На правах рукописи

Засухина Елена Семеновна

Быстрое автоматическое дифференцирование в задачах оптимального управления

Специальность 01.01.09 - Дискретная математика и математическая кибернетика

Автореферат

диссертации на соискание ученой степени кандидата физико-математических наук

Москва – 2007

Работа выполнена в Вычислительном центре
им. А.А. Дородницына Российской академии наук

Научный руководитель: доктор физико-математических наук
Зубов Владимир Иванович

Официальные оппоненты: доктор физико-математических наук, профессор
Лотов Александр Владимирович

кандидат физико-математических наук, доцент
Бирюков Александр Гаврилович

Ведущая организация: Институт системного анализа РАН

Защита состоится « 22 » февраля 2007 г. в 16 часов на заседании диссертационного совета Д 002.017.02 Вычислительного центра им. А.А. Дородницына Российской академии наук по адресу: 119991, Москва, ул. Вавилова, 40.

С диссертацией можно ознакомиться в библиотеке Вычислительного центра им. А.А. Дородницына Российской академии наук.

Автореферат разослан « 18 » января 2007 г.

Ученый секретарь
диссертационного совета
доктор физико-математических наук В.В. Рязанов

Общая характеристика работы

Актуальность темы. Одним из самых распространенных и плодотворных подходов к решению задач оптимального управления является сведение их к задачам нелинейного программирования. Среди методов численного решения таких задач градиентные методы часто оказываются наиболее эффективными. Еще более привлекательными в смысле нахождения решения с высокой точностью являются методы, использующие вторые производные. При этом следует учесть, что точность и время вычисления производных существенно влияют на эффективность алгоритмов оптимизации в целом.

В ВЦ РАН под руководством Ю. Г. Евтушенко разработана методология быстрого автоматического дифференцирования (БАД-методология), позволяющая с единых позиций определять градиенты не только функций, заданных в явном виде, но и функций, не имеющих явного аналитического представления. Примером могут служить функции, возникающие в результате дискретной аппроксимации задач оптимального управления процессами, описываемыми обыкновенными дифференциальными уравнениями и уравнениями с частными производными. Этот подход был успешно применен при решении ряда задач оптимального управления.

Однако, с ужесточением требований на точность нахождения решения, а также, как хорошо известно, в случае "овражного"вида минимизируемой функции возникает настоятельная необходимость применения в вычислительном процессе метода Ньютона.

Цель работы. Цель работы заключается в разработке эффективного метода вычисления вторых производных функций в задачах оптимального управления, что позволяет эффективно применить метод Ньютона и получить решение с высокой точностью.

Научная новизна работы состоит в следующем.

1. В диссертационной работе с помощью обобщенной БАД-методологии получены точные формулы для вычисления вторых производных сложных функций в виде, удобном для

их применения в дискретизированных задачах оптимального управления.

2. Установлено, что после вычисления градиента функции решение сопряженного уравнения для определения множителей Лагранжа, связанных с вычислением вторых производных, не представляет трудности, т. к. основная матрица этого уравнения и основная матрица сопряженной системы, уже решенной при определении градиента функции, являются транспонированными друг другу.

3. Количество сопряженных скалярных уравнений, которые необходимо решить для определения вторых производных целевой функции, есть линейная функция от размерности переменной управления.

4. Для дискретизированной с помощью метода Рунге-Кутты произвольного порядка задачи оптимального управления процессом, описываемым обыкновенными дифференциальными уравнениями, установлен вид сопряженных систем уравнений, указан способ их решения.

5. Для описанной выше конечномерной задачи формула для вычисления вторых производных целевой функции приведена в окончательном виде, пригодном для практического применения.

6. Анализ связи организации многошаговых процессов с видом сопряженных систем и со структурой матриц-слагаемых, входящих в формулу для вычисления вторых производных целевой функции, приводит к выводу о том, что использование другого метода дискретной аппроксимации не приведет к принципиальным изменениям вида сопряженных систем и способа их решения, а также алгоритма рациональных вычислений вторых производных по полученным формулам.

7. Расчеты пяти конечномерных задач, полученных в результате дискретной аппроксимации методом Рунге-Кутты раз-

личного порядка конкретной задачи оптимального управления, с помощью метода Ньютона и градиентным методом показали, что по скорости сходимости к решению, по точности получаемого решения, по количеству итераций, приводящих к решению, и, по времени, затрачиваемому на нахождение решения, – по всем этим показателям метод Ньютона оказывается предпочтительнее градиентного метода.

8. Метод Ньютона демонстрирует преимущество по всем перечисленным выше показателям перед градиентным методом не только при проведении расчетов с "хорошим"начальным приближением, но даже в том случае, когда поиск решения начинается с произвольного приближения, не являющегося хорошим для метода Ньютона.

