

На правах рукописи

УДК 519.8

Марковцев Денис Анатольевич

МЕТОД ГЛАДКИХ ШТРАФНЫХ ФУНКЦИЙ В
ЗАДАЧАХ ПАРАМЕТРИЧЕСКОГО
ПРОГРАММИРОВАНИЯ

01.01.09 – дискретная математика и математическая
кибернетика

АВТОРЕФЕРАТ

диссертации на соискание ученой степени
кандидата физико-математических наук

Москва – 2012

Работа выполнена на кафедре высшей математики Московского физико-технического института (государственного университета)

Научный руководитель: доктор технических наук, профессор
Умнов Александр Евгеньевич

Официальные оппоненты: Дикусар Василий Васильевич,
д.ф.-м.н., профессор,
Вычислительный центр
им. А.А. Дородницына РАН,
ведущий научный сотрудник,

Бирюков Александр Гаврилович,
к.ф.-м.н., доцент,
кафедра математических основ управления
Московского физико-технического института
(государственного университета), доцент

Ведущая организация: Институт проблем управления
им. В.А.Трапезникова РАН

Защита состоится _____ 2012 г. в ____ часов на заседании диссертационного совета Д 212.156.05 при Московском физико-техническом институте (государственном университете) по адресу 141700, Московская обл., г. Долгопрудный, Институтский пер., д. 9, ауд. 903 КПМ.

С диссертацией можно ознакомиться в библиотеке Московского физико-технического института (государственного университета).

Автореферат разослан _____ 2012 г.

Ученый секретарь
диссертационного совета

Федько Ольга Сергеевна

Общая характеристика работы

Актуальность темы. Разработанные к настоящему времени методы решения задач математического программирования в общем случае не гарантируют получения решения при приемлемом уровне затрат вычислительных ресурсов таких как требуемая оперативная или энергонезависимая память, процессорное время. Поэтому с методологической точки зрения представляется актуальным дальнейшее исследование возможных модификаций как постановок задач, так и алгоритмов поиска их решений, позволяющих снижать совокупные затраты вычислительных усилий и ресурсов. Один из таких подходов, традиционно называемый *параметрическим программированием*, основан на разделении множества переменных исходной задачи на две части: собственно *переменные* и *параметры*, значения которых при необходимости могут быть фиксированы. Этот подход является предметом исследования в диссертационной работе.

Цель диссертационной работы. Цель исследования - разработка метода решения задач параметрического программирования, основанного на использовании гладких штрафных функций для построения сглаженных аппроксимаций зависимостей решения задач математического программирования от параметров.

В соответствии со сформулированной целью исследования в диссертационной работе рассматриваются следующие задачи.

1. Анализ теоретических и практических аспектов, как обосновывающих целесообразность использования схем параметрического программирования, так и ограничивающих возможности их применения.
2. Формулировка и обоснование подхода, позволяющего преодолевать принципиальные затруднения, возникающие при использовании методов параметрического программирования.
3. Построение двухуровневых алгоритмов решения задачи параметрического программирования, исследование их сходимости, а также необходимых и достаточных условий оптимальности получаемых решений.

4. Разработка методов решения задач параметрического программирования, обладающих свойством линейности по всем переменным при фиксированных значениях параметров.
5. Практическая реализация предлагаемых подходов и исследование ее эффективности при решении конкретных задач математического программирования.

Научная новизна. Основные результаты работы являются новыми и заключаются в следующем.

1. В работе рассмотрен основанный на методе штрафных функций алгоритм сглаживания зависимости решения задачи математического программирования от ее параметров. Исследованы свойства процедур сглаживания для достаточно широкого класса штрафных функций. Предложен и обоснован алгоритм решения задач параметрического программирования.
2. Получены необходимые и достаточные условия существования решения двухуровневой системы задач. Обоснована сходимость итерационного процесса решения двухуровневой системы задач. Исследована сходимость метода гладких штрафных функций к решению двойственной задачи.
3. Исследована схема использования классических методов гладкой оптимизации для решения задач математического программирования, допускающих параметрическую линейаризацию, при помощи квадратичной штрафной функции. Показана практическая эффективность подхода на серии моделей.

Практическая ценность. Модели, методы и алгоритмы, разработанные в диссертации, применялись в Московском физико-техническом институте (государственном университете) для решения различных задач моделирования экономических процессов.

Апробация. Основные положения диссертационной работы докладывались, обсуждались и получили одобрение специалистов на научных семинарах кафедр высшей математики и математических основ управления Московского физико-технического института (государственного университета) (2007-2012), научных семинарах в Вычислительном центре им. А.А. Дородницына РАН (2012) и в Институте проблем управления им. В.А. Трапезникова РАН (2012).

Публикации. По теме диссертации опубликовано 8 печатных работ, из них 5 [1, 2, 3, 6, 8] – из списка, рекомендованного ВАК РФ.

Объем и структура работы. Диссертация состоит из введения, четырех глав, заключения и приложения. Список использованных источников включает 44 наименования. Текст диссертации содержит 120 страниц.

Содержание работы

Во введении обоснована актуальность исследуемых проблем, сформулированы цели и задачи диссертационной работы, изложены основные результаты, выносимые на защиту, характеристика их научной новизны, практической значимости и апробации полученных результатов.

