Орагнизатор:

Институт Систем Информатики
им. А. П. Ершова СО РАН

Спонсоры:

Microsoft Research

НЦИТ УниПро

Перспективы систем информатики

Пятая международная конференция
памяти академика А. П. Ершова

Перспективы систем информатики

Международное совещание по интервальной математике и методам распространения ограничений

8-9 июля 2003
Новосибирск
Академгородок

PSI'03

Пятая международная конференция
ПЕРСПЕКТИВЫ СИСТЕМ ИНФОРМАТИКИ

8-9 июля 2003 года, Новосибирск, Академгородок, Россия

Рабочее совещание

Интервальная математика и методы распространения ограничений

Доклады и тезисы

Настоящий сборник трудов составлен из кратких аннотаций и полных текстов докладов, представленных на международное рабочее совещание «Интервальная математика и методы распространения ограничений» (ИМРО'03) проходившее в Новосибирском Академгородке 8–9 июля 2003 года под крышей Новосибирского Центра Информационных Технологий «УниПро». Совещание организовано Институтом систем информатики им. А. П. Ершова Сибирского Отделения РАН. Содержание сборника охватывает широкий круг тем, принадлежащих собственно интервальной математике и молодым быстро развивающимся методам распространения ограничений.

Сборник адресован исследователям, инженерам, аспирантам и студентам, которые изучают или применяют интервальные методы и технику распространения ограничений, а таже всем интересующимся этими перспективными и увлекательными областями знаний.

Председатель рабочего совещания

Семенов А. Л.
Институт систем информатики им. А. П. Ершова СО РАН (Новосибирск)

Программный комитет

Кашеварова Т. П.
Институт систем информатики им. А. П. Ершова СО РАН (Новосибирск)

Лакеев А. В.
Институт динамики систем и теории управления СО РАН (Иркутск)

Шарый С. П.
Институт вычислительных технологий СО РАН (Новосибирск)

Andrei Ershov Fifth International Conference
PERSPECTIVES OF SYSTEM INFORMATICS

8–9 July 2003, Novosibirsk, Akademgorodok, Russia

Workshop
"Interval Mathematics and Constraint Propagation Methods"

Proceedings

This volume comprises of both abstracts and full texts of the talks presented at International Workshop “Interval Mathematics and Constraint Propagation Methods'” (IMCP'03) held in Novosibirsk, Akademgorodok, on July 8-9, 2003, at Novosibirsk Center of Information Technologies “UniPro”. The workshop is organized by A. P. Ershov Institute of Informatics Systems of Siberian Branch of Russian Academy of Sciences. The content covers a wide range of topics from interval mathematics proper and young rapidly developing areas of constraint propagation methods: interval algebraic problems, differential equations and dynamic systems with interval uncertainty, applications of interval methods to natural sciences and engineering, constraint propagation / constraint satisfaction techniques in continuous and discrete optimization, constructing cooperative solvers for complex and mixed-structure problems.

The book is addressed to researchers, engineers, postgraduates and students, who study and apply interval methods and constraint propagation techniques, as well as to all those interested in the above promising and fascinating scientific areas.

Workshop chair

Alexander Semenov
A. P. Ershov Institute of Informatics System of SB RAS (Novosibirsk)

Program committee

Tamara Kashevarova
A. P. Ershov Institute of Informatics System of SB RAS (Novosibirsk)

Anatoly Lakeyev
Institute of System Dynamics and Control Theory of SB RAS (Irkutsk)

Sergey Shary
Institute of Computational Technologies of SB RAS (Novosibirsk)

Содержание

Предисловие

Многие задачи в различных областях естествознания и техники представляются как сложные системы алгебраических уравнений, а также как системы с неточными данными. Часто требуется получить гарантированные решения этих задач, то есть решения, заведомо находящиеся около истинного решения, а не искаженные аппроксимациями и ошибками округления приближенные решения. Для получения гарантированных решений используются методы интервальной математики, а также интенсивно развивающиеся в последние годы интервальные методы распространения ограничений.

Целью данного совещания было желание собрать вместе ученых, занимающихся исследованиями в области интервальной математики и распространения ограничений для взаимного обмена мнениями и совместного обсуждения существующих проблем. Мы считаем, что такое общение будет способствовать взаимному обогащению идеями и приведет к более широкому использованию методов интервальной математики в методах распространения ограничений и наоборот. В настоящее время в России существует всего два места (Москва и Новосибирск), где ведутся работы по интервальным методам распространения ограничений. Мы надеемся, что данное совещание послужит толчком к более широкому применению этого подхода.

Следует отметить, что в России такое взаимодействие началось около двадцати лет назад и, вероятно, было первым в мире. С начала 90-х годов работы по интервальным методам распространения ограничений присутствовали на всех российских конференциях по интервальной математике, однако данное совещание является первым российским мероприятиям, специально посвященным этим двум подходам, взаимно дополняющим друг друга и расширяющим возможности получения гарантированных решений сложных практических задач. Мы ожидаем, что конференции с такой тематикой станут регулярными и будут привлекать большее количество участников.

На Совещании представлены доклады, отражающие различные аспекты интервальной математики и методов распространения ограничений:

- решение интервальных линейных систем уравнений;
- решение нелинейных алгебраических систем уравнений;
- исследование дифференциальных уравнений и динамических систем с интервальной неопределенностью;
- приложения методов интервальной математики и распространения ограничений;
- решение оптимизационных задач с использованием методов распространения ограничений;
- вопросы построения кооперативных решателей.

Труды Совещания включают все принятые доклады, в то время как Программа Совещания отражает только доклады, представленные авторами лично. Доклады, которые не были представлены, находятся в конце сборника.

А.Л. Семенов
С.П. Шарый

Программа совещания

Вторник, 8 июля

9:20–9:50 Регистрация участников совещания в холле здания НЦИТ «УниПро».
9:50–10:00 Открытие совещания.

10:00–11:30 Сессия *обзорных докладов*. Председатель – *Кинит Н.В.*

10:00–10:30 *Шарый С.П.* Решение интервальных линейных систем со связями
10:30–11:00 *Соколова С. П.* Иммунокомпьютинг для сложных интервальных систем
11:00–11:30 *Семенов А.Л.* Методы распространения ограничений: основные концепции

11:30–11:45 Перерыв на чай и кофе.

11:45–13:15 Сессия “*Приложения*”. Председатель – *Перцев Н.В.*

11:45–12:05 *Бозоров М. Б.* Интервальный алгоритм определения статического давления грунта на жесткие круглые трубы в высокой насыпи в несколько ниток
12:05–12:35 *Кинит Н.В., Петрунько Н.Н.* Возможности описания схемы соединения интервальной электрической цепи
12:35–12:55 *Колдаков В.В., Панов Н. В.* Программный комплекс для графического представления процесса и результатов работы интервальных алгоритмов
12:55–13:15 *Проскурин А.В., Сагалаков А.М.* О спектральной задаче Орра-Зоммерфельда

13:15–14:30 Обед в столовой ВЦ

14:30 – 16:30 Сессия “*Интервальные алгебраические задачи*”. Председатель – *Шарый С.П.*

14:30–14:50 *Джаныбеков Б.С., Шарый С.П.* Оптимальное внешнее оценивание множеств решений интервальных линейных систем
14:50–15:20 *Назин С.А., Поляк Б.Т.* Интервальная техника в задаче параметрического оценивания
15:20–15:40 *Носков С.И.* Точечная аппроксимация множества решений интервальной системы линейных алгебраических уравнений
15:40–16:10 *Пушков С.Г., Кривошапко С.Ю.* О проблеме реализации в пространстве состояний для интервальных динамических систем
16:10–16:30 *Стрельникова Е.А.* Построение интервальных оболочек множества решений ИСЛАУ

16:30–16:45 Перерыв на чай и кофе.

16:45 – 18:35 Сессия “*Распространение ограничений*”. Председатель – *Семенов А.Л.*

16:45–17:15 *Бревнов Е.В.* Интервальный подход к решению оптимизационной задачи удовлетворения ограничений
17:15–17:45 *Кашеварова Т.П.* О решениях систем нелинейных уравнений
17:45–18:15 *Клейменов А.Е.* Построение кооперативных решателей на основе методов распространения ограничений
18:15–18:35 *Бозоров М. Б., Бердиев Б.Х.* К вопросу сравнения по эффективности интервальных алгоритмов решения систем нелинейных уравнений

19:00–22:00 Товарищеский ужин

Среда, 9 июля

9:00 – 10:50. Сессия “*Дифференциальные уравнения и динамические системы*”. Председатель – *Кашеварова Т.П.*

9:00–9:30 *Рогалев А.Н.* Задачи практической (интервальной) устойчивости с заданной областью предельных отклонений
9:30–10:00 *Ивлев Р.С.* Экспоненциальная устойчивость одного класса нелинейных интервальных динамических систем
10:00–10:30 *Перцев Н.В.* Построение областей притяжения устойчивых решений дифференциальных уравнений с помощью монотонного метода и М—матриц
10:30–10:50 *Утюбаев Г.Ш.* Об интервальных методах для задач с неопределенностями из выпуклых множеств

10:50–11:05 Перерыв на чай и кофе.

11:05 – 12:35. Сессия “*Задачи оптимизации*”. Председатель – *Соколова С.П.*.

11:05–11:35 *Ахмеров Р.Р.* Интервально-аффинная арифметика с управляемой точностью
11:35–12:05 *Ершов А.Г.* Гарантированно субоптимальные решения задач линейной оптимизации
12:05–12:35 *Петров Е. С.* Символьно-интервальная эвристика для минимизации при краевых ограничениях

12:40–13:20 Общая дискуссия и закрытие совещания

13:20–14:30 Обед в столовой ВЦ

14:30–16:00 Демонстрация матобеспечения для интервальных вычислений

Program of the Workshop

Tuesday, July 8

9:20–9:50 Registering the participants of the workshop in the hall of NCIT “UniPro”
9:50–10:00 Opening the workshop

10:00–11:30. Session “*Survey presentations*”. Chairman: *Nikolai V. Kinsht*

10:00–10:30 *Sergey P. Shary* Solving interval linear systems with tied data
10:30–11:00 *Svetlana P. Sokolova* Immunocomputing for complex interval systems
11:00–11:30 *Alexander L. Semenov* Constraint propagation methods: basic concepts

11:30–11:45 Tea/coffee break

11:45–13:15. Session “*Applications*”. Chairman: *Nikolai V. Pertsev*

11:45–12:05 *Mamurjon B. Bozorov* An interval algorithm for determining soil static pressure on multilined stiff round pipes in a high embankment
12:05–12:35 *Nikolai V. Kinsht, Nataliya N. Petrunko* On possibilities to describe a connection diagram of an interval electrical circuit
12:35–12:55 *Vladimir V. Koldakov, Nikita V. Panov* A software for graphically representing the process and results of running some interval algorithms
12:55–13:15 *Alexander V. Proskurin, A.M. Sagalakov* On spectral Orr-Sommerfeld problem

13:15–14:30 Lunch

14:30–16:30. Session “Interval algebraic problems”. Chairman: Sergey P. Shary

14:30–14:50 *Bakyt S. Janybekov, Sergey P. Shary* Optimal outer estimation of the solution sets to interval linear systems
14:50–15:20 *Sergey A. Nazin, Boris T. Polyak* Interval technique in the parameter estimation problem
15:20–15:40 *Sergey I. Noskov* A point approximation of the solution set to interval system of linear algebraic equation
15:40–16:10 *Sergey G. Pushkov, Svetlana Yu. Krivoshapko* On the state space realization problem for interval dynamic systems
16:10–16:30 *Elena A. Strelnikova* Constructing interval hulls of the solution sets to interval linear algebraic systems

16:30–16:45 Tea/coffee break

16:45–18:35. Session “*Constraint propagation*”. Chairman: *Alexander L. Semenov*

16:45–17:15 *Evgenii V. Brevnov* Interval approach to the solution of the interval constraint satisfaction problem
17:15–17:45 *Tamara P. Kashevarova* On the solutions to systems of nonlinear equations
17:45–18:15 *Alexander E. Kleimenov* Constructing cooperative solvers on the basis of constraint propagation methods
18:15–18:35 *Mamurjon B. Bozorov, Bakhriddin Kh. Berdiev* Towards the efficiency comparison of interval algorithms solving nonlinear equations systems

19:00–22:00 Friendly party.

Wednesday, July 9

9:00–10:50. Session “*Differential equations and dynamical systems*”. Chairwoman: *Tamara P. Kashevarova*

9:00–9:30 *Alexey N. Rogalev* Problems of practical (interval) stability with a given set of marginary deviations
9:30–10:00 *Ruslan S. Ivlev* Exponential stability of a class of nonlinear interval dynamical systems
10:00–10:30 *Nikolai B. Pertsev* Construction of the attracting sets for stable solutions of differential equations with the use of monotone technique and M-matrices
10:30–10:50 *Gabdulrakhim Sh. Utyubaev* On interval methods for differential problems with data uncertainty in the form of convex sets

10:50–11:05 Tea/coffee break

11:05–12:35. Session “*Optimization problems*”. Chairwoman: *Svetlana S. Sokolova*

11:05–11:35 *Ramil R. Akhmerov* Interval-affine arithmetic with controlled accuracy
11:35–12:05 *Alexei G. Ershov* Guaranteed suboptimal solutions for linear optimization problems
2:05–12:35 *Evgenii S. Petrov* A symbolic-interval heuristic for minimization under bound constraints

12:40–13:20 General discussion and closing the workshop

13:20–14:30 Lunch

14:30–16:00 Demonstrating software for interval computations

Метод дробления параметров для интервальных линейных систем со связями

С.П. Шарый*

Институт вычислительных технологий СО РАН
Россия, 630090 г. Новосибирск, просп. Лаврентьева, 6
shary@ict.nsc.ru

Аннотация Работа посвящена развитию *методов дробления параметров* (PPS-методов) для оптимального (точного) внешнего покоординатного оценивания множеств решений интервальных линейных систем уравнений, на элементы которых наложены дополнительные связи специального вида.

1 Введение

Станем говорить, что задана *интервальная величина* (интервальный параметр), если имеется переменная, изменяющаяся в пределах некоторого интервала. Интервальные величины $x_1 \in \mathbf{x}_1$, $x_2 \in \mathbf{x}_2$, …, $x_n \in \mathbf{x}_n$ назовём *независимыми* (несвязанными), если кортеж из соответствующих переменных $(x_1, x_2, \ldots, x_n)$ принимает любые значения из декартова произведения интервалов их изменения $\mathbf{x}_1$, $\mathbf{x}_2$, …, $\mathbf{x}_n$, т.е. из бруса $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$. В противном случае интервальные величины называются *зависимыми* или *связанными*. Будем говорить также, что на них *наложены связи*, если имеются в виду какие-то соотношения для x_1, x_2, …, x_n в виде равенств, неравенств и т.п.

Конкретный вид связанности (зависимости) интервальных величин удобно представлять наглядно графически на чертеже, изображающем множество всевозможных значений кортежа $(x_1, x_2, \ldots, x_n)$ на фоне декартова произведения $\mathbf{x}_1 \times \mathbf{x}_2 \times \ldots \times \mathbf{x}_n$. Мы будем называть такие чертежи *диаграммами связанности* (диаграммами зависимости).

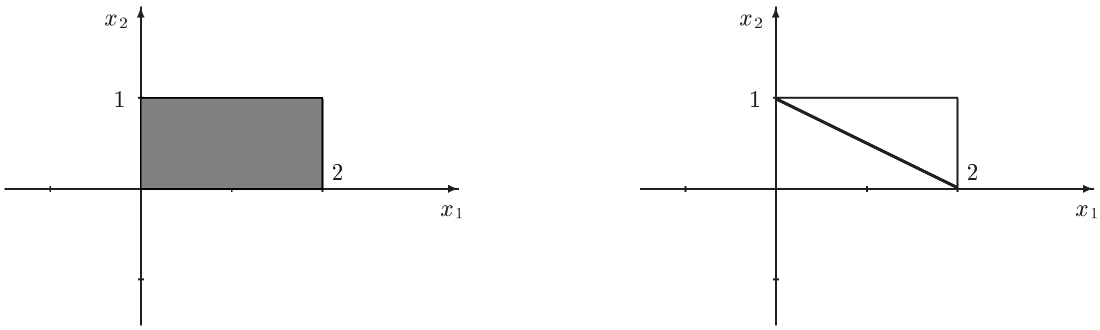


Рис. 1. Диаграммы связанности для независимых (слева) и связанных (справа) интервальных величин.

К примеру, пусть переменные x_1 и x_2 изменяются так, что имеет место равенство

$$x_1 + 2x_2 = 2. \tag{1}$$

При варьировании x_1 в пределах $[0, 2]$ переменная x_2 также изменяется в пределах $[0, 1]$, а множество их совместных значений — пар (x_1, x_2) — "заметает" диагональ в прямоугольнике $([0, 2], [0, 1])^\top$. Диаграмма

* Работа поддержана грантом Президента России №НШ–2314.2003.1.

связанности интервальных величин $x_1 \in [0, 2]$ и $x_2 \in [0, 1]$ выглядит в этом случае так, как изображено на правом чертеже Рис. 1, а для тех же самых, но независимых $x_1 \in [0, 2]$ и $x_2 \in [0, 1]$ диаграмма связанности представлялась бы полным прямоугольником на левом чертеже Рис. 1.

