Let further $Z \subset H$ be as follows:

$$Z = \{(x, u) \in H; \dot{x} = Ax + Bu, x \text{ is absolutely continuous}, x(0) = 0\}.$$

13

Theorem 4.3. *Suppose that the pair (A, B) is stabilizable. Then Z is a closed vector subspace in H. Then*

$$Z^\perp = \left\{ \begin{bmatrix} \dot{p} + A^T p \\ B^T p \end{bmatrix} : p \in L_2^n[0, +\infty), p \text{ is absolutely continuous}, \dot{p} \in L_2^n[0, +\infty) \right\}$$

We are now in the position to describe the optimal solution to (4.4).

Theorem 4.4. *Suppose that the conditions of theorem 4.2 are satisfied. Then the problem (4.4)-(4.5) has a unique solution which can be described as follows.*

There exists a stabilizing solution K_0 to the Riccati equation (4.8). Then the matrix $C = -(\tilde{A} + LK_0)$ is antistable, $(K_0 B - \Sigma_{12})\Sigma_{22}^{-1}v + y \in L_2^n[0, +\infty)$.

Let ρ be a unique solution from $L_2^n[0, +\infty)$ of the system of differential equations

$$\dot{\rho} = C^T \rho + (K_0 B - \Sigma_{12})\Sigma_{22}^{-1}v + y \tag{4.9}$$

(which exists according to Theorem 4.1), x is the solution to the system of differential equations

$$\dot{x} = (\tilde{A} + LK_0)x + L\rho - B\Sigma_{22}^{-1}v, \quad x(0) = x_0, \tag{4.10}$$

$$p = K_0 x + \rho, \quad u = \Sigma_{22}^{-1}(B^T p - v - \Sigma_{21}x).$$

Now back to the problem of finding NT-direction for $T = \infty$. Solving (3.21) is reduced to solving an algebraic Riccati equation (4.8), systems of differential equations (4.9),(4.10) and a system of $(m+1)\times(m+1)$ linear algebraic equations on each iteration.

4.2 Discrete-time Multi-criteria linear quadratic control problem

In this section, we consider a discrete-time formulation for the problem (4.1)-(4.2). For simplicity, let us introduce some useful notations.

Let x denote a sequence $\{x_k\} \subset \mathbb{R}^n$ for $k = 0, \cdots, \infty$. We say that $x \in l_2^n(\mathbb{N})$ if $\sum_{i=1}^\infty \|x_i\|^2 < \infty$ where $\|\cdot\|$ is a norm induced by an inner product $<,>$ in $\mathbb{R}^n$. Let $(x, u) \in l_2^n(\mathbb{N}) \times l_2^m(\mathbb{N})$. Then the Discrete-time Multi-criteria Linear Quadratic Control Problem takes the form:

$$\max_{i=1,\ldots,m} \sum_{k=0}^\infty \{< (x_k - \varphi_k^{(i)}), Q_i(x_k - \varphi_k^{(i)}) > + < (u_k - \phi_k^{(i)}), R_i(u_k - \phi_k^{(i)}) >\} \quad \to \quad \min$$

$$\tag{4.11}$$

$$x_{k+1} = Ax_k + Bu_k, \ k = 0, 1, \cdots, \quad x_0 = \gamma_0,$$

where $(\varphi^{(i)}, \phi^{(i)}) \in H = l_2^n(\mathbb{N}) \times l_2^m(\mathbb{N})$ for $i = 1, \cdots, m$, A, B, Q_i, R_i for $i = 1, \cdots, m$ are matrices of appropriate sizes. We assume that Q_i's and R_i's are positive definite.

It is easy to see that the problem (4.11) can be rewritten into the second-order cone programming problem (3.1)- (3.3).

Observe now the inner product in H has the following form:

$$< (x, y), (u, v) >_H = \sum_{k=0}^{\infty} \{< x_k, u_k > + < y_k, v_k >\}. \qquad (4.12)$$

The vector subspace Z now takes the form:

$$Z = \{(x, u) \in H : \ x_{k+1} = Ax_k + Bu_k, k = 0, 1, \cdots, \ x_0 = 0\}$$

The constructions of a initial feasible point for primal and dual problems are similar to the constructions of a feasible points in the continuous case. In the primal, we solve a discrete LQ problem via discrete Algebraic Riccati equation. In the dual, we take the same choice as in the continuous case.

Now the problem (3.21) is discrete-time linear-quadratic control problem with linear term.

4.2.1 Discrete-time linear-quadratic control problem with linear term

In this section, we describe a complete solution to discrete linear-quadratic control problem with a linear term. In order to derive the solution, we need the following auxiliary results.

Theorem 4.5. *Let A be a d-stable matrix (ie. all eigenvalues of A lie inside a unit circle on a on a complex plane.) Consider the following system of linear difference equation.*

$$x_k = Ax_{k+1} + f_k, \quad k = 0, 1, \cdots \qquad (4.13)$$

where $f = \{f_k\} \in l_2^n(\mathbb{N})$. Then there exists a unique solution $\mathcal{L}(f)$ of (4.13) such that $\mathcal{L}(f) \in l_2^n(\mathbb{N})$. Moreover the map $f \to \mathcal{L}(f)$ is linear and bounded. Explicitly,

$$x_k = \mathcal{L}(f)_k = \sum_{r=0}^{\infty} A^r f_{r+k}, \quad k = 0, 1, \cdots. \qquad (4.14)$$

Proof Since A be a stable matrix, there exists $S = S^T \geq 0$, such that

$$A^T SA - S = -I \qquad (4.15)$$

(see e.g. [12]). Let $w_k = < x_k, Sx_k >$. Obviously, $w_k \geq 0, \forall k$. Then using (4.13) and (4.15), we obtain the following:

$$
\begin{aligned}
\triangle w_{k+1} &= w_{k+1} - w_k \\
&= < x_{k+1}, Sx_{k+1} > - < x_k, Sx_k > \\
&= < x_{k+1}, Sx_{k+1} > - < (Ax_{k+1} + f_k), S(Ax_{k+1} + f_k) > \\
&= < x_{k+1}, Sx_{k+1} > - < x_{k+1}, A^T SAx_{k+1} > - \\
&\quad 2 < Ax_{k+1}, Sf_k > - < f_k, Sf_k > \\
&= < x_{k+1}, x_{k+1} > -2 < Ax_{k+1}, Sf_k > - < f_k, Sf_k > \\
&\geq \|x_{k+1}\|^2 - 2\|x_{k+1}\|\|A^T S\|\|f_k\| - \|S\|\|f_k\|^2.
\end{aligned}
$$

