

На правах рукописи

СОКОЛИНСКАЯ Ирина Михайловна



УДК 519.6

**СИНТЕЗ МЕТОДОВ ОПТИМИЗАЦИИ
И ДИСКРИМИНАНТНОГО АНАЛИЗА В
МАТЕМАТИЧЕСКИХ МОДЕЛЯХ ЭКОНОМИКИ**

05.13.18 - математическое моделирование,
численные методы и комплексы программ

Автореферат
диссертации на соискание ученой степени
кандидата физико-математических наук

Челябинск – 2006

Работа выполнена на кафедре вычислительной математики Челябинского государственного университета.

Научный руководитель: доктор физико-математических наук, профессор
МАЗУРОВ Владимир Данилович.

Официальные оппоненты: академик РАН, доктор физико-математических наук, профессор
ЕРЕМИН Иван Иванович;
кандидат физико-математических наук, доцент
ЗЫКИНА Анна Владимировна.

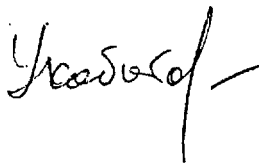
Ведущая организация: Вычислительный центр им. А.А. Дородницына РАН

Защита состоится 19 октября 2006 г. в 13 часов
на заседании диссертационного совета Д 212.296.02 при Челябинском государственном университете по адресу: 454021, Челябинск,
ул. Бр. Кашириных, 129.

С диссертацией можно ознакомиться в библиотеке Челябинского государственного университета

Автореферат разослан 29 августа 2006 г.

Ученый секретарь
диссертационного совета,
доктор физ.-мат. наук, профессор



Ухоботов В.И.

Общая характеристика работы

Актуальность темы. В практике экономико-математического моделирования часто встречаются задачи, при решении которых приходится учитывать большое число взаимосвязанных факторов, включая плохо формализуемые¹. Процесс решения плохо формализуемой задачи включает в себя преобразование ее формулировки путем уточнений и упрощений. В результате этого процесса мы получаем формализованную задачу, имеющую некоторое отношение к исходной постановке задачи. Во многих случаях удается построить формализованную задачу в виде задачи линейного программирования. Однако решение формализованной задачи может сколь угодно сильно отличаться от решения исходной задачи в силу наличия неформализованных ограничений.

Помимо формальных методов и моделей для решения плохо формализуемой задачи используются средства неформального анализа, включающие в себя суждения экспертов для учета так называемых «внемодельных» факторов. При этом человеческий компонент в процессе экспертизы может постепенно, с помощью обучения, заменяться машинным компонентом путем использования экспертных систем или нейросетевых программ. Однако при решении задач линейного программирования большой размерности эти подходы могут оказаться неэффективными, если их применять для решения всей задачи в целом. Это объясняется тем, что трудно адекватно настроить (обучить) экспертную систему или нейронную сеть в случае, когда система ограничений содержит тысячи линейных неравенств и десятки тысяч переменных. В таких ситуациях перспективным является подход, основанный на сочетании методов линейной оптимизации с процедурами экспертного оценивания. Алгоритм решения задачи в этом случае строится как итерационный процесс, в ходе которого шаг за шагом происходит уточнение ограничений, относящихся к неформализованной части задачи. Приближенные решения исходной задачи, получаемые в ходе итерационного процесса, формируют множество образцов, квалифицируемых экспертом. Применяя к этому множеству образцов процедуру дискриминантного анализа, мы получаем разделяющую поверхность, аппроксимирующую неформализованные ограничения. В соответствии с этим *актуальной* является задача разработки, анализа и реализации на ЭВМ алгоритмов решения задач линейного программирования с неформализованными ограничениями путем синтеза методов дискриминантного анализа и линейной оптимизации.

Цель и задачи исследования. Цель данной работы состояла в разработке и исследовании методов и алгоритмов решения задач линейного программирования с неформализованными ограничениями и их реализации в виде программного комплекса. Для достижения этой цели необходимо было решить следующие задачи:

1. Построить математическую модель задачи *линейного программирования с неформализованными ограничениями* (ЛПНО) и описать общий подход к ее решению.

¹ Еремин И. И., Мазуров Вл. Д. Вопросы оптимизации и распознавания образов. - Свердловск: Сред.-Урал. кн. изд-во, 1979.

2. На базе данного подхода разработать метод решения задачи ЛПНО, допускающий реализацию в виде алгоритма для ЭВМ, и исследовать сходимость этого метода.
3. На основе описанного метода разработать и исследовать алгоритм ЛП-ДА (Линейное Программирование – Дискриминантный Анализ), соединяющий алгоритмы линейного программирования с алгоритмами дискриминантного анализа.
4. Спроектировать и реализовать программный комплекс для решения задач ЛПНО, использующий предложенные методы и алгоритмы.
5. Провести вычислительные эксперименты для анализа эффективности предложенного подхода.
6. Разработать и реализовать параллельную версию алгоритма ЛП-ДА. Провести вычислительные эксперименты на кластерной системе для исследования ускорения и расширяемости параллельного алгоритма.

Методы исследования. В исследованиях, проводимых в диссертационной работе, используется аппарат теории математического программирования и распознавания образов, применяются методы дискриминантного анализа и линейной оптимизации.

Научная новизна работы заключается в следующем:

- 1) предложен метод решения задачи линейного программирования с неформализованными ограничениями, базирующийся на синтезе методов дискриминантного анализа и методов линейной оптимизации, и доказана сходимость генерируемой им итерационной последовательности к точному решению задачи ЛПНО;
- 2) предложен оригинальный метод осцилляций, позволяющий существенно ускорить сходимость последовательности приближений к точному решению задачи ЛПНО;
- 3) разработан новый алгоритм, соединяющий симплекс-метод и метод линейной коррекции, и предложена его параллельная версия;
- 4) выполнена реализация алгоритма решения задач ЛПНО в виде программного комплекса для персональных компьютеров и многопроцессорных систем на основе стандарта MPI-2.