Связанность переменных является весьма распространённым явлением в окружающем нас мире, но в традиционном интервальном анализе задачи со связанными переменными практически не рассматривались. Это было вызвано, прежде всего, тем обстоятельством, что подобные задачи гораздо более сложны, чем обычные постановки с независимыми входными данными, и, кроме того, классическая интервальная арифметика и другие интервальные арифметики, которые являются основным инструментом интервального анализа, нацелены на обработку именно независимых величин.

Предмет представляемой работы — задачи оценивания множеств решений интервальных систем линейных алгебраических уравнений (ИСЛАУ), имеющих вид

$$\left\{\begin{array}{ccccccccc} \mathbf{a}_{11}x_1 & + & \mathbf{a}_{12}x_2 & + & \ldots & + & \mathbf{a}_{1n}x_n & = & \mathbf{b}_1, \\ \mathbf{a}_{21}x_1 & + & \mathbf{a}_{22}x_2 & + & \ldots & + & \mathbf{a}_{2n}x_n & = & \mathbf{b}_2, \\ & \vdots & & & \ddots & & & \vdots & \\ \mathbf{a}_{n1}x_1 & + & \mathbf{a}_{n2}x_2 & + & \ldots & + & \mathbf{a}_{nn}x_n & = & \mathbf{b}_n, \end{array}\right. \tag{2}$$

с интервалами $\mathbf{a}_{ij}$ и $\mathbf{b}_i$, или, в краткой форме,

$$\mathbf{A}x = \mathbf{b} \tag{3}$$

с интервальной $n \times n$-матрицей $\mathbf{A} = (\,\mathbf{a}_{ij}\,)$ и интервальным n-вектором правой части $\mathbf{b} = (\,\mathbf{b}_i\,)$, причём на элементы матрицы будут наложены связи некоторого частного вида. Мы подробно рассмотрим случай интервальных симметричных линейных систем, а на некоторые другие типы ИСЛАУ со связями — имеющими кососимметричные, персимметричные, ганкелевы, тёплицевы или циркулянтные матрицы — наши рассуждения переносятся с минимальными модификациями.

Интервальной симметричной линейной системой будем называть интервальную систему линейных алгебраических уравнений (2)–(3) с матрицей $\mathbf{A}$, симметричной относительно главной диагонали, и такой, что в её пределах также рассматриваются только вещественные матрицы $A \in \mathbf{A}$, обладающие свойством симметричности, $A = A^\top$. Для интервальных систем уравнений существуют различные определения решений и множеств решений, для которых, в свою очередь, возможны те или иные способы оценивания, но ниже мы ограничимся так называемым *объединённым множеством решений* для (2)–(3), которое образовано всевозможными решениями x точечных систем $Ax = b$, когда матрица A и вектор b пробегают $\mathbf{A}$ и $\mathbf{b}$ соответственно, подчиняясь наложенным на них связям. Объединённое множество решений интервальной симметричной линейной системы определяется строго как

$$\Xi_{sym}(\mathbf{A}, \mathbf{b}) = \bigl\{\, x \in \mathbb{R}^n \mid (\exists\, A \in \mathbf{A})(\exists\, b \in \mathbf{b})(\, A = A^\top \text{ и } Ax = b\,)\,\bigr\}, \tag{4}$$

и мы будем называть его просто *множеством решений* интервальной линейной системы (2)–(3), так как другие множества решений нами не исследуются.

В этой заметке нас интересует нахождение для множества решений возможно более узких *внешних* оценок в виде объемлющих множеств, в качестве которых берутся *брусы* — прямоугольные параллелепипеды в пространстве $\mathbb{R}^n$ со сторонами, параллельными координатным осям, т.е. геометрические образы интервальных векторов. Таким образом, мы решаем задачу:

$$\boxed{\begin{array}{c}\text{Найти (по возможности меньший) брус } \mathbf{U}, \\ \text{содержащий множество решений } \Xi_{sym}(\mathbf{A}, \mathbf{b}) \\ \text{интервальной линейной системы } \mathbf{A}x = \mathbf{b}.\end{array}} \tag{5}$$

В свою очередь, задача (5) может быть редуцирована к следующей покомпонентной форме:

$$\boxed{\begin{array}{c}\text{Для интервальной линейной системы уравнений } \mathbf{A}x = \mathbf{b} \text{ найти} \\ \text{как можно более точные оценки для } \min\{\, x_\nu \mid x \in \Xi_{sym}(\mathbf{A}, \mathbf{b})\,\} \\ \text{снизу и для } \max\{\, x_\nu \mid x \in \Xi_{sym}(\mathbf{A}, \mathbf{b})\,\} \text{ сверху, } \nu = 1, 2, \ldots, n.\end{array}} \tag{6}$$

При этом в постановке (6) можно даже ограничиться требованием вычисления только минимума $\min\{\, x_\nu \mid x \in \varXi_{sym}(\mathbf{A},\mathbf{b})\,\}$, поскольку

$$\max\{\, x_\nu \mid x \in \varXi_{sym}(\mathbf{A},\mathbf{b})\,\} = -\min\{\, x_\nu \mid x \in \varXi_{sym}(\mathbf{A},-\mathbf{b})\,\}.$$

Ниже мы предполагаем невырожденность интервальной матрицы $\mathbf{A}$. Наконец, используемые нами обозначения соответствует недавно принятому проекту международного стандарта [19].

2 Методы дробления параметров для интервальных линейных систем

Изучение интервальных линейных систем со связями и первые попытки решения для них задачи внешнего оценивания (5) восходят к началу 90-х годов прошлого века [14,16]. Предложенный Алефельдом и Майером в [14] интервальный метод Холесского является естественным интервальным расширением вещественного метода Холесского (известного также как "метод квадратного корня"), и численные эксперименты демонстрируют его невысокое качество. Далее в большом цикле работ [7,8,9,10,11,12,13] и других Алефельд, Крейнович и Майер обратились к теоретическому исследованию ИСЛАУ с симметричными, кососимметричными, тёплицевыми и даже более общими видами связей, дав характеризацию соответствующих множеств решений. Они установили, в частности, что границы множеств решений интервальных симметричных и кососимметричных линейных систем в общем случае составлены из кусков гиперплоскостей и поверхностей второго порядка (в отличие от кусочно-плоских границ множеств решений ИСЛАУ с независимыми данными). В более общем случае аффинных связей на элементы матрицы и правой части множество решений ИСЛАУ является, как было показано, в [10,11], полуалгебраическим множеством.

Для точной характеризации множеств решений ИСЛАУ со связями специального вида Алефельд, Крейнович и Майер разработали метод исключения, родственный процессу исключения Фурье-Моцкина для систем линейных неравенств. В силу огромной трудоёмкости этого метода он имеет, главным образом, теоретическое значение: для 3×3-системы, например, результирующая система неравенств, описывающая пересечение $\varXi_{sym}$ с одним ортантом пространства $\mathbb{R}^n$ имеет аж 44 неравенства [12]. Таким образом, задача практического вычисления оценок множеств решений ИСЛАУ с симметричными, кососимметричными и т.п. матрицами со связанными элементами стоит весьма остро.

Для уточнения внешних оценок множеств решений интервальных систем уравнений автором в работах [4,6,22] были предложены и развиты так называемые *методы дробления параметров*. Они основаны на часто используемой в интервальном анализе идее адаптивного (т.е. подстраивающегося под задачу и текущий результат) дробления интервальных входных данных. Для интервальных линейных систем с независимыми величинами на входе вычислительная схема методов дробления параметров выглядит особенно просто, так как экстремальные оценки множеств решений достигаются на крайних матрицах и правых частях (теорема Никеля). Дробление интервальных параметров системы сводится в этом случае, фактически, к распадению интервала на его концы.

Но для систем со связанными параметрами этот факт уже неверен. Например, для интервальной симметричной линейной системы

$$\begin{pmatrix} 1 & [0,1] \\ [0,1] & [-4,-1] \end{pmatrix} x = \begin{pmatrix} [0,2] \\ [0,2] \end{pmatrix}, \tag{7}$$

Алефельд, Крейнович и Майер показали в [12], что её множества решений имеют вид, изображенный на Рис. 2, и, кроме того,

$$\max\{\, x_1 \mid x \in \varXi_{sym}(\mathbf{A},\mathbf{b})\,\} = 1 + \sqrt{2},$$

причём этот максимум достигается на вещественной системе

$$\begin{pmatrix} 1 & \sqrt{2}-1 \\ \sqrt{2}-1 & -1 \end{pmatrix} x = \begin{pmatrix} 2 \\ 2 \end{pmatrix},$$

матрица которой не является крайней (при этом $x_2 = -1$). Таким образом, алгоритмы метода дробления параметров, описанные в [4,6,22], неприменимы напрямую для уточнения оценок множеств решений интервальных линейных систем со связями. Тем не менее, путем несложной модификации исходной схемы мы можем исправить это положение, получив в результате методы для вычисления точных (оптимальных) оценок множеств решений ИСЛАУ со связями.

Основная идеи этой модификации просты и естественны:

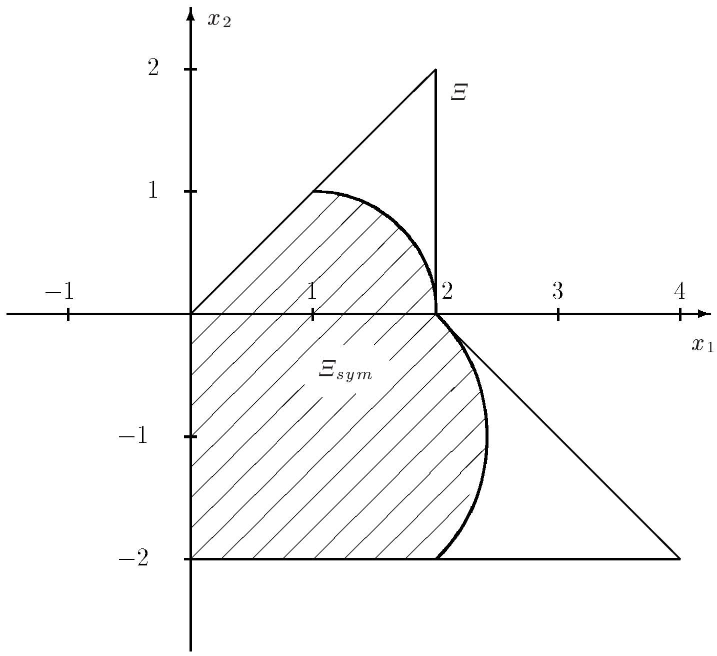


Рис. 2. Множество решений интервальной симметричной линейной системы уравнений (7) в сравнении с множеством решений такой же системы без связей.

– мы отказываемся от поконцевого дробления элементов интервальной системы и дробим, как это обычно делается, интервальные параметры системы на подинтервалы ненулевой длины, в объединении дающие исходный дробимый интервал,
– мы дробим интервальную систему так, чтобы получающиеся системы-потомки соответствовали связям, накладываемым на систему.

Например, если рассматривается интервальная симметричная система уравнений, то в единичном акте дробления должны одновременно дробиться симметричные относительно главной диагонали элементы матрицы, чтобы получающиеся ИСЛАУ снова имели интервальные симметричные системы. Если матрица исходной системы была интервальной кососимметричной (персиммметричной, ганкелевой, тёплицевой, циркулянтной и т.п.), то и системы-потомки на каждом шаге алгоритма должны порождаться так, чтобы иметь кососимметричную (персиммметричную, ганкелеву, тёплицеву, циркулянтную и т.п.) матрицу.

Приступая к реализации наших задумок, обозначим, аналогично тому, как это сделано в [4,6,22],

Encl — какой-нибудь фиксированный метод внешнего оценивания множеств решений ИСЛАУ (мы будем называть его *базовым*);

Encl $(\mathbf{Q}, \mathbf{r})$ — получаемый с его помощью интервальный вектор внешней оценки для множества решений $\Xi_{sym}(\mathbf{Q}, \mathbf{r})$ системы $\mathbf{Q}x = \mathbf{r}$, т. е. *Encl* $(\mathbf{Q}, \mathbf{r}) \in \mathbb{IR}^n$ и

$$Encl\,(\mathbf{Q}, \mathbf{r}) \supseteq \Xi_{sym}(\mathbf{Q}, \mathbf{r});$$

$\Upsilon(\mathbf{Q}, \mathbf{r})$ — нижний конец ν-ой компоненты (для заданного фиксированного номера $\nu \in \{1, 2, \ldots, n\}$) внешней интервальной оценки множества решений $\Xi_{sym}(\mathbf{Q}, \mathbf{r})$, получаемой методом *Encl*, т. е.

$$\Upsilon(\mathbf{Q}, \mathbf{r}) := \underline{\big(Encl\,(\mathbf{Q}, \mathbf{r})\big)_\nu}. \tag{8}$$

Подчеркнем, что, коль скоро $\Xi_{sym}(\mathbf{A}, \mathbf{b}) \subseteq \Xi(\mathbf{A}, \mathbf{b})$, то в качестве базовых методов можно брать традиционные и хорошо разработанные методы внешнего оценивания множеств решений ИСЛАУ без связей, описанные, к примеру, в [1,2,5,15,17,18,20,21]. Получается, что алгоритм оценивания множеств решений ИСЛАУ со связями конструируется из более простых и грубых методов решения ИСЛАУ с независимыми данными.

Мы потребуем лишь от базового метода *Encl* удовлетворения следующим условиям:

$$\boxed{\begin{array}{c}\text{оценка } \Upsilon(\mathbf{Q},\mathbf{r}) \text{ монотонна по включению}\\ \text{относительно матрицы } \mathbf{Q} \text{ и вектора } \mathbf{r}\,,\\ \text{т. е. для всех } \mathbf{Q}',\, \mathbf{Q}'' \in \mathbb{IR}^{n\times n} \text{ и } \mathbf{r}',\, \mathbf{r}'' \in \mathbb{IR}^{n}\\ \text{при } \mathbf{Q}' \subseteq \mathbf{Q}'' \text{ и } \mathbf{r}' \subseteq \mathbf{r}'' \text{ верно неравенство}\\ \Upsilon(\mathbf{Q}',\mathbf{r}') \geq \Upsilon(\mathbf{Q}'',\mathbf{r}'')\end{array}} \tag{C1}$$

и

$$\boxed{\begin{array}{c}\text{оценка } \Upsilon(\mathbf{Q},\mathbf{r}) \text{ является точной на вещественных}\\ \text{линейных алгебраических системах, т. е.}\\ \Upsilon(Q,r) = (\,Q^{-1}r\,)_\nu \text{ для всех } Q\in\mathbb{R}^{n\times n} \text{ и } r\in\mathbb{R}^n.\end{array}} \tag{C2}$$

Если базовый метод *Encl* является естественным интервальным расширением какого-нибудь вещественного метода (как, например, интервальный метод Гаусса), или, более общо, дерево Канторовича базового метода *Encl* своими узлами имеет только интервальные арифметические операции, то свойство (C1) очевидным образом выполняется в силу свойства монотонности интервальных арифметик по включению. Иначе, если в алгоритме базового метода встречаются неинтервальные операции, то свойство (C1) может и нарушаться. Проверку того, что данный конкретный базовый метод обладает свойством (C1), мы возлагаем на разработчиков программ.

Имеем

$$\min\{\, x_\nu \mid x\in \Xi_{sym}(\mathbf{A},\mathbf{b})\,\} = \bigl(\,\tilde{A}^{-1}\tilde{b}\,\bigr)_\nu$$

для некоторых точечной симметричной матрицы $\tilde{A} = (\,\tilde{a}_{ij}) \in \mathbb{R}^{n\times n}$ и точечного вектора $\tilde{b} = (\,\tilde{b}_i) \in \mathbb{R}^n$, составленных из представителей элементов матрицы $\mathbf{A}$ и вектора $\mathbf{b}$, причём по самому определению оценки Υ

$$\Upsilon(\tilde{A},\tilde{b}) \leq \bigl(\,\tilde{A}^{-1}\tilde{b}\,\bigr)_\nu.$$

Предположив, что в матрице $\mathbf{A}$ симметричные относительно главной диагонали элементы $\mathbf{a}_{ij}$ и $\mathbf{a}_{ij}$ имеют ненулевую ширину, обозначим

$\mathbf{A}'$ — матрицу, полученную из $\mathbf{A}$ заменой элементов $\mathbf{a}_{ij}$ и $\mathbf{a}_{ji}$ на $[\,\underline{\mathbf{a}}_{ij}, \operatorname{mid}\mathbf{a}_{ij}]$,
$\mathbf{A}''$ — матрицу, полученную из $\mathbf{A}$ заменой элементов $\mathbf{a}_{ij}$ и $\mathbf{a}_{ji}$ на $[\operatorname{mid}\mathbf{a}_{ij}, \overline{\mathbf{a}}_{ij}]$,
$\tilde{\mathbf{A}}'$ — матрицу, полученную из $\tilde{A}$ заменой элементов $\tilde{a}_{ij}$ и $\tilde{a}_{ji}$ на $[\,\underline{\mathbf{a}}_{ij}, \operatorname{mid}\mathbf{a}_{ij}]$,
$\tilde{\mathbf{A}}''$ — матрицу, полученную из $\tilde{A}$ заменой элементов $\tilde{a}_{ij}$ и $\tilde{a}_{ji}$ на $[\operatorname{mid}\mathbf{a}_{ij}, \overline{\mathbf{a}}_{ij}]$.

Интервальные системы линейных алгебраических уравнений $\mathbf{A}'x = \mathbf{b}$ и $\mathbf{A}''x = \mathbf{b}$, полученные из исходной системы путём рассечения пары симметричных относительно главной диагонали интервальных элементов пополам, мы будем называть *системами-потомками* для $\mathbf{A}x = \mathbf{b}$.

