Математика–2

ОБЩИЙ ПЛАН РЕШЕНИЯ ЗАДАЧ ОПТИМАЛЬНОГО УПРАВЛЕНИЯ
в непрерывном времени

Материал подготовил ассистент А. Тонис

Представим себе, что имеется динамическая система, которая эволюционирует (изменяет свое состояние с течением времени) по какому-то закону, причем мы можем оказывать влияние на этот процесс. Задача оптимального управления состоит в том, чтобы выбрать такое влияние на систему, при котором она бы вела себя наилучшим для нас образом (с точки зрения выбранного критерия оптимальности).

В качестве базовой задачи, которую будем исследовать, рассмотрим задачу оптимального управления вида

$$\int\limits_0^T f(x(t), u(t), t)\, dt \to \max_{u(\cdot)} \tag{1}$$

$$\text{s. t.} \quad \dot{x} = g(x, u, t); \tag{2}$$

$$u(t) \in U; \tag{3}$$

$$x(0) = x_0,\ x(T) = x_T. \tag{4}$$

Здесь обозначения трактуются следующим образом:

- $x(t) = (x_1(t), \ldots, x_n(t))^\top$ — *фазовая переменная* (state variable), определяющая состояние нашей динамической системы в каждый из моментов времени $t \in [0, T]$. Эволюция системы происходит согласно дифференциальному *уравнению связи* (2). Это, в частности, означает, что переменная x не может претерпевать мгновенных скачков и изменяется лишь постепенно. Для наших целей будет вполне достаточно считать, что $x(t)$ как функция времени является непрерывной и кусочно непрерывно дифференцируемой (ее первая производная непрерывна всегда кроме, может быть, конечного числа моментов, когда может претерпевать конечный скачок).

- $u(t) = (u_1(t), \ldots, u_k(t))^\top$ — *переменная управления* (control variable), оказывающая влияние на эволюцию динамической системы (от нее зависит правая часть уравнения связи (2)). Это та переменная, по которой надо максимизировать функционал (1). В отличие от фазовой переменной, переменная управления может меняться скачкообразно (установка нами того или иного управления не обладает инерцией, руль можно повернуть резко). Говоря формально, $u(\cdot)$ принадлежит классу кусочно непрерывных функций. Кроме того, в каждый момент времени $u(t)$ должно принадлежать множеству допустимых управлений U.

- Следует подчеркнуть, что в отличие от задач нелинейного программирования, изучавшихся в предыдущем модуле, мы теперь имеем дело с задачей *бесконечномерной* оптимизации. Действительно, задать режим управления $u(\cdot)$ — это значит определить $u(t)$ для *каждого* (или почти каждого) $t \in [0, T]$. В задаче (1) максимизируется функционал, т. е. функция от функций; его аргумент — не значение $u(t)$ при каком-то заданном t, а вся совокупность этих значений при $t \in [0, T]$, как бы весь график функции $u(t)$. Чтобы это подчеркнуть, обычно используют обозначение $u(\cdot)$ для u как аргумента функционала. Множество, по которому производится максимизация, т. е. множество всевозможных $u(\cdot)$, не укладывается ни в какое линейное пространство конечной размерности.

- Отображения $f : \mathbf{R}^n \times U \times [0, T] \to \mathbf{R}$ и $g : \mathbf{R}^n \times U \times [0, T] \to \mathbf{R}^n$ ($g = (g_1, \ldots, g_n)^\top$)предполагаются непрерывно дифференцируемыми по x. Заметим, что на всяком замкнутом ограниченном подмножестве $\mathbf{R}^n$ все g_i удовлетворяют условию Липшица, что гарантирует однозначную разрешимость системы дифференциальных уравнений (2).

- Граничные данные x_0 и x_T, а также сам конечный момент времени T, считаются фиксированными.

Предлагается следующий план решения задачи (1)–(4):

1. Выписать *гамильтониан* задачи (другое название — функция Понтрягина):

$$H(x,u,\pi,t) \stackrel{\mathrm{def}}{=} f(x,u,t) + \pi g(x,u,t), \tag{5}$$

   где $\pi = (\pi_1,\ldots,\pi_n)$ — новая, специально введенная *двойственная переменная* (costate variable). Как и x, π может меняться лишь плавно (принадлежит классу непрерывных кусочно непрерывно дифференцируемых функций). До некоторой степени H является аналогом функции Лагранжа, а π — аналогом множителей Лагранжа для ограничений (2). Поскольку уравнение (2) должно выполняться для всех $t \in [0,T]$, то этих ограничений, на самом деле, бесконечно много (континуум). Соответственно, множителей тоже бесконечно много: $\pi(t)$ надо определить при всех t, так что уместно обозначение $\pi(\cdot)$.