Обоснованность и достоверность результатов гарантируется результатами теоремы о неявной функции, обеспечена сравнением с результатами, полученными другими методами, широким тестированием численных результатов.

Практическая ценность работы заключается в том, что разработанный алгоритм вычисления вторых производных позволяет эффективно применять метод Ньютона при решении широкого круга задач оптимального управления, дискретизированных методом Рунге-Кутты произвольного порядка, и получать решение с высокой точностью.

На основе созданного алгоритма вычисления вторых производных сложных функций разработан комплекс компьютерных программ, позволяющий решать методом Ньютона задачу оптимального управления системой обыкновенных дифференциальных уравнений, заданную в стандартной форме, с произвольной правой частью и дискретизированную методом Рунге-Кутты произвольного порядка.

На защиту выносятся следующие положения:

1. Полученные формулы для вычисления вторых производных сложных функций, на переменные которых наложены связи.

2. Разработанный алгоритм вычисления вторых производных целевых функций в задачах оптимального управления, дискретизированных методом Рунге-Кутты произвольного порядка.

3. Эффективный алгоритм использования метода Ньютона с применением разработанного подхода к вычислению вторых производных при решении задач оптимального управления.

Апробация работы. Основные положения и результаты работы были доложены и обсуждены на следующих научных конференциях: Всероссийская конференция "Обратные и некорректно поставленные задачи (Москва, МГУ, 1998 г.); 3-я Международная конференция по быстрому автоматическому дифференцированию (Third International Conference on Automatic Differentiation: "From Simulation to Optimization 2000, Maison du Seminaire, Nice, France) (Франция, 2000г.); а также на научных семинарах отдела прикладных проблем оптимизации ВЦ РАН (Москва, 2001-2006); на научных семинарах отдела механики сплошных сред ВЦ РАН (Москва, 2006); на научном семинаре кафедры математических основ управления МФТИ (г. Долгопрудный, 2006).

Объём и структура. Диссертация состоит из введения, трех глав, заключения, списка используемых источников и приложения, состоящего из рисунков и таблиц. Полный объём работы, включая 17 таблиц, 64 рисунка и список литературы, насчитывающий 67 наименований, составляет 120 страниц.

Краткое содержание работы.

Во введении обоснована актуальность исследования, проведен анализ существующих работ в исследуемом направлении, дано общее описание решаемой задачи.

Первая глава посвящена выводу точных формул для вычисления вторых производных функции, на переменные которой наложены функциональные связи. В главе приводится математическая постановка задачи, демонстрируется, как с помощью обобщенной БАД-методологии можно вычислить градиент сложной

функции, обобщенная БАД-методология распространяется на случай вычисления вторых производных сложной функции, оценивается количество скалярных уравнений в сопряженном матричном уравнений, которое необходимо решить для определения вторых производных, полученные формулы для вычисления вторых производных функции адаптируются для случая многошаговых процессов.

Математическая постановка задачи.

Пусть $W : R^n \times R^r \to R^1$, $\Phi : R^n \times R^r \to R^n$ дважды непрерывно дифференцируемые отображения. Рассмотрим систему из n уравнений связи $\Phi(x, u) = 0_n$, где $x \in R^n$ и $u \in R^r$. Пусть матрица $\Phi_x^\top(x, u)$ неособенная всюду в интересующей нас области. Тогда по теореме о неявной функции существует дважды непрерывно дифференцируемая функция $x = x(u)$, такая что $\Phi(x(u), u) = 0_n$.

Задача состоит в нахождении вторых производных функции $\Omega(u) = W(x(u), u)$.

Чтобы определить с помощью обобщенной БАД-методологии градиент функции $\Omega(u)$, вводится функция Лагранжа $L(x, u, p) = W(x, u) + p^\top \Phi(x, u)$, где p - множитель Лагранжа, $p \in R^n$. Продифференцировав $L(x, u, p)$ по p и по x и приравняв нулю полученные производные, получим исходную систему уравнений связи и сопряженную систему уравнений

$$L_x(x, u, p) = W_x(x, u) + \Phi_x^\top(x, u)p = 0_n.$$

Выражение $L_u(x(u), u, p(u))$, где функции $x(u)$ и $p(u)$ удовлетворяют исходной и сопряженным системам уравнений, определяет градиент функции $\Omega(u)$

$$d\Omega(u)/du = W_u + \sum_{i=1}^{n} \Phi_u^i \cdot p^i.$$

Распространение обобщенной БАД-метологии на случай вычисления вторых производных функции $\Omega(u)$ производится путем определения производных каждой компоненты градиента этой функции с помощью описанной техники. Возникшие при этом сопряженные уравнения, собранные вместе, записываются в матричной

