ТЕОРІЯ ОПТИМАЛЬНИХ РІШЕНЬ

УДК 519.8

А.В. РУДЕНКО

ОБ ОДНОМ СВОЙСТВЕ ФУНКЦИИ БЕЛЛМАНА В ЗАДАЧЕ ОПТИМАЛЬНОГО БЫСТРОДЕЙСТВИЯ

Показано, что в двумерной задаче о тележке функция Беллмана $T = T(x, \dot{x})$ обладает свойством

$$T(k^2 x, k\dot{x}) = kT(x, \dot{x}),\ k > 0.$$

Получено также явное выражение

$$T = -\frac{2p_1 x_0 + p_2 \dot{x}_0}{p_1 \dot{x}_0 + \|p_2\|},$$

где (p_1, p_2) - начальное состояние сопряженной системы.

Формулировка задачи. В евклидовой плоскости $\boldsymbol{R}^2$ имеется дифференциальное уравнение второго порядка вида

$$\ddot{\boldsymbol{x}} = \boldsymbol{u}, \quad \|\boldsymbol{u}\| \le 1. \tag{1}$$

которое, согласно закону Ньютона, описывает динамику материальной точки массы $m = 1$ под действием изотропной силы $\boldsymbol{u}$.

Ищется оптимальное (программное) управление $\boldsymbol{u} = \boldsymbol{u}(t)$, которое переводит систему (1) из произвольного начального состояния $(\boldsymbol{x}_0, \dot{\boldsymbol{x}}_0)$ в конечное состояние $(\boldsymbol{x}_1, \dot{\boldsymbol{x}}_1) = (\boldsymbol{0}, \boldsymbol{0})$ за минимальное время T.

Принцип максимума утверждает, что это оптимальное управление существует, единственно и имеет вид

$$\boldsymbol{u} = \frac{\boldsymbol{p}_2 - \boldsymbol{p}_1 t}{\|\boldsymbol{p}_2 - \boldsymbol{p}_1 t\|}, \tag{2}$$

где $(\boldsymbol{p}_1, \boldsymbol{p}_2)$ - заранее неизвестное начальное состояние сопряженной системы

$$\dot{\boldsymbol{\psi}}_1 = \boldsymbol{0}, \quad \boldsymbol{\psi}_1(0) = \boldsymbol{p}_1,$$
$$\dot{\boldsymbol{\psi}}_2 = -\boldsymbol{\psi}_1, \quad \boldsymbol{\psi}_2(0) = \boldsymbol{p}_2.$$

Параметры $\boldsymbol{p}_1, \boldsymbol{p}_2$ определены с точностью до произвольного общего положительного множителя λ, что, очевидно, не влияет на выражение для оптимального управления $\boldsymbol{u}$ ввиду его однородности степени нуль по $(\boldsymbol{p}_1, \boldsymbol{p}_2)$.

Эти параметры находятся из краевых условий

$$\boldsymbol{x}(0) = \boldsymbol{x}_0, \quad \dot{\boldsymbol{x}}(0) = \dot{\boldsymbol{x}}_0,$$
$$\boldsymbol{x}(T) = \boldsymbol{0}, \quad \dot{\boldsymbol{x}}(T) = \boldsymbol{0},$$

которые в исходном интегральном виде записываются следующим образом:

$$\int_0^T \frac{\boldsymbol{p}_2 - \boldsymbol{p}_1 t}{\|\boldsymbol{p}_2 - \boldsymbol{p}_1 t\|} dt = -\dot{\boldsymbol{x}}_0, \tag{3}$$

$$\int_0^T \frac{(\boldsymbol{p}_2 - \boldsymbol{p}_1 t)t}{\|\boldsymbol{p}_2 - \boldsymbol{p}_1 t\|} dt = \boldsymbol{x}_0, \tag{4}$$

где фигурирует также заранее неизвестное время T .

Таким образом, задача нахождения параметров $(\boldsymbol{p}_1, \boldsymbol{p}_2)$ оптимального управления (2) на самом деле связана с нахождением совокупности $(\boldsymbol{p}_1, \boldsymbol{p}_2, T)$. Поэтому важно изучить свойства как $(\boldsymbol{p}_1, \boldsymbol{p}_2)$ в их зависимости от $(\boldsymbol{x}_0, \dot{\boldsymbol{x}}_0)$, так и свойства функции $T = T(\boldsymbol{x}_0, \dot{\boldsymbol{x}}_0)$.