Теоретическая ценность работы состоит в том, что в ней представлены доказательства теоремы об устойчивости задачи линейного программирования по неформализованным ограничениям и теоремы о сходимости метода, соединяющего дискриминантный анализ и рандомизацию с линейной оптимизацией. **Практическая ценность** работы заключается в том, что предложенный программный комплекс в сочетании с системами экспертного оценивания может быть использован для решения плохо формализуемых задач, возникающих в экономико-математическом моделировании.

Апробация работы. Основные положения диссертационной работы, разработанные модели, методы, алгоритмы и результаты вычислительных экспериментов докладывались автором на следующих международных и всероссийских научных конференциях:

- на Всероссийской научной конференции "Алгоритмический анализ неустойчивых задач" (2-6 февраля 2004 г., Екатеринбург);
- на Международной научной конференции "Системный Анализ и Информационные Технологии" САИТ-2005 (12-16 сентября 2005 г., Переславль-Залесский);
- на III Всероссийской научной конференции "Проблемы оптимизации и экономические приложения" (Омск, 11-15 июля 2006 г.).

Публикации. Основные научные результаты диссертации опубликованы в 5 печатных работах, приведенных в конце автореферата. Статья [2] опубликована в научном журнале «Pattern Recognition and Image Analysis», включенном ВАК в перечень журналов, в которых должны быть опубликованы основные результаты диссертаций на соискание ученой степени доктора наук. В статье [2] Вл.Д. Мазурову принадлежит общая постановка задачи; И.М. Соколинской принадлежат формальное описание алгоритма \mathcal{L} порождения образцов, формулировки и доказательства всех теорем и лемм.

Структура и объем работы. Диссертация состоит из введения, четырех глав, заключения, библиографии и приложения, в котором приводятся основные обозначения, используемые в диссертационной работе. Объем диссертации составляет 92 страницы, объем библиографии – 77 наименований.

Содержание работы

Во введении приводится обоснование актуальности темы, формулируются цели работы и кратко излагается содержание диссертации.

Первая глава, «Синтез дискриминантного анализа и линейной оптимизации», посвящена описанию и исследованию общего метода решения задач ЛПНО, сочетающего дискриминантный анализ, линейную оптимизацию и рандомизацию. Сначала рассматривается случай одного неформализованного ограничения, а затем показывается, как полученные результаты могут быть обобщены на случай произвольного количества ограничений.

Предположим сначала, что только одно ограничение не может быть формализовано. Данную ситуацию можно записать в виде следующей задачи *линейного программирования с неформализованным ограничением* (ЛПНО) в пространстве \mathbb{R}^n :

$$\max \{(c, x) \mid Ax \leq b, \bar{\varphi}(x) \leq 0\} \quad (1)$$

Здесь неравенство $\bar{\varphi}(x) \leq 0$ играет роль неформализованного ограничения в том смысле, что нам не известен вид функции $\bar{\varphi}$. Будем предполагать, что $\bar{\varphi}$ принадлежит к классу аффинных функций. Предположим также, что задача (1) имеет единственное решение:

$$\bar{x} = \arg \max \{(c, x) \mid Ax \leq b, \bar{\varphi}(x) \leq 0\}.$$

Неформализованное ограничение $\bar{\varphi}(x) \leq 0$ представляется двумя информационными ресурсами: *экспертом* и *набором образцов*. В качестве абстрактной модели эксперта введем функцию $\epsilon(x) = \text{sgn}(\bar{\varphi}(x))$. С помощью

функции эксперта ϵ мы можем для любого $x \in \mathbb{R}^n$ определить, удовлетворяет x неформализованному ограничению или нет, используя следующую систему правил:

$$\begin{cases} \epsilon(x) = 1 & \Leftrightarrow \bar{\varphi}(x) > 0; \\ \epsilon(x) = 0 & \Leftrightarrow \bar{\varphi}(x) = 0; \\ \epsilon(x) = -1 & \Leftrightarrow \bar{\varphi}(x) < 0. \end{cases} \quad (2)$$

Набор образцов $\mathfrak{B} = \{A, B\}$ определяется путем задания пары конечных непересекающихся множеств A и B из пространства \mathbb{R}^n , удовлетворяющих условиям

$$\begin{cases} \epsilon(x) = 1, \quad \forall x \in A; \\ \epsilon(x) = -1, \quad \forall x \in B. \end{cases} \quad (3)$$

Применяя методы дискриминантного анализа, мы можем найти для множеств A и B разделяющую функцию $\bar{\varphi} \in \Phi$ (Φ – класс всех аффинных функций), удовлетворяющую условиям

$$\begin{cases} \bar{\varphi}(x) > 0, & x \in A; \\ \bar{\varphi}(x) < 0, & x \in B \end{cases}$$

(существование $\bar{\varphi}$ следует из (2), (3) и предположения, что $\bar{\varphi} \in \Phi$). В качестве приближенного решения задачи ЛПНО мы можем теперь взять решение (если оно существует) следующей задачи линейного программирования, называемой *аппроксимационной* (ЛПА):

$$\bar{x} = \arg \max \{(c, x) \mid Ax \leq b, \bar{\varphi}(x) \leq 0\}.$$

Далее формулируется и доказывается теорема об устойчивости задачи линейного программирования по неформализованному ограничению:

Теорема 1. Задача ЛПНО (1) является устойчивой по $\bar{\varphi}$.

Эта теорема играет важную роль при решении проблемы сходимости.