Одним из возможных способов расширения множества задач математического программирования, поддающихся решению численными методами, является *параметризация* их постановки. Параметризация заключается в разделении множества переменных исходной задачи

$$\text{найти } \min_{x,u} f(x, u) \quad (1)$$

$$\text{при условиях } g_s(x, u) \geq 0, s = [1, m],$$

на два подмножества: собственно *переменных* x и *параметров* u , значения которых при необходимости могут быть зафиксированы. Такое разделение позволяет свести решение исходной задачи (1) к двухуровневой схеме. На первом уровне решается задача математического программирования

$$\text{найти } \min_x f(x, u) \quad (2)$$

$$\text{при условиях } g_s(x, u) \geq 0, s = [1, m], u = \text{fix},$$

а на втором

$$\begin{aligned} & \text{найти } \min_u f(x^*(u), u) & (3) \\ & \text{при условиях } g_s(x^*(u), u) \geq 0, s = [1, m], \end{aligned}$$

где $x^*(u)$ есть решение задачи (2) при фиксированном векторе параметров u .

В этом случае упрощение процедуры решения исходной задачи может быть обусловлено, как очевидно, *меньшей размерностью* задач (2) и (3) по сравнению с размерностью задачи (1), так и возможным *упрощением* их постановок. Примером такого упрощения является случай, когда за счет подходящего выбора параметризуемых переменных задача нижнего уровня оказывается линейной.

Условимся в дальнейшем задачу (3) называть задачей *верхнего уровня*, а задачу (2) будем называть задачей *нижнего уровня*. Пару задач *верхнего и нижнего уровня* будем называть *двухуровневой системой задач* или *задачей параметрического программирования*. В рассматриваемых задачах $x \in E^n$ - вектор переменных и $u \in E^l$ - вектор параметров являются элементами конечномерных евклидовых пространств, координаты которых $\|x\| = \left\| \begin{matrix} \xi_1 & \xi_2 & \cdots & \xi_n \end{matrix} \right\|^T$ и $\|u\| = \left\| \begin{matrix} \nu_1 & \nu_2 & \cdots & \nu_l \end{matrix} \right\|^T$. Предполагается, что функции $f(x, u), g_s(x, u), s = [1, m]$ имеют непрерывные частные производные до второго порядка включительно.

В первой главе «Задачи параметрического программирования: проблемы и методы их решения» описаны факторы, осложняющие решение двухуровневой системы задач (2)-(3) и проведен тематический обзор возможных методов преодоления отмеченных затруднений.

При любом фиксированном u задача (2) представляет собой обычную задачу математического программирования. Оптимальное решение $x^* = x^*(u)$ естественным образом зависит от u . При этом некоторые свойства зависимости $x^*(u)$ могут существенно усложнять решение задачи (3).

Действительно, зависимость $x^*(u)$ обладает рядом «неудобных» для вычислений свойств:

- невозможность (за исключением тривиальных случаев) аналитического представления, что препятствует решению в явном виде задач математического программирования, в которых $x^*(u)$ входит в условие;

- существование не для всех значений параметра u ;
- неоднозначность зависимости $x^*(u)$ для тех u , при которых задача (2) не имеет локально изолированных решений;
- недифференцируемость зависимости $x^*(u)$ даже в случае, когда у функций $f(x, u)$ и $g_s(x, u)$ существуют непрерывные частные производные достаточно высокого порядка.

Последнее свойство является следствием того, что условия задачи (2) содержат ограничения типа неравенство. Более того, не гарантируется не только желательная гладкость, но даже и непрерывность зависимости $x^*(u)$.

За последние пять десятков лет было выполнено большое число исследований, целью которых являлось преодоление отмеченных затруднений и построение алгоритмов решения двухуровневой системы задач (2)-(3). Среди этих исследований можно выделить следующие направления:

- методы, предполагающие возможность получения явного вида зависимости $x^*(u)$;
- двухуровневые схемы (типа bilevel programming), нацеленные на решение более общей задачи с различными функционалами на верхнем и нижнем уровне;
- методы недифференцируемой оптимизации, использующие субдифференциальное исчисление;
- методы, основанные на построении «сглаженных» аппроксимаций зависимости $x^*(u)$, лишенных отмеченных выше недостатков.

Большинство методов, относящихся к первым трем пунктам, продемонстрировали либо ограниченную область применимости, либо высокий уровень затрат необходимых вычислительных ресурсов. В то время как для методов, принадлежащих четвертой группе и основанных на использовании тейлоровских аппроксимаций зависимостей $x^*(u)$, известны отдельные примеры их эффективного использования для решения некоторых конкретных классов

задач (анализа чувствительности, многокритериальной оптимизации и декомпозиции, оптимального управления). Поэтому представляется целесообразным исследовать схему сглаживания в общем виде для класса задач (1).

Во второй главе «Метод сглаживающих штрафных функций и его обоснование» предложен метод сглаживания, который позволяет единообразно преодолевать, обусловленные особенностями зависимостей $x^*(u)$, затруднения и получать решения для достаточно широкого класса задач верхнего уровня.

Идея метода сглаживания заключается в замене зависимости $x^*(u)$ достаточно гладкой и определенной для всех $u \in \Omega \subseteq E^l$ и любых положительных r функцией $x^+(r, u)$, такой что $\lim_{r \rightarrow +0} x^+(r, u) = x^*(u)$. При этом в качестве аппроксимирующей зависимости $x^+(r, u)$ предлагается использовать решение задачи нижнего уровня, получаемое по методу гладких штрафных функций.