форме:

$$\Phi_{u^\top}(x,u) + \Phi_{x^\top}(x,u)\cdot\Lambda = 0_{nr},$$

где Λ есть $n\times r$-матрица, элементами которой являются новые (по сравнению с множителем Лагранжа p) сопряженные переменные, а формула для вычисления вторых производных функции $\Omega(u)$ выглядит следующим образом:

$$d^2\Omega(u)/du^2 = W_{x^\top u}\Lambda + (W_{x^\top u}\Lambda)^\top + W_{u^\top u} + \Lambda^\top W_{x^\top x}\Lambda +$$
$$+\sum_{i=1}^{n} p^i\cdot(\Phi^i_{u^\top u} + \Phi^i_{x^\top u}\Lambda + (\Phi^i_{x^\top u}\Lambda)^\top + \Lambda^\top\Phi^i_{x^\top x}\Lambda),$$

Основная матрица сопряженной системы относительно сопряженной переменной p и основная матрица сопряженного матричного уравнения относительно Λ – матрицы сопряженных переменных, входящих в формулу для вторых производных, являются транспонированными друг другу. Поэтому после вычисления градиента функции решение сопряженного матричного уравнения относительно матрицы Λ не представляет трудности.

Количество скалярных уравнений, которые необходимо решить для определения вторых производных функции $\Omega(u)$, равно nr, что в $(r+1)/2$ меньше количества уравнений, которые необходимо решить при определении вторых производных путем дифференцирования уравнений связи.

При рассмотрении многошаговых процессов все переменные x и u расщепляются на N переменных более низкой размерности.

Уравнения, описывающие процесс, расщепляются на N соотношений вида

$$x_i = F(i, X_i, U_i), \quad 1\le i\le N,$$

где X_i, U_i – заданные подмножества множеств $\{x_1,\dots,x_N\}$ и $\{u_1,\dots,u_N\}$ соответственно.

Во всех рассматриваемых матрицах множества строк и столбцов разделяются на N равных частей (возникают блок-строки и блок-столбцы), эти матрицы рассматриваются как матрицы блочных элементов. Вводится матрица B, разность между которой и

основной матрицей сопряженной системы относительно множителя Лагранжа p составляет единичную матрицу. Вводится также матрица C, $C = \Phi_u^\top$. Сопряженное матричное уравнение для определения матрицы Λ с введением матриц B и C представляется в виде:

$$C^\top + B^\top \cdot \Lambda = 0_{nr}.$$

Полученные формулы для нахождения вторых производных сложной функции адаптируются к случаю многошаговых процессов — записываются в терминах входных и выходных множеств.

Вторая глава посвящена нахождению вторых производных целевой функции в задачах оптимального управления процессами, описываемыми обыкновенными дифференциальным уравнениями.

Рассматривается задача оптимального управления процессом, описываемым системой обыкновенных дифференциальных уравнений

$$dx/dt = f(t, x, u, \xi), \quad T_1 \le t \le T_2, \quad x(T_1, x_0) = x_0,$$

где $f(t, x, u, \xi)$ – дважды непрерывно дифференцируемая функция, $x \in R^s$ – вектор состояний, управление u есть произвольная кусочно непрерывная функция переменной t со значениями в U, U – компакт в пространстве R^m, ξ – управляющий параметр, $\xi \in V \subset R^q$, $V-$компактное множество. Величины T_1, T_2, x_1 могут быть как фиксированными, так и свободными. Если они не фиксированы, то должны определяться из решения задачи оптимизации, и в этом случае их следует включить в вектор ξ.

Задача состоит в нахождении управляющей функции $u(t)$, $u(t) \in U$, $T_1 \le t \le T_2$, и управляющего вектора ξ, $\xi \in V$, которые минимизируют целевую функцию $W(x(T_2, x_0), \xi)$. Здесь $x(t, x_0)$ обозначает решение системы уравнений, описывающей управляемый процесс, при заданных $u(t)$ и ξ, удовлетворяющее начальному условию

$$x(T_1, x_0) = x_0.$$

Рассматриваются конечномерные задачи оптимизации, получаемые в результате дискретной аппроксимации этой задачи с по-

мощью схемы Эйлера, модифицированной схемы Эйлера и метода Рунге-Кутты порядка M. Для этих конечномерных задача устанавливается вид сопряженных систем, устанавливается структура матриц B и C.