Далее описывается алгоритм получения новых образцов, служащий основой итерационного процесса. При этом используются следующие обозначения. $D = \{x \mid Ax \leq b\}$ – многогранник, определяемый формализованной частью задачи ЛПНО. Предполагается, что D – телесное множество (то есть D содержит внутренние точки), и что D ограничен. Перепишем задачу ЛПНО в виде

$$\max \{(c, x) \mid x \in D; \bar{\varphi}(x) \leq 0\}. \quad (4)$$

Будем рассматривать случай, когда задача (4) имеет единственное решение:

$$\bar{x} = \arg \max \{(c, x) \mid x \in D; \bar{\varphi}(x) \leq 0\}.$$

Обозначим как *ЛПФ задачу линейного программирования с формализованной частью*:

$$\max \{(c, x) \mid x \in D\}. \quad (5)$$

Предполагаем, что задача (5) также имеет единственное решение:

$$\hat{x} = \arg \max \{(c, x) \mid x \in D\},$$

причем

$$\hat{x} \neq \bar{x}.$$

Пусть $\bar{P} = \{x \mid \bar{\varphi}(x) \leq 0\}$ – множество значений, удовлетворяющих неформализованному ограничению $\bar{\varphi}(x) \leq 0$. Так как $\bar{\varphi}$ является аффинной функцией, то \bar{P} представляет собой замкнутое полупространство. Допустимое множество значений задачи (4) будет иметь вид $D \cap \bar{P}$. Будем предполагать, что $D \cap \bar{P}$ содержит внутренние точки. Определим: $\bar{H} = \{x \mid \bar{\varphi}(x) = 0\}$ – гиперплоскость, определяемая неформализованным ограничением.

Алгоритм \mathcal{L} . Пусть заданы конечные множества A^0 и B^0 , такие, что $A^0 \subset D \setminus \bar{P}$, $B^0 \subset D \cap (\bar{P} \setminus \bar{H})$. Множества A^0 и B^0 задают начальный набор образцов \mathfrak{P}^0 . Будем предполагать, что $A^0 \neq \emptyset$ и $B^0 \neq \emptyset$. Алгоритм \mathcal{L} состоит из следующих шагов:

Шаг 0. $k := 0$.

Шаг 1 (дискриминантный анализ). Найдем аффинную функцию $\bar{\varphi}^k$, удовлетворяющую условиям

$$\forall x \in A^k : \bar{\varphi}^k(x) > 0, \quad \forall x \in B^k : \bar{\varphi}^k(x) < 0.$$

и условию *нормализации*: вектор нормали к гиперплоскости $\bar{H}^k = \{x \mid \bar{\varphi}^k(x) = 0\}$ должен быть коллинеарен вектору нормали к гиперплоскости \bar{H} .

Шаг 2 (порождение образцов). Получим новый образец как решение следующей задачи ЛП A^k :

$$\hat{x}^k = \arg \max \{(c, x) \mid x \in D, \bar{\varphi}^k(x) \leq 0\}.$$

Шаг 3 (экспертиза и пополнение \mathfrak{P}).

Если

$$(c, \hat{x}^k) \notin \{(c, x) \mid x \in A^k \cup B^k \cup \bar{H}\}, \quad (6)$$

то положим

$$A^{k+1} = \begin{cases} A^k, & \epsilon(\hat{x}^k) = -1 \\ A^k \cup \{\hat{x}^k\}, & \epsilon(\hat{x}^k) = 1 \end{cases}$$

$$B^{k+1} = \begin{cases} B^k \cup \{\hat{x}^k\}, & \epsilon(\hat{x}^k) = -1 \\ B^k, & \epsilon(\hat{x}^k) = 1 \end{cases}$$

Шаг 4 (рандомизация). Если условие (6) не выполняется, то выполним *процедуру рандомизации*. Пусть

$$\mathfrak{R}^p(C) = \{x + \tau_i^p(x) \mid \forall x \in C; i = 1, \dots, m\},$$

где $\tau_i^\rho(x)$ – случайный вектор в ρ -окрестности точки x . Будем называть ρ радиусом рандомизации, m – мощностью рандомизации ($m \geq 1$). Выберем ρ_k такое, что

$$0 < \rho_k < \min(\text{dist}(\mathbf{A}^k, \bar{H}), \text{dist}(\mathbf{B}^k, \bar{H}))$$

(здесь $\text{dist}(X, Y) = \inf \{\|x - y\| : x \in X, y \in Y\}$).

Положим:

$$\mathbf{A}^{k+1} = \mathbf{A}^k \cup \mathfrak{R}^{\rho_k}(\mathbf{A}^k), \mathbf{B}^{k+1} = \mathbf{B}^k \cup \mathfrak{R}^{\rho_k}(\mathbf{B}^k).$$

Шаг 5. $k := k + 1$.

Шаг 6. Перейти на шаг 1.

Применение алгоритма \mathfrak{L} для нахождения решения задачи ЛПНО (4) заключается в последовательном вычислении точек \tilde{x}^k , являющихся решением задач ЛПА^k:

$$\max \{(c, x) \mid x \in D, \tilde{\varphi}^k(x) \leq 0\}.$$

Процесс завершается при выполнении некоторого критерия, указывающего на то, что очередной \tilde{x}^k оказался в достаточной близости от точного решения \bar{x} .

Центральной частью первой главы является доказательство сходимости полученной последовательности приближенных решений к точному решению исходной задачи.

Теорема 2. Для последовательности $\tilde{x}^0, \tilde{x}^1, \dots, \tilde{x}^k$, определяемой алгоритмом \mathfrak{L} , имеем:

$$\{\tilde{x}^k\}_{k=0}^\infty \rightarrow \bar{x}.$$

Для доказательства теоремы 2 используются леммы 1-3.

