

Алгоритмы решения задач оптимального управления с терминальными ограничениями*

А. Ю. ГОРНОВ

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

A set of numerical methods for optimal control problems with terminal restrictions is discussed. Details of the algorithm constructions are described. Numerical technology based on the proposed algorithms was implemented as a software complex for dynamical optimization problems.

Введение

Задача оптимального управления (ЗОУ) с терминальными ограничениями занимает особое место в практике численного решения задач оптимизации динамических систем. К классу задач с терминальными ограничениями относятся популярнейшие задачи перевода системы из начальной точки в конечную, краевые экстремальные задачи, изопериметрические задачи вариационного исчисления и др. Кроме того, задачи более сложных типов (с фазовыми ограничениями, с промежуточными ограничениями, алгебро-дифференциальные и др.) могут быть тривиально сведены к терминальным задачам с помощью общеизвестных математических редукций.

Традиционно при реализации алгоритмов оптимизации динамических систем для учета терминальных ограничений используется метод штрафов. Применение этого простого и надежного метода, к сожалению, приводит к вычислительным схемам с серьезными проблемами при решении вспомогательных задач и большой потерей точности решения. В статье рассматриваются, кроме алгоритмов, основанных на методе штрафов, также и другие, более эффективные, подходы к построению численных алгоритмов для задач оптимального управления с терминальными ограничениями.

1. Задача оптимального управления с терминальными ограничениями (ЗОУТО)

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

*Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (гранты № 06-07-89215, 07-07-00265а, 08-07-00172) и РГНФ 07-02-12112в.

© Институт вычислительных технологий Сибирского отделения Российской академии наук, 2008.

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 присутствуют также ограничения вида $I_j(u) = \varphi_j(x(t_1)) = (\leq) 0$, $j = \overline{1, m}$. Все функции $\varphi_j(x(t_1))$, $j = \overline{0, m}$, предполагаются непрерывно дифференцируемыми по всем аргументам.

2. Условия оптимальности для ЗОУТО

Принцип максимума для ЗОУТО имеет следующий вид (см., например, [1]): если $u^*(t)$ — оптимальное управление, то существуют двойственные переменные y_j , $j = \overline{1, m}$, такие, что функция Гамильтона—Понтрягина $H(p, x, u, t)$ в точке $u^*(t)$ достигает максимума. Сопряженные переменные $p_i(t)$, $i = \overline{1, n}$ здесь должны удовлетворять сопряженной системе вида

$$\dot{p} = -\frac{\partial H(p, x, u, t)}{\partial x},$$

$$p(t_1) = -\frac{\partial \varphi_0(x(t_1))}{\partial x} - \sum_{j=1}^m y_j \frac{\partial \varphi_j(x(t_1))}{\partial x}.$$

Методы для задач такого типа должны, следовательно, кроме поиска управлений, близких к оптимальным, находить и двойственный вектор y , позволяющий получить решение с необходимой точностью.

3. Метод внешних штрафных функций

Метод штрафов является, по-видимому, старейшим из численных методов условной оптимизации (см., например, [2]). Идея его проста: заменить условную задачу несколькими безусловными, последовательность решений которых дает решение исходной задачи. Богатый опыт применения метода штрафов создал ему славу простого и надежного метода; хорошо известны и его недостатки — грубость расчетов и повышение трудности вспомогательных задач при движении к экстремуму. Создание пакетов программ оптимизации, реализующих, как правило, набор методов, привело к новому взгляду на класс методов штрафных функций как на средство, позволяющее получить хорошее начальное приближение для более точных методов.

Введем штрафной функционал вида $\bar{I}(u) = I_0(u) + \sum_{j=1}^m s_j^K I_j^2(u)$. Выбрав $u^0(t)$, зададим коэффициенты штрафа s_j^K , $j = \overline{1, m}$, и, используя один из методов для ЗОУСК, решим задачу $\min_{u \in U} \bar{I}(u) = \bar{I}(u^{K+1})$. Затем увеличим коэффициенты штрафа и снова решим ЗОУСК и т. д. Критерием остановки служит критерий для алгоритмов безусловной задачи. Величины $2s_j^K I_j(u^K)$, $j = \overline{1, m}$, являются оценками оптимального значения двойственных переменных y_j , на каждой итерации они запоминаются и, после окончания работы алгоритма, могут быть использованы в качестве первого приближения для других методов.

