A dynamic programming approach to the linear impulse control synthesis problem [1]

A. N. Daryin, A. B. Kurzhanski, A. V. Seleznev

Faculty of Computational Mathematics & Cybernetics
Moscow State (Lomonosov) University
1 Vorobievy gory, Moscow 119992, Russia

Abstract

The linear impulse control problem is studied via dynamic programming techniques. It is approached by introducing a hard bound on control, with a diameter of this bound tending to infinity. A new definition of a feedback strategy for such problem is proposed, which takes into account the fact that in actual systems control values are large but still bounded, thus justifying the use of additional hard bound. A special attention is given to solving two-dimensional problems. For such problems the structure and properties of optimal control are studied, and optimal feedback strategy is constructed.

Key words: Impulse Control, Dynamic Programming, Control Synthesis

1 Introduction

Consider the problem

$$\begin{cases} J(u(\cdot)) = \operatorname*{Var}_{[t_0,t_1]} U(\cdot) + \phi(x(t_1+0)) \to \inf, \\ \mathrm{d}x(t) = A(t)x(t)\,\mathrm{d}t + B(t)\,\mathrm{d}U(t), \quad t \in [t_0,t_1], \\ x(t_0-0) = x_0. \end{cases} \tag{1}$$

Email addresses: `daryin@mail.ru` (A. N. Daryin), `kurzhans@mail.ru` (A. B. Kurzhanski), `seleznev@hotbox.ru` (A. V. Seleznev).

[1] This work is supported by Russian fund for basic research (grant 03-01-00663). It has been realized within the programs "State support of the leading scientific schools" (1889.2003.1) and "Russian Universities — basic research" (UR.03.03.036).

Here $x(t) \in \mathbb{R}^n$, $U(\cdot) \in BV([t_0, t_1]; \mathbb{R}^m)$. $\phi(x)$ is a convex terminal function; in particular one may choose $\phi(x) = \delta(x \mid \{x_1\})$ to get a problem of moving the system from point x_0 at time t_0 to point x_1 at time t_1.

The problems of this kind have been studied thoroughly already (see [1–4]). They may be solved using the maximum principle for impulse control, and the solution is an open-loop control program. However in the present paper we shall be interested in expressing the solution in terms of dynamic programming and finding the optimal control in the form of feedback control.

2 The Dynamic Programming Approach

Denote the optimal value of the problem (1) through $V(t_0, x_0) = V(t_0, x_0; t_1, \phi(\cdot))$. This problem may be decomposed into a pair of subproblems: first, find the optimal right end x_1 of the trajectory, and second, find optimal control $U(\cdot)$ under condition $x(t_1+0) = x_1$. The solution of the latter is known [2,5], which allows the following representation:

$$V(t_0, x_0) = \inf_{x_1 \in \mathbb{R}^n} \left\{ \phi(x_1) + \sup_{p \in \mathbb{R}^n} \frac{\langle p, x_1 - X(t_1, t_0)x_0 \rangle}{\|B'(t)X'(t_1, \cdot)p\|_{C[t_0,t_1]}} \right\}. \tag{2}$$

Here $X(t, \tau)$ is the solution to the matrix differential equation $\partial X / \partial t = A(t)X$, $X(\tau, \tau) = I$.

We then define a semi-norm $\|p\|_{[t_0,t_1]} = \|B'(t)X'(t_1, \cdot)p\|_{C[t_0,t_1]}$ and rewrite (2) as

$$V(t_0, x_0) = \sup_{p \in \mathbb{R}^n} \left[\langle p, X(t_1, t_0)x_0 \rangle - \phi^*(p) - \delta\Big(p \;\Big|\; \mathcal{B}_{\|\cdot\|_{[t_0,t_1]}}\Big) \right]. \tag{3}$$

where $\mathcal{B}_{\|\cdot\|_{[t_0,t_1]}}$ is the unit ball in the introduced semi-norm. Relation (3) shows that $V(t, x)$ is a convex function and its conjugate is given by

$$V^*(t_0, p) = \phi^*(X'(t_0, t_1)p) + \delta\Big(X'(t_0, t_1)p \;\Big|\; \mathcal{B}_{\|\cdot\|_{[t_0,t_1]}}\Big). \tag{4}$$

Using (4) it is easy to prove the following result:

Theorem 1. *The value function $V(t, x; t_1, \phi(\cdot))$ of the problem (1) satisfies the optimality principle in the form of the semigroup property:*

$$V(t_0, x_0; t_1, \phi(\cdot)) = V(t_0, x_0; \tau, V(\tau, \cdot; t_1, \phi(\cdot))), \quad \tau \in [t_0, t_1],$$

and it is the viscosity solution [6] to the Hamilton–Jacobi–Bellman equation:

$$\min\{H_1(t,x,V_t,V_x),\ H_2(t,x,V_t,V_x)\} = 0, \tag{5}$$

$$V(t_1,x) = V(t_1,x;t_1,\phi(\cdot)).$$

$$H_1(t,x,\xi_t,\xi_x) = \xi_t + \langle \xi_x, A(t)x\rangle, \quad H_2(t,x,\xi_t,\xi_x) = \min_{u\in S_1}\langle \xi_x, B(t)u\rangle + 1.$$

Note that in general case $V(t_1,x;t_1,\phi(\cdot)) \le \phi(x)$, because from (4) it follows that

$$V^*(t_1,p) = \phi^*(p) + \delta(\,B(t_1)p \mid \mathcal{B}_1).$$

Due to (5), in any position (t,x) there are two possible cases for the control. Either $H_1 = 0$, and control may choose $dU(t) = 0$; or $H_1 > 0$, in which case necessarily $H_2 = 0$, and control has to jump in the direction $-B'(t)V_x$. The magnitude of the jump is to be chosen in such a way that after the jump again $H_1 = 0$.

3 The Double Constraint Approach

Let us introduce a hard bound on the control, $u(t) \in \mathcal{B}_\mu$, and consider the corresponding problem:

$$\begin{cases} J(u(\cdot)) = \displaystyle\int_{t_0}^{t_1} \|u(t)\|\,\mathrm{d}t + \phi(x(t_1)) \to \inf, \\ \quad \dot{x}(t) = A(t)x(t) + B(t)u(t), \quad t\in[t_0,t_1], \quad x(t_0) = x_0, \\ \|u(t)\| \le \mu. \end{cases} \tag{6}$$

The value function $V_\mu(t_0,x_0) = V_\mu(t_0,x_0;t_1,\phi(\cdot))$ of this problem is the viscosity solution to the Hamilton–Jacobi–Bellman equation

$$\frac{\partial V_\mu}{\partial t} + \min_{u\in\mu\mathcal{B}_1}\left\{\left\langle \frac{\partial V_\mu}{\partial x}, A(t)x(t) + B(t)u\right\rangle + \|u\|\right\} = 0 \tag{7}$$

with the initial condition $V_\mu(t_1,x) = \phi(x)$. Thus, except some degenerate cases control only takes values from $S_\mu \cup \{0\}$.

The solution of (6) is given by

$$V_\mu(t,x) = \sup_{p\in\mathbb{R}^n}\Big\{\langle p, X(t_1,t_0)x_0\rangle - \\ - \mu\int_t^{t_1} (\|B'(\tau)X'(t_1,\tau)p\| - 1)_+\,\mathrm{d}\tau - \phi^*(p)\Big\}, \tag{8}$$

and its conjugate function in x is

$$V_\mu^*(t,p) = \phi^*(p) + \mu \int_t^{t_1} \left(\|B'(\tau)X'(t_1,\tau)p\| - 1\right)_+ \mathrm{d}\tau. \tag{9}$$

Note that as μ tends to infinity, expressions (8), (9) turn into (3) and (4) respectively. It may further be shown under certain conditions that

$$0 \le V_\mu(t,x) - V(t,x) = O(\mu^{-1}).$$

The HJB equation (5) may be also derived in the limit from (7) as $\mu \to \infty$.

The optimal feedback control strategy is the minimizer in (7), and in points of the differentiability of $V_\mu(t,x)$ it may be written as follows:

$$\mathcal{U}_\mu^*(t,x) = \begin{cases} *0, \|\zeta\| < 1; \\ [0, -\mu\zeta], \|\zeta\| = 1; \\ \left\{-\mu \dfrac{\zeta}{\|\zeta\|}\right\}, \|\zeta\| > 1, \end{cases} \qquad \zeta = B'(t)\frac{\partial V_\mu}{\partial x}. \tag{10}$$

However, in (10) it is not possible to proceed to the limit as in (8) and (7). In particular, it is not clear what the “synthesized” system for the problem (1) will look like. To avoid this difficulty, we introduce the following definition of control synthesis for this problem.

Definition 2. A pair of functions $\mathfrak{U} = \{u(t,x;\mu),\, \theta(t,x;\mu)\}$, such that

$$\begin{aligned} &u(t,x;\mu) \in S_1 \cup \{0\}, && u(t,x;\mu) \underset{\mu\to\infty}{\to} u_\infty(t,x), \\ &\theta(t,x;\mu) \ge 0, && \mu\theta(t,x;\mu) \underset{\mu\to\infty}{\to} m_\infty(t,x), \end{aligned}$$

is called the *feedback control strategy for* (1).