Лемма 1. Пусть для всех достаточно больших k задача ЛПА^k

$$\tilde{\mu}^k = \max \{(c, x) \mid x \in D, \tilde{\varphi}^k(x) \leq 0\},$$

порождаемая алгоритмом \mathfrak{L} , имеет бесконечное множество решений $\bar{X}^k = \{x \mid x \in D \cap \bar{P}^k; (c, x) = \tilde{\mu}^k\}$. Тогда

$$\lim_{k \rightarrow \infty} \text{dist}(\bar{X}^k, \bar{H}) = 0 \Rightarrow \bar{x} = \bar{x}.$$

Лемма 2. В контексте алгоритма \mathfrak{L} справедливо следующее утверждение:

$$\lim_{k \rightarrow \infty} \text{dist}(\bar{H}^k, \bar{H}) = 0 \Rightarrow \{\tilde{x}^k\}_{k=0}^\infty \rightarrow \bar{x}.$$

Лемма 3. Применительно к алгоритму \mathfrak{L} хотя бы одно из следующих условий истинно:

$$\{\tilde{x}^k\}_{k=0}^\infty \rightarrow \bar{x};$$

$$\lim_{k \rightarrow \infty} \text{dist}(\text{co } \mathbf{A}^k, \text{co } \mathbf{B}^k) = 0$$

(co – выпуклая оболочка).

Реализация алгоритма \mathcal{L} непосредственно в том виде, как он был описан выше, на практике не всегда возможна. Прежде всего это относится к условию нормализации. Во многих случаях эксперт не в состоянии определить нормаль к гиперплоскости, задаваемой неформализованным ограничением. Следующая теорема показывает, что при определенных обстоятельствах неформализованным ограничением можно пренебречь.

Теорема 3. Пусть L^ε – плоский слой толщины $\varepsilon > 0$, серединой которого является гиперплоскость \bar{H} , задаваемая неформализованным ограничением. Обозначим через \mathcal{L}' алгоритм, получающийся из алгоритма \mathcal{L} путем снятия ограничения на $\bar{\varphi}^k$, связанного с выполнением условия нормализации. Тогда, если для алгоритма \mathcal{L}' выполняется условие

$$\forall k' : \exists k \geq k' : \bar{x}^k \in \mathbf{A}^k \cup \mathbf{B}^k \cup \bar{H},$$

то

$$\forall \varepsilon > 0 : \exists k' : \forall k \geq k' : (\bar{H}^k \cap D) \subset (L^\varepsilon \cap D).$$

Теорема 3 показывает, что, если при поиске приближенного решения с помощью алгоритма \mathcal{L}' регулярно выполняется рандомизация, то при k , стремящемся к бесконечности, \bar{H}^k «стремится стать параллельной» к \bar{H} , а это означает, что в данном случае мы можем пренебречь на практике условием нормализации.

Далее рассматривается случай нескольких неформализованных ограничений. Пусть имеется задача линейного программирования с несколькими неформализованными ограничениями (ЛПНО*) в пространстве \mathbb{R}^n :

$$\bar{x} \in \text{Arg max} \{ (c, x) \mid Ax \leq b, \bar{\varphi}_1(x) \leq 0, \dots, \bar{\varphi}_q(x) \leq 0 \}. \quad (7)$$

Здесь неравенства $\bar{\varphi}_1(x) \leq 0, \dots, \bar{\varphi}_q(x) \leq 0$ играют роль неформализованных ограничений в том смысле, что нам не известен вид функций $\bar{\varphi}_i$ ($i = 1, \dots, q$). Будем предполагать, что функции $\bar{\varphi}_i$ принадлежат к классу аффинных функций, и что задача (7) имеет единственное решение. Пусть $\bar{P}_i = \{x \mid \bar{\varphi}_i(x) \leq 0\}$ – множество значений, удовлетворяющих неформализованному ограничению $\bar{\varphi}_i(x) \leq 0$. Так как $\bar{\varphi}_i$ является аффинной функцией, то \bar{P}_i представляет собой замкнутое полупространство. Пусть

$\bar{P}^* = \bigcap_{i=1}^q \bar{P}_i$. Тогда допустимое множество задачи ЛПНО* будет иметь вид

$D \cap \bar{P}^*$. Будем предполагать, что $D \cap \bar{P}^*$ содержит внутренние точки.

Имеет место следующее обобщение теоремы 1 для случая нескольких неформализованных ограничений.

Теорема 4. Задача ЛПНО* (7) является устойчивой по $\bar{\varphi}_1, \dots, \bar{\varphi}_q$.

Решение задачи ЛПНО* может быть сведено к решению нескольких задач ЛПНО с одним неформализованным ограничением следующим образом. Рассмотрим следующую последовательность *независимых* задач ЛПНО, каждая из которых имеет только одно неформализованное ограничение

$$\text{ЛПНО}_1: \quad \bar{x}_1 = \arg \max \{(c, x) \mid Ax \leq b, \bar{\varphi}_1(x) \leq 0\};$$

...

$$\text{ЛПНО}_q: \quad \bar{x}_q = \arg \max \{(c, x) \mid Ax \leq b, \bar{\varphi}_q(x) \leq 0\}$$

(предполагается, что каждая из этих задач имеет единственное решение).

Используя алгоритм \mathcal{L} , мы можем построить для каждой задачи ЛПНО_i

($i = 1, \dots, q$) последовательность аффинных *разделяющих* функций $\bar{\varphi}_i^1, \dots, \bar{\varphi}_i^k$.

Сконструируем следующую задачу линейного программирования:

$$\text{ЛП}^k: \quad \hat{x}^k \in \text{Arg max} \{(c, x) \mid Ax \leq b, \bar{\varphi}_1^k(x) \leq 0, \dots, \bar{\varphi}_q^k(x) \leq 0\}. \quad (8)$$

Теорема 5. Пусть задача ЛПНО* (7) имеет единственное решение \bar{x} . Пусть каждая задача ЛПНО_i ($1 \leq i \leq q$) также имеет единственное решение \bar{x}_i , причем $\bar{x}_i \neq \hat{x}$, где $\hat{x} = \arg \max \{(c, x) \mid Ax \leq b\}$ – единственное решение задачи ЛПФ. Пусть \hat{x}^k – решение задачи ЛП^k. Тогда

$$\{\hat{x}^k\}_{k=0}^{\infty} \rightarrow \bar{x},$$

то есть при достаточно большом k в качестве приближенного решения задачи ЛПНО* (7) можно взять решение задачи ЛП^k (8).