15

But then it is easy to see that the following inequality is true.

$$2\|x_{k+1}\|\|A^T S\|\|f_k\| \leq \frac{1}{2}\|x_{k+1}\|^2 + 2\|A^T S\|^2\|f_k\|^2$$

Then

$$\triangle w_{k+1} \geq \frac{1}{2}\|x_{k+1}\|^2 - (2\|A^T S\|^2 + \|S\|)\|f_k\|^2$$

Then it follows that

$$w_{k+1} - w_0 = \sum_{r=1}^{k+1} \triangle w_r$$

$$\geq \sum_{r=1}^{k+1}(\frac{1}{2}\|x_r\|^2 - (2\|A^T S\|^2 + \|S\|)\|f_{r-1}\|^2).$$

Hence,

$$\sum_{r=1}^{k+1}\|x_r\|^2 \leq 2\sum_{r=1}^{k+1}(2\|A^T S\|^2 + \|S\|)\|f_{r-1}\|^2 + 2(w_{k+1} - w_0). \qquad (4.16)$$

Let now

$$x_k = \sum_{r=0}^{\infty} A^r f_{r+k}, \ k = 0, 1, \cdots, \qquad (4.17)$$

then we have that

$$
\begin{aligned}
Ax_{k+1} + f_k &= A\sum_{r=0}^{\infty} A^r f_{r+k+1} + f_k \\
&= \sum_{r=0}^{\infty} A^{r+1} f_{r+k+1} + f_k \\
&= \sum_{r=1}^{\infty} A^r f_{r+k} + f_k \\
&= \sum_{r=0}^{\infty} A^r f_{r+k} \\
&= x_k.
\end{aligned}
$$

Hence, (4.17) satisfies (4.13). Let us now consider the following inequilities:

$$
\begin{aligned}
\|x_k\| &= \|\sum_{r=0}^{\infty} A^r f_{r+k}\| \\
&\leq \sum_{r=0}^{\infty} \|A^r f_{r+k}\| \\
&\leq \sum_{r=0}^{\infty} \|A^r\|\|f_{r+k}\|.
\end{aligned}
$$

16

Then by Cauch-Schawrz inequality we have

$$\|x_k\|^2 \le \sum_{r=0}^{\infty} \|A^r\|^2 \sum_{r=0}^{\infty} \|f_{r+k}\|^2$$

$$\le \sum_{r=0}^{\infty} \|A\|^{2r} \sum_{s=0}^{\infty} \|f_s\|^2$$

for all k. But since A is stable and all norms are equivalent, by Householder's Theorem [17] we can assume without lost of generality that

$$\|A\| < 1.$$

Therefore,

$$\|x_k\|^2 < \infty \text{ for all } k.$$

Hence $\{x_k\}$ is bounded and as a result $\{w_k\}$ is also bounded. But since $x_0 = \sum_{r=0}^{\infty} A^r f_r$,

$$\|x_0\|^2 \le \sum_{r=0}^{\infty} \|A\|^{2r} \sum_{r=0}^{\infty} \|f_r\|^2$$

$$\le \frac{1}{1 - \|A\|^2} \sum_{r=0}^{\infty} \|f_r\|^2.$$

Then from (4.16) we have,

$$\sum_{r=0}^{k+1} \|x_r\|^2 \le 2(2\|A^T S\|^2 + \|S\| + \frac{1}{1 - \|A\|^2}) \sum_{r=0}^{k+1} \|f_r\|^2 + 2(w_{k+1} - w_0)$$

for all k. Since $\{w_k\}$ is bounded, we conclude from (4.18):

$$\{x_k\} \in l_2^n(\mathbb{N}).$$

Hence

$$\lim_{k \to \infty} x_k = 0.$$

As a result,

$$\lim_{k \to \infty} w_k = 0.$$

From (4.18), we now have

$$\sum_{r=0}^{\infty} \|x_r\|^2 \le 2(2\|A^T S\|^2 + \|S\| + \frac{1}{1 - \|A\|^2}) \sum_{r=0}^{\infty} \|f_r\|^2 - 2w_0. \qquad (4.18)$$

But since

$$w_0 = < x_0, S x_0 > \ge 0,$$

we have

$$\sum_{r=0}^{\infty} \|x_r\|^2 \quad \leq \quad 2(2\|A^T S\|^2 + \|S\| + \frac{1}{1 - \|A\|^2}) \sum_{r=0}^{\infty} \|f_r\|^2$$

$$\leq \quad C \sum_{r=0}^{\infty} \|f_r\|^2 \tag{4.19}$$

for some constant C.

But now suppose that $\{\tilde{x}_k\} \in l_2^n(\mathbb{N})$ be another solution to (4.13). Then

$$\begin{aligned} y_k &= x_k - \tilde{x}_k \\ &= A(x_{k+1} - \tilde{x}_{k+1}) \\ &= Ay_{k+1} \end{aligned}$$

for $k = 0, \cdots$. Since A is d-stable and $\{y_k\} \in l_2^n(\mathbb{N})$, one can easily see that $y_k = 0$ for all $k = 0, \cdots$.

