

## Алгоритмы решения задач оптимального управления с фазовыми ограничениями\*

А. Ю. Горнов

*Институт динамики систем и теории управления СО РАН, Иркутск, Россия*  
e-mail: gornov@icc.ru

Рассматривается набор алгоритмов решения задачи оптимального управления с ограничениями типа неравенства, наложенными на траектории системы на всем интервале времени. Предлагаются модификации известных методов, приводятся детальные конструкции алгоритмов, обсуждаются их вычислительные характеристики. Алгоритмы использованы в качестве базовых для построения многометодных вычислительных технологий при реализации нескольких программных комплексов, ориентированных на задачи оптимизации динамических систем.

*Ключевые слова:* задача оптимального управления, фазовые ограничения, численные методы.

### Введение

В практических задачах оптимального управления (ЗОУ), как правило, присутствуют ограничения не только на управляющие воздействия, но и на фазовые переменные. Традиционные подходы предполагают выделение двух классов ЗОУ с траекторными ограничениями: задачи с терминальными ограничениями (см. [1]), налагаемыми в конечный момент времени, и задачи с фазовыми (смешанными, интервальными) ограничениями, действующими на всем временном интервале. Задачи с промежуточными ограничениями, накладываемыми во внутренних точках интервала изменения независимой переменной, легко приводятся к одному из этих классов путем простых математических редукций и поэтому в отдельный класс не выделяются. Задачи с фазовыми ограничениями постоянно возникают в случаях, когда в процессе построения модели в систему дифференциальных уравнений не были внесены некоторые естественные требования: неотрицательности переменных, нахождения траекторий в некоторой “трубке”, отсечения областей нефизичности модели и др. К задачам с фазовыми ограничениями формально может быть отнесен интенсивно исследуемый в настоящее время класс алгебро-дифференциальных (в других источниках дифференциально-алгебраических) систем.

Традиционно при реализации алгоритмов оптимизации динамических систем для учета траекторных ограничений применяется метод штрафов (см., например, [2–4]). Будучи простым и удобным в использовании этот метод, как известно, имеет существенные вычислительные трудности при решении вспомогательных задач и значительную потерю точности решения. В настоящей статье рассматриваются другие подходы

---

\*Работа выполнена при финансовой поддержке РФФИ (грант № 09-07-00267), РГНФ № 09-02-00650 и Междисциплинарного интеграционного проекта СО РАН № 4.

© ИВТ СО РАН, 2010.

к построению численных алгоритмов для задач оптимального управления с фазовыми ограничениями и предложенные на их основе вычислительные методы.

## 1. Канонические постановки задач оптимального управления

Пусть имеется управляемый процесс, описываемый системой обыкновенных дифференциальных уравнений с начальными условиями  $\dot{x} = f(x(t), u(t), t)$ ,  $x(t_0) = x^0$ , определенный на интервале  $T = [t_0, t_1]$ . Здесь  $t$  — независимая переменная (чаще всего время);  $x(t)$  —  $n$ -вектор фазовых координат;  $u(t)$  —  $r$ -вектор управляющих функций;  $n$ -вектор-функция  $f(x(t), u(t), t)$  предполагается непрерывно-дифференцируемой по всем аргументам, кроме  $t$ . Начальный фазовый вектор  $x(t_0) = x^0$  задан. Допустимыми будем называть управляющие функции  $u(t) \in U \subset E^r$  для  $\forall t$ , где  $U$  — выпуклое замкнутое множество из соответствующего пространства.

Задача оптимального управления со свободным правым концом траектории (ЗОУСК) состоит в поиске вектор-функции  $u(t)$ , удовлетворяющей ограничениям и доставляющей минимум функционалу  $I_0(u) = \varphi_0(x(t_1))$ . В задаче оптимального управления с терминальными ограничениями (ЗОУТО) присутствуют также ограничения вида  $I_j(u) = \varphi_j(x(t_1)) = (\leq) 0$ ,  $j = \overline{1, m}$ . В задаче оптимального управления с фазовыми ограничениями (ЗОУФО) на траектории системы имеются фазовые ограничения типа неравенства  $I_j(u) = g_j(x(t), u(t), t) \leq 0$ ,  $j = \overline{m+1, mt}$ . Все функции  $\varphi_j(x(t_1))$ ,  $j = \overline{0, m}$ , и  $g_j(x(t), u(t), t)$ ,  $j = \overline{m+1, mt}$ , предполагаются непрерывно-дифференцируемыми по всем аргументам.