Формулировка теоремы 5 опирается на предположение, что все задачи ЛПНО_i имеют решение \bar{x}_i , не совпадающее с решением \hat{x} задачи ЛПФ. На практике это требование мы можем выполнить следующим образом. На первом этапе находим приближенные значения коэффициентов и свободных членов для всех неформализованных ограничений $\bar{\varphi}_i(x) \leq 0$, удовлетворяющих условию $\bar{x}_i \neq \hat{x}$ (каждое ограничение вычисляется *независимо* от других). Заметим, что если ни одно неформализованное ограничение не удовлетворяет условию $\bar{x}_i \neq \hat{x}$, то решением задачи ЛПНО будет \hat{x} . На втором этапе добавляем к системе ограничений задачи ЛПФ найденные на первом этапе *приближения* неформализованных ограничений и полагаем \hat{x} равным решению расширенной задачи ЛПФ. Из оставшихся неформализованных ограничений снова выбираем все, удовлетворяющие условию $\bar{x}_i \neq \hat{x}$, и находим их приближенный вид. Процесс заканчивается, когда либо все неформализованные ограничения будут найдены, либо останутся несколько ограничений, для которых $\bar{x}_i = \hat{x}$ (эти ограничения можно отбросить). Последняя расширенная задача ЛПФ даст приближенное решение исходной задачи ЛПНО*.

Во второй главе, «Алгоритм ЛП-ДА», предлагается вариант реализации метода, описанного в первой главе, в виде алгоритма, основанного на синтезе алгоритмов линейного программирования и дискриминантного анализа. Такую комбинацию мы обозначаем как *алгоритм ЛП-ДА*. На первом шаге формируется начальный набор образцов $\mathfrak{F} = \{A, B\}$ путем задания пары конечных непересекающихся множеств A и B из пространства \mathbb{R}^n , удовлетворяющих условиям

$$\begin{cases} \epsilon(x) = 1, & \forall x \in A; \\ \epsilon(x) = -1, & \forall x \in B. \end{cases}$$

На следующем шаге с помощью какого-либо метода дискриминантного анализа строится аффинная функция $\tilde{\varphi}$, разделяющая множества A и B :

$$\begin{cases} \tilde{\varphi}(x) > 0, & \forall x \in A; \\ \tilde{\varphi}(x) < 0, & \forall x \in B. \end{cases}$$

В качестве метода дискриминантного анализа могут фигурировать, например, метод линейной коррекции, метод комитетов или фейеровские методы.

Далее с использованием полученной разделяющей функции строится *аппроксимационная задача* линейного программирования (ЛПА):

$$\tilde{x} \in \text{Arg max} \{ (c, x) \mid Ax \leq b, \tilde{\varphi}(x) \leq 0 \}.$$

Аппроксимационная задача решается в блоке *оптимизации* одним из методов линейного программирования. Здесь может использоваться, например, симплекс-метод или методы фейеровского типа.

С помощью *критерия завершения* проверяется, является ли полученное решение \tilde{x} допустимым приближением к точному решению \bar{x} . Если точка \tilde{x} не удовлетворяет критерию завершения, то она направляется в блок *экспертизы*, где к ней применяется функция эксперта ϵ . Если $\epsilon(\tilde{x}) = 1$, то точка \tilde{x} добавляется в множество A , если $\epsilon(\tilde{x}) = -1$, то точка \tilde{x} добавляется в множество B . После этого описанный процесс повторяется применительно к расширенному набору образцов.

При проведении экспертизы и пополнении набора образцов могут возникать коллизии следующих трех типов:

1. $\epsilon(\tilde{x}) = 0$;
2. $\epsilon(\tilde{x}) = 1$ и $\tilde{x} \in A$ (образец уже присутствует в множестве A);
3. $\epsilon(\tilde{x}) = -1$ и $\tilde{x} \in B$ (образец уже присутствует в множестве B).

В этом случае к \tilde{x} применяется *процедура рандомизации*. Точки, полученные в результате процедуры рандомизации, вновь подвергаются экспертизе и используются для пополнения набора образцов.

Эффективную работу алгоритма ЛП-ДА обеспечивает набор образцов $\mathfrak{F} = \{A, B\}$, удовлетворяющий следующим условиям:

$$\begin{aligned} A &\subset D \setminus \bar{P}, & B &\subset D \cap (\bar{P} \setminus \bar{H}), \\ A &\neq \emptyset, & B &\neq \emptyset. \end{aligned}$$

При этом мы предполагаем, что множества $D \setminus \bar{P}$ и $D \cap (\bar{P} \setminus \bar{H})$ имеют внутренние точки. В этом случае аппроксимационная задача всегда будет иметь решение. В соответствии с этим можно использовать следующий алгоритм для формирования начального набора образцов. Мы находим все вершины многогранника D , применяем к ним функцию эксперта ϵ и формируем множества A и B по следующему правилу:

$$A = \{x \mid \epsilon(x) = 1\},$$

$$B = \{x \mid \epsilon(x) = -1\}.$$

Точки, для которых $\epsilon(x) = 0$, отбрасываются.

Для реализации алгоритма ЛП-ДА в виде программы нам необходим некоторый критерий близости решения аппроксимационной задачи к точному решению задачи ЛПНО. Этот критерий необходим для завершения итерационного процесса. Пусть $\bar{x}^0, \bar{x}^1, \dots, \bar{x}^k$ — последовательность решений аппроксимационных задач, полученных при работе алгоритма ЛП-ДА. Тогда в качестве искомого критерия может выступать условие вида

$$\sum_{i=1}^l \|\bar{x}^k - \bar{x}^{k-i}\| < \epsilon,$$

где $\epsilon > 0$ — заданное вещественное число, определяющее точность вычислений. Здесь l — фиксированное целое положительное число, являющееся параметром алгоритма ЛП-ДА.