2. Выписать *сопряженное уравнение*:

$$\dot{\pi} = -\frac{\partial H}{\partial x}. \tag{6}$$

3. Получить условие оптимальности по u, воспользовавшись *принципом максимума*, который гласит, что в оптимуме гамильтониан (5) должен достигать максимума по u при всех $t \in [0,T]$ и при фиксированных остальных переменных:

$$u(t) \in \operatorname{Arg}\max_{u' \in U} H(x(t),u',\pi(t),t) \qquad (t \in [0,T]). \tag{7}$$

   В частности, если функция H дифференцируема по u и максимум в (7) достигается *внутри* множества U, то условие первого порядка для максимизации в (7) имеет вид

$$\frac{\partial H}{\partial u} = 0. \tag{8}$$

   Следует помнить, однако, что максимум может достигаться и на границе U, и тогда в точке оптимума равенство (8) не выполнено. В таких случаях типична ситуация, когда оптимальное u задано кусочно, т. е. вычисляется по разным формулам при разных x,π,t. Скажем, при $(x,\pi,t) \in D_1$ будет одна формула, при $(x,\pi,t) \in D_2$ — другая и т. д. Если так, то все дальнейшие действия придется производить отдельно для каждой из областей $D_1, D_2, \ldots$, причем всякий раз надо проверять, действительно ли найденный кусок решения содержится в соответствующем D_i (точнее, требуется установить, при каких значениях времени, переменных и параметров это имеет место), и затем состыковать результаты (см. п. 6 ниже).

4. Разрешить (7) относительно u и подставить в системы дифференциальных уравнений (2) и (6), избавившись таким образом от явного участия в них переменной управления.

5. Получилась замкнутая система из $2n$ обыкновенных дифференциальных уравнений первого порядка с $2n$ неизвестными функциями (x,π). Надо ее решить, а если не выходит сделать это аналитически, то хотя бы провести качественный анализ, например, построить фазовую диаграмму в пространстве (x,π) или в других удобных координатах. Общее решение должно зависеть от $2n$ констант интегрирования.

6. Теперь осталось подключить $2n$ граничных условий (4) и с их помощью определить константы интегрирования, тем самым получив окончательное решение задачи (1)–(4).

   Впрочем, выполнение этого пункта несколько усложняется, если решение имеет кусочный вид, т. е. есть области $D_1, D_2, \ldots$, в которых общее решение задается по-разному (см. п. 3). В этом случае надо понять, через какие области проходит оптимальная траектория, и состыковать решения на границах областей. Что это означает? Пусть, например, есть две граничащих области D_1 и D_2, в которых зависимость от времени одномерной фазовой переменной x задается, соответственно, формулами

$$x = x^{(1)}(t,t^*,A) \qquad (x \in D_1); \tag{9}$$

$$x = x^{(2)}(t,t^*,B) \qquad (x \in D_2), \tag{10}$$

   где A и B — неизвестные константы интегрирования, а t^* — момент времени (тоже неизвестный), в который происходит переход из одной области в другую. В силу непрерывности x по t

$x(t^*)$ может вычисляться как с помощью (9), так и с помощью (10), результат будет один и тот же. Таким образом, верно равенство

$$x^{(1)}(t^*, t^*, A) = x^{(2)}(t^*, t^*, B). \tag{11}$$

Уравнение (11) позволяет исключить одну из трех констант A, B, t^*. Остальные две определяются из граничных условий (4). Остается только проверить, что переход из одной области в другую действительно происходит, т. е. $0 \leq t^* \leq T$.

7. Принцип максимума — это набор *необходимых* условий первого порядка на *локальный* максимум, поэтому, вышеописанный метод решения задачи (1)–(4), вообще говоря, дает нам управления, подозрительные на оптимум (некоторые из них могут не доставлять максимум функционалу). Вот *достаточное* условие, гарантирующее, что найденное решение $(x(\cdot), u(\cdot), \pi(\cdot))$ действительно является глобальным оптимумом: множество U должно быть выпуклым, а функция $H(x, u, \pi(t), t)$ должна быть вогнутой по совокупности переменных (x, u) при каждом $t \in [0, T]$ (здесь $\pi(t)$ задается найденным решением, а x и u могут меняться). Это довольно сильное требование (принцип максимума дает правильное решение и при более слабых предположениях), но во многих экономических приложениях оно выполнено.

Выше был намечен лишь один из возможных подходов к проблеме обработки условий оптимальности и стыковки кусочно заданных решений. Вы можете делать то же самое любым удобным для вас способом.

Помимо базовой задачи (1)–(4), в вашей практике могут встретиться другие варианты задачи оптимального управления в непрерывном времени. Рассмотрим некоторые из них.