Как известно, решением задачи нижнего уровня (2) при использовании метода штрафных функций является

$$x^+(r, u) = \arg \min_x \varepsilon(r, x, u) \quad (4)$$

где *вспомогательная функция* метода гладких штрафных функций выбирается следующего вида:

$$\varepsilon(r, x, u) = f(x, u) + \sum_{s=1}^m P(r, g_s(x, u)). \quad (5)$$

Если помимо стандартного свойства

$$\lim_{r \rightarrow +0} P(r, \alpha) = \begin{cases} +\infty, & \alpha < 0 \\ 0, & \alpha \geq 0, \end{cases} \quad (6)$$

достаточно гладкая *штрафная функция* $P(r, \alpha)$ удовлетворяет также и неравенствам

$$\frac{\partial P}{\partial \alpha} \leq 0 \quad \text{и} \quad \frac{\partial^2 P}{\partial \alpha^2} \geq 0, \quad \forall \alpha, \quad (7)$$

то $x^+(r, u)$ - точка минимума вспомогательной функции ε , находящаяся из условия стационарности

$$\operatorname{grad}_x \varepsilon(r, x^+, u) = o, \quad (8)$$

или

$$\operatorname{grad}_x f(x, u) + \sum_{s=1}^m \frac{\partial P(r, g_s(x, u))}{\partial g_s} \cdot \operatorname{grad}_x g_s(x, u) = o, \quad (9)$$

удовлетворяет условиям сглаживания зависимости $x^*(u)$.

Теоретическую основу предлагаемого подхода составляют использование локальных тейлоровских аппроксимаций и теоремы о неявных функциях. Пусть штрафная функция $P(r, \alpha)$ локально строго выпукла по α для любых $r > 0$ и, по крайней мере, дважды непрерывно дифференцируема по всем своим аргументам. Будем также предполагать, что задача нижнего уровня имеет ограниченное решение при всех допустимых значениях параметров. Тогда, в силу теоремы о неявных функциях и свойств метода штрафных функций, оказываются справедливыми следующие утверждения.

1. Решение уравнений (8), являющихся условиями стационарности функции (5), $x^+(r, u)$ существует и локально единственно для любых $u \in \Omega \subseteq E^l$ и $r > 0$, и потому зависимость $x^+(r, u)$ является функциональной.
2. Для зависимости $x^+(r, u)$ верно равенство $\lim_{r \rightarrow +0} x^+(r, u) = x^*(u)$.
3. Функция $x^+(r, u)$ является непрерывно дифференцируемой по всем своим аргументам.

Как известно, основное свойство метода штрафных функций состоит в том, что

$$\lim_{r \rightarrow +0} \varepsilon(r, x^+(r, u), u) = f(x^*(u), u), \quad (10)$$

где $x^*(u)$ – решение задачи (2). Поэтому в качестве целевого функционала, оптимизация которого позволяет получить оценку решения задачи верхнего уровня, можно принять

$$E(r, u) = \varepsilon(r, x^+(r, u), u). \quad (11)$$

Введем обозначение

$$\tilde{u}(r) = \arg \min_u E(r, u). \quad (12)$$

Пусть пара векторов x^* и u^* есть решение задачи (1), тогда справедливы равенства:

$$\begin{aligned} \lim_{r \rightarrow +0} \tilde{u}(r) &= u^*, \quad \lim_{r \rightarrow +0} x^+(r, u) = x^*(u) \text{ и} \\ \lim_{r \rightarrow +0} E(r, \tilde{u}(r)) &= \lim_{r \rightarrow +0} \varepsilon(r, x^+(r, \tilde{u}(r)), \tilde{u}(r)) = f(x^*, u^*), \end{aligned} \quad (13)$$

следующие из теорем 2, 3 и 4.

Теорема 2. Точка $\{\tilde{x}, \tilde{u}\}$ является стационарной точкой функционала $\varepsilon(r, x, u)$ тогда и только тогда, когда \tilde{u} – стационарная точка функционала $E(r, u)$.

Теорема 3. Точка $\{\tilde{x}, \tilde{u}\}$ является локальным изолированным минимумом функционала $\varepsilon(r, x, u)$ тогда и только тогда, когда точки \tilde{x} и \tilde{u} являются локальными изолированными минимумами функционалов $\varepsilon(r, x, \tilde{u})$ и $E(r, u)$ соответственно.

Теорема 4. Точка $\{x^*, u^*\} \in E^n \oplus E^l$ является локальным изолированным решением задачи (1) тогда и только тогда, когда x^* и u^* – соответственно локальные изолированные решения задач (2) для u^* и (3) соответственно.

С другой стороны, из теорем 1 и 6 следует, что $\forall s = [1, m]$

$$\begin{aligned} \lim_{r \rightarrow +0} \frac{\partial P(r, g_s(x^+(r, \tilde{u}(r)), \tilde{u}(r)))}{\partial g_s} &= -\lambda_s^*, \\ \lim_{r \rightarrow +0} \frac{\partial P(r, g_s(x^+(r, \tilde{u}(r)), \tilde{u}(r)))}{\partial g_s} g_s(x^+(r, \tilde{u}(r)), \tilde{u}(r)) &= 0. \end{aligned} \quad (14)$$

Что, в сочетании с условиями (9), позволяет убедиться в выполнении *необходимых условий оптимальности* решения задачи (3):

$$\begin{aligned} \forall s = [1, m] \quad \exists \lambda_s^* \geq 0 \quad \text{такие, что} \\ \lambda_s^* g_s(x^*, u^*) = 0, \quad \text{и} \\ \text{grad}_u f(x^*, u^*) + \sum_{s=1}^m \lambda_s^* \cdot \text{grad}_u g_s(x^*, u^*) = 0. \end{aligned} \quad (15)$$

Следует отметить, числа λ_s^* существуют всегда и совпадают со стандартными множителями Лагранжа, когда последние имеют конечное единственное значение.