Структура этих матриц в случае дискретизации методом Рунге-Кутты 4-го порядка имеет следующий вид (здесь и далее ненулевые блочные элементы отмечены черными кружочками):

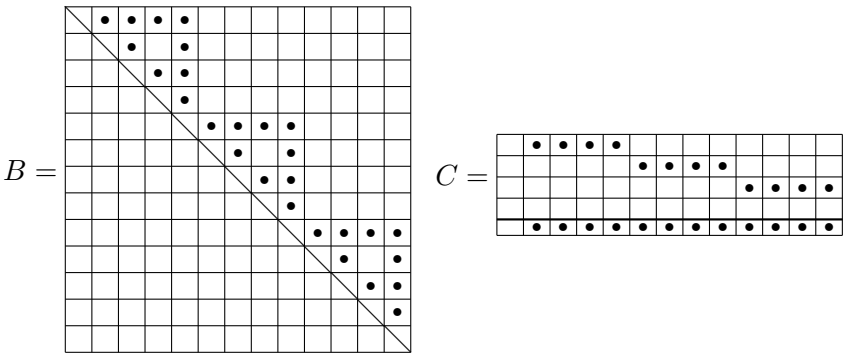

Рис. 1.

Матрицы B и C в условиях всех указанных аппроксимаций имеют сходную структуру. Матрица B имеет клеточно-наддиагональную структуру (случай схемы Эйлера) или несколько видоизменненную структуру, когда над диагональю лежит цепочка равнобедренных прямоугольных треугольников с катетом в M блочных элементов (случай метода Рунге-Кутты порядка M, $M > 1$). Поэтому сопряженные системы имеют простой вид. Сопряженная система для определения градиента функции решаются снизу-вверх методом “бегущего счета”, вычисление градиента с помощью БАД-методологии соответствует технике обратного дифференцирования. Сопряженные системы для определения вторых производных функции, напротив, решаются сверху вниз.

В матрица C в каждой блок-строке, за исключением последней (она соответствует управляющему параметру ξ и, вообще говоря, имеет отличную от остальных строк толщину) и предпоследней, содержится M ненулевых блочных элементов, расположен-

ных вплотную один к другому. В матрице эти ненулевые блоки образуют ступеньки шириной в M блоков. В случае схемы Эйлера эти ступеньки переходят в наддиагональную линию.

Анализ структур матриц B и C приводит к выводу о том, что решение сопряженного уравнения матрица Λ имеет такой вид (дискретизация методом Рунге-Кутты 4-го порядка):

$\Lambda =$

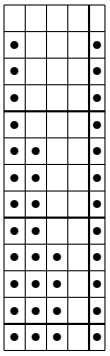

Рис. 2.

Высота ступенек в матрице Λ составляет M блочных элементов (M – порядок аппроксимации исходной задачи методом Рунге-Кутты), ширина – один блочный элемент. Из структуры матрицы Λ следует вывод о том, что количество уравнений, которое необходимо решить для нахождения вторых производных функции $\Omega(u)$, равно $(1 - s/n)(1 + q/r)nr/2$, что с увеличением N стремится к $nr/2$.

Для всех указанных конечномерных задач устанавливается структура всех слагаемых-матриц в формуле для вычисления матрицы вторых производных целевой функции. Из всех слагаемых в этой формуле выделяется множество пар, составленных из 1-го и 2-го слагаемых, а также из 2-го и 3-го слагаемых, стоящих за знаком суммы. Матрицы, составляющие каждую такую пару, являются транспонированными друг другу. Так как матрица вторых производных симметричная, то для ее определения достаточно найти ее нижний треугольник. Ввиду особой структуры матриц

внутри каждой такой пары количества вычислений при нахождении нижнего треугольника их суммы существенно сокращается. Кроме того, некоторые слагаемые (которые являются произведением нескольких матриц) из формулы оказываются нулевыми и не требуют вычислений. Указан рациональный путь вычислений по этой формуле. Для всех указанных дискретизаций исходной задачи формула для вычисления вторых производных приводится в окончательном виде, пригодном для практического использования.

В третьей главе приводятся, описываются и анализируются результаты численных расчетов. Все расчеты могут быть разделены на три группы.

К первой группе относятся расчеты, связанные с поиском оптимального управления решением уравнения Бюргерса. Рассматривалась начально-краевая задача для одномерного нестационарного уравнения Бюргерса. В качестве управления назначалась функция, определяющая значение решения на краях отрезка, а в качестве минимизируемого функционала – средне-квадратическое отклонение решения $y(x,t)$ начально-краевой задачи от некоторой предписанной (экспериментальной) функции $\hat{y}(x,t)$.