Therefore the system (4.13) may have only one solution $\{x_k\} \in l_2^n(\mathbb{N})$. Hence we can conclude that the linear map $f \to \mathcal{L}(f)$ is correctly defined and bounded. $\square$

Let now consider the following classical discrete-time LQ problem

$$\frac{1}{2} \sum_{k=0}^{\infty} < \left[\begin{array}{c} x_k \\ u_k \end{array} \right], \Sigma \left[\begin{array}{c} x_k \\ u_k \end{array} \right] > \quad \to \min, \tag{4.20}$$

where $x_{k+1} = Ax_k + Bu_k$, $x_0 = x^0 \in \mathbb{R}^n$ and

$$\Sigma = \left[\begin{array}{cc} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12}^T & \Sigma_{22} \end{array} \right].$$

It is well known the the following discrete algebraic Riccati equation plays a very important role in finding the optimal solution to the problem (4.20) [12]:

$$K = A^T K A - (A^T K B + \Sigma_{12})(\Sigma_{22} + B^T K B)^{-1}(A^T K B + \Sigma_{12}) + \Sigma_{11}. \tag{4.21}$$

Observe that (4.21) is defined under assumption that the matrix $\Sigma_{22} + B^T K B$ is invertible. Let

$$F = (A^T K B + \Sigma_{12})(\Sigma_{22} + B^T K B)^{-1}.$$

Then K_0 a symmetric solution to (4.21) is called a stabilizing if $\Sigma_{22} + B^T K B > 0$ and $A^T - FB^T$ is d-stable.

Theorem 4.6. *Suppose Σ_{22} is a positive definite matrix, (4.21) has a stabilizing solution. Then*

$$\frac{1}{2} \sum_{k=0}^{\infty} < \left[\begin{array}{c} x_k \\ u_k \end{array} \right], \Sigma \left[\begin{array}{c} x_k \\ u_k \end{array} \right] > > 0. \tag{4.22}$$

18

for all (x_k, u_k) satisfying

$$x_{k+1} = Ax_k + Bu_k, \tag{4.23}$$

$$x_0 = 0, \ (x, u) \in l_2^n(\mathbb{N}) \times l_2^m(\mathbb{N}).$$

Proof Let K be a stabilizing solution to (4.21), using (4.22) we obtain the following equalities:

$$
\begin{aligned}
& < x_{k+1}, Kx_{k+1} > - < x_k, Kx_k > \\
= & < (Ax_k + Bu_k), K(Ax_k + Bu_k) > - < x_k, Kx_k > \\
= & < x_k, (A^T K A - K)x_k > + < x_k, A^T K B u_k > + < u_k, B^T K A x_k > + \\
& < u_k, (B^T K B + \Sigma_{22})u_k > - < u_k, \Sigma_{22} u_k > .
\end{aligned}
$$

Let now $R_K = \Sigma_{22} + B^T K B$, and $S_K = A^T K B + \Sigma_{12}$ then using (4.21), we obtain:

$$
\begin{aligned}
& < x_{k+1}, Kx_{k+1} > - < x_k, Kx_k > \\
= & < x_k, (S_K R_K^{-1} S_K - \Sigma_{11})x_k > + < x_k, A^T K B u_k > \\
& + < u_k, B^T K A x_k > + < u_k, R_K u_k > - < u_k, \Sigma_{22} u_k > \\
= & < x_k, (S_K R_K^{-1} S_K)x_k > + < x_k, S_K u_k > + < u_k, S_K^T x_k > + < u_k, R_K u_k > \\
& - < x_k, \Sigma_{12} u_k > - < u_k, \Sigma_{12}^T x_k > - < u_k, \Sigma_{22} u_k > - < x_k, \Sigma_{11} x_k > \\
= & < (R_K^{-1} S_K x_k + u_k), R_K(R_K^{-1} S_K x_k + u_k) > - < \left[\begin{array}{c} x_k \\ u_k \end{array} \right], \Sigma \left[\begin{array}{c} x_k \\ u_k \end{array} \right] > .
\end{aligned}
$$

Then take summation with respect to k from $k = 0$ to $k = N$ to obtain:

$$
\begin{aligned}
\sum_{k=0}^{N} < \left[\begin{array}{c} x_k \\ u_k \end{array} \right], \Sigma \left[\begin{array}{c} x_k \\ u_k \end{array} \right] > \ = \ & \sum_{k=0}^{N} < (R_K^{-1} S_K x_k + u_k), R_K(R_K^{-1} S_K x_k + u_k) > \\
& + < x_0, Kx_0 > - < x_{N+1}, Kx_{K+1} > .
\end{aligned}
$$

Since

$$x_0 = 0 \text{ and } \lim_{N \to \infty} x_N = 0,$$

we have

$$
\sum_{k=0}^{\infty} < \left[\begin{array}{c} x_k \\ u_k \end{array} \right], \Sigma \left[\begin{array}{c} x_k \\ u_k \end{array} \right] > \ = \ \sum_{k=0}^{\infty} < (R_K^{-1} S_K x_k + u_k), R_K(R_K^{-1} S_K x_k + u_k) > .
$$

Hence, since $R_K > 0$, (4.22) is satisfied. Theorem is proved. $\qquad\square$

Remark Suppose that (A, B) is d-stabilizable and $\Sigma = I$ (identity matrix). Then (4.21) has a stabilizing solution (see e.g. [12]).

Theorem 4.7. *Let $H = l_2^n(\mathbb{N}) \times l_2^m(\mathbb{N})$. Suppose the pair (A, B) is d-stabilizable. Let X be defined as follows:*

$$X = \{(x, u) = (\{x_k\}, \{u_k\}) \in H; x_{k+1} = Ax_k + Bu_k, x_0 = 0 \in \mathbb{R}^n \ k = 0, 1, \cdots \}.$$

Let also

$$Z = \{(\{\xi_k\}, \{\eta_k\}); \qquad \xi_0 = A^T p_0, \ \xi_k = A^T p_k - p_{k-1} \ for \ k = 1, 2, \cdots ,$$
$$\eta_k = B^T p_k, \ for \ k = 0, 1, \cdots , \ \{p_k\} \in l_2^n(\mathbb{N})\}.$$

Then X is a closed vector subspace and Z is an orthogonal complement of X in H, i.e. $Z = X^\perp$.