Теорема 1. (Сходимость к двойственному решению)

Пусть задача математического программирования

$$\text{найти вектор переменных } x^* \in E^n, \quad (16)$$

доставляющий минимум функционалу $f(x)$

$$\text{при условиях } g_s(x) \geq 0, \quad s = [1, m],$$

решается методом штрафных функций. В качестве функции штрафа используется $P(r, \alpha)$, удовлетворяющая условиям (6) и (7), r – коэффициент штрафа. Функции $f(x), -g_s(x) \quad \forall s = [1, m]$ выпуклы, множество оптимальных точек задачи (16) компактно. Если к тому же $\{r_k\}$ – строго убывающая и стремящаяся к нулю последовательность положительных чисел, то в процессе решения получают точки

$$[x(r_k), \lambda(r_k)], \quad (17)$$

где

$$\lambda_s(r_k) = -\frac{\partial P(r_k, g_s(x(r_k)))}{\partial g_s}, \quad s = [1, m], \quad (18)$$

допустимые для двойственной задачи

$$\text{найти } \max_{\lambda} L(x, \lambda) \quad (19)$$

$$\text{при условиях} \quad (20)$$

$$\nabla_x L(x, \lambda) = 0, \quad (21)$$

$$\lambda_s \geq 0, \quad s = [1, m], \quad (22)$$

и также выполняются

$$\lim_{k \rightarrow \infty} L(x(r_k), \lambda(r_k)) = \min_R f(x) = f^* \quad (23)$$

и

$$\lim_{k \rightarrow \infty} \lambda_s(r_k) g_s(x(r_k)) = 0, \quad s = [1, m], \quad (24)$$

где

$$R = \{x | g_s(x) \geq 0, \quad s = [1, m]\} \quad (25)$$

и

$$L(x, \lambda) = f(x) - \sum_{s=1}^m \lambda_s g_s(x) \quad (26)$$

– функция Лагранжа.

Теорема 6. Пусть u^* есть решение задачи верхнего уровня, x^* – решение задачи нижнего уровня для $u = u^*$, функции $f(x)$, $-g_s(x)$, $s = [1, m]$ – локально выпуклы. Тогда существуют числа $\lambda_s^* \geq 0$, $s = [1, m]$ такие, что

$$\operatorname{grad}_u f(x^*, u^*) + \sum_{s=1}^m \lambda_s^* \cdot \operatorname{grad}_u g_s(x^*, u^*) = 0, \quad (27)$$

$$\lambda_s^* g_s(x^*, u^*) = 0. \quad (28)$$

Условия существования решения задачи верхнего уровня дает следующая теорема.

Теорема 5. Пусть допустимое множество $R = \{(x, u) \in E^n \times E^l \mid g_s(x, u) \geq 0, s = [1, m]\}$ исходной задачи (1) – компакт, тогда решение задачи верхнего уровня (3) существует.

Так как аналитическое представление вектор-функции $x^+(r, u)$ получить невозможно, то для практического решения задачи верхнего уровня естественно использовать стандартную итеративную процедуру построения последовательных приближений к искомому решению u^* в пространстве E^l , k -й шаг которой состоит из следующего набора операторов:

1. для некоторого исходного приближения u_k решается задача нижнего уровня,
2. затем оцениваются как величина σ_k , так и направление шага w_k ,
3. выполняется переход к точке u_{k+1} по формуле

$$u_{k+1} = u_k + \sigma_k w_k, \quad (29)$$

4. если в u_{k+1} условие окончания итерационного процесса не выполнено, то u_k присваивается значение u_{k+1} и делается переход к пункту 1.

При подходящем выборе w_k и σ_k процедура сойдется к точке $\tilde{u}(r)$, которая является приближенным решением задачи верхнего уровня.

Главным преимуществом описанной схемы является возможность использования классических численных методов, не требующих явного аналитического представления для $E(r, u)$.

Условия сходимости определены теоремой 7 и заключаются в дифференцируемости и ограниченности снизу функционала $E(r, u)$, градиент которого кроме того должен удовлетворять условию Липшица.

Теорема 7. Пусть функции $f(x, u)$, $g_s(x, u)$, $s = [1, m]$ и $P(r, \alpha)$ дважды непрерывно дифференцируемы на компакте $M \subset E^l$. Тогда градиент функционала (11) удовлетворяет условию Липшица на множестве M .

В свою очередь, поиск w_k и σ_k , гарантирующих сходимость процедуры (29), может потребовать нахождения локальных численных характеристик функционала $E(r, u)$, таких как: его значение, значения компонент градиента и матрицы Гесса, то есть, частных производных первого и второго порядка для некоторого фиксированного u . Эти характеристики могут быть найдены по формулам

$$E'_{\nu_j} = \frac{\partial \varepsilon}{\partial \nu_j}, \quad \forall j = [1, l] \quad (30)$$

и

$$E''_{\nu_j \nu_t} = \frac{\partial^2 \varepsilon}{\partial \nu_j \partial \nu_t} + \sum_{i=1}^n \frac{\partial^2 \varepsilon}{\partial \xi_i \partial \nu_j} \frac{\partial \xi_i^+}{\partial \nu_t}, \quad \forall j, t = [1, l]. \quad (31)$$

Необходимые для использования (31) значения

$$\frac{\partial \xi_i^+}{\partial \nu_t}, \quad \forall i = [1, n], \quad \forall j = [1, l] \quad (32)$$

можно найти, решив систему линейных уравнений:

$$\sum_{i=1}^n \frac{\partial^2 \varepsilon}{\partial \xi_i \partial \xi_j} \frac{\partial \xi_i^+}{\partial \nu_t} = -\frac{\partial^2 \varepsilon}{\partial \xi_j \partial \nu_t}, \quad \forall j, t = [1, l], \quad (33)$$

которая получается при дифференцировании (8) по всем компонентам $u \in E^l$.