Переход от интегральных соотношений к конечным. Все интегралы в системе (3), (4) берутся в элементарных функциях. Но делать это лучше не в координатах, а в бескоординатной форме, так как задача симметрична относительно вращений. Искомые параметры в итоге будут зависеть от инвариантов типа $\boldsymbol{x}_0^2, \boldsymbol{x}_0\dot{\boldsymbol{x}}_0, \dot{\boldsymbol{x}}_0^2$ (а не от всех четырех компонент векторов $\boldsymbol{x}_0, \dot{\boldsymbol{x}}_0$).

1. Умножим (3) скалярно на $(-\boldsymbol{p}_1)$ и обратим внимание на то, что возникающее в этом случае подынтегральное выражение есть в точности производная

$$\frac{d}{dt}\|\boldsymbol{p}_2 - \boldsymbol{p}_1 t\| = \frac{\boldsymbol{p}_2 - \boldsymbol{p}_1 t}{\|\boldsymbol{p}_2 - \boldsymbol{p}_1 t\|} \cdot (-\boldsymbol{p}_1). \tag{5}$$

Отсюда

$$\|\boldsymbol{p}_2 - \boldsymbol{p}_1 T\| - \|\boldsymbol{p}_2\| = \boldsymbol{p}_1 \dot{\boldsymbol{x}}_0,$$

или, что то же самое,

$$\boldsymbol{p}_1 \dot{\boldsymbol{x}}_0 + \|\boldsymbol{p}_2\| = \|\boldsymbol{p}_2 - \boldsymbol{p}_1 T\|. \tag{6}$$

Уравнение (6) представляет собой следствие универсального свойства гамильтониана $H = \boldsymbol{p}_1 \dot{\boldsymbol{x}}(t) + \|\boldsymbol{p}_2 - \boldsymbol{p}_1 t\|$ быть интегралом движения:

$$H = \boldsymbol{p}_1 \dot{\boldsymbol{x}}(t) + \|\boldsymbol{p}_2 - \boldsymbol{p}_1 t\| = \text{const}.$$

Поскольку это верно для всех $t \in [0, T]$ (и даже для всех $t \in \boldsymbol{R}$), то, в частности, $H(0) = H(T)$. Если теперь дополнительно учесть, что $\dot{\boldsymbol{x}}(T) = \boldsymbol{0}$, то мы приходим к уравнению (6). Поэтому его можно записать сразу, не выполняя интегрирования. Однако важно понимать его происхождение как результат проекции векторного уравнения (3) на направление $(\pm \boldsymbol{p}_1)$.

2. Оказывается, что в двумерном случае, кроме гамильтониана, в задаче о тележке, ввиду ее специфики, имеется еще один интеграл движения. Он интересен тем, что дает сразу же еще одно уравнение относительно $(\boldsymbol{p}_1, \boldsymbol{p}_2)$, причем линейное и без участия T.

Его легко обнаружить, если воспользоваться тем, что в плоскости $\boldsymbol{R}^2$ наряду со скалярным произведением $\boldsymbol{ab}$ пары векторов $(\boldsymbol{a}, \boldsymbol{b})$ можно ввести также их кососкалярное произведение $\boldsymbol{a} \wedge \boldsymbol{b}$, которое обладает свойством $\boldsymbol{a} \wedge \boldsymbol{b} = -\boldsymbol{b} \wedge \boldsymbol{a}$ и имеет следующий смысл (о кососкалярном произведении в $\boldsymbol{R}^{2n}$ см. [1, § 41]).

Сначала первый вектор поворачивается на 90° вперед (т.е. в положительном направлении), а затем берется его скалярное произведение со вторым. Получается ориентированная площадь параллелограмма, построенного на упорядоченной паре векторов $(\boldsymbol{a}, \boldsymbol{b})$. При этом $\boldsymbol{ab}$ и $\boldsymbol{a} \wedge \boldsymbol{b}$ не независимы и связаны соотношением $(\boldsymbol{ab})^2 + (\boldsymbol{a} \wedge \boldsymbol{b})^2 = \boldsymbol{a}^2 \boldsymbol{b}^2$, что следует из тригонометрического тождества $\cos^2\alpha + \sin^2\alpha = 1$.