Далее, так как

$$\tilde{\mathbf{A}}' \subseteq \mathbf{A}' \subseteq \mathbf{A}, \qquad\qquad \tilde{\mathbf{A}}'' \subseteq \mathbf{A}'' \subseteq \mathbf{A},$$

и $\tilde{b} \subseteq \mathbf{b}$, то условие (C1) имеет своим следствием неравенства

$$\Upsilon(\mathbf{A},\mathbf{b}) \leq \Upsilon(\mathbf{A}',\mathbf{b}) \leq \Upsilon(\tilde{\mathbf{A}}',\tilde{b})$$

и

$$\Upsilon(\mathbf{A},\mathbf{b}) \leq \Upsilon(\mathbf{A}'',\mathbf{b}) \leq \Upsilon(\tilde{\mathbf{A}}'',\tilde{b}).$$

Следовательно, беря почленный минимум от соответствующих частей неравенств, мы получим

$$\Upsilon(\mathbf{A},\mathbf{b}) \leq \min\bigl\{\,\Upsilon(\mathbf{A}',\mathbf{b}),\, \Upsilon(\mathbf{A}'',\mathbf{b})\,\bigr\} \leq \min\bigl\{\,\Upsilon(\tilde{\mathbf{A}}',\tilde{b}),\, \Upsilon(\tilde{\mathbf{A}}'',\tilde{b})\,\bigr\}. \tag{9}$$

Кроме того, поскольку матрица $\tilde{A}$ обязательно содержится либо в $\tilde{\mathbf{A}}'$, либо в $\tilde{\mathbf{A}}''$, то справедливо, по крайней мере, одно из неравенств

$$\varUpsilon(\tilde{\mathbf{A}}', \tilde{b}) \leq \varUpsilon(\tilde{A}, \tilde{b}) \qquad \text{или} \qquad \varUpsilon(\tilde{\mathbf{A}}'', \tilde{b}) \leq \varUpsilon(\tilde{A}, \tilde{b}),$$

так что

$$\min\{\, \varUpsilon(\tilde{\mathbf{A}}', \tilde{b}), \varUpsilon(\tilde{\mathbf{A}}'', \tilde{b}) \,\} \leq \varUpsilon(\tilde{A}, \tilde{b}) \leq \big(\, \tilde{A}^{-1}\tilde{b} \,\big)_\nu = \min\{\, x_\nu \mid x \in \varXi_{sym}(\mathbf{A}, \mathbf{b}) \,\}. \tag{10}$$

Сопоставление неравенств (9) и (10) приводит к соотношению

$$\varUpsilon(\mathbf{A}, \mathbf{b}) \leq \min\big\{\, \varUpsilon(\mathbf{A}', \mathbf{b}), \varUpsilon(\mathbf{A}'', \mathbf{b}) \,\big\} \leq \min\{\, x_\nu \mid x \in \varXi_{sym}(\mathbf{A}, \mathbf{b}) \,\},$$

и, как следствие, к следующему практическому рецепту:

решив две интервальных "системы-потомка" $\mathbf{A}'x = \mathbf{b}$ *и* $\mathbf{A}''x = \mathbf{b}$, *мы можем прийти, вообще говоря, к более точной оценке снизу для искомого значения* $\min\{\, x_\nu \mid x \in \varXi_{sym}(\mathbf{A}, \mathbf{b}) \,\}$ *в виде*

$$\min\big\{\, \varUpsilon(\mathbf{A}', \mathbf{b}), \varUpsilon(\mathbf{A}'', \mathbf{b}) \,\big\}.$$

Такой же эффект имеет и дробление в векторе правых частей $\mathbf{b}$ какого-нибудь интервального элемента $\mathbf{b}_i$ на подинтервалы $[\,\underline{\mathbf{b}}_i, \operatorname{mid} \mathbf{b}_i]$ и $[\,\operatorname{mid} \mathbf{b}_i, \overline{\mathbf{b}}_i]$, что может быть обосновано выкладками, совершенно аналогичными (9)–(10). Поэтому впредь для единообразия договоримся обозначать ИСЛАУ-"потомки", получающиеся из $\mathbf{A}x = \mathbf{b}$ дроблением пополам одного интервального элемента в матрице $\mathbf{A}$ либо в векторе $\mathbf{b}$, через $\mathbf{A}'x = \mathbf{b}'$ и $\mathbf{A}''x = \mathbf{b}''$.

Процедуру улучшения оценки для $\min\{\, x_\nu \mid x \in \varXi_{sym}(\mathbf{A}, \mathbf{b}) \,\}$ посредством дробления элементов интервальной системы (3) можно повторить по отношению к системам-потомкам $\mathbf{A}'x = \mathbf{b}'$ и $\mathbf{A}''x = \mathbf{b}''$, затем снова разбить потомков от $\mathbf{A}'x = \mathbf{b}'$ и $\mathbf{A}''x = \mathbf{b}''$ и снова улучшить оценку и т. д. Мы оформим подобный процесс последовательного улучшения оценки снизу для $\min\{\, x_\nu \mid x \in \varXi_{sym}(\mathbf{A}, \mathbf{b}) \,\}$ так, как это делается в широко известном в комбинаторной оптимизации "методе ветвей и границ" [3] и как это было адаптировано для интервальных методов глобальной оптимизации (см., к примеру, книги [17,18]):

- во-первых, организуем все интервальные системы $\mathbf{Q}x = \mathbf{r}$, которые возникают в процессе дробления исходной ИСЛАУ $\mathbf{A}x = \mathbf{b}$, вместе с их оценками $\varUpsilon(\mathbf{Q}, \mathbf{r})$ в некоторый список $\mathcal{L}$;
- во-вторых, дробить будем лишь ту интервальную систему $\mathbf{Q}x = \mathbf{r}$ из списка $\mathcal{L}$, которая обеспечивает рекордную (наилучшую) на данный момент оценку $\varUpsilon(\mathbf{Q}, \mathbf{r})$ для искомой величины $\min\{\, x_\nu \mid x \in \varXi_{sym}(\mathbf{A}, \mathbf{b}) \,\}$;
- в-третьих, в подвергаемой дроблению ИСЛАУ мы будем дробить лишь самый широкий из интервальных элементов.

Итак, в процессе выполнения алгоритма мы будем поддерживать список $\mathcal{L}$, состоящий из записей-троек вида

$$\big(\, \mathbf{Q}, \mathbf{r}, \varUpsilon(\mathbf{Q}, \mathbf{r}) \,\big), \tag{11}$$

где

$\mathbf{Q}$ — интервальная $n \times n$-матрица, $\mathbf{Q} \subseteq \mathbf{A}$,
$\mathbf{r}$ — интервальный n-вектор, $\mathbf{r} \subseteq \mathbf{b}$.

Кроме того, образующие $\mathcal{L}$ записи будут упорядочены по возрастанию значений оценки $\varUpsilon(\mathbf{Q}, \mathbf{r})$, а первую запись списка, так же как и соответствующие ИСЛАУ $\mathbf{Q}x = \mathbf{r}$ и оценку $\varUpsilon(\mathbf{Q}, \mathbf{r})$ (рекордную в списке), мы будем называть *ведущими* на данном шаге. Полный псевдокод получающегося нового алгоритма, который мы аналогично [6] также назовём *методом дробления параметров*, представлен в Таблице (где через "$\leftarrow$" обозначен оператор присваивания). Он отличается от метода дробления параметров, представленного в [4,22,6] способом порождения интервальных систем-потомков из ведущей интервальной системы и условием остановки.

То, насколько близкими окажутся результат работы алгоритма и искомый $\min\{\, x_\nu \mid x \in \varXi_{sym}(\mathbf{A}, \mathbf{b}) \,\}$, зависит, с одной стороны, от способа получения оценки $\varUpsilon(\mathbf{Q}, \mathbf{r})$, т. е. от выбранного нами базового метода

Простейший метод дробления параметров

для интервальных симметричных линейных систем

Вход

Интервальная симметричная линейная система $\mathbf{A}x = \mathbf{b}$.
Номер оцениваемой компоненты $\nu \in \{1, 2, \ldots, n\}$.
Метод *Encl*, формирующий оценку Υ по правилу (8).
Константа $\delta > 0$.

Выход

Оценка Z снизу для $\min\{\, x_\nu \mid x \in \Xi_{sym}(\mathbf{A}, \mathbf{b})\,\}$.

Алгоритм

присваиваем $\mathbf{Q} \leftarrow \mathbf{A}$ и $\mathbf{r} \leftarrow \mathbf{b}$;

вычисляем оценку $\upsilon \leftarrow \Upsilon(\mathbf{Q}, \mathbf{r})$;

инициализируем рабочий список $\mathcal{L} \leftarrow \{\, (\mathbf{Q}, \mathbf{r}, \upsilon)\,\}$;

`DO WHILE` (максимум ширины элементов в $\mathbf{Q}$ и $\mathbf{r}$ превосходит δ)

в матрице $\mathbf{Q} = (\,\mathbf{q}_{ij}\,)$ и векторе $\mathbf{r} = (\,\mathbf{r}_i\,)$ выбираем интервальный элемент $\mathbf{s}$, имеющий наибольшую ширину;

порождаем интервальные системы-потомки $\mathbf{Q}'x = \mathbf{r}'$ и $\mathbf{Q}''x = \mathbf{r}''$:

если $\mathbf{s} = \mathbf{q}_{kl}$ для некоторых $k, l \in \{\, 1, 2, \ldots, n\,\}$, то полагаем

$\mathbf{q}'_{ij} \leftarrow \mathbf{q}''_{ij} \leftarrow \mathbf{q}_{ij}$ для $(i, j) \neq (k, l)$ или $(i, j) \neq (l, k)$,

$\mathbf{q}'_{lk} \leftarrow \mathbf{q}'_{kl} \leftarrow [\, \underline{\mathbf{q}}_{kl}, \operatorname{mid} \mathbf{q}_{kl}\,]$, $\mathbf{q}''_{lk} \leftarrow \mathbf{q}''_{kl} \leftarrow [\, \operatorname{mid} \mathbf{q}_{kl}, \overline{\mathbf{q}}_{kl}\,]$,

$\mathbf{r}' \leftarrow \mathbf{r}'' \leftarrow \mathbf{r}$;

если $\mathbf{s} = \mathbf{r}_k$ для некоторого $k \in \{\, 1, 2, \ldots, n\,\}$, то полагаем

$\mathbf{Q}' \leftarrow \mathbf{Q}'' \leftarrow \mathbf{Q}$, $\mathbf{r}'_k \leftarrow [\, \underline{\mathbf{r}}_k, \operatorname{mid} \mathbf{r}_k\,]$, $\mathbf{r}''_k \leftarrow [\, \operatorname{mid} \mathbf{r}_k, \overline{\mathbf{r}}_k\,]$,

$\mathbf{r}'_i \leftarrow \mathbf{r}''_i \leftarrow \mathbf{r}_i$ для $i \neq k$;

вычисляем оценки $\upsilon' \leftarrow \Upsilon(\mathbf{Q}', \mathbf{r}')$ и $\upsilon'' \leftarrow \Upsilon(\mathbf{Q}'', \mathbf{r}'')$;

удаляем из $\mathcal{L}$ бывшую ведущей запись $(\mathbf{Q}, \mathbf{r}, \upsilon)$;

заносим записи $(\mathbf{Q}', \mathbf{r}', \upsilon')$ и $(\mathbf{Q}'', \mathbf{r}'', \upsilon')$ в список $\mathcal{L}$, сохраняя его упорядоченность по возрастанию третьего поля;

обозначаем первую запись списка через $(\mathbf{Q}, \mathbf{r}, \upsilon)$;

`END DO`

$Z \leftarrow \upsilon$;

для решения ИСЛАУ-потомков, а с другой — от обусловленности точечных систем, образующих последнюю ведущую систему (её можно оценивать в процессе выполнения алгоритма). В частности, для того, чтобы при $\delta \to 0$ вычисленное алгоритмом значение стремилось бы к $\min\{\, x_\nu \mid x \in \Xi_{sym}(\mathbf{A}, \mathbf{b})\,\}$), необходимо и достаточно выполнения условия (С2). Если же в исходной ИСЛАУ суммарная ширина интервальных элементов оказывается "большой" в сравнении с δ, то, как правило, простейший метод дробления параметров не будет прорабатывать до конца, и потому целесообразней рассматривать его как итеративную уточняющую процедуру.

3 Тест на монотонность

Пусть дана интервальная система линейных алгебраических уравнений $\mathbf{Q}x = \mathbf{r}$, и нам известны

$$\frac{\partial x_\nu(\mathbf{Q}, \mathbf{r})}{\partial q_{ij}} \qquad \text{и} \qquad \frac{\partial x_\nu(\mathbf{Q}, \mathbf{r})}{\partial r_i}$$

— интервальные расширения соответствующих производных

$$\frac{\partial x_\nu(Q, r)}{\partial q_{ij}} \qquad \text{и} \qquad \frac{\partial x_\nu(Q, r)}{\partial r_i}$$

от ν-ой компоненты вектора решения системы уравнений $Qx = r$ по ij-ому элементу матрицы $Q = (\, q_{ij})$ и i-ому элементу вектора $r = (r_i)$, взятых с учётом наложенных на систему связей. Если интервальные $n \times n$-матрица $\tilde{\mathbf{Q}} = (\, \tilde{\mathbf{q}}_{ij})$ и n-вектор $\tilde{\mathbf{r}} = (\, \tilde{\mathbf{r}}_i)$ образованы из элементов

$$\tilde{\mathbf{q}}_{ij} = \begin{cases} [\, \underline{\mathbf{q}}_{ij}, \underline{\mathbf{q}}_{ij}], & \text{при} \quad \dfrac{\partial x_\nu(\mathbf{Q}, \mathbf{r})}{\partial q_{ij}} \geq 0, \\[2ex] [\, \overline{\mathbf{q}}_{ij}, \overline{\mathbf{q}}_{ij}], & \text{при} \quad \dfrac{\partial x_\nu(\mathbf{Q}, \mathbf{r})}{\partial q_{ij}} \leq 0, \\[2ex] \mathbf{q}_{ij}, & \text{при} \ \ \mathrm{int}\, \dfrac{\partial x_\nu(\mathbf{Q}, \mathbf{r})}{\partial q_{ij}} \ni 0, \end{cases} \tag{12}$$

$$\tilde{\mathbf{r}}_i = \begin{cases} [\, \underline{\mathbf{r}}_i, \underline{\mathbf{r}}_i], & \text{при} \quad \dfrac{\partial x_\nu(\mathbf{Q}, \mathbf{r})}{\partial r_i} \geq 0, \\[2ex] [\, \overline{\mathbf{r}}_i, \overline{\mathbf{r}}_i], & \text{при} \quad \dfrac{\partial x_\nu(\mathbf{Q}, \mathbf{r})}{\partial r_i} \leq 0, \\[2ex] \mathbf{r}_i, & \text{при} \ \ \mathrm{int}\, \dfrac{\partial x_\nu(\mathbf{Q}, \mathbf{r})}{\partial r_i} \ni 0, \end{cases} \tag{13}$$

где "int" обозначает внутренность интервала, то, очевидно,

$$\min\{\, x_\nu \mid x \in \Xi_{sym}(\tilde{\mathbf{Q}}, \tilde{\mathbf{r}})\,\} = \min\{\, x_\nu \mid x \in \Xi_{sym}(\mathbf{Q}, \mathbf{r})\,\}.$$

А поскольку количество существенно интервальных (с ненулевой шириной) элементов в $\tilde{\mathbf{Q}}$ и $\tilde{\mathbf{r}}$ может быть, вообще говоря, меньшим, чем в $\mathbf{Q}$ и $\mathbf{r}$, то, переходя от исходной ИСЛАУ $\mathbf{Q}x = \mathbf{r}$ к системе $\tilde{\mathbf{Q}}x = \tilde{\mathbf{r}}$, мы упрощаем задачу вычисления $\min\{\, x_\nu \mid x \in \Xi_{sym}(\mathbf{Q}, \mathbf{r})\,\}$.

Ранее в интервальном анализе в ряде численных методик уже использовались производные решений систем линейных уравнений по элементам матрицы и правой части (см., к примеру, Главу 17 книги [1]). Выведем формулы для этих производных с учётом симметричности матрицы системы уравнений.