На практике использование предельных соотношений (13) ограничено ухудшением обусловленности уравнения (9) при уменьшении абсолютной величины r – «врожденном» недостатке метода гладких штрафных функций. Уменьшение вначале вызывает рост затрат вычислительных ресурсов на реализацию процесса (29), а затем и вовсе приводит к нарушению его сходимости. В случае, когда процедура (29) далека от завершения, различие между $x^+(r, u)$ и $x^*(u)$ не играет существенной роли, и вычисления можно выполнять при достаточно большом r , отложив его уменьшение до завершающей стадии этой процедуры.

Если требуемая точность решения уравнения (8) при этом все же не будет достигнута, то погрешность можно понизить при помощи тейлоровской аппроксимации:

$$x^+(r + \Delta r, u) = x^+(r, u) + \frac{\partial x^+}{\partial r} \Delta r + o(\Delta r), \quad (34)$$

из которой следует оценка при $\Delta r \rightarrow -r$ сверху:

$$x^*(u) = x^+(r, u) - \frac{\partial x^+}{\partial r} r + o(r). \quad (35)$$

Компоненты вектор-функции $\frac{\partial x^+(r, u)}{\partial r}$ могут быть найдены по теореме о неявной функции из системы линейных уравнений:

$$\sum_{i=1}^n \frac{\partial^2 \varepsilon}{\partial \xi_i \partial \xi_j} \frac{\partial \xi_i^+}{\partial r} = -\frac{\partial^2 \varepsilon}{\partial \xi_j \partial r}, \quad \forall j = [1, n], \quad (36)$$

получаемой при дифференцировании уравнения (8) по параметру r .

В третьей главе «Реализация метода сглаживающих штрафных функций» исследуются проблемы решения задач вида (1), допускающих параметрическую линеаризацию. То есть таких, что задача нижнего уровня (2) оказывается при фиксации параметров задачей линейного программирования с

$$\begin{aligned} f(x, u) &= \sum_{i=1}^n c_i(u) \xi_i, \\ g_s(x, u) &= \sum_{i=1}^n a_{si}(u) \xi_i - b_s(u). \end{aligned} \quad (37)$$

Основным преимуществом параметрического подхода в этом случае является возможность использовать для поиска вектора $x^*(u)$ высокоэффективный и надежный симплекс-метод. Вектор $x^*(u)$, в свою очередь, можно использовать в качестве начального приближения в процедуре решения уравнения (8).

В качестве штрафной функции выберем

$$P(r, \alpha) = \begin{cases} \frac{\alpha^2}{2r}, & \alpha < 0 \\ 0, & \alpha \geq 0, \end{cases} \quad (38)$$

Сходимость итерационной процедуры (29) при использовании штрафной функции такого типа обосновывается утверждениями теоремы 11. Процедура сглаживания, как и в общем случае, сводится к нахождению вектора $x^+(r, u)$, а также

$$\text{grad}_u E(r, u) \text{ и } \text{hess}_u E(r, u),$$

значения компонент которых вычислены в точке $x^+(r, u)$.

В рассматриваемом случае вспомогательный функционал задачи (2) выпуклый, что позволяет получить связь процедуры сглаживания с двойственной задачей. Обозначим

$$\Lambda^+(r, u) = -\frac{\partial P(r, g_s(x^+(r, u), u))}{\partial g_s}, \quad s = [1, m], \quad (39)$$

тогда условие стационарности примет вид, совпадающий по форме с условием стационарности функции Лагранжа для задачи нижнего уровня

$$\text{grad}_x f(x^+(r, u), u) - \sum_{s=1}^m \Lambda^+(r, u) \cdot \text{grad}_x g_s(x^+(r, u), u) = o, \quad (40)$$

или, с учетом линейности задачи (2),

$$c_i(u) - \sum_{s=1}^m \Lambda_s^+(r, u) a_{si}(u) = o, \quad i = [1, n]. \quad (41)$$

При этом уравнение (40) имеет в силу свойств метода штрафных функций локально единственное решение $x^+(r, u)$. Из теоремы 1 вытекает, что существуют

$$\lim_{r \rightarrow +0} \Lambda_s^+(r, u) = \Lambda_s^*(u) \geq 0, \quad s = [1, m]. \quad (42)$$

Следовательно, будет справедливо равенство

$$\text{grad}_x f(x^*(u), u) - \sum_{s=1}^m \Lambda_s^*(u) \cdot \text{grad}_x g_s(x^*(u), u) = o, \quad (43)$$

и числа $\Lambda_s^*(u), s = [1, m]$ совпадают со значениями множителей Лагранжа задачи нижнего уровня в случае, когда последние существуют и единственны. В общем же случае, числа $\Lambda_s^*(u), s = [1, m]$ являются одним из решений системы уравнений

$$\text{grad}_x f(x^*(u), u) - \sum_{s \in J(u)} \lambda_s^*(u) \cdot \text{grad}_x g_s(x^*(u), u) = o, \quad (44)$$

где $J(u)$ – множество индексов *активных ограничений* задачи нижнего уровня, то есть ограничений, для которых $g_s(x^*(u), u) = 0$. Система уравнений (44) не всегда однозначно разрешима относительно неотрицательных $\lambda_s^*(u)$. Это может иметь место, например, в случае, когда число активных ограничений больше, чем n . Равенство (43) позволяет из множества активных ограничений выделить особое подмножество $J^+(u) \subseteq J(u)$, такое что

$$J^+(u) = \{s = [1, m] | \Lambda_s^*(u) > 0\}. \quad (45)$$

Из свойств квадратичной штрафной функции $P(r, \alpha)$ при $r \rightarrow +0$ следует

$$\begin{aligned} g_s(x^+(r, u), u) &< 0, & s \in J^+(u), \\ g_s(x^+(r, u), u) &\geq 0, & s \notin J^+(u). \end{aligned} \quad (46)$$

Таким образом для ограничений с индексами $s \in J^+(u)$ метод штрафных функций генерирует точки $x^+(r, u)$ вне допустимой области, в то время как для ограничений с индексами $s \notin J^+(u)$ получаются точки внутри допустимой области. Поэтому ограничения, индексы которых принадлежат множеству $J^+(u)$, логично назвать *ре-активными*. Если ограничение не является реактивным, то значение и дифференциальные характеристики функционала $E(r, u)$ от него не зависят, и эти ограничения можно исключить из рассмотрения в точке u .