Рассмотрим выражение

$$\Lambda(t) = \boldsymbol{p}_1 \wedge \boldsymbol{x}(t) + (\boldsymbol{p}_2 - \boldsymbol{p}_1 t) \wedge \dot{\boldsymbol{x}}(t)$$

и убедимся, что его производная тождественно равна нулю. Действительно,

$$\frac{d}{dt}\Lambda(t) = \boldsymbol{p}_1 \wedge \dot{\boldsymbol{x}}(t) + (-\boldsymbol{p}_1 \wedge \dot{\boldsymbol{x}}(t)) + (\boldsymbol{p}_2 - \boldsymbol{p}_1 t) \wedge \ddot{\boldsymbol{x}}(t) =$$
$$= (\boldsymbol{p}_2 - \boldsymbol{p}_1 t) \wedge \frac{(\boldsymbol{p}_2 - \boldsymbol{p}_1 t)}{\|\boldsymbol{p}_2 - \boldsymbol{p}_1 t\|} = 0,$$

поскольку всегда $\boldsymbol{a} \wedge \boldsymbol{a} \equiv 0$. Отсюда следует, что

$$\Lambda(t) = \boldsymbol{p}_1 \wedge \boldsymbol{x}(t) + (\boldsymbol{p}_2 - \boldsymbol{p}_1 t) \wedge \dot{\boldsymbol{x}}(t) = \text{const}.$$

В частности, $\Lambda(0) = \Lambda(T)$. Но $\Lambda(T) = 0$ в силу того, что $\boldsymbol{x}(T) = \boldsymbol{0}$, $\dot{\boldsymbol{x}}(T) = \boldsymbol{0}$. Поэтому и $\Lambda(0) = 0$. Это приводит к линейному однородному уравнению

$$\boldsymbol{p}_1 \wedge \boldsymbol{x}_0 + \boldsymbol{p}_2 \wedge \dot{\boldsymbol{x}}_0 = 0, \tag{7}$$

связывающему векторы $\boldsymbol{p}_1, \boldsymbol{p}_2$ и не содержащему время T.

3. Продолжим переход от интегральных соотношений (3), (4) к конечным. Умножим (4) скалярно на $(-\boldsymbol{p}_1)$. В результате получим

$$\int_0^T \frac{(\boldsymbol{p}_2 - \boldsymbol{p}_1 t)}{\|\boldsymbol{p}_2 - \boldsymbol{p}_1 t\|} \cdot (-\boldsymbol{p}_1 t) = -\boldsymbol{p}_1 \boldsymbol{x}_0 .$$

Для вычисления этого интеграла заметим, что подынтегральное выражение, как следует из (5), имеет вид

$$t\frac{d}{dt}\|\boldsymbol{p}_2-\boldsymbol{p}_1 t\| = \frac{(\boldsymbol{p}_2-\boldsymbol{p}_1 t)}{\|\boldsymbol{p}_2-\boldsymbol{p}_1 t\|}\cdot(-\boldsymbol{p}_1 t).$$

Такой интеграл вычисляется по частям:

$$\int_0^T t\frac{d}{dt}\|\boldsymbol{p}_2-\boldsymbol{p}_1 t\|dt = t\|\boldsymbol{p}_2-\boldsymbol{p}_1 t\|\Big|_0^T - \int_0^T\|\boldsymbol{p}_2-\boldsymbol{p}_1 t\|dt.$$

Для нахождения интеграла от $\|\boldsymbol{p}_2-\boldsymbol{p}_1 t\|$ поступим следующим образом, используя непосредственно специфику задачи.

Рассмотрим функцию

$$L(t)=\boldsymbol{p}_1\boldsymbol{x}(t)+(\boldsymbol{p}_2-\boldsymbol{p}_1 t)\dot{\boldsymbol{x}}(t)$$

и вычислим ее производную $\dot{L}(t)$. Имеем

$$\dot{L}(t)=\boldsymbol{p}_1\dot{\boldsymbol{x}}+(-\boldsymbol{p}_1\dot{\boldsymbol{x}})+(\boldsymbol{p}_2-\boldsymbol{p}_1 t)\ddot{\boldsymbol{x}}(t)=$$

$$=(\boldsymbol{p}_2-\boldsymbol{p}_1 t)\frac{(\boldsymbol{p}_2-\boldsymbol{p}_1 t)}{\|\boldsymbol{p}_2-\boldsymbol{p}_1 t\|}=\|\boldsymbol{p}_2-\boldsymbol{p}_1 t\|.$$

Следовательно,

$$\int_0^T\|\boldsymbol{p}_2-\boldsymbol{p}_1 t\|dt=\int_0^T\dot{L}(t)dt=L(T)-L(0).$$