4. Метод линеаризации

Идея методов линеаризации проста и естественна: в окрестности имеющегося управления задача линеаризуется, и, с использованием решения линейной задачи, строится новое приближение к оптимуму нелинейной. Существует два подхода к построению методов линеаризации, отличающихся ограничениями на вариацию. Подход, предполагающий введение во вспомогательную задачу ограничения по норме $\|\delta u\| \leq c$, где δu — вариация оптимизируемых переменных, (здесь c — параметр алгоритма), весьма глубоко проработан в известной монографии [1]. Другой способ — введение в функционал штрафной добавки $(\delta u, \delta u)$ — рассматривается в [3]. Ниже приводится вариант метода линеаризации для ЗОУТО, построенный по типу методов Пшеничного.

Сформулируем линейную задачу в вариациях:

$$\begin{cases} \int_{t_0}^{t_1} \nabla I_0(u^K) \delta u dt \rightarrow \min, \\ I_j(u^K) + \int_{t_0}^{t_1} \nabla I_j(u^K) \delta u dt = 0, \quad j = \overline{1, m}. \end{cases}$$

Запишем для этой задачи функцию Лагранжа, добавив слагаемое штрафа:

$$L(\delta u, y) = \int_{t_0}^{t_1} \left[\frac{1}{2\nu} (\delta u, \delta u) + \nabla I_0(u^K) \delta u + \sum_{j=1}^m y_j \nabla I_j(u^K) \delta u \right] dt.$$

Здесь ν — алгоритмический параметр, y — двойственные переменные. Из условия оптимальности $\frac{\partial L}{\partial (\delta u)} = 0$, выразив вариацию $\delta u^K = - \left(\nabla I_0(u^K) + \sum_{j=1}^m y_j \nabla I_j(u^K) \right) \nu$ и подставив ее в уравнения линейной задачи, получим условия для определения y_j , $j = \overline{1, m}$, в виде системы линейных уравнений $Ay = b$, где

$$a_{ij} = \nu \int_{t_0}^{t_1} \nabla I_i(u^K) \nabla I_j(u^K) dt, \quad b_j = -I_j(u^K) + \nu \int_{t_0}^{t_1} \nabla I_0(u^K) \nabla I_j(u^K) dt, \quad i, j = \overline{1, m}.$$

Решив систему и получив y_j , $j = \overline{1, m}$, вычислим вариацию δu^K . Далее, следуя Б.Н. Пшеничному, сформулируем недифференцируемую функцию выигрыша

$$F_S(u^K + \alpha \delta u^K, \beta) = I_0(u^K + \alpha \delta u^K) + \beta \max |I_j(u^K + \alpha \delta u^K)|, \quad j = \overline{1, m}.$$

Штрафной коэффициент β будем вычислять с помощью двойственных переменных y : $\beta = 2 \sum_{j=1}^m |y_j|$. Сходимость алгоритмов такого типа для задачи математического программирования строго обоснована в [3].

5. Метод модифицированной функции Лагранжа

Практика использования методов типа последовательной безусловной оптимизации привела к следующей схеме комбинирования различных алгоритмов: на первом этапе применяются методы с широкой областью сходимости и невысокими характеристиками

точности — “грубые” — метод внешних штрафных функций, метод параметризации целевой функции и т. п. Результатом их работы является некоторая оценка оптимальных значений искомым переменных — как прямых, так и двойственных. Затем, на втором этапе, с использованием имеющегося приближения, “включаются” более точные методы, например, метод модифицированной функции Лагранжа (см. [4]). Обычно остается открытым естественный вопрос — когда “грубыми” методами получено достаточно хорошее приближение и можно переходить ко второму этапу.

Рассмотрим схему метода множителей Лагранжа, объединяющую оба этапа в рамках одного алгоритма. Схема представляет собой модификацию способа пересчета двойственных переменных и коэффициентов штрафа, описанного в [5, с. 131] для задачи математического программирования.

Рассмотрим модифицированную функцию Лагранжа следующего вида:

$$M(u, y, s) = I_0(u) + \sum_{j=1}^m y_j I_j(u) + \sum_{j=1}^m s_j I_j^2(u).$$

Здесь s_j , $j = \overline{1, m}$, — коэффициенты штрафа (или, в другой терминологии, коэффициенты регуляризации).

Алгоритм имеет следующий вид.