Функция $\hat{y}(x,t)$ определялась из решения прямой задачи при выбранном управлении. Это управление представляло на левом конце отрезка линейную функцию времени, изменяющуюся от 1 до 0, а на правом конце – линейную функцию времени, изменяющуюся от -1 до 0. Задача определения оптимального управления $\bar{\Psi}_{opt}$ решалась с использованием градиентов, вычисленных с помощью обобщенной БАД-методологии. В качестве начального управления, с которого начинается процесс оптимизации функционала, выбиралось $\bar{\Psi}_{init}$, полученное из выбранных граничных условий путем наложения на них гауссовского белого шума (со средне-квадратическим отклонением 0.3). На рис. 3 представлены начальное управление $\bar{\Psi}_{init}$ (пунктирная линия) и оптимальное управление $\bar{\Psi}_{opt}$. Найденное оптимальное управление $\bar{\Psi}_{opt}$ отличается от истинного значения не более, чем на 10^{-9}.

Подход, при котором для непрерывной прямой задачи строит-

ся ей сопряженная, определяется градиент функционала, а затем обе задачи и градиент функционала аппроксимируются, приводит к успеху только в том случае, когда дискретные аппроксимации прямой задачи, сопряженной задачи и градиента функционала являются согласованными. На рис. 4 представлены результаты расчетов, когда при использовании этого подхода эти аппроксимации были не согласованы. Наибольшее отличие оптимального управления от истинного (порядка 22%) наблюдается в окрестности начального момента времени $t = 0$.

В отличие от этого подхода БАД-методология автоматически обеспечивает согласованность аппроксимаций прямой и сопряженной задач и при выбранной аппроксимации прямой задачи позволяет вычислить точный градиент функционала.

Ко второй группе относятся расчеты, целью которых являлось оценить отношение времени, затрачиваемого на вычисление вторых производных сложной функции, ко времени вычисления самой функции. Для этого рассматривались три задачи оптимального управления процессами, описываемыми системами обыкновенных дифференциальных уравнений. Эти задачи дискретизировались по схеме Эйлера, для полученных задач вычислялись целевые функции и гессианы. Проводились замеры времен, затрачиваемых на вычисление целевой функции и гессиана. Расчеты проводились для сеток с различным количеством узлов, чтобы оценить, как изменяется отношение времени вычисления гессиана t_{hess} ко времени вычисления функции t_f при изменении размерности r переменной управления и как это отношение изменяется при переходе от одной задачи к другой.

Расчеты показали, что значение $(t_{hess}/t_f)/r$ незначительно изменяется в зависимости от r, но существенно зависит от конкретной задачи. Так, для задачи 3 это соотношение в среднем равно 0.3, для задачи 2 – 0.78, а для задачи 1 – 1.88. Полученные в ходе расчетов зависимости $(t_{hess}/t_f)/r$ от r демонстрируются на рис. 5. Таким образом, полученные результаты позволяют сделать заключение о том, что время, затрачиваемое на вычисление вто-

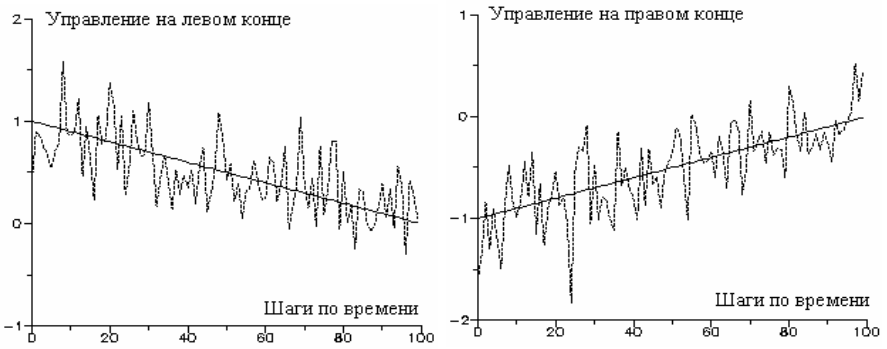

Рис. 3.

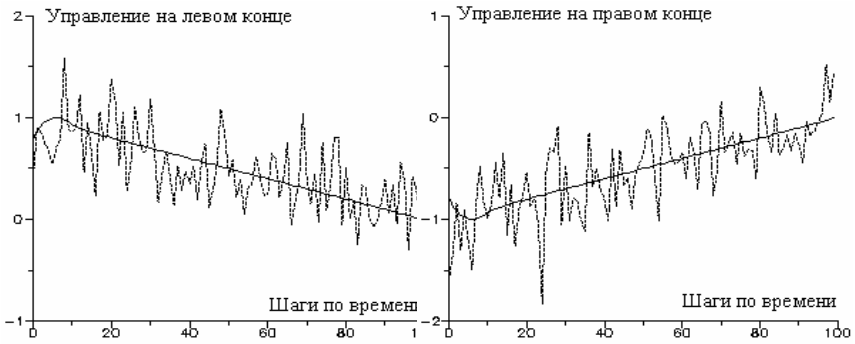

Рис. 4.