Пусть k и l — некоторые фиксированные индексы, такие что $1 \leq k \leq l \leq n$. Распишем систему уравнений $Qx = r$ в виде

$$\sum_{j=1}^{n} q_{ij} x_j = r_i, \qquad i = 1, 2, \ldots, n, \tag{14}$$

и продифференцируем по q_{kl}. Учитывая, что

$$\frac{\partial}{\partial q_{kl}}\left(q_{ij}x_j \right) = \frac{\partial q_{ij}}{\partial q_{kl}} x_j + q_{ij}\frac{\partial x_j}{\partial q_{kl}},$$

где

$$\frac{\partial q_{ij}}{\partial q_{kl}} = \begin{cases} 0, \text{если } (i,j) \neq (k,l), \\ 1, \text{если } (i,j) = (k,l), \end{cases}$$

мы получим из (14)

$$\begin{cases} \displaystyle\sum_{j=1}^{n} q_{ij}\frac{\partial x_j}{\partial q_{kl}} = 0, \;\text{ если } i \neq k \text{ и } i \neq l, \\ \displaystyle\sum_{j=1}^{n} q_{ij}\frac{\partial x_j}{\partial q_{kl}} + x_l = 0, \;\text{ если } i = k, \\ \displaystyle\sum_{j=1}^{n} q_{ij}\frac{\partial x_j}{\partial q_{kl}} + x_k = 0, \;\text{ если } i = l. \end{cases}$$

Таким образом, если

$$\frac{\partial x}{\partial q_{kl}} = \left(\frac{\partial x_1}{\partial q_{kl}}, \ldots, \frac{\partial x_n}{\partial q_{kl}} \right)^{\top},$$

то

$$Q \cdot \frac{\partial x}{\partial q_{kl}} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ -x_l \\ 0 \\ \vdots \\ 0 \\ -x_k \\ 0 \\ \vdots \\ 0 \end{pmatrix} \begin{array}{l} \\ \\ \\ \longleftarrow\; k\text{-е место} \\ \\ \\ \\ \longleftarrow\; l\text{-е место.} \\ \\ \\ \\ \end{array}$$

По этой причине

$$\frac{\partial x}{\partial q_{kl}} = Q^{-1} \cdot \Big(0, \ldots, 0, -x_l, 0, \ldots, 0, -x_k, 0, \ldots, 0 \Big)^{\top},$$

и если $Y = (y_{ij})$ — обратная матрица для Q, то производные решения вещественной симметричной линейной системы $Qx = r$ по элементам матрицы даются формулами

$$\frac{\partial x_\nu}{\partial q_{kl}} = -y_{\nu k}x_l - y_{\nu l}x_k.$$

Дифференцирование уравнений (14) по r_k приводит к более простым соотношениям

$$\begin{cases} \displaystyle\sum_{j=1}^{n} q_{ij}\frac{\partial x_j}{\partial r_k} = 0, \;\text{ если } i \neq k, \\ \displaystyle\sum_{j=1}^{n} q_{ij}\frac{\partial x_j}{\partial r_k} = 1, \;\text{ если } i = k. \end{cases}$$

Таким образом, если

$$\frac{\partial x}{\partial r_k} = \left(\frac{\partial x_1}{\partial r_k}, \ldots, \frac{\partial x_n}{\partial r_k} \right)^{\top},$$

то

$$Q \cdot \frac{\partial x}{\partial r_k} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \longleftarrow \quad k\text{-е место.}$$

По этой причине

$$\frac{\partial x}{\partial r_k} = Q^{-1} \cdot \Big(\; 0, \ldots, 0, 1, 0, \ldots, 0 \; \Big)^{\top},$$

и если $Y = (\, y_{ij})$ — обратная матрица для Q, то производные решения вещественной симметричной линейной системы $Qx = r$ по элементам вектора правой части даются формулами

$$\frac{\partial x_\nu}{\partial r_k} = y_{\nu k}.$$

Следовательно, если $\mathbf{Y} = (\, \mathbf{y}_{ij})$ — "обратная интервальная матрица" для $\mathbf{Q}$, т. е. внешняя интервальная оценка для множества обратных матриц из Q,

$$\mathbf{Y} \supseteq \{\, Q^{-1} \mid Q \in \mathbf{Q} \,\},$$

а $\mathbf{x}_k$ и $\mathbf{x}_l$ — k-ая и l-ая компоненты некоторого интервального вектора $\mathbf{x} \supseteq \Xi_{sym}(\mathbf{Q}, \mathbf{r})$, то мы можем принять следующие интервальные оценки производных

$$\frac{\partial x_\nu(\mathbf{Q}, \mathbf{r})}{\partial q_{kl}} = -\mathbf{y}_{\nu k}\mathbf{x}_l - \mathbf{y}_{\nu l}\mathbf{x}_k, \qquad\qquad \frac{\partial x_\nu(\mathbf{Q}, \mathbf{r})}{\partial r_k} = \mathbf{y}_{\nu k}.$$

Для интервальных линейных систем с кососимметричными матрицами аналогичные интервальные оценки производных, как нетрудно убедиться, имеют вид

$$\frac{\partial x_\nu(\mathbf{Q}, \mathbf{r})}{\partial q_{kl}} = -\mathbf{y}_{\nu k}\mathbf{x}_l + \mathbf{y}_{\nu l}\mathbf{x}_k, \qquad\qquad \frac{\partial x_\nu(\mathbf{Q}, \mathbf{r})}{\partial r_k} = \mathbf{y}_{\nu k}.$$

Вычисление "обратной интервальной матрицы" для $\mathbf{Q}$ можно выполнить как решение задачи внешнего оценивания объединённого множества решений интервального матричного уравнения

$$\mathbf{Q}Y = I, \qquad I \text{ — единичная матрица,}$$

применив, к примеру, n раз (для каждого столбца) тот же самый метод внешнего оценивания *Encl*, который выбран базовым для всего алгоритма. Соответствующие инструкции — пересчёт обратной интервальной матрицы и проверку на монотонность с возможным последующей заменой некоторых интервальных элементов на их концы по формулам (12)–(13) — следует выполнять сразу после порождения интервальных систем потомков в основном цикле алгоритма.

4 Стратегия дробления

В простейшем методе дробления параметров мы рассекали на каждом шаге самый широкий из интервальных элементов ведущей ИСЛАУ. Можно ли путём специального выбора элемента для дробления обеспечить наиболее значительное улучшение целевой функции на каждом шаге алгоритма? Строгая и точная оптимизация методов дробления параметров в таком виде трудна и вряд ли целесообразна в полном объёме. Мы будем решать этот вопрос, руководствуясь следующими эвристическими соображениями: величина произведения ширины (или радиуса) интервального элемента на модуль интервального расширения соответствующей производной решения может служить, в некотором смысле, мерой того, как бисекция элементов из $\mathbf{Q}$ либо $\mathbf{r}$ влияет на $\min\{\, x_\nu \mid x \in \Xi_{sym}(\mathbf{Q}, \mathbf{r}) \,\}$ и размеры объединённого множества решений ИСЛАУ.

Далее, оценки объединённого множества решений ИСЛАУ, получаемые по большинству существующих методов, являются тем более точными, чем меньше размеры этого множества решений. С подобными базовыми методами уменьшение размеров множества решений $\Xi_{sym}(\mathbf{Q},\mathbf{r})$ должно приводить к аналогичному и сравнимому по величине изменению в оценке $\Upsilon(\mathbf{Q},\mathbf{r})$. При этом, следовательно, требование наиболее быстрого улучшения целевой функции за один шаг метода дробления параметров становится, по существу, эквивалентным условию наиболее быстрого уменьшения размеров множества решений при дроблении ведущей ИСЛАУ.

Учитывая сделанные выше заключения, мы рекомендуем дробить ведущие ИСЛАУ по элементам, на которых достигается максимальная из величин

$$\left|\frac{\partial x_\nu(\mathbf{Q},\mathbf{r})}{\partial q_{ij}}\right|\cdot \operatorname{wid}\mathbf{q}_{ij}, \qquad \left|\frac{\partial x_\nu(\mathbf{Q},\mathbf{r})}{\partial r_i}\right|\cdot \operatorname{wid}\mathbf{r}_i, \qquad \operatorname{wid}\ \text{— ширина интервалов,} \tag{15}$$

$i, j \in \{1, 2, \dots, n\}$, т. е. по элементам, на которых достигается максимум произведения оценки производной решения на ширину интервала, а не просто по самым широким элементам ИСЛАУ. Стратегия дробления, требующая максимизации величин (15), была впервые рассмотрена автором в статье [22].

5 Численный пример

В качестве примера применения разработанного нами подхода рассмотрим интервальную симметричную линейную систему

$$\begin{pmatrix} [1, 10] & [0, 1] \\ [0, 1] & [-4, -1] \end{pmatrix} x = \begin{pmatrix} [0, 2] \\ [0, 2] \end{pmatrix}, \tag{16}$$

которая отличается от системы (7) из работы [12] единственным элементом $\mathbf{a}_{11}$ в матрице. Кажется, что Алефельд, Крейнович и Майер намеренно взяли его точечным, так как иначе выполненные ими вручную выкладки (и без того весьма длинные), приводящие к точному описанию формы множества решений системы (7) ещё бы значительно усложнились. Мы не стеснены подобными ограничениями и легко можем вычислить оценки множеств решений этой системы с существенно интервальным элементом $\mathbf{a}_{11}$.

Метод дробления параметров, в котором в качестве базового метода использовался интервальный метод Гаусса-Зейделя (описание которого можно найти, к примеру, в [18,20]), снабженный проверкой монотонности (из §3) и модифицированной стратегией выбора дробимого элемента (из §4), надёжно находит внешнюю оценку множества решений этой ИСЛАУ в виде

$$\begin{pmatrix} [-7.61868\,e{-}8,\ 2.41421] \\ [-2.00000,\ 1.00000] \end{pmatrix}$$

(мы сохранили по шесть значащих цифр результата). При этом брус начальной оценки, вычисленный с помощью процедуры Хансена-Рона [21], оказался равным

$$\begin{pmatrix} [-14.9333,\ 23.7333] \\ [-44.6667,\ 38.0000] \end{pmatrix},$$

т.е. весьма грубым. Для нахождения нижних оценок множества решений системы (16) мы затратили по каждой координате по 1000 итераций, и уточнение этих оценок шло исключительно за счет бисекции элементов ИСЛАУ, хотя (любопытный факт!) диагностика свидетельствует, что тест на монотонность (12)–(13) почти на каждом шаге сужал те или иные интервальные элементы в системах-потомках, появляющихся по ходу работы алгоритма. Но, как ни странно, это не привело к повышению эффективности всего алгоритма в целом. Что касается верхних оценок, то здесь тесты монотонности сыграли свою положительную роль, и количества шагов метода дроблений параметров оказались равными всего 40 по первой компоненте и 33 по второй.

Как видим, оценки множества решений системы (16) не изменились в сравнении с системой (7), т.е. интервализация элемента $\mathbf{a}_{11}$ никак не повлияла на них.

Отметим в заключение, что при проведении вычислительных экспериментов с подобного сорта задачами одной из основных трудностей является в настоящее время отсутствие тестовых примеров.

Список литературы

1. Алефельд Г., Херцбергер Ю. *Введение в интервальные вычисления.* – М.: Мир, 1987.
2. Калмыков С. А., Шокин Ю. И., Юлдашев З. Х. *Методы интервального анализа.* – Новосибирск: Наука, 1986.
3. Пападимитриу Х., Стайглиц К. *Комбинаторная оптимизация. Алгоритмы и сложность.* – М.: Мир, 1985.
4. Шарый С. П. Новый класс алгоритмов для оптимального решения интервальных линейных систем // Конференция "Актуальные проблемы прикладной математики", Саратов, 20 – 22 мая 1991 г. – Саратов, 1991. – С. 113 – 119.
5. Шарый С. П. Алгебраический подход во "внешней задаче" для интервальных линейных систем // *Вычислительные Технологии.* – 1998. – Т. 3, №2. – С. 67–114.
6. Шарый С.П. Оптимальное внешнее оценивание множеств решений интервальных систем уравнений. Часть 1 // *Вычислительные Технологии.* – 2002. – Т. 7, №6. – С. 90 – 113.
   Шарый С.П. Оптимальное внешнее оценивание множеств решений интервальных систем уравнений. Часть 2 // *Вычислительные Технологии.* – 2003. – Т. 8, №1. – С. 84 – 109.
7. Alefeld G., Kreinovich V., Mayer G. Symmetric linear systems with perturbed input data // *Numerical Methods and Error Bounds*, Alefeld G. and Herzberger J., eds. – Berlin: Akademie Verlag, 1996. – P. 16 – 22.
8. Alefeld G., Kreinovich V., Mayer G. The shape of the symmetric solution set // *Applications of Interval Computations*, Kearfott R.B. and Kreinovich V., eds. – Dordrecht: Kluwer, 1996. – P. 61 – 79.
9. Alefeld G., Kreinovich V., Mayer G. On the shape of the symmetric, persymmetric, and skew-symmetric solution set // *SIAM J. Matrix Anal. Appl.* – 1997. – Vol. 18. – P. 693 – 705.
10. Alefeld G., Kreinovich V., Mayer G. The shape of the solution set for systems of interval linear equations with dependent coefficients // *Mathematische Nachrichten.* – 1998. – Bd. 192. – P. 23 – 36.
11. Alefeld G., Kreinovich V., Mayer G. A comment on the shape of the solution set for systems of interval linear equations with dependent coefficients // *Reliable Computing.* – 2001. – Vol. 7, No. 3. – P. 275 – 277.
12. Alefeld G., Kreinovich V., Mayer G. On symmetric solution sets // *Computing Supplement 16*, Herzberger J., ed. – Wien, New York: Springer, 2003. – P. 1 – 23.
13. Alefeld G., Kreinovich V., Mayer G. On the solution sets of particular classes of linear interval systems // Journal of Computational and Applied Mathematics. – 2003. – Vol. 152, No. 1–2. – P. 1 – 15. (электронная версия доступна в Интернете на `http://www.cs.utep.edu/vladik/1996/abstr96.html`).
14. Alefeld G., Mayer G. The Cholesky method for interval data // *Linear Algebra and its Applications.* – 1993. – Vol. 194. – P. 161 – 182.
15. Gay D. M. Solving interval linear equations // *SIAM J. Numer. Analysis.* – 1982. – Vol. 19, №4. – P. 858 – 870.
16. Jansson Ch. Interval linear systems with symmetric matrices, skew-symmetric matrices, and dependencies in the right hand side // *Computing.* – 1991. – Vol. 46. – P. 265 – 274.
17. Hansen E. *Global optimization using interval analysis.* – New York: Marcel Dekker, 1992.
18. Kearfott R. B. *Rigorous global search: Continuous problems.* – Dordrecht: Kluwer, 1996.
19. Kearfott R.B., Nakao M.T., Neumaier A., Rump S.M., Shary S.P., van Hentenryck P. Standardized notation in interval analysis // *Reliable Computing.* – в печати (электронная версия доступна в Интернете на `http://www.mat.univie.ac.at/~neum/software/int`).
20. Neumaier A. *Interval methods for systems of equations.* – Cambridge: Cambridge University Press, 1990.
21. Neumaier A. A simple derivation of Hansen-Bliek-Rohn-Ning-Kearfott enclosure for linear interval equations // *Reliable Computing.* – 1999. – Vol. 5, №2. – P. 131 – 136.
22. Shary S. P. A new class of algorithms for optimal solution of interval linear systems // *Interval Computations.* – 1992. – №2(4). – P. 18 – 29.

Иммунокомпьютинг для сложных интервальных систем

С. П. Соколова, Л.А. Соколова

Санкт-Петербургский институт информатики и автоматизации РАН
Россия, г. Санкт-Петербург
sokolova_sv@mail.ru Lucy@sokolova.fsnet.co.uk

В докладе предложено дальнейшее развитие подхода иммунокомпьютинга на класс сложных систем с параметрической неопределенностью интервального типа. С использованием правил и терминологии интервальной математики разработаны процедуры сингулярного разложения (SVD) интервальных матриц, обучения с экспертом и самообучения, классификации и представления результатов исследования в пространстве образов. Статья включает примеры интервальных иммунных систем для мониторинга проблемы чумы и информационной безопасности.

Иммунокомпьютинг (ИК) как новый подход к вычислениям был предложен в работах [7, 8]. В его основе лежит биологический прототип иммунных сетей и строгие математические понятия формального протеина и формальной иммунной сети. В частности, подход к распознаванию образов на основе иммунокомпьютинга использует понятие сингулярного разложения матриц для определения экстремальных значений энергии связи между формальными протеинами. Этот подход с успехом применялся для информационной безопасности компьютерных сетей, интеллектуальных систем охраны сложных объектов, мониторинга биологических систем (на примере проблемы чумы в Средней Азии) и т.д.

В докладе представлены результаты дальнейшего развития подхода иммунокомпьютинга на класс интервальных систем. Разработаны вычислительные алгоритмы сингулярного разложения произвольных интервальных матриц на основе центрового и адаптивного подходов, введено понятие интервальной энергии связи между формальными протеинами. Разработаны вычислительные процедуры обучения с экспертом и самообучения для сложных объектов с интервальными данными и параметрами, классификации и группировки на плоскости вычисленных значений энергии связи.

Разработаны реальные системы иммунокомпьютинга для интервальной биологической системы (чумной триады: носитель – переносчик – чумный микроб) и системы информационной безопасности в компьютерной сети. На основе введенного понятия индекса риска возникновения чумы [6] разработана вычислительная процедура слежения за динамикой изменения этого индекса. Представлены результаты прогнозирования эпизоотий чумы для ланшафтно-экологических районов Прибалхашского природного очага чумы и результаты идентификации вторжений в компьютерную сеть.