Следует отметить, что при предельном переходе понятие реактивных ограничений сохраняет свой смысл и для выпуклой задачи (2) общего вида в случае произвольной штрафной функции, удовлетворяющей условиям (6)-(7).

Рассмотрим теперь алгоритм решения двухуровневой задачи (2)-(3) при условиях (37).

Алгоритм решения задачи нижнего уровня. Входными данными для алгоритма является точка $u_k \in E^l$. На начальном этапе решается исходная линейная задача нижнего уровня:

$$\text{найти } x^*(u_k) = \arg \min_x f(x, u_k) \quad (47)$$

$$\text{при условиях } g_s(x, u_k) \geq 0, s = [1, m].$$

Если задача совместна, то сформируем множество $J(u_k)$, включающее индексы ограничений, активных в точке $x^*(u_k)$. Если же задача несовместна, то включим во множество $J(u_k)$ индексы всех ограничений задачи (47).

Теперь сформируем вспомогательную функцию:

$$\varepsilon(r, x, u_k) = f(x, u_k) + \frac{1}{2r} \sum_{s \in J(u_k)} \Theta_s(-g_s(x, u_k)) g_s^2(x, u_k) \quad (48)$$

Эта функция является выпуклой в E^n в силу линейности целевого функционала и выбранной функции штрафа, поэтому множество точек минимума вспомогательной функции не пусто. Чтобы выбрать единственную точку $x^+(r, u_k)$ может понадобиться выполнить регуляризацию вспомогательной функции.

Если размер множества $J(u_k)$ не превышает n , и градиенты ограничений из $J(u_k)$ – линейно независимы, то в силу теоремы 1 можно выделить множество $J^+(u_k)$, решив двойственную к (47) задачу. В этом случае $x^+(r, u_k)$ является решением системы линейных уравнений:

$$\frac{\partial f(x, u_k)}{\partial \xi_i} + \frac{1}{r} \sum_{s \in J^+(u_k)} g_s(x, u_k) \frac{\partial g_s(x, u_k)}{\partial \xi_i} = 0, \quad i = [1, n]. \quad (49)$$

Если же размер множества $J(u_k)$ больше n , то необходимо найти $x^+(r, u_k)$ как минимум вспомогательной функции (48). Это можно сделать, например, с помощью метода Ньютона. При решении вспомогательной задачи необходимо следить, чтобы в процессе решения не изменилось множество активных ограничений. Вычислив $x^+(r, u_k)$, можно выделить множество индексов $J^+(u_k)$.

Решением задачи нижнего уровня для заданного u_k является пара: $x^+(r, u_k)$, $J^+(u_k)$.

Алгоритм решения задачи верхнего уровня. Перейдем к решению задачи верхнего уровня. Сформируем вспомогательную функцию:

$$E(r, u) = f(x^+(u), u) + \frac{1}{2r} \sum_{s \in J^+(u_k)} g_s^2(x^+(u), u). \quad (50)$$

В зависимости от выбранного метода безусловной минимизации может потребоваться вычисление производных функции $E(r, u)$. Производные находятся по формулам (30) - (33). Вычислив производные функции $E(r, u)$ в точке u_k находим направление спуска σ_k и величину шага ω_k . Новая точка в пространстве параметров определяется по формуле:

$$u_{k+1} = u_k + \sigma_k \omega_k. \quad (51)$$

Если условие окончания процесса выполнено, то найдено оптимальное решение. В противном случае выполняется переход к следующей итерации.

В четвертой главе «Практическое применение метода сглаживающих штрафных функций» рассмотрены классы математических моделей, для исследования которых возможно применение метода сглаживания в двухуровневых параметрических задачах вида (2)-(3), а также описан пример практического использования предложенного алгоритма.

Подход, сводящий задачу (1) и двухуровневой системе задач (2)-(3), может быть использован для достаточно широкого класса практически важных задач. В работе рассмотрены соответствующие постановки.

- Для задачи оптимизации формы множества Парето в многокритериальных моделях, основанных на схеме минимизации рассогласования критериев (методов класса reference point). Следует отметить, что в том случае задача (2)-(3) является трехуровневой, поскольку постановка задачи (2) содержит в своем описании оптимальные значения целевых функций, решаемых независимо друг от друга, однокритериальных задач.
- Для задачи оптимизации на логически (и, быть может, физически) распределенной системе математических моделей. Для задач этого класса характерно использование различных целевых функционалов на нижнем и верхнем уровнях, что является обобщением постановки (2)-(3), хотя схема сглаживания остается той же самой.
- Для задач на неполных математических моделях, то есть моделях, содержащих формализованные описания не всех факторов, существенно влияющих на функционирование исследуемого объекта. Неполнота модели приводит к необходимости изменения содержательных постановок задач (2)-(3). Например, задача верхнего уровня сводится к построению многокомпонентной метрики, позволяющей количественно оценивать степень близости множества допустимых и множества целевых состояний.
- Для параметрически линеаризуемых задач оптимального управления,

то есть задач, в которых путем подходящего выбора параметризуемых переменных, задача нижнего уровня (2) оказывается линейной.