Однако $L(T)=\boldsymbol{p}_1\boldsymbol{x}(T)+(\boldsymbol{p}_2-\boldsymbol{p}_1 T)\dot{\boldsymbol{x}}(T)=0$, $L(0)=\boldsymbol{p}_1\boldsymbol{x}_0+\boldsymbol{p}_2\dot{\boldsymbol{x}}_0$. Поэтому

$$\int_0^T\|\boldsymbol{p}_2-\boldsymbol{p}_1 t\|dt=-(\boldsymbol{p}_1\boldsymbol{x}_0+\boldsymbol{p}_2\dot{\boldsymbol{x}}_0).$$

Отсюда

$$\int_0^T t\frac{d}{dt}\|\boldsymbol{p}_2-\boldsymbol{p}_1 t\|dt = T\|\boldsymbol{p}_2-\boldsymbol{p}_1 T\| + (\boldsymbol{p}_1\boldsymbol{x}_0+\boldsymbol{p}_2\dot{\boldsymbol{x}}_0).$$

Таким образом,

$$-\boldsymbol{p}_1\boldsymbol{x}_0 = T\|\boldsymbol{p}_2-\boldsymbol{p}_1 T\| + (\boldsymbol{p}_1\boldsymbol{x}_0+\boldsymbol{p}_2\dot{\boldsymbol{x}}_0),$$

или, что то же самое,

$$T\|\boldsymbol{p}_2-\boldsymbol{p}_1 T\| = -(2\boldsymbol{p}_1\boldsymbol{x}_0+\boldsymbol{p}_2\dot{\boldsymbol{x}}_0).$$

Теперь воспользуемся соотношением (6), которое позволяет заменить радикал $\|\boldsymbol{p}_2-\boldsymbol{p}_1 T\|$ на не содержащее T выражение $\boldsymbol{p}_1\dot{\boldsymbol{x}}_0+\|\boldsymbol{p}_2\|$. Окончательно имеем

$$T=-\frac{2\boldsymbol{p}_1\boldsymbol{x}_0+\boldsymbol{p}_2\dot{\boldsymbol{x}}_0}{\boldsymbol{p}_1\dot{\boldsymbol{x}}_0+\|\boldsymbol{p}_2\|}. \tag{8}$$

4. Необходимо найти еще одно, четвертое уравнение, недостающее для определения совокупности параметров $(\boldsymbol{p}_1, \boldsymbol{p}_2, T)$. Оно оказывается трансцендентным (чего, к сожалению, нельзя избежать) и имеет вид

$$-\frac{\boldsymbol{p}_1 \wedge \dot{\boldsymbol{x}}_0}{\boldsymbol{p}_1 \wedge \boldsymbol{p}_2} = \frac{1}{\|\boldsymbol{p}_1\|} \ln \frac{\|\boldsymbol{p}_1\| \|\boldsymbol{p}_2 - \boldsymbol{p}_1 T\| - \boldsymbol{p}_1(\boldsymbol{p}_2 - \boldsymbol{p}_1 T)}{\|\boldsymbol{p}_1\| \|\boldsymbol{p}_2\| - \boldsymbol{p}_1 \boldsymbol{p}_2}. \tag{9}$$

Основной результат. Представим полученные конечные соотношения в виде системы относительно $(\boldsymbol{p}_1, \boldsymbol{p}_2, T)$, дополнив ее условием нормировки $\|\boldsymbol{p}_1\| = 1$:

$$\boldsymbol{p}_1 \dot{\boldsymbol{x}}_0 + \|\boldsymbol{p}_2\| = \|\boldsymbol{p}_2 - \boldsymbol{p}_1 T\|,$$

$$\boldsymbol{p}_1 \wedge \boldsymbol{x}_0 + \boldsymbol{p}_2 \wedge \dot{\boldsymbol{x}}_0 = 0,$$

$$T = -\frac{2\boldsymbol{p}_1 \boldsymbol{x}_0 + \boldsymbol{p}_2 \dot{\boldsymbol{x}}_0}{\boldsymbol{p}_1 \dot{\boldsymbol{x}}_0 + \|\boldsymbol{p}_2\|},$$

$$-\frac{\boldsymbol{p}_1 \wedge \dot{\boldsymbol{x}}_0}{\boldsymbol{p}_1 \wedge \boldsymbol{p}_2} = \frac{1}{\|\boldsymbol{p}_1\|} \ln \frac{\|\boldsymbol{p}_1\| \|\boldsymbol{p}_2 - \boldsymbol{p}_1 T\| - \boldsymbol{p}_1(\boldsymbol{p}_2 - \boldsymbol{p}_1 T)}{\|\boldsymbol{p}_1\| \|\boldsymbol{p}_2\| - \boldsymbol{p}_1 \boldsymbol{p}_2},$$

$$\|\boldsymbol{p}_1\| = 1. \tag{10}$$

Единственность решение этой системы гарантируется принципом максимума Понтрягина и является ее важным свойством. Из этого вытекает полезное следствие, позволяющее понизить размерность задачи при ее численном решении.