1 Математическая основа

Мы будем использовать следующие обозначения: R^m, $R^{m\times m}$ – соответственно пространство вещественных векторов U с m компонентами и пространство вещественных $m\times m$ матриц A. IR^m — множество интервальных векторов $[U]$

$$[U] = [U_-, U_+] = \{U\in R^m \,/ U_- \le U \le U_+\},\ U_- \le U_+$$

с m компонентами и $IR^{m\times m}$ – множество интервальных матриц

$$[A]=[A_-, A_+] = \{A \in R^{m\times m} \,/ A_- \le A \le A_+\},\ A_- \le A_+,$$

все неравенства понимаются в покомпонентном смысле. Вещественные интервальные компоненты интервальной матрицы $[A]$ являются:$([a_{ij}],\ i,j=1, \quad ,m) =([a_{ij}^-,\ a_{ij}^+]),\ a_{ij}^-,\ a_{ij}^+$ — нижняя и верхняя границы интервалов. Для $U \subseteq R^m$ *интервальная оболочка* $\square(U)$ определяется как

$$\square(U) = \cap \{[\upsilon]\in I\,R^m \,/\, U\subseteq [\upsilon]\}.$$

Пусть задана вещественная интервальная матрица $[A]\in IR^{m\times n}$. Мы понимаем интервальную матрицу $[A]$ как множество вещественных матриц размерности $m \times n$, для которых

$$\{A|\ (\forall\ A \in [A])\}.$$

Согласно [7] любой единичный вектор U с m вещественными компонентами

$$U= [u_1,\ \dots,\ u_m]^T,\ \ UU^T = 1,$$

может быть рассмотрен как особый вид формального пептида (ФП) с $m-1$ звеньями. Энергия связи между любой парой таких формальных пептидов $\{U, V\}$, размерности $(m\times 1)$ и $(n\times 1)$, соответственно, для заданной $[A]\in IR^{m\times n}$ определяется интервальной билинейной формой:

$$[\omega]=-U^T[A]V, \qquad (1)$$

где $[\omega]\in IR^1$, $[\omega] = [\omega_-,\ \omega_+] = \{\omega \in R^1\ /\omega_- \le \omega \le \omega_+\}$, $\omega_- \le \omega_+$. Энергия связи $\omega \in R^1$ для $\forall\ A\ \in [A]$ определяется билинейной формой:

$$\omega = -\,U^TAV, \qquad (2)$$

где $A\in R^{n\times m}$. Как известно, экстремальные значения билинейной формы (2) определяются с помощью сингулярного разложения матриц (SVD) матрицы $A\in R^{n\times m}$:

$$A\ = s_1U_1V_1^T+ s_2U_2V_2^T\ +\ \dots\dots+s_pU_pV_p^T, \qquad (3)$$

где s_i, $i = 1,\dots,\ p$ – сингулярные числа матрицы $\forall\ A\ \in [A]$, U_i, V_i, $i = 1,\dots,\ p$ – левые и правые сингулярные векторы, p – ранг матрицы A. Эти сингулярные числа и векторы удовлетворяют следующим соотношениям:

$$s_1 \ge\ s_2\ge\dots s_p \ge 0,\ \ s_i = U_i^TAV_i,\ V_i^TV_i = 1,\ U_i^TU_i = 1,\ i = 1,\dots,p.$$

Пусть $[A]\ \in IR^{\ n\times m}$. Для данной интервальной матрицы $[A]\ \in IR^{\ n\times m}$, предполагаем, что интервальные параметры $[s_1] \in IR$, $[U]\in IR^m$, $[V]\in IR^n$, обладают следующими свойствами:

- $[s_1]$ содержит максимальное сингулярное число каждой матрицы $\forall A\ \in [A]$;
- для каждого из этих сингулярных чисел интервальные параметры $[U]\in\ IR^m$, $[V]\in\ IR^n$ содержат, по крайней мере, один соответствующий левый и правый сингулярный векторы.

Ниже мы рассмотрим два подхода вычисления сингулярных чисел интервальной матрицы $[A]$: центровой подход и адаптивный подходы.

Центровой подход

Предположим, что интервальная матрица $[A]$ может быть представлена как $[A]\ =[A^C-\Delta,\ A^C+\Delta]$, $A^C\ =\ mid\ ([A])$, где операция взятия средины интервального числа $[a_{ij}]$ представляется в виде $mid[a_{ij}] = a_{ij}^-\ +0.5(a_{ij}^-\ - a_{ij}^+)$ и понимается в покомпонентном смысле. $A^C\ \in R^{\ n\times m}$, $\Delta \ge 0$, $\Delta\ \in R^{\ n\times m}$. Сингулярные числа интервальной матрицы $[A]$ могут быть определены путем вычисления граничных собственных значений симметрической интервальной матрицы $[B]= [A]^T[A]$ [3]. Однако, такой подход существенно завышает реальные интервалы. Последние лучше определять непосредственно путем решения проблемы собственных значений для матриц AA^T, $A\ \in [A]$ [3].

Пусть данная центровая матрица $A^C\ \in R^{\ n\times m}$ находится для интервальной матрицы $[A]$ следующим образом: $[A] =\{A:\ /A-A^C/ \le \Delta A\}$. (Здесь и в дальнейшем, абсолютное значение $/\cdot/$ и знаковое неравенство sign '$\le$' понимается в покомпонентном смысле).

Множество сингулярных чисел интервальной матрицы $[A]$ представляется в виде множества Σ:

$$\Sigma = \{s_i{:}\ A^TAx = s_i^2x,\ x\neq 0,\ A\in [A]\}.$$

В [3] решение для множества Σ было получено в аналитической форме. Пусть матрица S_1^i является диагональной матрицей со знаковыми компонентами элементов x^i, матрица S_2^i представляет знаки элементов матрицы $2A^Cx^i\ +\delta Ax^i$, где $/\delta Ax^i/ < 2/A^Cx^i/$. Тогда квадраты сингулярных чисел s_i^2 матрицы $A^C\ +\ \delta A$, $\forall\ /\delta A/ \le \Delta A$, располагаются в интервале

$$[s_i^2((A^c)^TA^C - 2(S_1^i\Delta A^TS_2^iA^C)_{sym} + S_1^i\Delta A^T\Delta AS_1^i),\ s_i^2((A^c)^TA^C + 2(S_1^i\Delta A^TS_2^iA^C)_{sym} + S_1^i\Delta A^T\Delta AS_1^i)],$$

где B_{sym} представляет симметрическую часть матрицы B.

Адаптивный подход

Случай 1. Пусть $[A] \in IR^{n\times m}$ – интервальная матрица. Тогда сингулярные числа $[A]$ являются собственными значениями следующей матрицы $[B] \in IR^{(n+m)\times(n+m)}$, $[B] = [B_1, B_2]$, где B_1, B_2 – являются столбцами, $B_1 = [0, [A]]^T$, $B_2 = [[A]^T, 0]^T$.

Множество собственных значений интервальной матрицы $[B]$ представлено в виде Σ_1:

$$\Sigma_1 = \{(\lambda, x) \in R^{n+m+1} / Bx = \lambda x,\ x \neq 0,\ B \in [B]\}.$$

Используем следующую функцию [4]:

$$F\{(\Delta X, \Delta\Lambda)^T\} = -R\ \{([B]-\lambda E)x,\ Ix-1\}^T + \{E-RG\}\{(\Delta X, \Delta\Lambda)^T\},$$

где $\Delta X \in IR^{n+m}$, $\Delta\Lambda \in IR^1$; E, I – единичная матрица и вектор,

$$G = \{G_1, G_2\},\ G_1 = \{([B]-\lambda E),\ I^T\},\ G_2 = \{-\Delta X - x,\ 0\}^T.$$

В [4] доказано, что если включение

$$F\{(\Delta X, \Delta\Lambda)^T\} \subseteq (\Delta X, \Delta\Lambda)^T, \qquad (5)$$

выполняется, тогда пара $[\lambda]$, $[x]$ интервальной матрицы $[B]$ удовлетворяет следующим включениям

$$[\lambda] \in \lambda + \Delta\Lambda,\ [x] \in x + \Delta X.$$

Притягивающая точка аппроксимации λ_i, $i = 1, \ldots, (n+m)$ для собственных значений интервальной матрицы $[B] \in IR^{(n+m)\times(n+m)}$ может быть получена с использованием смещенного QR алгоритма. Для этого воспользуемся итеративной процедурой

$$(\Delta X, \Delta\Lambda)^{(k+1)} = F(\Delta X, \Delta\Lambda)^{(k)},\ k = 0,1,2,\ldots,$$
$$\Delta X^{(0)} = x,\ \Delta\Lambda = \lambda.$$

На каждом шаге $k = 0,1,2,\ldots$ этой процедуры с использованием смещенного QR алгоритма проводятся вычисления до тех пор, пока включение *(5)* выполняется.

Случай 2. Пусть $[A] \in IR^{n\times m}$ – интервальная матрица. Множество максимальных сингулярных чисел, правых и левых сингулярных векторов для интервальной матрицы $[A]$ представляется в виде Σ^P:

$$\Sigma^P = \{(s_1, U, V) \in R^{m+n+1},\ s_1 \in R^1,\ U \in R^m,\ V \in R^n:\ AU = sV,\ A^T V = sU,\ V^T V = 1,\ U^T U = 1,\ \forall A \in [A],\ A \text{ со свойством } P\}, \qquad (6)$$

где P является некоторым фиксированным свойством, таким как симметричность, кососимметричность и т. д. К примеру, пусть интервальная матрица $[A] \in IR^{m\times m}$ с $[a_{ij}] = [a_{ji}]$ для $i,j = 1,\ldots,m$. Множество матриц

$$\{A^{sym}\} = \{A \in R^{m\times m} / A \in [A],\ A \text{ symmetric}\}$$

называется симметричной интервальной матрицей. $\{A^{sym}\} \notin IR^{m\times m}$ не является интервальной матрицей в обычном смысле. $\{A^{sym}\} \subseteq [A]$ и $\{A^{sym}\} = [A]$ если и только если $a_{ij}^- = a_{ij}^+$ для $i,j = 1,\ldots, m$, $i \neq j$.

Вышеприведенное множество решений Σ^P для симметрической интервальной матрицы $[A]$ представляется в виде Σ^{sym}

$$\Sigma^{sym} = \{(s_1, U, V) \in R^{m+n+1},\ s_1 \in R^1,\ U \in R^m,\ V \in R^n:\ AU = sV,\ A^TV = sU,\ V^TV = 1,\ U^TU = 1,\ A = A^T \in [A]\}. \quad (7)$$

Сингулярное разложение *(SVD)* интервальной матрицы включает следующие процедуры:

Процедура 1. Определяем множество *(4)* максимальных сингулярных чисел s_1 и соответствующие правые и левые сингулярные векторы для вещественной точечной матрицы $\forall A \in [A]$.

Для определения максимального сингулярного числа s_1, левого U и правого V сингулярных векторов, соответствующих s_1, использована достаточно простая и надежная вычислительная процедура и метод исчерпывания [6]:

$$U^T_{(k+1)} = V^T_{(k)}A,\ V_{(k+1)} = AU_{(k+1)}$$
$$s_k = U_k^T A V_k, \qquad |s_{k+1} - s_k| \le \varepsilon, \qquad (8)$$

где $k=0,1,2,...$ – номер итерации, ε — заданная точность вычисления. Можно показать, что для произвольного единичных векторов $V_{(0)}$, $U_{(0)}$ итерации по схеме *(8)* сходятся в общем случае к сингулярным векторам U, V соответствующим максимальному сингулярному числу $s_1 = V^TAU$.

Вычислим интервальную оболочку для Σ^P *(6)*

$$\Box\Sigma^P = \{[s_1], [U_1], [V_1]\}.$$

затем, используя метод исчерпывания на p шаге вычислительной процедуры, формируем следующую матрицу

$$A_{(p)} = A_{(p-1)} - s_{p-1}U_{p-1}V_{p-1}^T \qquad (9).$$

Ее максимальное сингулярное число s_p и соответствующие левый и правый сингулярные векторы U_p, V_p вычисляются по схеме *(8)*. Вычисляем интервальную оболочку для $\Sigma^P{}_p$

$$\Box\Sigma^P{}_p = \{[s_p], [U_p], [V_p]\}.$$

Процедура 2. Пусть задана интервальная матрица $[A]$. Сформируем следующую матрицу $B \in R^{m+n+2}$ для $\forall A \in [A]$ [3]:

$$B = \{B_1, B_2, B_3, B_4\}, \qquad (10)$$

где B_i, $i = 1, 2, 3, 4$ столбцы соответствующих размерностей

$$B_1 = [A, -sI_n, 2U^T, 0]^T,\ B_2 = [-sI_m, A^T, 0, 2V^T]^T,$$
$$B_3 = [-V, 0, 0, 0]^T, \quad B_4 = [0, -U, 0, 0]^T.$$

Введем векторы $\nu \in R^{m+n+2}$, $r \in R^{m+n+2}$ и $f(\nu) \in R^{m+n+2}$ следующим образом

$$\nu = [\Delta U, \Delta V, \Delta s, \Delta s]^T,\ r = [sV - AU,\ sU - A^TV,\ 1-U^TU,\ 1-V^TV]^T,$$
$$f(\nu) = [\Delta s\Delta U, \Delta s\Delta V, -\Delta U^T\Delta U, -\Delta V^T\Delta V]^T, \qquad (11)$$

где $A \in [A]$, s, U, V –соответственно, максимальное сингулярное число, правый и левый сингулярный вектор, ошибки соответственно равны

$$\Delta s = s - s',\ \Delta U = U - U',\ \Delta V = V - V',\ \Delta s \in R^1.$$

Пусть $\nu \in R^{m+n+2}$, и $[\nu] \in IR^{m+n+2}$. Как показано в [3] матрица B является несингулярной. Предположим, что, L является аппроксимацией инверсии B или точная инверсия B в себя. Рассмотрим интервальную функцию

$$[F] = Lr + (E\text{-}LB)[v] + Lf([v]) \qquad (12)$$

(E – единичная матрица соответствующей размерности). Если включение

$$[F'] \subseteq [v'], \text{ для } [v'] \subseteq IR^{m+n+2} \qquad (13)$$

выполняется, тогда интервальное сингулярное число $[s]$, правый и левый сингулярные векторы $[U]$, $[V]$ интервальной матрицы $[A]$ удовлетворяют включениям

$$[s] \in s + \Delta s', \ [U] \in U + \Delta U', \ [V] \in V + \Delta V'. \qquad (14)$$

Для определения *(14)* воспользуемся итеративной процедурой

$$(\Delta v^{(k+1)}) = F(\Delta v^{(k)}),\ k = 0,1,2,\ldots,\ \Delta v^{(0)} = v, \ (15)$$

до тех пор, пока *(13)* выполняется.

На основе вышеприведенных результатов разработан вычислительный алгоритм сингулярного разложения интервальных матриц. Представленные результаты по сингулярному разложению интервальных матриц были использованы при разработке вычислительных процедур по распознаванию образов для иммунокомпьютинга сложной биологической системы (на примере проблемы чумы) и информационной безопасности.

2 Иммунокомпьютинг для сложной биологической системы

Природные очаги чумы в Республике Казахстан покрывают территорию в 130 миллионов гектаров и за последние 50 лет считаются наиболее активными очагами чумы в мире. В последнее время в связи с активной разработкой полезных ископаемых (нефть и газ) на территории природных очагов чумы резко возросла возможность выноса инфекции за их пределы.

Процесс эпизоотии чумы является сложной многокомпонентной динамической системой, которая характеризуется взаимодействием отдельных подсистем чумной эпизоотической триады: носитель – переносчик – чумный микроб [1, 2].

Состояние носителей характеризуется агрегатированными факторами влияния, множеством характеристик на биоценотическом, популяционном и организменном уровнях. Это множество включает следующие основные факторы: численность носителей в различные сезоны, возростная и половая структура популяции носителя, воспроизводство и показатель смертности и т.д.

Данные, характеризующие состояние членов эпизоотической триады классифицируются по сезонам (весна, лето, осень). Состояние блох – переносчиков чумного микроба – определяется [1] их численностью, сезонной активностью при нападении на животных и т.д., а также характеризуется различными индексами: обилия, доминирования и т.д. Состояние чумного микроба характеризуется следующими свойствами [2]: морфология колоний, глицерин, ферментация, денитрификация/нитрификация, присутствие антигена Фракции 1, вирулентность на белых мышах и морских свинках и т.д.

Погода, географические и пространственные характеристики представляют индексы внешних факторов. К примеру, погода характеризуется следующими показателями: температура воздуха, почвы и количество осадков, различные гидротермальные коэффициенты, направление и скорость ветра и т.д.

Фрагмент базы данных для мониторинга проблемы чумы представлен в Таблице 1.

Таблица 1: Фрагмент базы данных для мониторинга проблемы чумы

Год	**1976**	**1977**	**1978**	**....**	**1987**	**1988**
p1	[2.5, 3.3]	[1.3,1.8]	[5.2,6.1]		[4.0,5.3]	[13.5,15.8]
p2	[0.3, 0.7]	[4.1,5.4]	[29,-32]		[4.4,4.9]	[16.1,18.6]
p3	[1.9, 2.3]	[1.1,1.4]	[3.4,3.9]		[1.7,1.8]	[4.8,5.0]
:	:	:	:	:	:	:
p40	33.9	12.0	34.2		36.2	15.3
p41	1.3	10.0	11.7		0.7	7.7
p42	5.0	6.0	8.7		1.0	25.7
p43	0.3	0.0	0.0		1.3	12.0
p44	0.0	0.0	0.7		1.0	1.0
p45	5.0	4.0	2.7		2.3	1.0

Фрагмент включает 45 параметров p1-p45 таких как: численность носителя на гектар p1 (осень), p2 (весна), численность инфицированных носителей на гектар p3 (осень), p4 (весна),…, количество осадков– p40(сентябрь), высота снежного покрова — p41 (январь) – 43(март), p44 (ноябрь) – p45(декабрь). Как представлено в Таблице 1 такие параметры как численность и количество инфицированных носителей и переносчиков на гектар являются интервальными.

С использованием подхода иммунокомьютинга разработаны вычислительные процедуры обучения с экспертом и самообучения для сложных систем с интервальными данными и параметрами, классификации и группировки на плоскости вычисленных значений энергии связи.