Proof Let $(x, u) \in X$ and $(\{\xi_k\}, \{\eta_k\}) \in Z$. Then

$$< (x, u), (\{\xi_k\}, \{\eta_k\}) >_H$$

$$= \ < x_0, A^T p_0 > + < u_0, B^T p_0 > + \lim_{n \to \infty} \sum_{k=1}^{n} (< x_k, A^T p_k - p_{k-1} > + < u_k, B^T p_k >)$$

$$= \ < x_1, p_0 > + \lim_{n \to \infty} \sum_{k=1}^{n} (< Ax_k + Bu_k, p_k > - < x_k, p_{k-1} >)$$

$$= \ < x_1, p_0 > + \lim_{n \to \infty} \sum_{k=1}^{n} (< x_{k+1}, p_k > - < x_k, p_{k-1} >)$$

$$= \ \lim_{n \to \infty} < x_{n+1}, p_n >$$

$$= \ 0.$$

Next we show that given $(\{\psi_k\}, \{\phi_k\}) \in H$, it admits the following representation:

$$\psi_0 = x_0 - [A^T p_0] \tag{4.24}$$

and

$$\psi_k = x_k - [A^T p_k - p_{k-1}] \tag{4.25}$$

for $k = 1, 2, \cdots$

$$\phi_k = u_k - B^T p_k \tag{4.26}$$

for $k = 0, 1, \cdots$, where $(\{x_k\}, \{u_k\}) \in X$ and $\{p_k\} \in l_2^n(\mathbb{N})$. This will imply easily that $Z = X^\perp$ and both X and Z are closed.

We look for p_k in the form

$$p_{k-1} = -Kx_k + \rho_k \tag{4.27}$$

Then by (4.25), we have

$$p_{k-1} = \psi_k - x_k + A^T p_k$$

where $K = K^T$ is a symmetric matrix. Substituting this into (4.27), we get

$$\rho_k = A^T p_k + \psi_k - x_k + Kx_k. \tag{4.28}$$

20

But then (4.26) implies:
$$u_k = B^T p_k + \phi_k. \tag{4.29}$$

Our goal now is to eliminate p_k. By using (4.27)
$$\begin{aligned}
p_k &= -Kx_{k+1} + \rho_{k+1} \\
&\quad -K(Ax_k + Bu_k) + \rho_{k+1}.
\end{aligned}$$

By substituting this into (4.29), we have
$$\begin{aligned}
u_k &= B^T(-K(Ax_k + Bu_k) + \rho_{k+1}) + \phi_k \\
&= -B^T KAx_k - B^T KBu_k + B^T \rho_{k+1} + \phi_k.
\end{aligned}$$

Hence,
$$(I + B^T KB)u_k = -B^T KAx_k + B^T \rho_{k+1} + \phi_k.$$

Hence, we have
$$u_k = -(I + B^T KB)^{-1} B^T KAx_k + (I + B^T KB)^{-1}(B^T \rho_{k+1} + \phi_k). \tag{4.30}$$

Using
$$x_{k+1} = Ax_k + Bu_k$$

we obtain
$$\begin{aligned}
x_{k+1} &= Ax_k - B(I + B^T KB)^{-1} B^T KAx_k \\
&\quad +B(I + B^T KB)^{-1}(B^T \rho_{k+1} + \phi_k). \tag{4.31}
\end{aligned}$$

But then since
$$p_k = -Kx_{k+1} + \rho_{k+1},$$

we have
$$\begin{aligned}
p_k &= -KAx_k + KB(I + B^T KB)^{-1} B^T KAx_k \\
&\quad -KB(I + B^T KB)^{-1}(B^T \rho_{k+1} + \phi_k) + \rho_{k+1}.
\end{aligned}$$

Then from (4.28) we have the following
$$\begin{aligned}
\rho_k &= A^T p_k + \psi_k - x_k + Kx_k \\
&= \psi_k - x_k + Kx_k - A^T KAx_k + A^T KB(I + B^T KB)^{-1} B^T KAx_k \\
&\quad -A^T KB(I + B^T KB)^{-1}(B^T \rho_{k+1} + \phi_k) + A^T \rho_{k+1}.
\end{aligned}$$

Recall, now, if Σ is the identity matrix then according to Remark which follows Theorem 4.6 the corresponding discrete algebraic Riccati (4.21) equation has a stabilizing solution. In this case, $\Sigma_{11} = I_n$, $\Sigma_{22} = I_m$, $\Sigma_{12} = 0$ i.e.
$$K = A^T KA - A^T KB(I + B^T KB)^{-1} A^T KB + I \tag{4.32}$$

21

If we choose K to be stabilizing solution, we obtain

$$\rho_k = A^T(I - KB(I + B^TKB)^{-1}B^T)\rho_{k+1} - A^TKB(I + B^TKB)^{-1}\phi_k + \psi_k \quad (4.33)$$

It is known that if K_0 is a stabilizing solution to DARE (4.32), then the matrix $A^T(I - K_0B(I + B^TK_0B)^{-1}B^T)$ is d-stable. Then by the Theorem 4.5, there exists a unique solution $\{\rho_k\} \in l_2^n(\mathbb{N})$ to (4.33). Hence by using (4.31) where $x_0 = 0$, (4.30) and then (4.27) we have shown that any $(\{\psi_k\}, \{\phi_k\}) \in H$ admits a unique decomposition as a sum of elements in X and $X^\perp$. Hence the Theorem is proved. $\qquad\square$

Now we are in the position to describe a solution of the LQ-problem on a semi-infinite interval.

$$\frac{1}{2}\sum_{k=0}^{\infty} < \left[\begin{array}{c} x_k \\ u_k \end{array}\right], \Sigma \left[\begin{array}{c} x_k \\ u_k \end{array}\right] > + \sum_{k=0}^{\infty} < \left[\begin{array}{c} x_k \\ u_k \end{array}\right], \left[\begin{array}{c} \psi_k \\ \phi_k \end{array}\right] > \ \to \min, \quad (4.34)$$

where

$$x_{k+1} = Ax_k + Bu_k + g_k, \ x_0 = x^0 \in R^n, \ \{g_k\} \in l_2^n(\mathbb{N}), \ (x, u) \in H. \quad (4.35)$$