Апробация алгоритма решения выполнена для задачи оценки эффективности управления инвестициями на рынках ценных бумаг, относящейся к классу нелинейных задач оптимального управления со смешанными ограничениями и допускающей параметрическую линейризацию вида (2)-(3). Полученная двухуровневая система состояла из высокоразмерной линейной задачи (2) и низкоразмерной нелинейной невыпуклой задачи (3), для решения которой использовался модифицированный метод Ньютона.

К параметризуемым переменным исходной задачи относились характеристики доходности ценных бумаг и затрат на обслуживание кредитных линий. Переменные нижнего уровня включали как фазовые переменные: объем портфеля ценных бумаг, объем привлеченных инвестиций, текущий баланс операции, так и управления: скорости купли-продажи ценных бумаг и привлечения-возврата заемных средств. Число периодов, граничные и краевые значения фазовых переменных и управлений являлись глобальными параметрами.

Маргинальные оценки эффективности управления пакетом ценных бумаг определялись в процессе минимизации расстояния между множеством допустимых и целевых состояний, первое из которых задавалось уравнениями динамики и ограничениями на фазы и управления, а второе – условием максимизации итогового баланса счета операции.

Было решено более пятидесяти вариантов задачи как для различных априорных сценариев динамики условий рынка, так и разных значений глобальных параметров.

Формальное описание апробационной модели выполнено при помощи величин: N (количество периодов), $k = [1, N]$ (номер периода), $x_a(k)$, $x_d(k)$, $x_e(k)$, $y_d(k)$, $y_e(k)$, $y_c(k)$, $y_p(k)$. Эти величины при любом $k = [1, N - 1]$ связываются следующими динамическими соотношениями:

$$x_d(k + 1) - x_d(k) = y_d(k), \quad (52)$$

$$x_e(k + 1) - x_e(k) = y_e(k), \quad (53)$$

$$x_a(k + 1) - x_a(k) = -y_c(k)x_d(k) + y_p(k)x_e(k) + y_d(k) - y_e(k). \quad (54)$$

В данной системе разностных уравнений фазовыми переменными являются $x_a(k)$, $x_d(k)$, $x_e(k)$, линейно входящими управлениями $y_d(k)$ и $y_e(k)$, а нелинейно входящими управлениями $y_c(k)$, $y_p(k)$. Фазовые переменные и управления должны удовлетворять следующим ограничениям:

$$x_a(k) \geq 0; \quad \forall k \in [1, N], \quad (55)$$

$$x_d(k) \geq 0; \quad \forall k \in [1, N], \quad (56)$$

$$x_e(k) \geq 0; \quad \forall k \in [1, N], \quad (57)$$

$$g_d \leq y_d(k) \leq g_D, \quad (58)$$

$$g_e \leq y_e(k) \leq g_E, \quad \forall k \in [1, N], \quad (59)$$

где g_d , g_D , g_e , g_E – константы.

Начальные условия:

$$x_a(1) = x_d(1) = x_e(1) = 0. \quad (60)$$

Конечные условия:

$$x_d(N) = x_e(N) = 0. \quad (61)$$

Наконец, целевой функцией модели является $x_a(N) \rightarrow \max$.

К управлениям модели, нелинейно входящим в динамические соотношения и подлежащим оптимизации на верхнем уровне, относятся

$$y_c(k), y_p(k), \quad \forall k = [1, N].$$

Для рассматриваемой модели требовалось найти оптимальные $y_c^*(k)$ и $y_p^*(k)$ в классе кусочно-постоянных функций вида:

$$y_c(k) = \begin{cases} g_c, & k < n_1; \\ g_c + \nu_1, & n_1 \leq k < n_2; \\ g_c, & n_2 \leq k < n_3; \\ g_c + \nu_2, & n_3 \leq k < n_4; \\ g_c, & n_4 \leq k; \end{cases} \quad y_p(k) = \begin{cases} g_p, & k < m_1; \\ g_p + \nu_1, & m_1 \leq k < m_2; \\ g_p, & m_2 \leq k < m_3; \\ g_p + \nu_2, & m_3 \leq k < m_4; \\ g_p, & m_4 \leq k. \end{cases} \quad (62)$$

Причем в качестве варьируемых параметров были выбраны ν_1 и ν_2 , в то время как остальные параметры имели постоянные значения.

Были выполнены численные расчеты для одномерной и двумерной задачи верхнего уровня. Рассмотрены варианты задачи с количеством периодов

$N = 60, 150, 300$. Для каждого варианта задачи указаны: значения фиксированных параметров; зависимость управлений $y_c(k), y_p(k)$ от варьируемых параметров; график зависимости вспомогательной функции верхнего уровня от значения варьируемых параметров при различных коэффициентах штрафа; таблицы решения задачи верхнего уровня при различных коэффициентах штрафа. Характерные графики, иллюстрирующие процесс решения задачи верхнего уровня, изображены на рисунках 1 и 2.