Definition 3. Fix a control strategy $\mathfrak{U}$, number $\mu > 0$ and a partition $t_0 = \tau_0 < \tau_1 < \ldots < \tau_s = t_1$ of interval $[t_0, t_1]$. An *approximating motion* of system (1) is the solution to the differential equation

$$\begin{aligned} \tau_i^* &= \tau_i \wedge \theta(\tau_{i-1}, x_\Delta(\tau_{i-1}); \mu), \\ \dot{x}_\Delta(\tau) &= \mu B(t) u(\tau_{i-1}, x_\Delta(\tau_{i-1}); \mu), \quad \tau_{i-1} < \tau < \tau_i^*, \\ x_\Delta(\tau_i) &= x_\Delta(\tau_i^*). \end{aligned}$$

Number $\sigma = \max\{\tau_i - \tau_{i-1}\}$ is the *diameter* of the partition.

Definition 4. *Constructive motion* of system (1) under feedback control $\mathfrak{U}$ is a piecewise continuous function $x(t)$, which is the pointwise limit of approximating motions $x_\Delta(t)$ as $\mu \to \infty$ and $\sigma \to 0$.

Example 5. Suppose that current position of (1) is $(\bar{t}, \bar{x})$, and from (5) it follows that control has a jump $\bar{r}\delta(t-\bar{t})$. Then the corresponding feedback strategy values $\bar{u} = u(\bar{t}, \bar{x}; \mu)$, $\bar{\theta} = \theta(\bar{t}, \bar{x}; \mu)$ is to be chosen in such a way that the following equality would hold:

$$B(t)\bar{r} = \mu \left[\int_{\bar{t}}^{\bar{t}+\bar{\theta}} B(t)\,\mathrm{d}t \right] \bar{u}.$$

In the limit this yields

$$u(\bar{t}, \bar{x}; \mu) \underset{\mu\to\infty}{=} \frac{\bar{r}}{\|\bar{r}\|}, \qquad \mu\theta(\bar{t}, \bar{x}; \mu) \underset{\mu\to\infty}{=} \|\bar{r}\|.$$

That is, an impulse $\bar{r}\delta(t-\bar{t})$ is (approximately) replaced by a plateau of magnitude μ, direction $\bar{r}$ and duration $\mu^{-1}\|\bar{r}\|$.

4 The Two-dimensional Case

In this section we shall study a special case of (1), namely two-dimensional stationary system with a scalar control:

$$\begin{cases} \mathrm{d}x(t) = Ax(t)\,\mathrm{d}t + b\,\mathrm{d}U(t), \quad t \in [t_0, t_1], \\ x(t_0 - 0) = x_0, \quad x(t_1 + 0) = x_1, \\ x(t) \in \mathbb{R}^2, \quad b \in \mathbb{R}^2, \quad U(t) \in \mathbb{R}^1, \\ \operatorname*{Var}_{[t_0,t_1]} U(\cdot) \le \mu. \end{cases} \tag{11}$$

For such systems time-optimal and energy-optimal control problems are considered. It is possible to construct an explicit form of optimal control and to study so-called *switching surfaces* on which optimal control has a jump and controlled trajectory is discontinuous. These results are based on geometry of attainability sets and on the theorem about impulse control structure.

Theorem 6 (Structure of attainability and solvability domains). *Let $\mathcal{X}_\mu[t; t_0, \mathcal{X}_0]$ be the attainability set for system* (11) *and $\mathcal{W}_\mu[t; t_1, \mathcal{X}_1]$ be the solvability set for this system. Then these sets can be represented in the following form:*

$$\mathcal{X}_\mu[t; t_0, \mathcal{X}_0] = X(t, t_0)\mathcal{X}_0 + \mu \operatorname{conv} \bigcup_{\tau\in[t_0,t]} X(t,\tau)B(\tau)\mathcal{B}_1,$$
$$\mathcal{W}_\mu[t; t_1, \mathcal{X}_1] = X(t, t_1)\mathcal{X}_1 + \mu \operatorname{conv} \bigcup_{\tau\in[t,t_1]} X(t,\tau)B(\tau)\mathcal{B}_1.$$

Theorem 7 (Control structure). *Consider* $x_0 \in \mathcal{W}_\mu[t_0;t_1,\{x_1\}]$. *Then there exists an impulse control* $U(t)$ *such that*

$$\frac{\mathrm{d}U}{\mathrm{d}t}(t) = \sum_{i=1}^{k} h_i\delta(t-s_i), \quad h_i \in \mathbb{R}^m, \quad s_i \in [t_0,t_1],$$

which translates system (11) *from the state* $\{t_0-0,x_0\}$ *to the state* $\{t_1+0,x_1\}$ *and*

$$\operatorname*{Var}_{[t_0,t_1]} U(\cdot) = \sum_{i=1}^{k} \|h_i\| \le \mu.$$

Here k *is a number of impulses, and* $k \le n$ *where* n *is a dimension of system. For systems on plane it is possible to control with only two impulses.*