Благодарности. Авторы выражают благодарность за поддержку исследований проект № 2200 р МНТЦ "Разработка математических моделей иммунных сетей для информационной безопасности ", Европейскую комиссию в рамках контракта № ICA2-CT-2000-10048 INCO-COPERNICUS "Эпидемиологическое исследование чумы в Средней Азии на основе пространственно-временной динамики ".

Список литературы

1. Агеев В.С. Паразитарные контакты носителей в устье речных долин пустынной зоны Казахстана и их влияние на эпизоотию чумы. – Саратов, 1975. С. 3–19.
2. Айкимбаев А.М. и др. *Наблюдение за эпидемиологией чумы в Урал-Эмба и Устюрт автономных очагах.* – Алма-Ата: Гылым, 1994.
3. Alefeld G. Rigorous Error Bounds for Singular Values of a Matrix Using the Precise Scalar Product // Computerarithmetic, E. Kaucher, U. Kulish and Ch.Ullrich, eds. – Stuttgart: Teubner, 1987. – P. 9-30.
4. Mayer G. Enclosures for eigenvalues and eigenvectors // Computer Arithmetic and Enclosure Methods, L. Atanassova and J. Herzberger, eds. – Amsterdam: North-Holland, Elsevier Science Publishers B. V., 1992.
5. Shary S.P. Solving the linear interval tolerance problem // *Mathematics and Computers in Simulations*. – 1995. – Vol. 39. – P. 53 – 85.
6. Соколова Л. А. Индекс риска для мониторинга проблемы чумы // *Труды СПИИРАН*. – 2002. – Т. 3. – (Санкт-Петербург, СПИИРАН, 2002).
7. Tarakanov A.O. Information Security with Formal Immune Networks // Proc. of the Methods, Models and Architectures for Network Security (MMM-ACNS 2001), St. Petersburg, 2001. – P. 115–126.
8. Tarakanov A., Skormin V. and Sokolova S. *Immunocomputing: Principles and Applications.* – New York: Springer, 2003 (in press).

Методы распространения ограничений: основные концепции

Семенов А.Л.

Институт систем информатики им. А.П. Ершова СО РАН Новосибирск, пр. ак. Лаврентьева, 6
semenov@iis.nsksu

Аннотация. В работе рассматриваются основные понятия и алгоритмы программирования в ограничениях — одной из областей искусственного интеллекта. Даются понятия совместности для дискретных и непрерывных областей и приводятся алгоритмы достижения совместности. Кратко описаны три системы на базе этого подхода.

1 Введение

Распространение ограничений (иногда также называемое удовлетворением ограничениям или программированием в ограничениях) является одной из наиболее интенсивно развивающихся областей искусственного интеллекта, связанной с решением разнообразных задач. Многие проблемы в различных приложениях могут быть сформулированы как задачи удовлетворения ограничений, хотя люди, не являющиеся специалистами в этой области, не всегда могут это увидеть. Представление же проблемы в виде задачи распространения ограничений позволяет применять для ее решения наряду со специальными методами прикладной области достаточно эффективные и универсальные методы решения задач распространения ограничений. В настоящее время техника распространения ограничений все чаще используется как основа для решения различных прикладных задач, таких как временные рассуждения, задачи ресурсного и календарного планирования, математическое и инженерное моделирование, задачи на графах и т.д. Поэтому естественным является большое внимание, уделяемое исследованию и методов решения задач удовлетворения ограничений. В рамках этих исследований получено достаточно много интересных и важных результатов, а также создан ряд прикладных систем, успешно применяемых для решения широкого спектра прикладных задач.

В методах распространения ограничений выделяют два класса методов — методы для задач с дискретными переменными, называемые распространением ограничений над дискретными областями [5,6,7,9,11-15,17], и методы для задач с вещественными переменными, называемые распространением ограничений над непрерывными областями [1,3,4,8,10,18,19,]. Для каждого из этих классов разработан свой набор эффективных алгоритмов и построены соответствующие прикладные системы [2,16,18,20]. В настоящее время методы распространения ограничений все чаще объединяются с другими подходами, порождая новые эффективные алгоритмы. Например, распространение ограничений над конечными областями используется в комбинации с методами исследования операций и линейного программирования, а распространение ограничений над непрерывными областями использует методы интервальной математики.

Говоря неформально, задача удовлетворения ограничений — это задача, состоящая из конечного числа переменных, каждая из которых связана с некоторой областью ее определения (дискретной или непрерывной), и конечного множества ограничений, которые ограничивают множества значений переменных, принимаемых одновременно. Решение задачи удовлетворения ограничений состоит в присваивании значения каждой переменной так, чтобы одновременно удовлетворялись все ограничения. Целью может являться как нахождение одного такого набора (любого или с заданным свойством), так и нахождение всех наборов.

В данной статье будут рассмотрены основы теории распространения ограничений, ее основные алгоритмы и примеры использования данного подхода для решения прикладных проблем. Статья организована следующим образом. Во второй главе мы дадим основные сведения — базовые определения и алгоритмы для методов распространения ограничений над конечными областями. Третья глава будет посвящена понятиям и алгоритмам для задач над непрерывными областями. В четвертой главе мы рассмотрим некоторые наиболее известные системы на основе методов распространения ограничений. В заключение будет обсуждена дополнительная информация о методах удовлетворения ограничений.

2. Методы распространения ограничений над конечными областями

2.1 Формальное определение задачи удовлетворения ограничений

Определение 2.1. *Областью определения* переменной называется множество всех возможных значений, которые могут быть присвоены этой переменной. В дальнейшем мы будем говорить просто область переменной и обозначать область переменной *x* через D_x .

Если область переменной содержит только числовые значения, то эта переменная называется *числовой переменной*. В качестве области числовых переменных обычно рассматривается кольцо целых чисел ***Z***, точнее ее конечные подмножества. Если область переменной содержит два значения, то такая переменная называется *булевой переменной* и ее значениями считают 0 и 1. Если область переменной есть множество перечислимых значений (например, дни недели, цвета и т.д.), то переменная называется *символьной переменной*.

Определение 2.2. *Меткой* называется пара переменная-значение, которая определяет присваивание некоторого значения переменной. Метка, определяющая присваивание значения *v* переменной *x*, обозначается $<x,v>$. Метка $<x,v>$ осмысленна только в том случае, когда $v \in D_x$.

Определение 2.3. *Составной меткой* называется одновременное присваивание значений некоторому множеству переменных. Составная метка, присваивающая значения $v_1,\ldots,v_n$ соответственно переменным $x_1,\ldots,x_n$, обозначается $\left(< x_1,v_1 >,\ldots,< x_n,v_n >\right)$. Здесь также полагается, что $v_i \in D_{x_i}$ для $i=1,\ldots,n$.

Составная метка рассматривается как множество, поэтому порядок переменных в ней не важен, и она не может содержать двух одинаковых меток.

Определение 2.4. *k-составной меткой* называется составная метка, присваивающая *k* значений одновременно *k* переменным.

Определение 2.5. Пусть *m* и *n* — целые числа и $m \le n$. Будем говорить, что *m*-составная метка *M* есть *проекция n*-составной метки *N*, если все метки из *M* являются метками в *N*. Проекция обозначается *Pr(N,M)*.

Например, $(<a,1>,<c,3>)$ является проекцией $(<a,1>,<b,2>,<c,3>)$.

Определение 2.6. *Ограничением* на некотором множестве переменных *S* называется набор составных меток для всех переменных множества S. Будем обозначать ограничение C_S.

Определение 2.7. О*граничение называется k-арным*, если его переменные образуют множество из *k* элементов. На практике особо выделяются случаи $k=1$ и $k=2$. В первом случае ограничение называется *унарным*, а во втором — *бинарным*.

Определение 2.8. Пусть переменные составной метки *N* совпадают с переменными ограничения *C*. Будем говорить, что *метка N удовлетворяет ограничение C* тогда и только тогда, когда *N* является элементом *C*:

Определение 2.9. Пусть даны составная метка *N* и ограничение *C*, такие, что переменные ограничения *C* являются подмножеством переменных составной метки *N*. Будем говорить, что *метка N удовлетворяет ограничение C* тогда и только тогда, когда проекция *N* на переменные *C* является элементом *C*.

Другими словами, говоря, что *N* удовлетворяет *C*, мы подразумеваем, что если *C* есть ограничение на множестве переменных $\{x_1,\ldots,x_k\}$ или его подмножестве, то метки из *N* для соответствующих переменных образуют составную метку, которая удовлетворяет *C*.

Определение 2.10. *Задача удовлетворения ограничений* (ЗУО) — это тройка ***P = (X, D, C),*** обозначаемая **CSP(P)**, где

X — конечное множество переменных $\{x_1,\ldots,x_k\}$,
D — функция, отображающая каждую переменную из *X* на множество объектов произвольного типа:
D**: **X –> {конечное множество объектов некоторого типа}

Будем рассматривать D_{x_i} как множество объектов, отображенных из x_i функцией ***D***. Эти объекты называются значениями переменной x_i, а множество D_{x_i} — областью x_i.

C — конечное (возможно пустое) множество ограничений на произвольном подмножестве переменных из ***X***, то есть ***C*** -это множество наборов составных меток.

Говоря неформально, основная цель в задачах удовлетворения ограничений состоит в присваивании каждой переменной некоторого значения таким образом, чтобы одновременно удовлетворялись все ограничения. Формально это можно сформулировать следующим образом.

Определение 2.11. *Решением задачи удовлетворения ограничений* называется составная метка для всех переменных задачи, которая удовлетворяет все ограничения задачи.

Определение 2.12. Задача удовлетворения ограничений называется *разрешимой*, если она имеет хотя бы одно решение.

Все задачи удовлетворения ограничений можно разделить на следующие категории:

1) Задачи, в которых достаточно найти одно любое решение;
2) Задачи, в которых требуется найти все решения;
3) Задачи, в которых надо найти оптимальное решение, где оптимальность определяется в соответствии с некоторым критерием. В таких задачах часто бывает достаточно найти почти оптимальное решение.

2.2 Сведение задач и совместность

Одной из базовых техник решения задач удовлетворения ограничений является сведение. Сведением называется класс методов преобразования одних задач к другим, которые легче решать или о которых известно, что они неразрешимы. Несмотря на то, что одно только сведение задач обычно не приводит к решению, оно может облегчить последующие действия.

Определение 2.13. Две ЗУО называются *эквивалентными*, если они имеют одно и то же множество переменных и одно и то же множество решений.

Определение 2.14. Задача *P=(X,D,C) сводится* к задаче *P'=(X',D',C')* если

а) *P* и *P'* эквивалентны;

б) каждая область переменных в *D'* является подмножеством области в *D*;

в) C' является более ограничивающей или также ограничивающей как *C*, (то есть все составные метки, которые удовлетворяют *C'*, будут удовлетворять *C*).

Так как мы рассматриваем ограничения как множества составных меток, то сведение задачи означает удаление элементов из ограничений тех составных меток, которые не появляются ни в одном решении. Если ограничения рассматриваются как функции, то сведение задачи означает изменение функции ограничения.

Определение 2.15. *Значение* в области является *лишним*, если оно не является частью ни одного решения:

Заметим, что удаление "лишних" значений из областей переменных не меняет множество решений задачи.

Определение 2.16. *Составная метка* в ограничении является *лишней*, если она не является проекцией ни одного решения:

Удаление "лишних" меток из ограничений также не меняет множество решений задачи.

Техника сведения преобразует ЗУО к эквивалентной более легкой задаче путем уменьшения размеров областей переменных и ограничений в исходной задаче. Сведение задач включает две возможных цели:

a. удаление лишних значений из областей определения переменных;
b. сужение ограничений так, чтобы меньшее число составных меток удовлетворяли им.

Если в результате область какой-либо переменной или какое-либо ограничение оказались сведены к пустому множеству, то задача является неразрешимой.

Сведение задачи предполагает умение определять лишние значения и лишние составные метки. Эта информация может быть выведена из ограничений. Например, если задано ограничение $x > y$, то мы можем удалить из области переменной x все значения, которые меньше наименьшего значения переменной y, а из области переменной у — все значения, которые больше наибольшего значения переменной х.

Существует ряд концепций совместности, которые помогают в определении лишних значений и составных меток. Эти понятия основываются на том, что если значение в области переменной или составная метка в ограничение нарушает ограничения, то можно показать, что они являются лишними. Далее мы рассмотрим несколько наиболее используемых видов совместностей.

Определение 2.17. *Выражением из ограничений* на множестве переменных *S* в ЗУО *P*, обозначаемое *CE(S, P)*, называется набор всех соответствующих ограничений из *P* на *S* и его подмножествах.

Определение 2.18. Будем говорить, что *составная метка CL удовлетворяет* выражение из ограничений *CE*, если *CL* удовлетворяет все ограничения в *CE*.

Определение 2.19. *k-составная метка CL k-удовлетворяет* выражение из ограничений *CE*, если и только если *CL* удовлетворяет все ограничения в *CE*.

Определение 2.20. ЗУО *CSP(X,D,C)* называется *k-удовлетворяемой*, если и только если для всех подмножеств из *k* переменных из *X* существует множество меток для них, которое удовлетворяет все соответствующие ограничения в *CE(X, (X,D,C))*.

Определение 2.21. Задача удовлетворения ограничений *CSP(X,D,C)* с *n* переменными называется *удовлетворяемой*, если она является *n*-удовлетворяемой.

Определение 2.22. ЗУО *CSP(X,D,C)* является *1-совместной*, если и только если каждое значение в каждой области удовлетворяет унарные ограничения на соответствующую переменную. ЗУО является *k-совместной* для $k > 1$, если и только если все *(k-1)*-составные метки, которые удовлетворяют все соответствующие ограничения, могут быть расширены включением какой-нибудь дополнительной переменной до *k*-составной метки, которая *k*-удовлетворяет все соответствующие ограничения.

Стоит отметить, что 1-удовлетворяемая задача не обязана быть 1-совместной. Например, это будет не так в случае, когда некоторые значения в некоторых областях нарушают ограничения на эту переменную. 1-совместная задача также может быть 1-неудовлетворяемой. Примером такой ситуации является задача, в которой области некоторых переменных пустые, а все значения в непустых областях удовлетворяют ограничения на соответствующие переменные.

Если для всех переменных *x* в ЗУО мы удалим все значения, которые не удовлетворяют ограничению C_x, то полученная ЗУО будет эквивалентна исходной задаче. Это является результатом того, что при этом никакое решение не может быть удалено или добавлено (любое значение, которое не появляется в C_x, не может появиться в полном решении). Полученная ЗУО будет 1-совместной по определению.

Бинарные задачи составляют очень широкий класс и задачи общего вида часто можно свести к бинарным. Поэтому для бинарных задач был разработано несколько особых видов совместностей и, так как каждой бинарной ЗУО можно сопоставить граф, в этих совместностях используется терминология теории графов.

Определение 2.23. ЗУО называется *совместной в узле (node consistency)* если и только если для всех переменных все значения из их областей удовлетворяют унарным ограничениям на данную переменную. Для обозначения совместной в узле задачи *P* используется запись *NC(P)*.

Определение 2.24. Дуга *(x,y)* в графе, представляющем ЗУО, называется *совместной по дугам* (обозначается AC), если и только если для каждого значения *a* из области D_x, которое удовлетворяет ограничение на *x*, существует значение из области D_y, которое совместно с $<x,a>$.

Определение 2.25. ЗУО называется *совместной по дугам* (arc consistency), если и только если каждая дуга в графе ограничений является совместной.

Другими словами, ЗУО является совместной по дугам, если и только если для каждой переменной *x*, для каждой метки $<x, a>$, которая удовлетворяет ограничения на *x*, существует значение *b* для каждой переменной *y*, такое что составная метка $(<x,a>,<y,b>)$ удовлетворяет всем ограничениям на *x* и на *y*. Это полностью совпадает с определением 2-совместности.

Имеет место следующая теорема, свидетельствующая о полезности сведения задачи к эквивалентной задаче, совместной по дугам.

Теорема 1. Поиск в ЗУО может выполняться без использования возвратов, если граф задачи является деревом и задача является совместной в узлах и по дугам.

Определение 2.26. Путь $(x_0,x_1,\ldots,x_m)$ в графе ограничений для некоторой ЗУО называется *совместным по путям* (PC- path consistency), если и только если для любой 2-составной метки $(<x_0,v_0>,<x_m,v_m>)$, которая удовлетворяет все ограничения на x_0 и x_m, существует метка для каждой из переменных от x_1 до x_{m-1}, так что удовлетворяется каждое бинарное ограничение на соседние переменные в пути.

Заметим, что определение совместности по путям не требует, чтобы значения $v_0,v_1,\ldots,v_m$ удовлетворяли все ограничения в выражении из ограничений $CE(\{x_0,x_1,\ldots,x_m\},(X,D,C))$. Например, если переменные x_i и x_j не являются соседними в пути, то 2-составная метка $(<x_i,v_i>,<x_j,v_j>)$ не обязана удовлетворять ограничение $C_{x_i x_j}$.

Определение 2.27. ЗУО называется *совместной по путям*, если и только если каждый путь в ее графе ограничений является совместным.

Другими словами, ЗУО является совместной по путям, если и только если выполняется следующее условие: если для всех пар переменных x и y составная метка *(<x,a>,<y,b>)* удовлетворяет ограничения на x и y, то для каждой переменной z существует метка *<z,c>*, такая что 3-составная метка *(<x,a>,<y,b>,<z,c>)* удовлетворяет все ограничения на x, y, z.