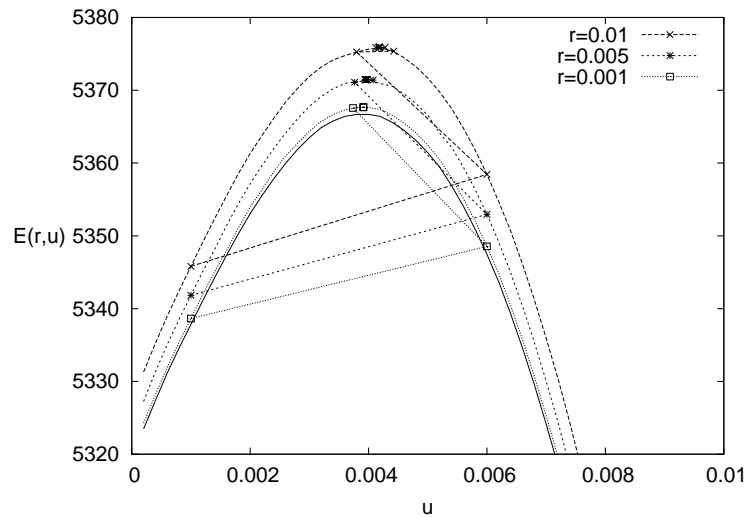


Рис. 1. Траектории движения в пространстве параметров при решении задачи об инвестициях, один параметр, 150 периодов.

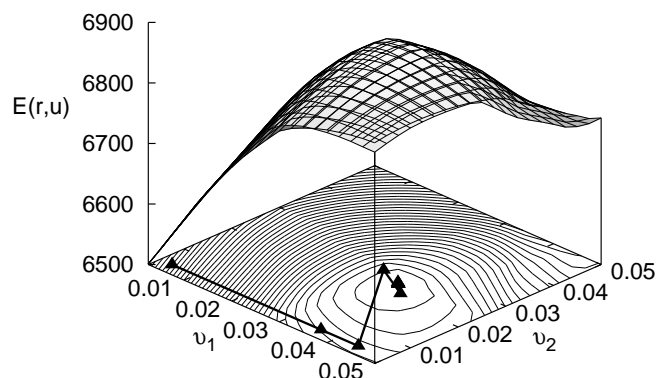


Рис. 2. Траектории движения в пространстве параметров при решении задачи об инвестициях, два параметра, 60 периодов.

В заключении изложены основные результаты работы.

В приложении приведено описание практической реализации одной из версий предложенного алгоритма, использующей современные средства программирования.

Основные результаты работы

1. Предложен и обоснован алгоритм решения двухуровневой системы задач, основанный на схеме сглаживания зависимости решения задачи математического программирования от параметров.
2. Получены необходимые и достаточные условия существования решения двухуровневой системы задач. Обоснована сходимость итерационного процесса решения двухуровневой системы задач. Доказана теорема о сходимости к двойственному решению задачи математического программирования при использовании метода гладких штрафных функций.
3. Исследована и обоснована возможность использования классических методов гладкой оптимизации для решения двухуровневой системы задач.
4. Исследованы особенности использования квадратичной штрафной функции при решении задачи математического программирования, допускающей параметрическую линейризацию. Описан алгоритм решения и выполнена компьютерная реализация предлагаемого подхода.
5. Предлагаемые алгоритмы апробированы на ряде моделей анализа эффективности управления инвестициями на рынке ценных бумаг.

Публикации автора по теме диссертации

1. *Марковцев Д. А.* Об использовании дискретных периодических моделей в задачах с неполной информацией // Труды Института системного анализа Российской академии наук. Динамика неоднородных систем. — М.: КомКнига, 2006. — Т. 10(2). — С. 51–56.

2. *Марковцев Д. А., Умнов Е. А.* Использование метода штрафных функций для решения задач параметрического программирования // Труды Института системного анализа Российской академии наук. Динамика линейных и нелинейных систем. — М.: КомКнига, 2006. — Т. 25(1). — С. 181–187.
3. *Марковцев Д. А., Умнов А. Е., Умнов Е. А.* Параметрическая оптимизация для систем математических моделей // Труды Института системного анализа Российской академии наук. Динамика неоднородных систем. — М.: Издательство ЛКИ, 2007. — Т. 31(1). — С. 42–50.
4. *Марковцев Д. А., Умнов А. Е., Умнов Е. А.* Об одной квазиклассической схеме решения задач математического программирования // Современные проблемы фундаментальной и прикладной математики. — М.: МФТИ, 2008. — С. 99–112.
5. *Умнов А. Е., Марковцев Д. А., Умнов Е. А.* Задача параметрического программирования: Учебно-методическое пособие. — М.: МФТИ, 2008, 44 с.
6. *Марковцев Д. А., Умнов А. Е., Умнов Е. А.* Параметрический анализ нелинейной модели управления инвестициями на рынке ценных бумаг // Труды Института системного анализа Российской академии наук. Динамика неоднородных систем. — М.: Книжный дом ЛИБРОКОМ, 2009. — Т. 42(2). — С. 195–202.
7. *Марковцев Д. А.* Об обосновании метода параметризации в задачах математического программирования // Современные проблемы фундаментальной и прикладной математики. — М.: МФТИ, 2012. — С. 80–88.
8. *Марковцев Д. А.* Условия сходимости итерационного процесса решения задач параметрического программирования методом гладких штрафных функций // Труды Московского физико-технического института. — М.: МФТИ, 2012. — Т. 4(1). — С. 50–54.

Марковцев Денис Анатольевич

МЕТОД ГЛАДКИХ ШТРАФНЫХ ФУНКЦИЙ В ЗАДАЧАХ
ПАРАМЕТРИЧЕСКОГО ПРОГРАММИРОВАНИЯ

АВТОРЕФЕРАТ

Подписано в печать 27.02.2012 Формат 60 ' 84 1/16. Усл. печ. л. 1,0.

Тираж 100 экз. Заказ № 624.

Федеральное государственное бюджетное образовательное учреждение
высшего профессионального образования "Московский физико-технический институт
(государственный университет)"

Отдел оперативной полиграфии "Физтех-полиграф"

141700, Московская обл., г. Долгопрудный, Институтский пер., 9