## 2. Специфика задач оптимального управления с фазовыми ограничениями

В отличие от задач оптимального управления с терминальными ограничениями подавляющее большинство ЗОУФО формулируется с ограничениями в виде неравенств. Появление в задаче фазовых ограничений-равенств типа  $g(x, u, t) = 0$  чаще всего говорит о методической незавершенности модели; в такой ситуации уместно предварительно рассмотреть целесообразность сохранения в формулировке ЗОУ такого типа критерия. Если ограничения-равенства необходимо учитывать, можно воспользоваться одним из способов редукции задачи либо применить метод штрафования. Фазовые ограничения-равенства в силу вышесказанного не включены в каноническую постановку задачи, и в дальнейшем будут рассматриваться только методы учета фазовых неравенств.

Задачи оптимального управления с фазовыми ограничениями типа  $g(x, t) \leq 0$  или  $g(x, u, t) \leq 0$  всегда считались сложными. Существенным фактором при решении таких задач является инерционность управляемой системы (индекс ограничения). При большой инерционности оптимальные траектории содержат на временном интервале участки “предварительного подхода” к ограничению (“сингулярная мера” при касании), на которых любые методы сходятся плохо. В терминах аппроксимирующей задачи математического программирования фазовое ограничение после поточечной дискретизации переходит в набор ограничений большой размерности. Многие аппроксимации задач с фазовыми ограничениями могут трактоваться как переопределенные задачи оптимизации, так как число ограничений в них может сильно превышать число переменных. Кроме того, фазовые ограничения могут быть причиной большого числа “паразитических” локальных экстремумов, возникающих во вспомогательных экстремальных задачах.

Возможности методов, хорошо изученные теоретически, существенно различаются по критериям точности, надежности и эффективности. Метод внешних штрафных функционалов несмотря на многочисленные критические замечания обладает (см. [5]) рядом очень привлекательных теоретических свойств и может быть весьма полезен на начальных этапах расчетов. Методы модифицированных функционалов Лагранжа (см., например, [6]) значительно более точные и могут расходиться при отсутствии хороших начальных приближений по прямым и двойственным переменным. Области сходимости методов ньютоновского типа еще меньше, однако при хороших приближениях они позволяют получать высокоточные решения. Таким образом, возникает естественная комбинация методов (мультиметодная схема), комплексные свойства которой превышают возможности каждого из методов, в нее входящих. Для задач оптимального управления с фазовыми ограничениями такая комбинация выглядит следующим образом: метод внешних штрафных функционалов — метод модифицированного функционала Лагранжа — метод приведенного градиента [7].

### 3. Метод внешних штрафных функционалов

При отсутствии априорной информации о задаче “метод штрафов” является наиболее разумным подходом на начальных стадиях расчетов, так как в силу своей простоты не вызывает никаких непрогнозируемых вычислительных эффектов.

Для сохранения свойства непрерывности второй производной, как и в задачах с терминальными ограничениями, используются кубические штрафные функции. Штрафом за нарушение ограничений  $g(x, u, t) \leq 0$  является добавление к целевому функционалу слагаемого

$$\int_{t_0}^{t_1} \left\{ \begin{array}{l} g^3(x, u, t), \quad g(x, u, t) > 0 \\ 0, \quad g(x, u, t) \leq 0 \end{array} \right\} dt,$$

производимое путем увеличения числа фазовых переменных:

$$\dot{x}_{n+1} = \left\{ \begin{array}{l} g^3(x, u, t), \quad g(x, u, t) > 0 \\ 0, \quad g(x, u, t) \leq 0 \end{array} \right\}, \quad x_{n+1}(t_0) = 0.$$

Тогда суммарный функционал будет иметь вид  $\bar{I}(u) = I_0(u) + s \cdot x_{n+1}(t_1)$ , где  $s$  — коэффициент штрафа. В целях экономии памяти при реализации алгоритмов можно не вводить дополнительную переменную в явном виде и вычислять штрафные добавки, используя одну из формул численного интегрирования. Схемы алгоритмов внешних штрафных функционалов тривиальны и различаются только стратегиями увеличения штрафного коэффициента  $s$ .