Theorem 4.8. *Suppose there exists a stabilizing solution K to DARE (4.21): Then*

$$\{f_k\} \in l_2^n(\mathbb{N})$$

where

$$f_k = (A^TKB + \Sigma_{12})(\Sigma_{22} + B^TKB)^{-1}\phi_k + \Sigma_{12}(\Sigma_{22} + B^TKB)^{-1}B^TKg_k - \psi_k.$$

Furthermore, let $\{\rho_k\}$ be a unique solution from $l_2^n(\mathbb{N})$ of the system

$$\rho_k = C\rho_{k+1} + f_k$$

where

$$C = A^T - (A^TKB + \Sigma_{12})(\Sigma_{22} + B^TKB)^{-1}B^T.$$

Then the solution $(\{x_k\}, \{u_k\})$ to LQ-problem on a semi-infinite interval (4.34) can be described by the following recurrent relations:

$$x_{k+1} = C^Tx_k + B(\Sigma_{22} + B^TKB)^{-1}(B^T\rho_{k+1} - B^TKg_k - \phi_k) \quad (4.36)$$

$$\begin{aligned} u_k \ = \ & -(\Sigma_{22} + B^TKB)^{-1}(B^TKA + \Sigma_{12})x_k \\ & +(\Sigma_{22} + B^TKB)^{-1}(B^T\rho_{k+1} - B^TKg_k - \phi_k). \end{aligned} \quad (4.37)$$

Proof By Theorem 4.5, the functional (4.34) restricted to subspace X is convex. Hence necessary and sufficient optimality condition for (4.34),(4.35)takes the form

$$\Sigma \left[\begin{array}{c} x_k \\ u_k \end{array}\right] + \left[\begin{array}{c} \psi_k \\ \phi_k \end{array}\right] = \left[\begin{array}{c} A^Tp_k - p_{k-1} \\ B^Tp_k \end{array}\right]$$

22

when (x, u) satisfy (4.35) for some $\{p_k\} \in l_2^n(\mathbb{N})$. Moreover, we are looking for $\{p_k\}$ in the form

$$p_k = -Kx_{k+1} + \rho_{k+1}$$

where K is a stabilizing solution to DARE (4.21). We finish the proof exactly as in Theorem 4.7. $\qquad\square$

5 Some extensions

5.1 Linear-quadratic control problem with quadratic constraints

Consider the following optimization problem:

$$\|W_0(y - y_0)\|_H \to \min, \tag{5.1}$$

$$\|W_i(y - y_i)\|_H \le t_i, \ i = 1, ..., m \tag{5.2}$$

$$y \in c + Z. \tag{5.3}$$

With our choice of H, Z and other parameters, we arrive at the linear-quadratic control problem with quadratic constraints (see e.g [7, 20]). We first consider the problem of checking the feasibility of (5.1) - (5.3):

$$t \to \min, \tag{5.4}$$

$$\|W_i(y - y_i)\|_H \le t + t_i, \ i = 1, ..., m \tag{5.5}$$

$$y \in c + Z. \tag{5.6}$$

To cast (5.4) - (5.6) in the conic form, consider

$$\Lambda : \mathbb{R} \times Z \to (\mathbb{R} \times H) \times (\mathbb{R} \times H) \times ...(\mathbb{R} \times H), \text{ m-times such that}$$

$$\Lambda(t, y) = (t, W_1 y), (t, W_2 y), ..., (t, W_m y),$$

$a = ((1, 0), (0, 0), ..., (0, 0))$, $b = ((t_1, W_1(c - y_1)), ..., (t_m, W_m(c - y_m)))$ and $X = \Lambda(\mathbb{R} \times Z)$. Then we can rewrite (5.4)-(5.6) in the form of (2.1),(2.2). Observe that the only difference between (2.1),(2.2) and our conic formulation is the choice of b. It is quite obvious that the dual will be of the form

$$\sum_{i=1}^m t_i r_i + \sum_{i=1}^m < W_i(c - y_i), u_i >_H \to \min, \tag{5.7}$$

$$\sum_{i=1}^m W_i^* u_i \in Z^\perp, \ \ \|u_i\| \le r_i, \ \ i = 1, ..., m, \tag{5.8}$$

$$\sum_{i=1}^m r_i = 1. \tag{5.9}$$

Thus, the whole numerical scheme remain the same for (5.4)-(5.6)(the only difference is in the choice of initial feasible solution).

Returning to (5.1)-(5.3) consider the following;

$$\widetilde{\Lambda} : \mathbb{R} \times Z \to (\mathbb{R} \times H) \times (\mathbb{R} \times H) \times ...(R \times H), \ (m+1)\text{-times such that}$$

$$\widetilde{\Lambda}(t, y) = (t, W_0 y), (0, W_1 y), ..., (0, W_m y),$$

$$\widetilde{X} = \widetilde{\Lambda}(\mathbb{R} \times Z),$$

$$a = ((1, 0), (0, 0), ..., (0, 0)),$$

$$b = ((0, W_0(c - y_0)), (t_1, W_1(c - y_1)), ..., (t_m, W_m(c - y_m))).$$

We can rewrite (5.1)-(5.3) in the conic form

$$< a, z > \to \min,$$

$$z \in (b + \widetilde{X}) \cap \overline{\Omega}.$$

One can easily see that the dual will have the following form

$$\sum_{i=1}^{m} t_i r_i + \sum_{i=0}^{m} < W_i(c - y_i), u_i >_H \to \min, \tag{5.10}$$

$$\sum_{i=0}^{m} W_i^* u_i \in Z^{\perp}, \ \ \|u_i\| \le r_i, \ \ i = 1, ..., m, \tag{5.11}$$

$$\|u_0\| \le 1. \tag{5.12}$$

It is quite easy to work out a scheme for the computation of the Nesterov-Todd direction following the ideas developed in [10].

6   Numerical results

In this section we consider several examples for both continuous and discrete-time formulations. The major difference between the two methods is in solving linear-quadratic control problem with linear term. For continuous-time case for finite interval (respectively, infinite interval), solving linear-quadratic control problem with linear term requires solving differential Riccati equation (continuous algebraic Riccati equation) and systems of linear differential equations. For discrete-time case, we have to solve discrete Riccati equation and systems of difference equations.