Существует теорема, сводящая установление совместности по путям к совместности путей длины 2.

Теорема 2. ЗУО является совместной по путям тогда и только тогда, когда все пути длины 2 совместны.

Заметим, что совместность в узле совпадает с 1-совместностью, совместность по дугам — с 2-совместностью, а совместность по путям эквивалентна 3-совместности в случае бинарных ЗУО.

2.3 Соотношения между совместностью и удовлетворяемостью

Возникает вопрос: как совместность соотносится с удовлетворяемостью, в частности, является ли k-совместность необходимым или достаточным условием для k-удовлетворяемости или удовлетворяемости задачи вообще? Как доказано в [17], k-совместность не является ни необходимым, ни достаточным условием для удовлетворяемости задач удовлетворения ограничений. В частности в примере выше было показано, что ЗУО, которые 1-совместны, не обязаны быть 1-удовлетворяемы. Вместе с тем, существуют расширения понятия совместности, в частности сильная совместность, достижение которой гарантирует удовлетворяемость ЗУО, если она является 1-удовлетворяемой, что ведет к практическим результатам по решению ЗУО.

Таким образом, k-совместность не является ни необходимым, ни достаточным условием для удовлетворяемости ЗУО. В частности выше было показано, что ЗУО, которые 1-совместны, не обязаны быть 1-удовлетворяемы. Вместе с тем, существуют расширения понятия совместности, в частности сильная совместность, достижение которой гарантирует удовлетворяемость ЗУО, если она является 1-удовлетворяемой.

Таким образом, рассмотренные виды совместности являются *локальными совместностями* — они обеспечивают выполнение каждого из ограничений по отдельности, но не обеспечивают решения задачи в целом. Получение решения задачи гарантирует глобальная совместность.

Определение 2.28. ЗУО $P=(X, D, C)$ называется *глобально совместной* тогда и только тогда, когда каждое значение любой переменной задачи принадлежит решению задачи, т.е. для каждого значения из области каждой переменной существует составная метка, удовлетворяющая все ограничения задачи.

Тем самым, если нам дана некоторая ЗУО, то сведя ее к эквивалентной глобально совместной задаче, мы решим исходную проблему. Однако нахождение глобально совместной задачи в общем случае является NP-полной проблемой и поэтому практически не применимо. Тем не менее, сведение задачи к эквивалентной локально совместной задаче путем определения лишних значений и/или составных меток и их последующего удаления приводит к задаче, имеющей те же решения, что и исходная, но которая может быть решена легче.

Хотя одно сведение задач редко приводит к решениям, оно может разными путями помочь решать ЗУО. Такое сведение может быть использовано как препроцессирование перед применением других техник для нахождения решений. Оно также может быть использовано во время поиска для уменьшения пространства поиска после каждого означивания переменных. Иногда сведение может помочь сделать поиск безвозвратным. Суммируя сказанное, из комбинирования сведения задач с поиском можно получить следующие выгоды:

1. Уменьшение пространства поиска. Так как размер пространства поиска измеряется произведением размеров всех областей в задаче, то сведение задач может помочь уменьшить пространство поиска за счет уменьшения размеров областей отдельных переменных.
2. Избежание повторного поиска в "бесполезных" поддеревьях Лишние значения и составные метки представляют ветви и пути, которые ведут к поддеревьям, не содержащим решений. Если лишние значения и составные метки могут быть удалены в процессе сведения задачи, то это поможет избежать повторного поиска в этих пустых поддеревьях.
3. Определение неразрешимых задач. Если правильный алгоритм сведения задач выдает в качестве результата ЗУО, в которой по крайней мере одна область сведена к пустому множеству, то можно сделать вывод, что задача неразрешима. В этом случае не требуется дальнейших усилий для нахождения решений.

Сведение задач часто называется достижением совместности. Под достижением свойств определенной совместности данной ЗУО мы понимаем сведение задачи посредством удаления лишних значений из областей и лишних составных меток в ограничениях таким образом, что это свойство совместности выполняется в сведенной задаче. Например, процедура, которая "достигает совместность по дугам" в ЗУО — это процедура, которая берет ЗУО P и возвращает ЗУО P', такую что P и P' эквивалентны и $AC(P')$ истинно. Свойства совместности определены таким образом, что полученные ЗУО эквивалентны исходным задачам.

2.4. Алгоритмы достижения совместности в узлах и по дугам

В этом разделе мы рассмотрим базовые алгоритмы достижения двух видов совместностей. Алгоритм достижения совместности по путям является достаточно громоздким, поэтому мы его рассматривать не будем.

Достижение совместности в узлах. Этот алгоритм позволяет удалить из областей определения всех переменных задачи значения, которые не удовлетворяют унарным ограничениям. Достижение совместности в узлах тривиально. Все, что требуется сделать — это просмотреть каждый элемент в каждой области и проверить удовлетворяет ли это значение унарным ограничениям на эту переменную. Все значения, которые нарушают унарные ограничения, удаляются из этих областей. Алгоритм, обеспечивающий совместность в узлах и называющийся **NC-1**, может быть записан следующим образом:

Алгоритм *NC-1(X,D,C)*
/* $D_x = D_x \cap \{x : C_x\}$
Для каждой переменной $x \in X$ выполнить:
Для каждого значения $v \in D_x$ выполнить:
Если метка <*x,v*> не удовлетворяет ограничение C_x, то D_x =- D_x — {*v*};
Возврат с *(X,D,C)*;

После завершения алгоритма исходная задача будет сведена к задаче. которая является совместной в узлах. Пусть *a* — максимальный размер областей и *n* — число переменных в задаче. Так как каждое значение проверяется один раз, то временная и емкостная сложности алгоритма равны $O(a{\cdot}n)$.

Алгоритм достижения совместности по дугам. В отличие от алгоритмов достижения совместности в узлах, этот алгоритм более сложен и в настоящее время существует ряд алгоритмов достижения совместности по дугам [5,11-15,19]. Самый простой алгоритм называется AC-1 [11,12], а самый эффективный AC-7 [5]. Наиболее используемым из этого множества является алгоритм AC-3 [12], который также применяется и в алгоритмах установления совместности для непрерывных областей. Поэтому мы здесь рассмотрим только этот алгоритм.

Алгоритмы достижения совместности по дугам в общем случае удаляют больше лишних значений из областей, чем алгоритм **NC-1**. Самым первым алгоритмом достижения совместности по дугам был алгоритм фильтрации Вольца. Основное действие в алгоритме достижения совместности по дугам — это установление совместности для каждой дуги. Очевидно, что дугу (x_i, x_j) можно сделать совместной, удаляя из D_{x_i} все значения, для которых не существует значений в D_{x_j}, при которых удовлетворяется соответствующее ограничение. При удалении таких значений заведомо не выкидывает ни одного решения исходной задачи. Процедура установления совместности для одной дуги может быть сформулирована следующим образом:

Процедура *REVISE_DOMAIN(x, y)*
DELETED <- ложь;
Для каждого значения $a \in D_x$ выполнить:
Если в области D_y не существует значения *b*, такого что метка (<*x,a*>,<*y,b*>) удовлетворяет ограничение C_{xy},, то
$D_x = D_x$ — {*a*};
DELETED <- истина
Возврат с DELETED

Для того, чтобы сделать граф ограничений совместным по дугам, недостаточно выполнить процедуру REVISE для каждой дуги однажды. Если эта процедура уменьшила область определения для некоторой переменной *x*, то каждая дуга *(x, y)*, проверенная на предыдущем шаге, должна быть подвергнута этой процедуре еще раз, т.к. некоторые значения из области D_y могут оказаться уже несовместными с оставшимися значениями в полученной области D_x. Поэтому мы должны к ней снова применить процедуру REVISE_DOMAIN. Приводимый ниже алгоритм **AC-3** решает эту проблему наиболее эффективным образом.

Алгоритм *AC-3(X,D,C)*
Выполнить алгоритм *NC-1(X,D,C)*;
Q <- $\{(x,y) \mid C_{xy} \in C\}$
Пока $Q \neq \{\}$ выполнять:
Выбираем какую-либо дугу *(x,y)* из *Q* и удаляем ее из *Q*;
Если *REVISE_DOMAIN(x,y)* = истина, то $Q = Q \cup \{(z,x) \mid C_{zx} \in C, x \neq z, z \neq y\}$
Возврат с *(X,D,C)*;

Алгоритм **AC-3** выполняет повторную проверку только тех дуг, на которые могло повлиять предыдущее применение процедуры REVISE_DOMAIN. Условие $y \neq z$ при формировании множества Q следует из того, что ни для одного из удаленных процедурой REVISE_DOMAIN(x,y) значений области D_x не существует значения в D_y, обеспечивающего совместность по этой дуге. Временная сложность этого алгоритма равна $O(a^3n^2)$, то есть является полиномиальной по мощности областей переменных и числу переменных.

Более совершенные алгоритмы, начиная с **АС-4** [13], используют более сложные структуры данных, чтобы уменьшить объем проверок удовлетворяемости ограничений. Для достижения совместности по путям также существует класс алгоритмов, называемых **PC** алгоритмами (от **PC-1** до **PC-4**) [6,12,13], но они являются достаточно трудоемкими и результаты сведения, полученные с помощью этих алгоритмов, не являются достаточно удовлетворительными, чтобы компенсировать расходы на работу алгоритма. Поэтому алгоритмы достижения совместности по путям применяются весьма редко на практике и имеют в основном теоретический интерес.

3 Методы распространения ограничений над непрерывными областями

Задачи распространения ограничений над непрерывными областями обычно называются численными задачами удовлетворения ограничений и обозначают ЧЗУО или NCSP. Задачи этого класса очень часто встречаются в реальных приложениях, например, при моделировании физических явлений, при описании химических процессов, в системах моделирования и проектирования. Как правило, эти задачи представляют собой совокупность уравнений и неравенств, связывающих некоторые непрерывные переменные, хотя иногда ограничения могут задаваться в виде таблиц, а также включать целочисленные переменные. Очень часто реальные системы являются недоопределенными или переопределенными. Поэтому применение классических вычислительных методов или методов компьютерной алгебры для решения таких задач оказывается неудачным и приходится прибегать к упрощению проблемы. Методы, основанные на применении техники распространения ограничений, могут оказаться единственным подходом к решению реальных задач. В случае задач с непрерывными переменными таким аппаратом является интервальное распространение ограничений.

3.1 Особенности задач распространения ограничений над непрерывными областями

Задачи распространения ограничений над непрерывными областями определяются аналогично случаю дискретных областей (определение 2.10). Однако для непрерывных областей есть особенности в определении понятия значение переменной. Ниже этот вопрос будет обсужден подробнее. Кроме того, мы будем считать, что ограничения для числовых ЗУО задаются в виде алгебраических уравнений и неравенств (хотя в общем случае они могут быть заданы таблицами, функциями и т.д.).

Так как мы рассматриваем вещественные (непрерывные) переменные, то следует уточнить понятие "значение" в определении 2.10 для числовых ЗУО и в дальнейшем изложении. Когда x_i является целочисленной переменой, то соответствующая область D_i является подмножеством целых чисел и значением переменной является некоторый элемент этого подмножества. Если же x_i является непрерывной переменной, то соответствующая область содержит бесконечное число элементов, являющихся вещественными числами, большинство из которых не представимо в компьютере. Так как на компьютере можно представить только конечное множество вещественных чисел (это множество называется множеством чисел с плавающей запятой и в общем случае для разных видов компьютеров оно разное), то тем самым в практических приложениях бесконечная область значений непрерывной переменной заменяется конечной. Мы будем обозначать множество чисел с плавающей запятой через **FP**. Таким образом, для вещественных чисел, не входящих в множество **FP**, мы используем аппроксимацию элементами из **FP**. Обычно принимается, что некоторое вещественное число a представлено либо числом a^+, либо числом a^-, где a^- — максимальный элемент из **FP**, такой что $a^- \leq a$, а a^+ — соответственно минимальный элемент из **FP**, такой что $a^+ \geq a$.

Пусть мы имеем ограничение $C(x_1,...,x_n)$, заданное в виде отношения равенства

$$f(x_1,...,x_n) = 0.$$

Очевидно, что определенная аппроксимация значений переменных приводит к появлению ошибок вычислений значения левой части отношения и, тем самым, к вопросу, когда ограничение C справедливо. Одно из решений проблемы состоит в ведении понятия точности вычислений ε, которая задается пользователем.

Определение 3.1. Будем говорить, что *отношение C на множестве непрерывных переменных справедливо* для множества значений $(a_1,...,a_n)$, если $abs(f(a_1,...,a_n)) < \varepsilon$, где ε- некоторое заданное малое число.

Аналогичным образом можно определить справедливость ограничений для отношений больше и меньше. Но такое определение справедливости отношений не позволяет адекватно использовать понятия совместности, введенные в предыдущей главе, так как мы должны будем связывать их с точностью вычислений. Более точным и надежным с вычислительной точки зрения является представление вещественного числа а интервалом

$[a^-, a^+]$. Таким образом, мы аппроксимируем бесконечное множество значений конечным, заменяя одним элементом бесконечное множество значений, лежащих в данном интервале и не являющихся элементами из **FP**. Таким образом, мы используем описанные в предыдущей главе методы интервальной математики для определения значений переменных и нахождения значений ограничений, то есть в этом случае результатом вычисления значения левой части отношения будет некоторый интервал. Для интервального представления вещественных чисел изменим определение 3.1 следующим образом.

Определение 3.2. Будем говорить, что отношение C на множестве непрерывных переменных справедливо для множества значений $(a_1,...,a_n)$, где каждое a_i есть интервал, если $f(a_1,...,a_n) \supset 0$.

Таким образом, мы заменили бесконечную область значений переменной x_i на конечную, элементами которой являются определенные выше интервалы. Поэтому далее под значением непрерывной переменной будем понимать, как и в случае целочисленных переменных, некоторый элемент конечной области. Однако, говоря об удовлетворении ограничений, мы будем иметь в виду определение 3.2.

Определение 3.3. Решением численной ЗУО $P=(X,D,C)$ называется множество $(a_1,...,a_n)$ значений переменных $x_1,...,x_n$, такое что $a_i \in D_i$ и все ограничения из C удовлетворяются.

Использование интервалов для представления значений переменных и методов интервальной математики для установления совместностей привело к тому, что распространение ограничений над непрерывными областями называется *интервальным распространением ограничений*.

3.2 Совместность численных задач удовлетворения ограничений

В предыдущем разделе было рассмотрено несколько видов совместностей, в частности, совместность в узлах, совместность по дугам и совместность по путям. На практике наиболее эффективным оказалось понятие совместности по дугам, различные алгоритмы достижения которого широко используются в разных системах. . Для интервальных методов распространения ограничений существует несколько обобщений понятия совместности по дугам, в частности, 2B-совместность [4,7], совместность на брусе (box consistency) [3], совместность по границам (bound consistency) [18] и ряд других. Ниже будут даны определения наиболее используемых видов совместностей для численных ЗУО и дано их сравнение.

Определение 3.4. Пусть даны две ЧЗУО $P = (X,D,C)$ и $P' = (X, D', C)$, то есть задачи с одним и тем же множеством переменных и одним и тем же множеством ограничений. Будем говорить, что *задача P меньше задачи P'*, если $D \subseteq D'$ (то есть для всех i $D_i \subseteq D_i'$).

Определение 3.5. Будем говорить, что *совместность C_1 слабее совместности C_2*, если задача P_2, сведенная из задачи P к совместности C_2, меньше, чем задача P_1, сведенная из P к совместности C_1.

Определение 3.6. Пусть имеется некоторое подмножество S множества ***R***. *Интервальной оболочкой* множества S (обозначаемой □) называется минимальный интервал I, такой что $S \subseteq I$.

Определение 3.7. Пусть дана ЧЗУО $P = (X, D, C)$ и $c \in C$ есть k-местное ограничение над переменными $(x_1,\dots,x_k)$. Будем говорить, что *ограничение c является 2B-совместным*, если и только если

$$\forall i,\ \mathbf{x}_i = \square\{x_i \mid \exists x_1 \in \mathbf{x}_1,\dots, \exists x_{i-1} \in \mathbf{x}_{i-1}, \exists x_{i+1} \in \mathbf{x}_{i+1},\dots, \exists x_k \in \mathbf{x}_k \text{ такие что } c(x_1,\dots,x_{i-1},x_i,x_{i+1},\dots,x_k) \text{ справедливы}\}$$

Определение 3.8. Пусть дана ЧЗУО $P = (X, D, C)$. Будем говорить, что *задача P является 2B-совместной*, если и только если все ее ограничения являются 2B-совместными.

Можно показать, что 2B-совместность слабее совместности по дугам. Рассмотрим численную ЗУО:

$P = (\{x_1,x_2\}, \{[1,4],[-2,2]\}, \{x_1 = x_2*x_2\})$. Задача P является 2B-совместной, но не является совместной по дугам, так как в области переменной x_1 не существует значения, удовлетворяющего ограничение при $x_2 = 0$.