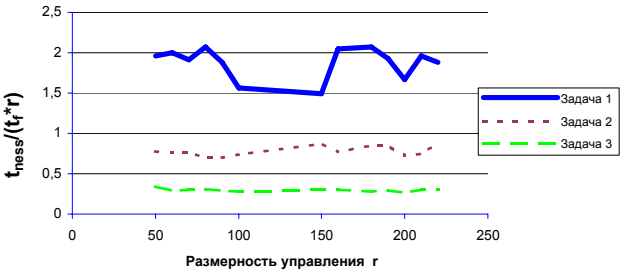

Рис. 5.

рых производных целевой функции, по отношению ко времени, затрачиваемому на вычисление самой функции, с увеличением r растет линейным образом.

К третьей группе относятся расчеты 5 конечномерных задач оптимизации. Эти задачи получены в результате дискретной аппроксимации одной из трех упомянутых выше задач оптимального управления с помощью методов Рунге-Кутты 1-го, 2-го, 3-го и 4-го порядков (рассматривалось два варианта дискретизации методом Рунге-Кутты 2-го порядка). Обозначим эти задачи как ДЭ, ДР-К 2.1, ДР-К 2.2, ДР-К 3 и ДР-К 4 соответственно. Расчеты всех задач проводились методом Ньютона с использованием полученных формул для вычисления вторых производных, а также градиентным методом.

Расчеты обоими методами проводились с произвольно выбранным начальным приближением и с "хорошим"начальным приближением. Под хорошим начальным приближением понималось начальное приближение, при котором решение могло быть найдено стандартным методом Ньютона, когда смещение на каждой итерации равнялось ньютоновскому вектору. В случае произвольно выбранного начального приближения использовался модифицированный метод Ньютона, когда до достижения хорошего начального приближения шаг вдоль ньютоновского направления определялся как точка минимума функции, полученной в результате интерполяции целевой функции вдоль выбранного направления с помощью кубических сплайнов. Эта же процедура выбора шага использовалась при расчетах градиентным методом.

Чтобы оценить влияние количества точек, определяющих интерполирующий сплайн, было проведено 4 варианта расчетов, когда интерполяция выполнялась по 5, 10, 20 и 40 точкам.

С целью нахождения хорошего начального приближения было проведено предварительное исследование: поиск решения с начальным приближением $u(t) = 0$ (которое хорошим начальным приближением не является, так как поиск решения методом Ньютона с этим начальным приближением, приводил к расхождению итерационного процесса) проводился модифицированным мето-

дом Ньютона, в котором на каждой итерации шаг α вдоль ньютоновского направления определялся в результате описанной процедуры одномерной оптимизации.

На начальных итерациях значения α были очень далеки от единицы ($10^{-2} \div 10^{-1}$), затем с увеличением номера итерации все чаще появлялись α, близкие к единице. Момент, когда, начиная с некоторой итерации, α принимало значение единица и оставалось таковым вплоть до получения решения с точностью $10^{-5} \div 10^{-7}$, принимался за момент переключения оптимизационного процесса на стандартный метод Ньютона, а управление, полученное на предшествующей итерации, – за хорошее начальное приближение.

Приближение, полученное после проведения 7 итераций, выполненных модифицированным методом Ньютоном в задаче Р-К 4 с начальным приближением $u(t) = 0$, оказалось хорошим для всех расчетных задач, кроме ДЭ. При этом одномерная оптимизация производилась с использованием сплайнов, построенных по 40 точкам.

В задаче ДЭ, использовалось свое хорошее начальное приближение, полученное после 8 итераций, выполненных модифицированным методом Ньютоном при решении этой же задачи с начальным приближением $u(t) = 0$ и с применением сплайнов, построенных по 40 точкам.

Изучался вопрос: не окажется ли градиентный метод более эффективен, чем модифицированный метод Ньютона, при нахождении хорошего начального приближения? С этой целью для каждой схемы дискретизации с помощью градиентного метода производилось то же количество итераций, которое потребовались модифицированному методу Ньютона для нахождения хорошего начального приближения. Полученное управление $u_{gr}(t)$, условно называемое "градиентным"начальным приближением, принималось в качестве начального приближения при решении задачи оптимизации методом Ньютона. Анализ расчетов показал, что градиентное начальное приближение не позволяет "включить"стандартный метод Ньютона на начальных итерациях.