The numerical experiments confirm fast convergence to optimal solution and good approximation of the continuous formulations by discrete formulations.

The algorithms are implemented in MATLAB and integrations of the ordinary differential equations are carried out with ODE45, a MATLAB ODE solver function for nonstiff problems.

6.1 Example 1: Multi-criteria Linear Quadratic Control Problem

We consider Multi-criteria Linear Quadratic Control Problem in both continuous and discrete cases. For continuous formulations, we first consider the problem (4.1),(4.2) (i.e. problem on the finite interval $[0, T]$) with $m = n = 3$ and the following data:

	1	2	3	4	5	6
x_{1i}	e^{-3t}	e^{-3t}	$\sin(t)e^{-4t}$	e^{-2t}	$\cos(t)e^{-4t}$	$\cos(t)e^{-3t}$
x_{2i}	$\cos(t)e^{-4t}$	$\sin(t)e^{-4t}$	e^{-3t}	$\sin(t)e^{-5t}$	e^{-3t}	e^{-4t}
x_{3i}	$\cos(t)e^{-4t}$	$\cos(t)e^{-4t}$	$\cos(t)e^{-4t}$	$\sin(t)e^{-5t}$	e^{-2t}	e^{-3t}
u_{1i}	$\frac{1}{t^2+5}$	$\frac{1}{t^2+3}$	$\frac{1}{t^2+3}$	$\frac{1}{t^2+5}$	$\frac{1}{t^2+8}$	$\frac{1}{t^2+8}$
u_{2i}	$\frac{1}{t^2+8}$	$\frac{1}{t^2+8}$	$\frac{1}{t^2+8}$	$\frac{1}{t^2+8}$	$\frac{1}{t^2+6}$	$\frac{1}{t^2+8}$
u_{3i}	$\frac{1}{t^2+6}$	$\frac{1}{t^2+6}$	$\frac{1}{t^2+6}$	$\frac{1}{t^2+6}$	$\frac{1}{t^2+6}$	$\frac{1}{t^2+6}$

	7	8
x_{1i}	$\cos(t)e^{-4t}$	$\sin(t)e^{-3t}$
x_{2i}	e^{-4t}	e^{-4t}
x_{3i}	e^{-3t}	e^{-3t}
u_{1i}	$\frac{1}{t^2+7}$	$\frac{1}{t^2+8}$
u_{2i}	$\frac{1}{t^2+8}$	$\frac{1}{t^2+6}$
u_{3i}	$\frac{1}{t^2+6}$	$\frac{1}{t^2+6}$

$$A = \begin{bmatrix} 2 & 5 & 2 \\ 4 & 2 & 3 \\ 2 & 3 & 2 \end{bmatrix}, B = \begin{bmatrix} 3 & 2 & 4 \\ 5 & 2 & 2 \\ 3 & 4 & 2 \end{bmatrix}, \alpha_0 = \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}$$

$Q_1 = \text{diag}(169, 196, 225)$, $Q_2 = \text{diag}(169, 169, 9)$, $Q_3 = \text{diag}(196, 121, 25)$,
$Q_4 = \text{diag}(225, 81, 49)$, $Q_5 = \text{diag}(196, 49, 81)$, $Q_6 = \text{diag}(289, 25, 121)$,
$Q_7 = \text{diag}(324, 9, 169)$, $Q_8 = \text{diag}(361, 121, 225)$,

$R_1 = \text{diag}(196, 144, 196)$, $R_2 = \text{diag}(144, 9, 16)$, $R_3 = \text{diag}(169, 225, 36)$,
$R_4 = \text{diag}(196, 49, 64)$, $R_5 = \text{diag}(225, 81, 100)$, $R_6 = \text{diag}(256, 121, 144)$,
$R_7 = \text{diag}(289, 169, 196)$, $R_8 = \text{diag}(324, 225, 36)$,

Then by applying a primal-dual algorithm based on NT-direction, we obtain the following table of solutions for different "T" for $\epsilon = 0.01$

T	1	2	3	4	5
$Opt\ Value$	2035	2071	2075	2076	2077
$Number\ of\ Iterations$	6	6	6	6	6

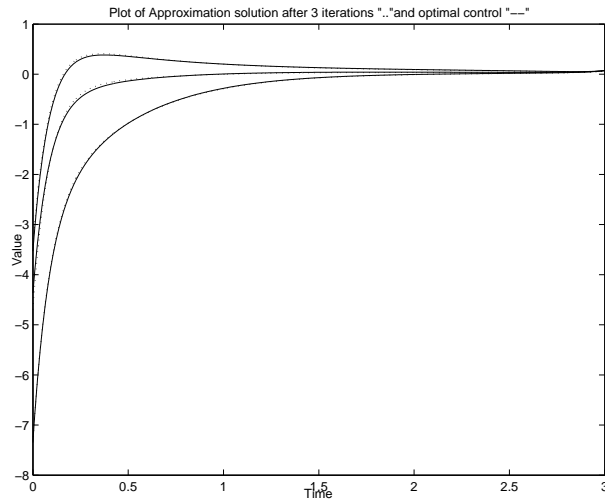| T	6	7	8	9	10
$Opt\ Value$	2077	2077	2077	2077	2077
$Number\ of\ Iterations$	6	6	6	6	6

Figure 1: Plot of approximation of optimal control and Optimal control: T = 3

Note also that the approximation of the optimal control approaches to the optimal control only after a few iterations. In fact, only after 3-4 iterations, we obtain a very good approximation to the optimal solution. Figure 1 and 2 are graphs of the approximate optimal solution after three iterations and the optimal solution. The dotted line is the plot of the approximation and the solid line is the actual optimal solution (the actual optimal solution means the solution obtained in the last iteration).