- **Задача со свободным правым концом.** Пусть ограничение на правом конце $x(T) = x_T$ (см. (4)) снято. Тогда множество допустимых планов управления становится шире (теперь подходят и те управления, в результате которых $x(T) \neq x_T$). Следовательно, может измениться и оптимальное управление. В качестве критерия оптимальности для правого конца выступает *условие трансверсальности*:

  $$\pi(T) = 0 \tag{12}$$

  Оно заменяет жесткое граничное условие $x(T) = x_T$ в системе уравнений для определения констант интегрирования (см. выше).

- **Задача с односторонним ограничением на правом конце.** Пусть, например, x — одномерная величина и ограничение на правом конце имеет вид $x(T) \geq a$. Тогда соответствующее условие трансверсальности является в каком-то смысле аналогом условий дополняющей нежесткости. А именно, нужно, чтобы выполнялись условия

  $$\begin{aligned} &x(T) - a \geq 0; \\ &\pi(T) \geq 0; \\ &\pi(T)(x(T) - a) = 0, \end{aligned} \tag{13}$$

  т. е. либо ограничение $x(T) \geq a$ активно (выполняется как равенство), и тогда соответствующий множитель неотрицателен, либо оно неактивно, и тогда выполняется (12).

- **Задача с терминантом.** Иногда вместо (1) решают задачу максимизации функционала

  $$\int_0^T f(x(t), u(t), t)\, dt + \psi(x(T)) \to \max_{u(\cdot)}, \tag{14}$$

  где $\psi(x)$ — вогнутая функция от x (разумеется, $x(T)$ не фиксировано). Такую задачу следует решать так же, как задачу без терминального члена $\psi(x(T))$ (в частности, слагаемое $\psi(x(T))$ не включается в гамильтониан), с той лишь разницей, что вместо (12) используется условие трансверсальности

  $$\pi(T) = \psi'(x(T)). \tag{15}$$

Если дополнительно вводится ограничение $x(T) \geq a$, то (12) надо заменить на

$$\begin{aligned} & x(T) - a \geq 0; \\ & \pi(T) - \psi'(x(T)) \geq 0; \\ & \Big(\pi(T) - \psi'(x(T))\Big)(x(T) - a) = 0. \end{aligned} \tag{16}$$

- **Условия трансверсальности для левого конца** $t = 0$ выглядят аналогично и получаются из условий для правого конца после замены $x(T)$ на $x(0)$ и $\pi(T)$ на $-\pi(0)$.

- **Задача с бесконечным горизонтом планирования.** Если в (1) $T = \infty$, то под решением такой задачи с бесконечным горизонтом планирования обычно понимается предел решений задач с конечным T при $T \to \infty$ (если он существует). Соответственно трактуются и условия трансверсальности. Так, для ограничения $x(\infty) \geq a$ условия трансверсальности имеют вид

$$\begin{aligned} & \lim_{t\to\infty} (x(t) - a) \geq 0; \\ & \lim_{t\to\infty} \pi(t) \geq 0; \\ & \lim_{t\to\infty} \pi(t)(x(t) - a) = 0. \end{aligned} \tag{17}$$

- **Задача с переменным горизонтом планирования.** Иногда (например, в задачах оптимального быстродействия) момент окончания процесса T является одной из переменных и подлежит оптимизации. Соответствующее условие трансверсальности (условие оптимального выбора T) имеет вид

$$H(x(T), u(T), \pi(T), T) = 0. \tag{18}$$

Впрочем, условие (18) является лишь необходимым и не гарантирует оптимальности даже для выпуклых задач (см. достаточное условие, приведенное выше). Вот если левая часть (18) убывает по T, то действительно найден максимум.

- **Задача с фазовыми ограничениями.** Пусть, например, в задаче (1)–(4) фазовая переменная x является одномерной и на нее наложено дополнительное ограничение $x(t) \geq a,\ t \in [0, T]$ (т. е., фактически, целый континуум ограничений). Тогда придется ввести континуум множителей Лагранжа $\lambda(\cdot)\quad (\lambda : [0, T] \to R)$, добавить к гамильтониану еще одно слагаемое $\lambda(x - a)$ и в дальнейшем работать уже вот с таким видоизмененным гамильтонианом. При этом, как всегда, должны выполняться условия дополняющей нежесткости

$$\forall t \in [0, T] \qquad \begin{cases} x(t) - a \geq 0; \\ \lambda(t) \geq 0; \\ \lambda(t)(x(t) - a) = 0. \end{cases} \tag{19}$$

Подробнее — см. лекции В. З. Беленького (РЭШ, 2001) и задачник А. И. Сотскова и Г. В. Колесника (РЭШ, 2002).