Для разрешения коллизий, возникающих при проведении экспертизы, мы используем следующую процедуру рандомизации. Определим случайное отображение $\tau^\rho : \mathbb{R}^n \rightarrow \mathbb{R}^n$, обладающее свойством: $\|\tau^\rho(x) - x\| < \rho$. Отображение $\tau^\rho(x)$ порождает случайные точки внутри сферы радиуса ρ с центром в x . Вещественное число $\rho > 0$ является параметром алгоритма ЛП-ДА, задающим радиус рандомизации. Процедура рандомизации заключается в построении с помощью отображения $\tau^\rho(\bar{x})$ множества точек $\mathfrak{X}_m^\rho = \{\hat{x}^1, \dots, \hat{x}^m\}$, удовлетворяющих условию

$$\hat{x}^i \notin A \cup B \cup \bar{H}, \quad \|\hat{x}^i - \bar{x}\| < \rho \quad (i = 1, \dots, m).$$

Целое положительное число m является параметром алгоритма ЛП-ДА, задающим мощность рандомизации. Все точки множества \mathfrak{X}_m^ρ подвергаются экспертизе и используются для пополнения набора образцов на текущей итерации.

Описывается оригинальный метод осцилляций, позволяющий во многих случаях значительно повысить эффективность алгоритма ЛП-ДА. Суть метода осцилляций состоит в том, что на каждой итерации алгоритма ЛП-ДА кроме аппроксимационной задачи решается следующая дополнительная задача линейного программирования (ЛПД):

$$\tilde{x} \in \text{Arg min} \{c, x \mid Ax \leq b, \varphi(x) \geq 0\}$$

Решение \tilde{x} дополнительной задачи подвергается процедуре экспертизы и, в случае отсутствия коллизий, добавляется к набору образцов $\mathfrak{X} = \{A, B\}$.

При возникновении коллизии к точке \bar{x} применяется процедура рандомизации.

В заключительном разделе главы предлагается *система допусков*, обеспечивающая корректную работу алгоритма при возникновении погрешностей вычислений, связанных с неточным представлением вещественных чисел в ячейках памяти ЭВМ.

В третьей главе, «Программный комплекс ЛП-ДА», рассматривается архитектура программного комплекса для решения задач линейного программирования с неформализованными ограничениями на базе использования алгоритма ЛП-ДА. Предлагается *модульная структура* программного комплекса, использующая технологию

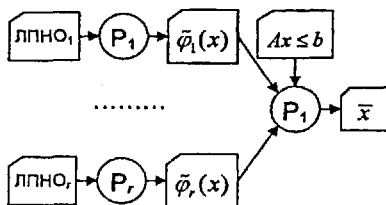


Рис. 1. Схема параллельного алгоритма.

картриджей, которая позволяет расширять программный комплекс путем добавления новых методов линейной оптимизации и дискриминантного анализа. Описывается *оригинальный формат входных данных LPNC* для представления задач ЛПНО, базирующийся на стандарте MPS. Рассматривается *параллельная реализация* комплекса для многопроцессорных систем. Схема параллельного алгоритма решения задачи ЛПНО приведена на Рис. 1. На первом шаге задача ЛПНО, имеющая r неформализованных ограничений, разбивается на r задач ЛПНО₁, ..., ЛПНО_r. Каждая из этих задач независимо решается методом ЛП-ДА. Из вычисленных приближений неформализованных ограничений и формализованной части задачи ЛПНО строится полная система ограничений. На втором шаге полученная таким образом задача решается одним из методов линейного программирования. В результате с некоторым приближением мы получаем \bar{x} – решение исходной задачи ЛПНО. При реализации параллельной версии алгоритма ЛП-ДА нахождение каждого неформализованного ограничения оформлялось в виде независимого процесса P_i ($i = 1, \dots, r$). Вычисленные ограничения передавались некоторому выделенному процессору-координатору P_1 , который формировал полную задачу линейного программирования и решал ее каким-либо методом линейной оптимизации. При реализации параллельной версии алгоритма ЛП-ДА мы использовали передачу сообщений на основе односторонних коммуникаций MPI-2. Проектирование программы выполнялось в соответствии в моделью SIMD.

В рамках диссертационного исследования была выполнена реализация прототипа программного комплекса ЛП-ДА на языке Си. В качестве блока оптимизации был реализован симплекс-метод, в качестве блока дискриминантного анализа был реализован метод линейной коррекции. Работа эксперта моделировалась явным заданием неформализованных ограничений. Общий объем кода на языке Си составил около 3000 строк.

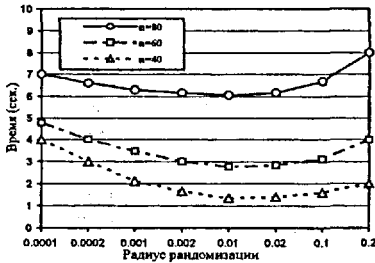


Рис. 2. Зависимость времени решения задачи от радиуса рандомизации ρ .

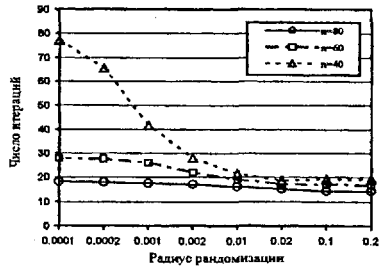


Рис. 3. Зависимость числа итераций от радиуса рандомизации ρ .

В четвертой главе, «Компьютерный анализ алгоритма ЛП-ДА», приводится описание реальных и искусственных задач, использованных для проверки эффективности предложенных подходов, методов и алгоритмов. Описываются и обсуждаются результаты экспериментов, проведенных на персональном компьютере и на высокопроизводительном вычислительном кластере.