Our next goal is to use the discrete formulations to estimate the optimal solution and optimal value for the above problem. The construction is quite straightforward. First,we choose a sufficiently large $K > 0$, then we discretize the partition $[0, T]$ in to K sub-intervals uniformly. Next we let $h = \frac{T}{K}$. Hence, our targets in the discrete formulation can be easily computed by the following simple procedure: for state components

$$\begin{bmatrix} x_k^{(1i)} \\ x_k^{(2i)} \\ x_k^{(3i)} \end{bmatrix} = \begin{bmatrix} x_{1i}(kh) \\ x_{2i}(kh) \\ x_{2i}(kh) \end{bmatrix},$$

for $i = 1, \cdots, m$, and $k = 0, 1, \cdots, K$. Similarly, for control components

$$\begin{bmatrix} u_k^{(1i)} \\ u_k^{(2i)} \\ u_k^{(3i)} \end{bmatrix} = \begin{bmatrix} u_{1i}(kh) \\ u_{2i}(kh) \\ u_{2i}(kh) \end{bmatrix},$$
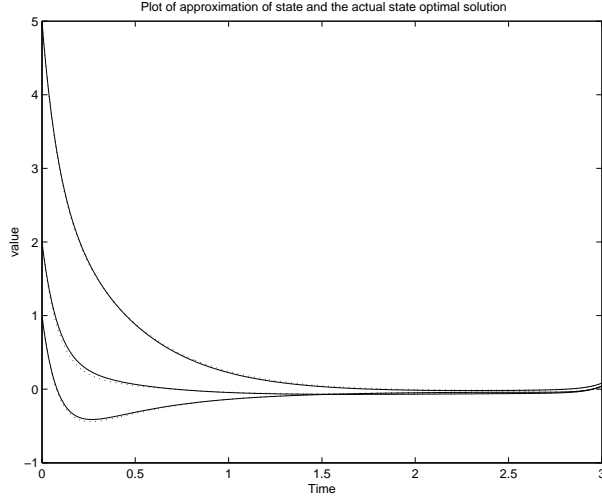
for $i = 1, \cdots, m$, and $i = 0, 1, \cdots, K$.

26

Figure 2: Plot of approximation of optimal state and Optimal state: T = 3

The corresponding discrete-time problem will have the form

$$
\max_{i=1,\dots,m} \sum_{k=1}^{\infty} \{< (x_k - \varphi_k^{(i)}), \tilde{Q}_i(x_k - \varphi_k^{(i)}) > + < (u_k - \phi_k^{(i)}), \tilde{R}_i(u_k - \phi_k^{(i)}) >\} \to \min
$$

(6.1)

$$
x_{k+1} = \tilde{A}x_k + \tilde{B}u_k, x_0 = \gamma_0
$$

where $\tilde{Q}_i = hQ_i$, $\tilde{R}_i = hR_i$, $\tilde{A} = (I + hA)$ and $\tilde{B} = hB$.

We choose $K = 300$. Then using above formulation, we obtain a optimal solution which is very close to the optimal solution to the continuous case. As it is shown in the Figure 3-4, we take the difference pointwise of the optimal solutions obtained from the continuous and discrete formulation. The result is quite good. The two optimal solutions are almost identical.
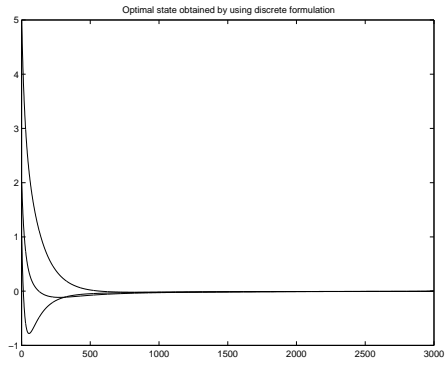
6.2 Example 3: Discrete-time linear-quadratic control problem with quadratic constraints

In this section we solve linear-quadratic control problem with quadratic constraints using discrete formulation. Let us first consider the checking feasibility problem (5.4)-(5.6) in the discrete formulation of the linear-quadratic control problem with quadratic constraints where Z, c, W_i, y_i for $i = 1, ..., m$ will be the same as in the Example 1, and t_i be given by the following table.

i	1	2	3	4	5	6	7	8
t_i	41	20	39	30	32	38	43	45

It is easy to see that there exists a feasible solution to the problem (5.1)-(5.3) if and only if the optimal value of the problem (5.4)-(5.6) is negative.

3.1) Discrete-time optimal state

Optimal state obtained by using discrete formulation

3.2) Difference between discrete-time and continuous-time Optimal state

Pointwise difference of state spaces between discrete and continuous formulations
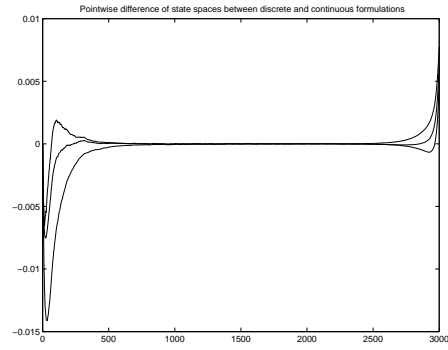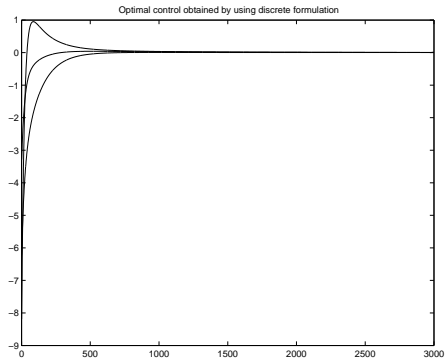
Figure 3: Plot of discrete-time optimal state for Multi-criteria Linear Quadratic Control Problem.