1. Выбираем начальное приближение $u^0(t)$, y_j^0 , s_j^0 , $j = \overline{1, m}$.
2. С имеющимися y^K , s^K , используя $u^K(t)$ в качестве начального приближения, решаем задачу

$$\min_{u \in U} M(u, y^K, s^K) = M(u^{K+1}, y^K, s^K).$$

3. Вычисляем новое приближение по двойственным переменным

$$y_j^{K+1} = y_j^K + 2s_j^K I_j(u^{K+1}), \quad j = \overline{1, m}.$$

4. Определяем относительное изменение двойственных переменных

$$\Delta y_{\max} = \max \frac{|y_j^{K+1} - y_j^K|}{1 + |y_j^{K+1}|}, \quad j = \overline{1, m}.$$

5. Если $\Delta y_{\max} > \Delta$ (Δ — параметр), то увеличиваются все коэффициенты штрафа

$$s_j^{K+1} = 10s_j^K, \quad j = \overline{1, m}.$$

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

$$s_j^{K+1} = \begin{cases} s_j^K 10, & |I_j(u^{K+1})| > 0.25 |I_j(u^K)|, \\ s_j^K, & |I_j(u^{K+1})| \leq 0.25 |I_j(u^K)|. \end{cases}$$

7. Переходим на шаг 2.

Алгоритм завершен.

Таким образом, сначала каждый пересчет двойственных переменных сопровождается ростом коэффициентов штрафа — “работает” метод внешних штрафных функций; затем, по мере стабилизации y , штрафы увеличиваются только при “упрямых” ограничениях; и, в конце концов, s_j , $j = \overline{1, m}$, стабилизируются и алгоритм превращается в метод модифицированной функции Лагранжа. Попутно в автоматическом режиме решается вопрос получения двойственных приближений и установки приемлемых значений коэффициентов регуляризации. Такой алгоритм может оказаться особенно удобным в случаях, если необходимо решать задачи в автоматическом режиме — без вмешательства пользователя в процесс вычислений.

Критерием останова в такого типа алгоритмах служат критерии выхода во вспомогательной ЗОУСК при одновременном выполнении условия стабилизации двойственных переменных: $\Delta y_{\max} \leq \Delta$. Обоснование сходимости методов модифицированной функции Лагранжа для задачи математического программирования можно найти, например, в книге [5].

6. Метод точных дифференцируемых штрафных функций

Методы точных дифференцируемых штрафных функций [5, с. 224] сравнительно недавно изобретены для задачи математического программирования.

Рассмотрим недифференцируемую точную штрафную функцию $N(u, c) = I_0(u) + c \max |I_j(u)|$, $j = \overline{1, m}$. Для алгоритмов последовательной безусловной минимизации, построенных на основе такой функции, по аналогии с конечномерной оптимизацией, можно ожидать выполнения следующего свойства: существует $\bar{c} > 0$ такое, что для получения решения с требуемой точностью достаточно только один раз минимизировать $N(u, \bar{c}) = I_0(u) + \bar{c} \max |I_j(u)|$. Этот чрезвычайно привлекательный конструкционный элемент (обычно для получения решения приходится решать десятки и сотни подобных задач!), к сожалению, обладает и принципиальным дефектом: вспомогательные задачи становятся негладкими. Алгоритмы для негладких задач оптимального управления развиты значительно слабее, чем для гладких, и пока не удается найти на этом пути эффективные вычислительные схемы.

Попытки преодолеть барьер негладкости, сохранив при этом эффективное свойство метода, привели к созданию точных дифференцируемых штрафных функций [6, 7]. Основная идея этого подхода заключается в построении механизма высокоточной оценки двойственных переменных с использованием производных функционалов ограничений. Рассмотрим функцию свертки типа

$$T(u, y, c) = I_0(u) + \sum_{j=1}^m y_j(u) I_j(u) + c \sum_{j=1}^m I_j^2(u),$$

где $y(u) = -[\nabla K(u)' \nabla K(u)]^{-1} \nabla K(u)' \nabla I_0(u)$, $\nabla K(u) = \{\nabla I_1(u), \dots, \nabla I_m(u)\}$.

Для эффективности такого метода при решении ЗОУТО требуется:

- 1) знание порогового значения \bar{c} ;
- 2) наличие гладких вторых производных функционалов в задаче;
- 3) невырожденность $\nabla K(u)$ в точке оптимума.