### 4. Метод модифицированного функционала Лагранжа

Для более точного учета фазовых ограничений может быть использован метод модифицированного функционала Лагранжа. Впервые подобные конструкции для задач оптимального управления были исследованы в работах Ф.Л. Черноусько в 1980-х гг. Рассмотрим модификацию этого популярного метода, ориентированную на ЗОУФО.

Введем функционал

$$L(u, \lambda, s) = I_0(u) + \int_{t_0}^{t_1} \left[ \frac{1}{2 \cdot s} (\max\{0, \lambda(t) + s \cdot g(x, u, t)\})^2 - \lambda^2(t) \right] dt,$$

где  $\lambda(t)$  — множители Лагранжа (точнее, двойственные переменные), зависящие в данном случае от времени,  $s$  — коэффициент штрафа. Как и в предыдущем случае, строится метод типа последовательной безусловной минимизации, в котором формула пересчета двойственных переменных имеет вид  $\lambda^{K+1}(t) = \lambda^K(t) + s \cdot \max\{0, g(x^K(t), u^K(t), t)\}$ . Необходимо отметить, что коэффициент штрафа носит для данного метода совершенно другой смысл, чем для метода внешних штрафных функционалов, являясь не столько основным средством достижения точности решения, сколько коэффициентом регуляризации, придающим свойство квазимонотонности процессу сходимости по двойственным переменным. Вычислительной практикой хорошо подтверждается теоретически известный факт, что для получения решения данным методом не требуется устремлять коэффициент штрафа  $s$  к бесконечности, достаточно достичь некоторого его порогового значения, после чего уточнение решения производится за счет изменения двойственных переменных  $\lambda(t)$ .

Для ЗОУФО алгоритм имеет следующий вид.

0. Задаются алгоритмические параметры:

$\Delta_\lambda$  — коэффициент допустимого изменения двойственных переменных;

$\beta$  — коэффициент увеличения штрафов;

$\gamma$  — коэффициент допустимого изменения невязок.

1. Выбираются начальное приближение по прямым и двойственным переменным и начальные значения коэффициентов штрафа  $u^0(t)$ ,  $\lambda_j^0(t)$ ,  $s_j$ ,  $j = \overline{1, mt}$ .

2. С имеющимися  $\lambda^K(t)$ ,  $s^K$ , используя  $u^K(t)$  в качестве начального приближения, решается задача

$$\min_{u \in U} L(u(t), \lambda^K(t), s^K) = L(u^{K+1}(t), \lambda^K(t), s^K).$$

3. Вычисляется новое приближение по двойственным переменным

$$\lambda_j^{K+1}(t) = \lambda_j^K(t) + s_j^K \cdot \max\{0, g(x^{K+1}(t), u^{K+1}(t), t)\}, \quad t \in T, \quad j = \overline{1, mt}.$$

4. Определяется относительное изменение двойственных переменных

$$\Delta \lambda_{\max} = \max \frac{|\lambda_j^{K+1}(t) - \lambda_j^K(t)|}{1 + |\lambda_j^{K+1}(t)|}, \quad t \in T, \quad j = \overline{1, mt}.$$

5. Если  $\Delta \lambda_{\max} > \Delta_\lambda$ , то увеличиваются все коэффициенты штрафа:

$$s_j^{K+1} = \beta \cdot s_j^K, \quad j = \overline{1, mt}.$$

6. Если  $\Delta \lambda_{\max} \leq \Delta_\lambda$ , увеличиваются те коэффициенты штрафа, которые не позволили существенно уменьшить соответствующие невязки:

$$s_j^{K+1} = \begin{cases} s_j^K \cdot \beta, & \max\{0, g(x^{K+1}(t), u^{K+1}(t), t)\} > \gamma \cdot \max\{0, g(x^K(t), u^K(t), t)\}, \\ s_j^K, & \max\{0, g(x^{K+1}(t), u^{K+1}(t), t)\} \leq \gamma \cdot \max\{0, g(x^K(t), u^K(t), t)\}, \end{cases}$$

$$j = \overline{1, mt}.$$

7. Производится переход на шаг 2.

Алгоритм завершен.

Сценарий работы алгоритма легко просматривается из его конструкции: на первом этапе корректируются как двойственные переменные, так и коэффициенты штрафа, при этом увеличиваются только те компоненты вектора штрафных коэффициентов, которые соответствуют наиболее “трудным” фазовым ограничениям. Затем, на втором этапе расчетов, все  $s_j^K$ ,  $j = \overline{1, mt}$ , фиксируются и дальнейшие коррекции касаются только  $\lambda_j^K(t)$ ,  $j = \overline{1, mt}$ .