4.1) Discrete-time optimal control

Optimal control obtained by using discrete formulation

4.2) Difference between discrete-time and continuous-time Optimal control

Pointwise difference in control spaces between discrete and continuous formulations
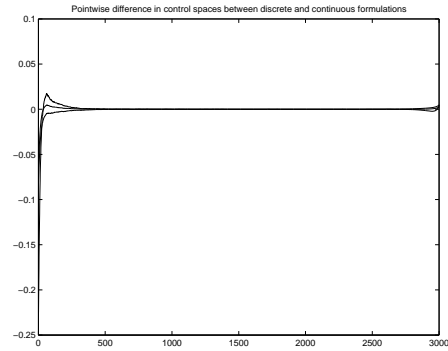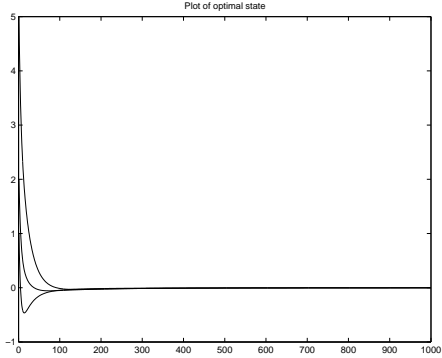
Figure 4: Plot of discrete-time optimal control for Multi-criteria Linear Quadratic Control Problem.

28

5.1) Plot of optimal state for T = 20.   5.2) Plot of optimal control for T = 20.
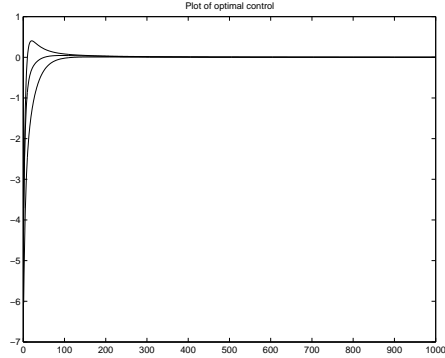

Figure 5: Plot of Discrete-time linear-quadratic control problem with quadratic constraints.

Then by applying a primal-dual algorithm described in [10] based on NT-direction for $T = 3$, we obtain the following data of optimal value t after some number of iterations.

#iterations	1	2	3	4	5	6	7	8
Opt Value	1.87	15.91	1.86	0.91	-0.16	-0.42	-0.49	-0.50

As it is shown in the above table, after the forth iteration we obtain a feasible solution to the linear-quadratic control problem with quadratic constraints(5.1)-(5.3). Once we obtain a feasible solution for the optimization problem (5.1)-(5.3), we then can apply primal-dual algorithms to solve the problem. The calculations for the NT-direction for this problem are quite similar to the calculation for (6.1). The only differences are in the definition of a closed vector subspace X and its complement $X^{\perp}$. The optimal solution to the problem (5.1)-(5.3) for $T = 20$ based on NT-direction is shown in the Figure 5.

29

References

[1] F. Alizadeh and D. Goldfarb, *Second-order cone programming*, ISMP 2000, Part 3 (Atlanta, GA). Math. Program. 95(2003), no.1, Ser. B, 3-51.

[2] A. Ben-tal and A. Nemirovski, *Lecture on Modern Convex Optimization*, SIAM, Philadelphia, 2001.

[3] L. Faybusovich, *Euclidean Jordan algebras and interior-point algorithms*, Positivity, Vol. 1(1997), pp. 331-357.

[4] L. Faybusovich, *A Jordan Algebric approach to potential-reduction algorithms*, Mah. Zeitshrift, vol. 239 (2002) 117-129.

[5] L. Faybusovich, *Linear systems in Jordan algebras and primal-dual interior point algorithms*, Journal of Computational and Applied Mathematics, Vol. 86(1997), pp.149-175.

[6] L. Faybusovich and R. Arana, *A long-step Primal-dual algorithm for symmetric programming problem* Lie theory and its applications in control (Wurzburg, 1999). Systems control Lett. 43(2001), no.1 3-7.

[7] L. Faybusovich and J.B. Moore, *Infinite-Dimensional quadratic Optimization: Interior-Point Methods and control Applications*, Vol.36(1997), p.p. 43-66.

[8] L. Faybusovich and T. Mouktonglang, *Finite-rank Pertubation of the Linear-Quadratic Control Problem.* Proceedings of the American Control Conference, Denver, Colorado, June 4-6, 2003. pp. 5347-5350.

[9] L. Faybusovich and T. Mouktonglang, *Linear-quadratic control problem with a linear term on semiinfinite interval:theory and applications.* Preprint (2003).

[10] L.Faybusovich and T.Tsuchiya, *Primal-dual Algorithms and Infinite-dimensional Jordan Algebras of finite rank.* New trends in optimization and computational algorithms(NTOC 2001)(Kyoto). Math. Progam. 97 (2003), no.3, Ser. B, 471-493.

[11] H. Kwakernaak and R. Sivan, *Linear Optimal Control Systems*, New York, Wiley Interscience (1972).

[12] P. Lancaster and L. Rodman, *Algebraic Riccati Equations*, Oxford Science Publications, 1995.

[13] M. Muramatsu, *On a commutative class of search direction for linear programming over symmetric cones*, Journal of Optimization Theory and its Applications, Vol.112, No.3(2002)595-625.

[14] R.D.C. Monteiro and T. Tsuchiya, *Polynomial convergence of primal-dual algorithms for the second-order cone program base on the MZ-family of direction* Mathematical Programming, 88(1), 61-83, 2000.

[15] Y.E. Nesterov and M. Todd, *Self-scaled barriers and interior-point methods for convex programming*, Mathematics of Operations Research, vol.22(1997),pp.1-42.

[16] S. Schmieta and F. Alizadeh, *Extensions of primal-dual interior point algorithms to symmetric cone*, Math. Program. 96(2003), no. 3, Ser. A, 409-438.

[17] D. Serre, *Matrices Theory and Applications*, Springer, pp. 66.

[18] T. Tsuchiya, *A convergence analysis of the scaling-invariant primal-dual path-following algorithms for second-order cone programming* Optimization Methods and Software, 10/11, 141-182,1999.

[19] S. Wright, *Primal-dual Interior Point Algorithms*, SIAM Publications, Philidelphia, Pennsylvania, USA, 1997.

[20] V.A. Yakubovich, *Nonconvex optimization problem: The infinite-horizon linear-quadratic control problem with quadratic constraints.* System and Control Letters, 19:13-22,1992.