Как показали расчеты задач ДЭ, ДР-К 2.1, ДР-К 2.2, ДР-К 3 и

ДР-К 4, с увеличением точности аппроксимации исходной задачи время, затрачиваемое на получение решения, увеличивается. При этом количество итераций, приводящих к решению с заданной точностью, изменяется незначительно. Сравнение оптимальных значений функционала, полученных для различных аппроксимаций, позволяет сделать вывод о том, что дискретизация по схеме Эйлера недостаточно точно аппроксимирует исходную задачу.

Метод Ньютона оказался гораздо эффективнее градиентного метода в смысле получения решения с высокой точностью: в то время как метод Ньютона приводит к решению с точностью 10^{-12} (при всех схемах дискретной аппроксимации), для градиентного метода точность 10^{-7} в получении решения (при всех схемах дискретной аппроксимации) оказывается недостижимой. Точность определяется C-нормой градиента целевой функции.

Анализ результатов расчетов, кроме того, показал, что при всех схемах интегрирования исходной системы с использованием различных количеств точек, по которым строится интерполирующий сплайн, с хорошим и произвольно выбранным начальным приближением,

- по количеству итераций, приводящих к решению;
- по времени, затрачиваемому на нахождение решения;
- и, наконец, по скорости сходимости итерационного процесса к решению, –

по всем этим показателям метод Ньютона обладает бесспорным преимуществом перед градиентным методом.

Преимущество метода Ньютона перед градиентным методом по времени, затрачиваемому на получение решения, демонстрирует табл. 1, в которой содержится отношение времен t_{gr} и t_N нахождения минимума целевого функционала градиентным методом и методом Ньютона соответственно. Это отношение t_{gr}/t_N существенно зависит от способа дискретизации исходной задачи, точности получения решения и количества точек, определяющих интерполирующий сплайн. При требуемой точности получения решения

Таблица 1: Отношение t_G/t_N.
(хорошее начальное приближение)

Количество точек		5	10	20	40
Э	10^{-3}	8.8	8.4	18.1	37.4
	10^{-5}	18.0	16.2	33.8	66.1
Р-К 2.1	10^{-3}	14.1	3.1	21.5	49.1
	10^{-5}	22.8	4.7	37.5	80.8
Р-К 2.2	10^{-3}	16.0	3.1	25.1	58.4
	10^{-5}	27.2	5.0	44.5	95.7
Р-К 3	10^{-3}	17.9	3.6	27.8	63.2
	10^{-5}	29.5	5.8	47.7	102.5
Р-К 4	10^{-3}	19.3	4.3	29.1	72.2
	10^{-5}	31.6	7.0	50.6	106.4

10^{-3} диапазон изменения этого отношения составляет от 3.1 до 72.2. При требуемой точности 10^{-5} этот диапазон составляет от 4.7 до 106.4.

Отношение времен t_{gr} и t_N, затрачиваемых на нахождение решения с произвольным начальным приближением градиентным методом и методом Ньютона, в зависимости от схемы дискретизации исходной задачи и количества точек, по которым строится интерполирующий сплайн, приводится в табл. 2. Это отношение изменяется при точности получения решения 10^{-3} от 1.7 до 14, а при точности 10^{-5} — от 2.6 до 28.5.

Преимущество по времени, затрачиваемому на нахождение решения, выглядит в данном случае более скромно, чем в случае расчетов задачи с хорошим начальным приближением. Это и понятно. Действительно, до достижения хорошего приближения каждая итерация модифицированного метода Ньютона продолжается более длительное время, чем в случае собственно метода Ньютона, так как включает в себя и время, необходимое для проведения одномерной оптимизации для получения значения α, в то время как стандартный метод Ньютона сразу полагает значение α рав-

Таблица 2: Отношение t_G/t_N.
(начальное приближение $u = 0.$)

Количество точек		5	10	20	40
Э	10^{-3}	4.4	3.1	8.5	7.5
	10^{-5}	8.7	6.8	16.2	24.39
Р-К 2.1	10^{-3}	6.7	1.9	12.7	12.0
	10^{-5}	12.8	3.0	24.8	22.0
Р-К 2.2	10^{-3}	7.3	1.7	13.3	10
	10^{-5}	14.5	2.6	28.5	20.2
Р-К 3	10^{-3}	7.6	1.8	14.0	13.0
	10^{-5}	15.0	2.8	26.8	26.6
Р-К 4	10^{-3}	10.0	1.9	13.9	9.5
	10^{-5}	18.2	3.2	27.5	23.5

ным единице.

В **заключении** приведены основные результаты, полученные в диссертации.

Основные результаты работы.

1. Обобщенная БАД-методология распространена на случай вычисления вторых производных сложной функции.