## 5. Нелинейный метод приведенного градиента

Другим способом получения высокоточных решений ЗОУФО является применение методов типа “приведенного градиента”. Основная идея данного подхода заключается в предварительном аналитическом или итеративном поиске управления, позволяющего найти траектории, точно скользящие вдоль фазового ограничения. Из равенства  $g(x, u, t) = 0$  аналитически находится один из компонентов управления как функция фазовых переменных, остальных переменных управления и времени:  $\bar{u}_p(t) = \pi(x, u_j, t)$ ,  $j = \overline{1, r}$ ,  $j \neq p$ . Если управления явно отсутствуют в формуле для ограничения (т.е. для ограничений вида  $g(x, t) \leq 0$ ), можно взять от уравнения  $g(x, t) = 0$  несколько производных по времени и, привлекая уравнения системы, получить искомым аналитический вид  $\bar{u}_p(t)$ . В общем случае, не привлекая пользователя к аналитической работе с моделью, для поиска  $\bar{u}_p(t)$  можно применить численные методы решения нелинейных уравнений, хотя такой подход, конечно, значительно влияет на время расчетов и точность результата. Далее строится двухуровневый алгоритм, верхний уровень которого составляет итерационное уточнение  $J_g(t)$ ,  $t \in [t_0, t_1]$ , участков активности ограничения  $g(x, u, t) \leq 0$ , а нижний — оптимизацию системы при управлениях, фиксированных на участках  $J_g(t)$ . Свойства рассматриваемого алгоритма, как и других высокоточных методов, не являются регулярными. Во-первых, возможен локальный “конфликт ограничений”, который выражается в том, что получаемые на итерациях алгоритма  $\bar{u}_l(t)$  не удовлетворяют ограничениям  $u \in U$ . Выходом из такой ситуации может стать конструирование методов, допускающих локальное нарушение ограничения  $u \in U$  и обеспечивающих попадание управлений в допустимый параллелепипед только в конце итеративного процесса. Во-вторых, генерируемая двухуровневым алгоритмом последовательность значений суммарного функционала в общем случае не является монотонной, при каждом изменении  $J_g(t)$  происходит скачкообразное изменение управления, так как вариация первого уровня принципиально не обладает свойствами малости. Это свойство алгоритма можно считать критическим, поскольку чаще всего в таком виде не наблюдается сходимости последовательности управлений к оптимальному. Для получения свойства квазимонотонности последовательности улучшаемых управлений и, следовательно, корректно работающего алгоритма требуется регуляризация процесса. В качестве регуляризатора можно использовать стандартный квадратичный штрафной функционал, коэффициент штрафа в котором должен сразу выбираться довольно большим. Вышеописанный подход легко обобщается на случай двусторонних и многомерных фазовых ограничений и позволяет сконструировать эффективные высокоточные алгоритмы для ЗОУФО широкого класса.

Ниже рассмотрен алгоритм интегрирования прямой системы, применяемый при реализации нелинейного метода приведенного градиента.

1. Выбирается  $u^K(t)$ ,  $t \in T$ .
2. Вычисляется вспомогательное управление  $\bar{u}_p(t) = \pi(x, u_j, t)$ ,  $j = \overline{1, r}$ ,  $j \neq p$ .
3. В цикле по  $\tau$  от  $t_0$  до  $t_1$  с шагом  $h$ :
  - 3.1. Выбирается  $v(\tau) = u^K(\tau)$ ;
  - 3.2. Делается один шаг метода интегрирования, получается  $x(\tau + h)$ ;
  - 3.3. Если  $g(x(\tau + h), v(\tau), \tau) \leq 0$ , то производится переход на шаг 4 (см. ниже);
  - 3.4. Если  $g(x(\tau + h), v(\tau), \tau) > 0$ , то корректируется управление; полагается  $v(\tau) = \bar{u}_p(\tau)$ ;
  - 3.5. Снова делается один шаг метода интегрирования; вычисляется  $\bar{x}(\tau + h)$ .
4. Заканчиваются операции на шаге  $\tau$ .

Алгоритм завершен.