Для проведения экспериментов была сконструирована масштабируемая модельная задача *Mod-n*, содержащая n переменных $2(n-1)$ формализованных ограничений и одно неформализованное. В первой серии экспериментов была исследована зависимость количества итераций и времени решения задачи *Mod-n* от радиуса рандомизации ρ . Эксперименты проводились для значений n , равных 40, 60 и 80. Результаты представлены на Рис. 2 и Рис. 3. Графики на Рис. 2 показывают, что время решения задачи слабо зависит от радиуса рандомизации, достигая минимума приблизительно при значении $\rho = 0.01$. Графики на Рис. 3 показывают, что количество итераций для задач небольшой размерности ($n = 40$) может значительно уменьшаться с увеличением радиуса рандомизации, принимая горизонтально-асимптотический характер для значений $\rho \geq 0.01$. Для задач большой размерности ($n = 80$) влияние радиуса рандомизации на количество итераций становится незначительным. Проведенные эксперименты показывают, что в большинстве случаев хорошим выбором является значение $\rho = 0.01$.

Во второй серии экспериментов исследовалась зависимость количества итераций и времени решения задачи *Mod-n* от мощности рандомизации m . Эксперименты проводились для значений n , также равных 40, 60 и 80. Графики на Рис. 4 показывают, что время решения задачи также слабо зависит от мощности рандомизации, достигая минимума приблизительно при значении $m = 30$. Графики на Рис. 5 показывают, что количество итераций для всех трех размерностей уменьшается с ростом m , принимая горизонтально-асимптотический характер для значений $m > 20$. При этом для задач большой размерности ($n = 80$) влияние мощности рандомизации на количество итераций также становится незначительным. Проведенные эксперименты показывают, что в большинстве случаев хорошим выбором является значение $m = 30$.

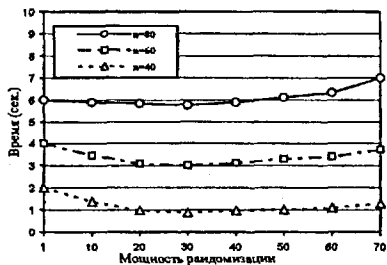


Рис. 4. Зависимость времени решения задачи от мощности рандомизации m .

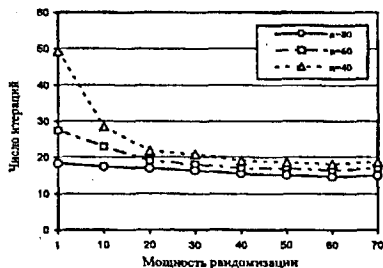


Рис. 5. Зависимость числа итераций от мощности рандомизации m .

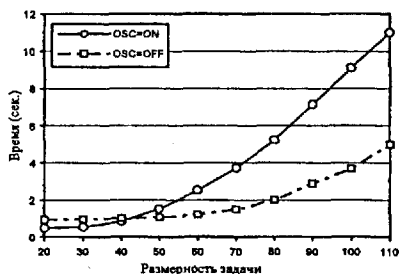


Рис. 6. Зависимость времени решения задачи от размерности n .

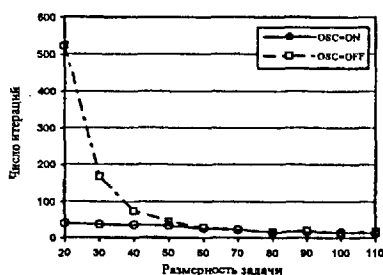


Рис. 7. Зависимость числа итераций от размерности n .

В следующей серии экспериментов исследовалась эффективность применения метода осцилляций при решении задач различной размерности. Были проведены вычислительные эксперименты с задачей $Mod-n$, в ходе которых размерность задачи n варьировалась в диапазоне от 20 до 110. В первом эксперименте метод осцилляций был включен (OSC=ON), а во втором — выключен (OSC=OFF). Результаты этих экспериментов показывают (см. Рис. 6 и Рис. 7), что в обоих вариантах при увеличении размерности время решения задачи растет, а количество итераций уменьшается. Однако, если задача имеет размерность $n < 50$, вариант алгоритма, использующий метод осцилляций, позволяет получить решение при значительно меньшем количестве итераций, что дает возможность существенно сократить число обращений к эксперту.

Эффективность алгоритма ЛП-ДА была также проверена на реальной задаче управления металлургическим предприятием¹. Эта задача моделирует управление прокатным комплексом, включающим рельсобалочный стан, участок отделки фасонных профилей и термомою. Была проведена серия вычислительных экспериментов, в которых исследовались зависимость времени решения задачи и количества итераций от значения ϵ , используемого в

¹ Фролов В.Н. Оптимизация плановых программ при слабо согласованных ограничениях. -М.: Наука, 1986.

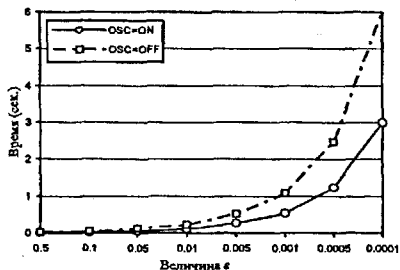


Рис. 8. Зависимость времени решения задачи от величины ϵ .

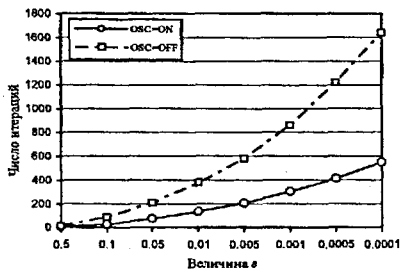


Рис. 9. Зависимость числа итераций от величины ϵ .

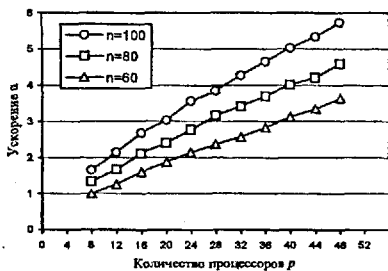


Рис. 10. Ускорение для Mod- n_{48} .