2. Получены точные формулы для вычисления вторых производных сложной функции. В эти формулы входят сопряженные переменные – множители Лагранжа, которые определяются в результате решения сопряженного линейного матричного уравнения.

3. Установлено важное свойство предлагаемого подхода: после вычисления градиента функции решение сопряженного уравнения для определения множителей Лагранжа, связанных с вычислением вторых производных, не представляет трудности, т. к. основная матрица этого уравнения и основная матрица сопряженной системы, уже решенной при опре-

делении градиента функции, являются транспонированными друг другу.

4. Полученные формулы для нахождения вторых производных сложной функции адаптированы к случаю многошаговых процессов.

5. Рассмотрена в общем виде задача оптимального управления процессом, описываемым обыкновенными дифференциальными уравнениями. Для конечномерной задачи оптимизации, полученной в результате дискретной аппроксимации этой задачи с помощью метода Рунге-Кутты произвольного порядка, установлен вид сопряженных уравнений и структура их основных матриц. Установлено, что сопряженные уравнения имеют простой вид и могут быть решены методом "бегущего счета". Анализ связи организации многошаговых процессов с видом сопряженных систем, структур основных матриц сопряженных систем позволяет сделать заключение о том, что использование другого метода дискретной аппроксимации не приведет к принципиальным изменениям вида сопряженных систем и способа их решения.

6. Для указанной конечномерной задачи оптимизации установлена структура всех матриц, входящих в формулу для вычисления матрицы вторых производных функции. Указан рациональный путь вычислений по этой формуле. Анализ связи организации многошаговых процессов с видом матриц, входящих в формулу для вычисления матрицы вторых производных целевой функции, позволяет сделать заключение о том, что использование другого метода дискретной аппроксимации не приведет к принципиальным изменениям структуры этих матриц.

7. Формулы для вычисления матрицы вторых производных целевой функции в этой задаче приведены в окончательном виде, пригодном для практического использования.

8. Расчеты 5 задач, полученных в результате дискретной аппроксимации конкретной задачи оптимального управления с помощью метода Рунге-Кутты различного порядка, методом Ньютона (с применением предлагаемого алгоритма вычисления вторых производных) и градиентным методом показали, что по скорости сходимости к решению, по точности получаемого решения, по количеству итераций, приводящих к решению, и, по времени, затрачиваемому на нахождение решения, — по всем этим показателям метод Ньютона оказывается предпочтительнее градиентного метода.

9. Показано, что метод Ньютона демонстрирует преимущество по всем перечисленным выше показателям перед градиентным методом даже в том случае, когда поиск решения начинается с произвольного приближения, не являющегося хорошим для метода Ньютона.

Список публикаций по теме диссертации

1. Ю.Г. Евтушенко, Е.С. Засухина, В.И. Зубов. О численном подходе к оптимизации решения задачи Бюргерса с помощью граничных условий // Журн. вычисл. матем. и ма-тем. физики. 1997. т. 37. №12. С.1449-1458

2. Ю.Г. Евтушенко, Е.С. Засухина, В.И. Зубов. Применение метода быстрого автоматического дифференцирования при решении обратных задач. Тезисы докладов конференции "Обратные некорректно поставленные задачи Москва, МГУ, 16-17 июня 1998 г. М: Изд-во МГУ. 1998. С. 31.

3. Yu. Evtushenko, E. Zasuhina, V. Zubov. The application of FAD-methodology for computing second order derivatives. Abstracts of Third International Conference on Automatic Differentiation: "From Simulation to Optimization June 19-23, 2000, Maison du Seminaire, Nice, France. Published by INRIA, Sophia Antipolis, France, P. 23.

4. Yu. Evtushenko, E. Zasuhina, V. Zubov. Fast AD technique and inverse problems. Abstracts of Third International Conference on Automatic Differentiation: "From Simulation to Optimization June 19-23, 2000, Maison du Seminaire, Nice, France. Published by INRIA,

Sophia Antipolis, France, P. 23-24.

5. Yu. Evtushenko, E. Zasuhina, V. Zubov. FAD Method to Compute Second Order Deriva-tives. In.: "Automatic Differentiation of Algorithms: from Simulation to Optimization New York: Inc. Springer- Verlag. 2002. P. 327-333.

6. Евтушенко Ю. Г., Засухина Е. С., Зубов В. И. Вычисление вторых производных сложной функции с помощью обобщенной БАД-методологии. М: ВЦ РАН, 2005.

7. Засухина Е. С. Применение быстрого автоматического дифференцирования для вычисления вторых производных сложных функций // Ж. вычисл. матем. и матем. физ. 2006. Т. 46. № 11. С. 1923-1949.