Выполнение первого условия тривиально достигается, как и в методе множителей Лагранжа, путем конструирования механизма итеративной подстройки значения c . Гладкость вторых производных есть обычное требование к постановке ЗОУ, позволяющее применять для решения вспомогательных задач методы со сверхлинейной скоростью сходимости. Выполнение третьего условия, к сожалению, не гарантировано, оно довольно часто нарушается, что подтверждается практикой расчетов. Тем не менее рассмотренный ниже алгоритм — достаточно эффективный инструмент оптимизации динамических систем и его целесообразно включать в пакеты программ для ЗОУ.

Алгоритм имеет следующий вид.

1. Выбираем начальное приближение $u^0(t)$, c^0 .
2. Вычисляем $y^K(u^K) = -[\nabla K(u^K)' \nabla K(u^K)]^{-1} \nabla K(u^K)' \nabla I_0(u^K)$.
3. С имеющимися y^K , c^K , используя $u^K(t)$ в качестве начального приближения, решаем задачу

$$\min_{u \in U} T(u, y^K, c^K) = T(u^{K+1}, y^K, c^K).$$

4. Вычисляем новое приближение по двойственным переменным

$$y^{K+1}(u^{K+1}) = -[\nabla K(u^{K+1})' \nabla K(u^{K+1})]^{-1} \nabla K(u^{K+1})' \nabla I_0(u^{K+1}).$$

5. Определяем относительное изменение двойственных переменных

$$\Delta y_{\max} = \max \frac{|y_j^{K+1} - y_j^K|}{1 + |y_j^{K+1}|}, \quad j = \overline{1, m}.$$

6. Если $\Delta y_{\max} > \Delta$ (Δ — параметр), то увеличивается коэффициент штрафа

$$c^{K+1} = 2c^K.$$

7. Переходим на шаг 2.

Алгоритм завершен.

Обоснование сходимости методов точной дифференцируемой штрафной функции можно найти, например, в книге [5].

7. Методы учета терминальных ограничений-неравенств

Задачи с терминальными ограничениями-неравенствами встречаются на практике значительно реже, чем задачи с равенствами. Тем не менее, чтобы не потерять общности постановки задачи, алгоритмы для учета таких ограничений обязательно должны быть в составе функционального наполнения программных комплексов.

Для задач с ограничениями-неравенствами $\varphi_j(x(t_1)) \leq 0$, $j = \overline{1, m}$, основная проблема — сохранение непрерывности вторых производных суммарного функционала. Это свойство требуется для эффективной работы методов со сверхлинейной скоростью сходимости, обычно применяемых для решения вспомогательных задач. Обычно эта проблема преодолевается введением кубических штрафных функций: $\bar{I}(u) = I_0(u) + \sum_{j=1}^m s_j^K \max\{0, I_j(u)\}^3$. Формулы пересчета штрафных коэффициентов аналогичны формулам, применяемым для ограничений-равенств.

Заключение

Предложенные модификации алгоритмов послужили 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), программный комплекс OPTCON-PARALL (1999–2002), вычислительный сервер OPTCON (2000–2003 гг.). С применением предложенных алгоритмов был решен ряд прикладных задач оптимизации из областей динамики полета, космонавигации, робототехники, электроэнергетики, экономики, экологии, социологии.

Список литературы

- [1] ФЕДОРЕНКО Р.П. Приближенное решение задач оптимального управления. М.: Наука, 1978. 488 с.
- [2] ЕРЕМИН И.И. Метод штрафов в выпуклом программировании // Докл. АН СССР. 1967. Т. 173, №4. С. 748–751.
- [3] ПШЕНИЧНЫЙ Б.Н. Метод линеаризации. М.: Наука, 1979. 208 с.
- [4] ЕВТУШЕНКО Ю.Г. Методы решения экстремальных задач и их применение в системах оптимизации. М.: Наука, 1982. 432 с.
- [5] БЕРТСЕКАС Д. Условная оптимизация и методы множителей Лагранжа. М.: Радио и связь, 1987. 400 с.
- [6] DiPILLO G., GRIPPO L. A new class of augmented Lagrangians in nonlinear programming // SIAM J. Control Optim. 1979. N 17. P. 618–628.
- [7] HAN S.P., MANGASARIAN O.L. Exact penalty function in nonlinear programming // Math. Programming. 1979. N 17. P. 251–269.

*Поступила в редакцию 28 августа 2007 г.,
в переработанном виде — 29 мая 2008 г.*