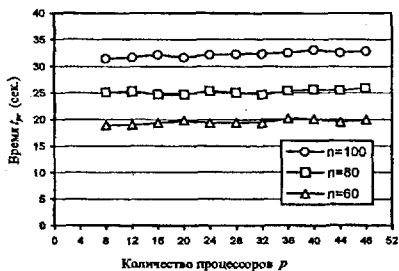


Рис. 11. Расширяемость для Mod- n_p .

критерии завершения. Проведенные эксперименты показали (см. Рис. 8 и Рис. 9), что метод осцилляций во всех случаях обеспечивает примерно двукратное превосходство по скорости сходимости.

В заключительном разделе главы исследуется масштабируемость параллельной версии алгоритма ЛП-ДА для варианта, использующего симплекс-метод и метод линейной коррекции. Для проведения экспериментов был использован вычислительный кластер с 52 процессорами Intel Xeon64 3,2 GHz. Результаты по исследованию ускорения представлены на Рис. 10. Эксперименты проводились для трех размерностей: $n = 60$, $n = 80$ и $n = 100$. Ускорение вычислялось по формуле $\alpha = \alpha_0 + t_p/t_8$, где t_8 – время, затраченное на решение модельной задачи размерности n с 48 неформализованными ограничениями на 8 процессорах, t_p – время, затраченное на решение этой же задачи на p процессорах, $\alpha_0 = (n - 60)/60$. Для каждой размерности варьировалось количество процессоров, используемых для решения задачи. Во всех трех случаях было получено ускорение, близкое к линейному. Результаты по исследованию расширяемости для параллельного алгоритма ЛП-ДА приведены на Рис. 11. Здесь t_{pr} обозначает время, затраченное на решение модельной задачи размерности n с r неформализованными ограничениями на p процессорах. При проведении эксперимента для каждой размерности варьировалось количество процессоров, используемых

для решения задачи. При этом количество неформализованных ограничений всегда равнялось количеству процессоров. Проведенные эксперименты показали, что во всех случаях параллельный алгоритм ЛП-ДА демонстрировал расширяемость, близкую к линейной.

В заключении суммируются основные результаты диссертационной работы, выносимые на защиту, приводятся данные о публикациях и апробациях автора по теме диссертации, и рассматриваются направления дальнейших исследований в данной области.

В приложении приводятся основные сокращения и обозначения, используемые в диссертационной работе.

Основные результаты диссертационной работы

На защиту выносятся следующие новые научные результаты.

1. Предложен и исследован итерационный метод для решения задач линейного программирования с неформализованными ограничениями, базирующийся на синтезе дискриминантного анализа и линейной оптимизации.
2. Предложен метод осцилляций, позволяющий существенно ускорить сходимость последовательности приближений к точному решению.
3. Разработан и исследован алгоритм ЛП-ДА, соединяющий методы линейного программирования и дискриминантного анализа.
4. Спроектирован и реализован прототип программного комплекса для решения задач линейного программирования с неформализованными ограничениями. Общий объем кода на языке Си составил около 3000 строк. Исходные тексты прототипа свободно доступны в Интернет по адресу: <http://life.susu.ru/lpno/>.
5. Спроектирована и реализована параллельная версия алгоритма ЛП-ДА с использованием пакета MPI-2. Исходные тексты параллельной версии свободно доступны в Интернет по адресу: <http://life.susu.ru/lpnoMPI/>.
6. Проведены вычислительные эксперименты на модельных и реальных задачах экономико-математического моделирования, подтверждающие эффективность предложенных алгоритмов, методов и подходов.

Публикации по теме диссертации

Основные результаты диссертации полностью опубликованы автором в следующих работах:

1. *Соколинская И.М.* Синтез симплекс-метода и метода линейной коррекции в задачах линейной оптимизации с неформализованными ограничениями // Вычислительные методы и программирование. –2005. –Том 6, №2. –С. 103-115.

2. *Mazurov V.I.D., Sokolinskaya I.M.* Discrimination analysis and randomization in linear optimization problems with not formalized restrictions // Pattern Recognition and Image Analysis. –2006. –Vol. 16, No. 2. –P. 170-178.
3. *Соколинская И.М.* Метод осцилляций в задачах линейного программирования с неформализованным ограничением // Алгоритмический анализ неустойчивых задач: Тез. докл. Всерос. науч. конф. (2-6 февраля 2004 г., Екатеринбург). –Екатеринбург: Изд.-во Урал. ун-та. -2004. -С. 302-303.
4. *Соколинская И.М., Соколинский Л.Б.* Программный комплекс для решения задач линейного программирования с неформализованными ограничениями // Первая Международная Конференция «Системный Анализ и Информационные Технологии» САИТ-2005 (12-16 сентября 2005 г., Переславль-Залесский, Россия): Труды конференции. В 2 т. Т. 2. –М.: КомКнига. -2005. -С. 286-292.
5. *Соколинская И.М., Соколинский Л.Б.* Параллельный алгоритм решения задачи линейного программирования с неформализованными ограничениями // Проблемы оптимизации и экономические приложения: Тезисы докладов международной конференции. –Омск: Омск. гос. ун-т. -2006. -С. 1.

Работа выполнялась при поддержке *Российского фонда фундаментальных исследований* (проекты 03-01-00565, 06-01-00380).

Подписано в печать 04.07.2006
Формат 60x84 1/16. Бумага офсетная.
Печать офсетная. Усл. печ. л. 1,0. Уч.-изд. л. 1,2.
Тираж 100 экз. Заказ № 76. Бесплатно

Челябинский государственный университет
454021 Челябинск, ул. Братьев Кашириных, 129

Полиграфический участок Издательского центра ЧелГУ
454021 Челябинск, ул. Молодогвардейцев, 576