Результатами работы такого алгоритма являются: а — управление, состоящее из кусков  $u^K(t)$  на участках временного интервала, где фазовые ограничения неактивны, и кусков  $\bar{u}_p(\tau)$  на участках, где фазовые ограничения нарушаются; б — соответствующая траектория, незначительно нарушающая фазовые ограничения. Построенное кусочно-непрерывное управление при условии улучшения целевого функционала используется на следующей итерации общего алгоритма оптимизации в качестве  $u^{K+1}(t)$ . Кроме мощного регуляризатора, метод нелинейного приведенного градиента для сходимости требует наличия хорошего начального приближения. В результате получаются траектории, аппроксимирующие оптимальные и нарушающие фазовые ограничения с точностью порядка погрешностей метода интегрирования.

## 6. Метод параметризации ограничений

Требования к точности удовлетворения фазовых ограничений в практических задачах редко бывают очень высокими. Во многих случаях, особенно на начальных стадиях работы с моделью и при проведении качественных исследований, вполне достаточно решать ЗОУФО довольно грубо. Уточнение решения, зачастую производимое с целью подтверждения адекватности модели и работоспособности алгоритма, редко приводит к значительному изменению качественной картины решения и значений оптимизируемого функционала. Задачи с высокими требованиями к точности удовлетворения ограничений, очевидно, являются более сложными, требуют специальных подходов и могут рассматриваться как специальный подкласс ЗОУФО; для таких задач хороший вычислительный эффект обычно дает параметризация ограничений.

Для ограничения  $g(x, u, t) \leq 0$  строится вспомогательный функционал качества, опирающийся на ограничение вида  $g(x, u, t) \leq \varepsilon_g$ , где  $\varepsilon_g$  — коэффициент параметризации. Далее решается последовательность из  $k$  задач при  $\varepsilon_g^0 < \varepsilon_g^1 < \dots < \varepsilon_g^k = 0$ , где начальным управлением в каждой последующей задаче является наилучшее приближенное решение, полученное в предыдущей. Начальное значение  $\varepsilon_g^0$  можно получить, решив одну вспомогательную задачу оптимального управления методом штрафных функционалов с небольшими штрафными коэффициентами и положив  $\varepsilon_g^0 = \max \{g(x(t), u(t), t), t \in [t_0, t_1]\}$ . Для учета параметризованных ограничений можно использовать один из описанных выше стандартных методов, например, алгоритм модифицированных функционалов Лагранжа. Такой комбинированный метод погружения исходной задачи в ряд

аппроксимирующих задач, как правило, позволяет значительно повысить точность удовлетворения фазовых ограничений и достичь регулярности при сходимости алгоритма.

## Заключение

Предложенные модификации алгоритмов стали основой функционального наполнения нескольких программных средств, реализованных автором для задач оптимального управления: пакет прикладных программ МАПР, блок “Нелинейное оптимальное управление” (1981–1989 гг.), программный комплекс OPTCON-DOS (1987–1990 гг.), программный комплекс OPTCON-PARALL (1999–2002 гг.), вычислительный сервер OPTCON (2000–2003 гг.), программный комплекс OPTCON-III (2004–2008 гг.). С применением предложенных алгоритмов был решен ряд прикладных задач оптимизации в следующих областях: динамика полета, космонавигация, робототехника, электроэнергетика, экономика, экология, социология, медицина.

## Список литературы

- [1] Горнов А.Ю. Алгоритмы решения задач оптимального управления с терминальными ограничениями // Вычисл. технологии. 2008. Т. 13, № 4. С. 44–50.
- [2] Федоренко Р.П. Приближенное решение задач оптимального управления. М.: Наука, 1978. 488 с.
- [3] Еремин И.И. Метод штрафов в выпуклом программировании // Докл. АН СССР. 1967. Т. 173, № 4. С. 748–751.
- [4] Евтушенко Ю.Г. Методы решения экстремальных задач и их применение в системах оптимизации. М.: Наука, 1982. 432 с.
- [5] Поляк Б.Т. Введение в оптимизацию. М.: Наука, 1983. 382 с.
- [6] Бертсекас Д. Условная оптимизация и методы множителей Лагранжа. М.: Радио и связь, 1987. 400 с.
- [7] Горнов А.Ю., Диваков А.О. Комплекс программ OPTCON для решения задач оптимального управления. Руководство пользователя. Иркутск: ИрВЦ СО АН СССР, 1990. 36 с.

*Поступила в редакцию 30 апреля 2009 г.*