Consider system (11) with time-optimal control problem, i.e. $t_1 - t_0 \to \min$. Without any loss of generality we shall take $x_1 = 0$. If it is possible to move this system from the state x_0 to the origin, it is also possible to do this only with two impulses. Thus the optimal control can be represented as

$$\frac{dU}{dt}(t) = h_1\delta(t-s) + h_2\delta(t-t_1), \quad h_1,h_2 \in \mathbb{R}^1, \quad s \in [t_0,t_1).$$

We will study a special case when $s = t_0$, i.e. when optimal control has an impulse at the first instant. The set of such states x^0 will be referred to as the *switching set* $\mathcal{J}_t[\mu]$. It is convenient to represent $\mathcal{J}_t[\mu]$ as a union of sets of equal impulses $[r(h_1,h_2)]$. By definition, when $x^0 \in [r(h_1,h_2)]$, optimal control from x^0 to the origin has impulses h_1 at the first time and h_2 at the last time. So that if $T_0(x^0)$ is an optimal time to transfer the system (11) from the state x^0 to the origin, set of equal impulses is

$$r(h_1,h_2) = \left\{ x \in \mathbb{R}^n \;\middle|\; x + h_1 b + h_2\, \mathrm{e}^{-T_0(x)A}\, b = 0 \right\}$$

Lemma 8. *All systems* (11) *are divided into two classes:*

(1) when matrix A *has real eigenvalues of different sign, and*
(2) all other systems.

In case 1) $r(h_1,h_2) = \left\{ -h_1 b - h_2\, \mathrm{e}^{-\tau A}\, b \;\middle|\; \tau \ge 0 \right\}$. *In case 2)*

$$r(h_1,h_2) = \begin{cases} -h_1 b - h_2\, \mathrm{e}^{-\tau A}\, b \tau \ge 0, h_1h_2 < 0; \\ \emptyset, h_1h_2 \ge 0. \end{cases}$$

Theorem 9. *Switching set structure The switching set* $\mathcal{J}_t[\mu]$ *is the union of equal impulses sets*

$$\mathcal{J}_t[\mu] = \bigcup_{\substack{|h_1|+|h_2|=\mu \\ h_1 \ne 0}} r(h_1,h_2)$$

After calculating the switching set, the optimal control synthesis is expressed in the following form. When $x_0 \in r(h_1, h_2)$, for some h_1 and h_2 such that $|h_1| + |h_2| = \mu$, then there is a jump $h_1\delta(t - t_0)$. Otherwise, that is when $x_0 \notin \mathcal{J}_t[\mu]$, there is no jump at time t_0.

Similar result can be proven for energy-optimal control problem. It simply follows from the theorem about switching set structure in case of time-optimal problem.

Theorem 10 (Switching set structure for energy-optimal control problem). *A switching set $\mathcal{J}_\mu[t]$ has the following representation:*

- *if $t > t^*$ then $\mathcal{J}_\mu[t] = \emptyset$;*
- *If $t \leq t^*$ then $\mathcal{J}_\mu[t] = \left\{-h_1 b - h_2\, \mathrm{e}^{-tA}\, b\right\}$ where h_1 and h_2 must have the same sign in case 2) from lemma 8 and arbitrary signs in case 1).*

The time parameter t^* may be calculated directly from system (11).

Control synthesis strategy is also similar to the one in case of time-optimal system.

References

[1] A. B. Kurzhanski, Control and Observation under Uncertainty, Nauka, Moscow, 1977, in Russian.

[2] N. N. Krasovski, The Theory of Motion Control, Nauka, Moscow, 1968, in Russian.

[3] A. B. Kurzhanski, Optimal systems with impulse controls, in: Differential Games and Control Problems [7].

[4] M. I. Gusev, On optimal control of generalized processes under non-convex state constraints, in: Differential Games and Control Problems [7].

[5] A. B. Kurzhanski, Yu. S. Osipov, On controlling linear systems through generalized controls, Differenc. Uravn. 5 (8), in Russian.

[6] M. G. Crandall, P.-L. Lions, Viscosity solutions of Hamilton–Jacobi equations, Transactions of American Mathematical Society 277 (1983) 1–41.

[7] Differential Games and Control Problems, UNC AN SSSR, Sverdlovsk, 1975.