Теорема. Пусть $(\boldsymbol{p}_1, \boldsymbol{p}_2, T)$ - единственное решение системы (10), соответствующее заданному начальному состоянию $(\boldsymbol{x}_0, \dot{\boldsymbol{x}}_0)$. Для произвольного положительного числа k рассмотрим преобразование вида $(\boldsymbol{x}_0, \dot{\boldsymbol{x}}_0) \mapsto (k^2 \boldsymbol{x}_0, k\dot{\boldsymbol{x}}_0)$ и обозначим $(\boldsymbol{p}_{1k}, \boldsymbol{p}_{2k}, T_k)$ единственное решение системы, соответствующее начальному состоянию $(k^2 \boldsymbol{x}_0, k\dot{\boldsymbol{x}}_0)$. Тогда

$$\boldsymbol{p}_{1k} = \boldsymbol{p}_1, \quad \boldsymbol{p}_{2k} = k\boldsymbol{p}_2, \quad T_k = kT. \tag{11}$$

Доказательство. Заменим сначала в системе (10) $(\boldsymbol{x}_0, \dot{\boldsymbol{x}}_0)$ на $(k^2 \boldsymbol{x}_0, k\dot{\boldsymbol{x}}_0)$ и, соответственно, $(\boldsymbol{p}_1, \boldsymbol{p}_2, T)$ на $(\boldsymbol{p}_{1k}, \boldsymbol{p}_{2k}, T_k)$. Непосредственная подстановка выражений (11) в полученную систему показывает, что $(\boldsymbol{p}_1, k\boldsymbol{p}_2, kT)$ является ее решением. В частности, $T(k^2 x, k\dot{x}) = kT(x, \dot{x})$, что не зависит от нормировки.

Заключение. Работа является развитием исследований [1, 2]. Полученные результаты показывают, что при численном нахождении $(\boldsymbol{p}_1, \boldsymbol{p}_2, T)$ нет необходимости рассматривать произвольные наборы $(\boldsymbol{x}_0, \dot{\boldsymbol{x}}_0)$. Выбрав, например, $k^{-1} = \|\dot{\boldsymbol{x}}_0\|$, вектор скорости $\dot{\boldsymbol{x}}_0$ всегда можно считать единичным.

О.В. Руденко

ПРО ОДНУ ВЛАСТИВІСТЬ ФУНКЦІЇ БЕЛЛМАНА В ЗАДАЧІ ОПТИМАЛЬНОЇ ШВИДКОДІЇ

Показано, що в двовимірній задачі про візок функція Беллмана $T = T(x, \dot{x})$ має властивість

$$T(k^2 x, k\dot{x}) = kT(x, \dot{x}),\ k > 0 .$$

Отримано також явний вираз

$$T = -\frac{2p_1 x_0 + p_2 \dot{x}_0}{p_1 \dot{x}_0 + \|p_2\|},$$

де (p_1, p_2) - початковий стан спряженої системи.

A.V. Rudenko

ON A PROPERTY OF BELLMAN FUNCTION IN A TIME OPTIMAL CONTROL PROBLEM

In the framework of a 2D Time Optimal Control Problem, the following property of Bellman function is obtained:

$$T(k^2 x, k\dot{x}) = kT(x, \dot{x}),\ k > 0.$$

The explicit expression

$$T = -\frac{2p_1 x_0 + p_2 \dot{x}_0}{p_1 \dot{x}_0 + \|p_2\|},$$

where (p_1, p_2) is an initial state of a conjugate system, is obtained, too.

1. *Арнольд В.И.* Математические методы классической механики. – М.: Наука, 1974. – 432 с.
2. *Руденко А.В.* Об одной задаче оптимального быстродействия // Теорія оптимальних рішень. – Киев: Ин-т кибернетики им. В.М.Глушкова НАН Украины, 2002. – №1. – С. 135-141.
3. *Руденко А.В.* Двумерная тележка: вариационная оценка снизу функции Беллмана // Теорія оптимальних рішень. – Киев: Ин-т кибернетики им. В.М. Глушкова НАН Украины, 2003. – №2. – С. 135-148.

Получено 04.04.2005