Алгоритм достижения 2B-совместности практически совпадает с приведенным выше алгоритмом ***AC-3*** достижения совместности по дугам. Он может быть представлен следующим образом:

```
Алгоритм IC(X,D,C)
  Выполнить алгоритм NC-1(X,D,C);
  Q <- C;
  Пока  Q ≠ {} выполнять:
      Выбираем какое-либо ограничение из Q и удаляем его из Q;
      D' <- narrow(c, D);
      Если D' ≠ D, то  D = D', Q = Q ∪ {c' ∈ C | Var(c) ∩ Var(c') ≠ {}}
```

Возврат с *(X,D',C)*;

Var(c) в алгоритме выше обозначает множество всех переменных в ограничении *c*. Самым важным шагом в алгоритме является изменение областей переменных с помощью оператора *narrow*. Одним из способов реализации этого оператора является разбиение множества ограничений на примитивные ограничения, которые представляют собой бинарные и тернарные ограничения, и вычисление интервальных значений одних аргументов этих ограничений на основе интервальных значений других аргументов. В [1] показано, что такое разбиение не приводит к потере решений или появлению новых решений исходной системы ограничений. Рассмотрим пример построения системы примитивных ограничений. Пусть мы имеем численную задачу удовлетворения ограничений $P = (\{x, y\}, D = (-\infty, +\infty) \times (-\infty, +\infty), C)$, где

$$C\colon\ e^{\sin(x)} + x^y = 0$$

Записав это уравнение в виде примитивных отношений, мы получим систему *C'*:

$$t_1 = \sin(x) \quad x = \arcsin(t_1)$$

$$t_2 = e^{t_1} \quad t_1 = \ln(t_2)$$

$$t_3 = x^y \quad x = e^{\ln(t_3)/y} \quad y = \frac{\ln(t_3)}{\ln(x)} \quad t_1 = -t_3$$

Таким образом, мы построили численную ЗУО, эквивалентную исходной:

$$P' = (\{x, y, t_1, t_2, t_3\}, D \times [-\infty, +\infty] \times [-\infty, +\infty] \times [-\infty, +\infty], C')$$

Таким образом, все ограничения новой системы имеют вид *x = f(y,z)* или *x = f(y)*. Оператор *narrow* для каждого ограничения *c'* данной системы записывается следующим образом:

$x' = x \cap f(y)$

где значениями всех переменных, входящих в ограничения, являются интервалы, и эти интервалы фактически задают области соответствующих переменных. Например, в результате выполнения первого ограничения область переменной t_1 будет сужена от интервала $[-\infty, +\infty]$ до интервала $[-1, +1]$, а в результате .выполнения шестого ограничения область переменной *x* станет $[0, +\infty]$.

Определение 3.9. Пусть дана ЧЗУО *P = (X, D, C)* и $c \in C$ есть *k*-местное ограничение над переменными $(x_1,\ldots,x_k)$. Будем говорить, что *ограничение c является совместным на брусе*, если для всех переменных x_i выполняются следующие соотношения:

$$c(x_1, \ldots, x_{i-1}, [\underline{x}_i, \underline{x}_i^{+}), x_{i+1}, \ldots, x_k)$$

$$c(x_1, \ldots, x_{i-1}, (\overline{x}_i^{-}, \overline{x}_i], x_{i+1}, \ldots, x_k)$$

Определение 3.10. Пусть дана ЧЗУО *P = (X, D, C)*. Будем говорить, что *задача P является совместной на брусе*, если и только если все ее ограничения являются совместными на брусе.

В работе [8] показано, что *совместность на брусе является более слабой, чем 2B-совместность*, то есть задача, которая 2B-совместна является также и совместной на брусе, но задача, которая совместна на брусе, может не быть 2B-совместной. Рассмотрим численную ЗУО:

$P = (\{x_1,x_2\}, \{[-1,1],[0,1]\}, \{x_1 + x_2 + x_2 = 0\})$. Задача *P* не является 2B-совместной по x_2 так как не существует значения x_1 такого, что ограничение удовлетворяется при $x_2 = 1$., но задача является совместной в брусе для x_2, так как для обоих граничных значений этой переменной ограничение задачи удовлетворяется.

Также можно показать, что 2B-совместность примитивной системы ограничений слабее, чем совместность в брусе исходной системы ограничений.

Достижение совместности на брусе может быть выполнено с помощью описанного выше алгоритма **IC**, в котором оператор сужения narrow реализован с помощью интервального однопеременного алгоритма Ньютона, описанного в предыдущей главе, и бисекции получаемых областей до достижения требуемой точности [18].

Аналогично определению совместностей более высокого порядка для ЗУО над непрерывными областями (например, совместности по путям), для численных ЗУО также можно ввести понятия k-совместностей для определенных выше совместностей. Наиболее практичными являются понятия 3B-совместности и совместности по границам, которые позволяют отбросить большие области лишних значений по сравнению с 2B-совместностью и совместностью на брусе. Однако достижение совместностей более высоких порядков требует существенного увеличения объема работ, но не всегда трудозатраты окупаются качеством полученных результатов.

Определение 3.11 Пусть дана ЧЗУО *P = (X, D, C)* и пусть $x \in X$. Обозначим:

$P_{[\underline{x},\underline{x}^+)}$ — задача, полученная из *P* заменой *x* на $[\underline{x}, \underline{x}^+)$

$P_{(\overline{x}^-,\overline{x}]}$ — задача, полученная из *P* заменой *x* на $(\overline{x}^-, \overline{x}]$

Область D_x называется 3В-совместной, если и только если сведение задач $P_{[\underline{x},\underline{x}^+)}$ и $P_{(\overline{x}^-,\overline{x}]}$, к 2В-совместным задачам не порождает задач с пустыми областями значений.

Определение 3.12 Численная ЗУО *P = (X, D, C)* называется 3В-совместной, если и только если все ее области являются 3В-совместными.

Доказано, что совместность в боксе слабее, чем 3В-совместность примитивной системы ограничений, то есть достижение 3В-совместности для примитивной системы ведет к совместности в брусе исходной системы.

Определение 3.13. Пусть дана ЧЗУО *P = (X, D, C)* и $c \in C$ есть *k*-местное ограничение над переменными $(x_1,\ldots,x_k)$. Будем говорить, что *ограничение c является совместным по границам*, если для всех переменных x_i сведение ограничения *c* с подставленными значениями границ переменной x_i, то есть ограничений

$$c(x_1,\ldots,x_{i-1},[\underline{x}_i,\underline{x}_i^+),x_{i+1},\ldots,x_k)$$
$$c(x_1,\ldots,x_{i-1},(\overline{x}_i^-,\overline{x}_i],x_{i+1},\ldots,x_k)$$

к совместным на брусе ограничениям не порождает пустых областей переменных задачи *P*.

Определение 3.14. Пусть дана ЧЗУО *P = (X, D, C)*. Будем говорить, что *задача P является совместной по границам*, если и только если все ее ограничения являются совместными по границам.

В [8] показано, что 3В-совместность примитивной системы ограничений слабее совместности по границам исходной задачи. Таким образом, можно выписать следующие соотношения совместностей для интервальных задач распространения ограничений ($C_1 \le C_2$ означает, что совместность C_2 слабее совместности C_1):

по-границам $\le$ 3В(примитивной) $\le$ на-брусе

2В $\le$ на-брусе $\le$ 2В(примитивной)

Данные соотношения показывают, что 2В-совместность для примитивной системы ограничений является самой слабой, а самой сильной является совместность по границам. Вместе с тем, установление 2В-совместности для примитивной системы ограничений может быть выполнено легче всего и, к тому же, оно применимо для любых систем ограничений, например, содержащих разрывные функции, целочисленные переменные (при наличии операций целочисленной интервальной арифметики) и другие типы данных. Для достижения совместности на брусе требуется применение метода Ньютона, что предполагает дифференцируемость функций или необходимости наклонов вместо производных. Для достижения 3В-совместности и совместности по границам требуется трудоемкое исследование областей переменных, выполняемое с помощью алгоритмов бисекции и соответствующих алгоритмов установления совместностей более низких порядков.

Следует отметить, что рассмотренные алгоритмы достижения совместностей для задач интервального распространения ограничений основаны на алгоритме **IC** — версии алгоритма **AC-3**. Вместе с тем, в литературе существуют ссылки на расширение алгоритма **AC-5** [14,15], которое может использоваться для непрерывных областей. Этот алгоритм используется в коммерческой системе ILOG [20] и его описание пока не доступно.

4. Системы, основанные на методах распространения ограничений

4.1 ILOG Solver [20]

ILOG является одной из крупнейших в мире фирм, разрабатывающих продукты с использованием методов распространения ограничений. Их основным продуктом в этой области является ILOG Solver — библиотека Си++ классов для решения комбинаторных задач и нахождения оптимального решения. Он основан на методах распространения ограничений и может также служить средством для интеграции с другими оптимизационными технологиями, включая линейное программирование, локальный поиск, определяемые пользователем эвристики и генетические алгоритмы.

ILOG Solver поддерживает работу с целыми, вещественными, логическими и множественными переменными и позволяет использовать линейные, нелинейные и логические ограничения. В ILOG Solver реализованы несколько так называемых глобальных ограничений, которые позволяют связывать в одном ограничении несколько переменных, и эти ограничения могут эффективно обрабатываться. При этом существенно уменьшается общее число ограничений по сравнению с использованием отдельных ограничений с той же семантикой. Глобальными ограничениями являются alldiff, distribute, table constraint, sequence.

Новые типы и ограничения на них, а также новые ограничения для существующих типов могут быть расширены разработкой новых C++ классов. Существуют метаограничения — конъюнкция и дизъюнкция существующих ограничений, и допускается определение ограничений с весами и упорядочивание ограничений (это также может рассматриваться как метаограничения). Решатель содержит возможность динамического добавления и удаления ограничений.

В качестве алгоритма распространения ограничений применяется модифицированный алгоритм совместности по дугам AC-5, который в отличие от базового алгоритма, работающего только с целыми числами, может применяться также и с вещественными, логическими и множественными переменными. По мнению авторов, этот алгоритм реализован очень тщательно, и они утверждают, что их алгоритм превосходит другие аналогичные реализации более чем в 2000 раз. Следует также отметить наличие параллельной версии ILOG Solver, что позволяет получать хороший выигрыш во времени при решении больших прикладных задач.

В ILOG Solver реализовано несколько наиболее эффективных алгоритмов поиска, включающих как стандартные стратегии (поиск в глубину, поиск сначала наилучшего), так и специализированные алгоритмы для задач удовлетворения ограничения над конечными областями (Limited Discrepancy Search, Depth Bounded Discrepancy Search). Пользователь может управлять процессом поиска, устанавливая параметры поиска, добавляя различные эвристики для сокращения областей, а также может создавать комбинации алгоритмов поиска. Пользователь может также выбирать требуемое решение (все, первое, лучшее и т.д.), изменять имеющееся решение с учетом новой информации, проверять имеющееся решение на оптимальность и т.д.

4.2 Numerica [18]

Numerica представляет собой одну из наиболее известных систем, разработанных на основе интервального распространения ограничений. Ее автор — Паскаль Ван Хентенрик, является автором ряда базовых концепций в этой области и им были разработаны основные алгоритмы достижения совместности на брусе и совместности по границам. Эти, а также другие специальные методы, и лежат в основе системы Numerica.

Numerica представляет собой систему с графическим пользовательским интерфейсом, доступную IBM PC совместимых на машинах. Система предоставляет входной язык, достаточно удобный для записи систем уравнений и неравенств. Numerica позволяет решать только системы с вещественными переменными, которые должны быть квадратными, то есть число переменных в них должно совпадать с числом уравнений. Все остальные уравнений, а также неравенства рассматриваются как дополнительные ограничения. При этом все уравнения решаемой системы должны быть дифференцируемы. Все это является следствием применения алгоритма достижения совместности в брусе, в котором используется интервальный алгоритм Ньютона.

Достоинствами Numerica является использование нескольких способов вычисления значений функций, включая разложение в ряд Тейлора, что позволяет получать достаточно точные границы при оценивании функций с интервальными параметрами. Также в Numerica используется специальная модификация алгоритма ветвей и границ, называемая методом ветвей и отсечений, которая позволяет весьма эффективно решать задачи глобальной оптимизации. Еще одним привлекательным свойством системы является доказательство существования корней, которое основано на использовании интервального метода Ньютона.

В [18] приводятся результаты численных экспериментов, показывающих, что Numerica может решать достаточно сложные задачи с хорошей эффективностью. В частности, она превосходит большинство численных пакетов для решения систем нелинейных уравнений, основанных как на классических, так и на интервальных методах. Наличие большого количества стандартных тестовых примеров, поставляемых с системой, наряду с приводимыми решениями и временами получения решений делает Numerica хорошим инструментом для тестирования эффективности пакетов для решения систем нелинейных уравнений.

4.3 UniCalc [2]

Решатель UniCalc представляет собой интегрированную среду с графическим пользовательским интерфейсом, доступную на большинстве платформ. Эта среда включает встроенный редактор, различные настройки, удобный входной язык для записи решаемых задач и средства графического отображения одномерных интервальных функций. UniCalc содержит несколько различных методов, в том числе некоторые виды символьных преобразований (упрощение, символьное дифференцирование, приведение подобных). Базовым вычислительным методом является один из вариантов методов распространения ограничений, называемый методом недоопределенных вычислений [1] и основанный на достижении 2B-совместности для примитивной системы ограничений, которая получена из исходной задачи. Ядро решателя, объединяющее все методы, называется вычислителем и может быть использовано отдельно. В настоящее время на базе этого вычислителя создано несколько программных продуктов, и он встроен в ряд промышленных систем.

UniCalc предназначен для решения задач, представленных системами алгебраических уравнений, неравенств и логических выражений. Решаемая система может быть переопределенной или недоопределенной, а параметры уравнений и неравенств могут быть заданы в виде интервалов допустимых значений. UniCalc позволяет проводить вычисления как с целыми, так и с вещественными переменными, причем они могут входить в систему одновременно. В результате вычислений находится покоординатная внешняя оценка всех действительных решений системы (в общем случае неоптимальная). Если заданная система не имеет действительных решений, то выдается сообщение об ее несовместности. В решатель также встроен алгоритм локализации корней, позволяющий выделять отдельные решения из множества всех решений.

Для локализации корней используется алгоритм, основанный на рекурсивном делении интервалов переменных и применении метода распространения ограничений. В этом алгоритме выбирается переменная, по которой начинается поиск корня, и задается направление — слева или справа. Выбором последовательности переменных, по которым осуществляется бисекция, можно управлять. В этом алгоритме сужение интервалов происходит как за счет их деления, так и за счет применения метода распространения ограничений. Корень считается найденным, если максимальная ширина интервалов по всем переменным не превышает заданную точность и система остается совместной. Базовый алгоритм локализации корней является основой для нескольких дополнительных алгоритмов, в которых используются различные эвристики для динамического выбора переменных разбиения и их комбинаций. В частности, это позволяет сводить исходную систему к 2B- и 3B-совместным задачам.

Заключение

В работе были рассмотрены основные понятия и алгоритмы программирования в ограничениях — одной из наиболее интенсивно развивающихся областей искусственного интеллекта. Этот подход применяется как для задач с дискретными (целочисленными, перечислимыми) переменными, так и для задач с непрерывными (вещественными) переменными. Для обоих случаев в работе были даны определения основополагающего понятия методов распространения ограничений — совместности и представлены некоторые алгоритмы достижения совместности, а для непрерывных областей также были приведены сравнения четырех видов специальных совместностей. Однако понятие совместности не ограничивается вариантами, рассмотренными в работе. В частности, для дискретных областей тоже существует целый ряд различных видов совместности — направленная, сильная, адаптивная и другие, которые также используются для решения задач. Для достижения различных совместностей разработаны специальные алгоритмы (в том числе параллельные), реализованные в прикладных системах.

В настоящее время основные усилия исследователей, работающих в этой области, направлены на разработку эффективных алгоритмов поиска решений. Эти алгоритмы представляют собой комбинацию различных подходов — методов распространения ограничений, алгоритмов поиска и специальных алгоритмов из предметных областей. Одно из направлений работ в этих исследованиях — создание кооперативных решателей, которые объединяют несколько различных подходов и используют их совместно при решении одной задачи, включая параллельную работу различных методов с обменом полученными результатами. Данное направление кажется очень перспективным для решения действительно сложных задач. Примером такого кооперативного решателя является решатель SibCalc [16].

Список литературы

1. Кашеварова Т.П., Семенов А.Л. Некоторые вопросы сходимости метода недоопределенных вычислений. В: *Проблемы представления и обработки не полностью определенных знаний*. — Москва–Новосибирск: РосНИИ ИИ, 1996, с. 31-37.
2. Babichev A.B., Kadyrova O.B., Kashevarova T.P., Leshchenko A.S., Semenov A.L. UniCalc, A Novel Approach to Solving Systems of Algebraic Equations. *Interval Computations* **2** (1993), pp. 29-47.
3. Benhamou F., McAllester D., Van Hentenryck P. CLP(Intervals) Revisited. In: *Proc. of ILPS'94,* Ithaca, NY, USA, 1994, pp. 124–138.
4. Benhamou F., Older W. Applying Interval Arithmetic to Real, Integer and Boolean Constraints. *Journal of Logic Programming* **32** (1997), pp.1-24.
5. Bessiere C. Using constraint metaknowledge to reduce arc consistency computation. *Artificial Intelligence* **65**, (1994), pp. 179-190.
6. Bessiere C., Freuder E.C., and Régin J.-R. Comments on Mohr and Henderson's path consistency algorithm. *Artificial Intelligence.* **107** (1999), pp. 125-148.
7. Cleary. J. Logical Arithmetic. *Future Comput. Syst*. **2** (2) (1987), pp. 125–149.
8. Collavizza H., Delobel F., Rueher M. A note on partial consistencies over continuous domains. In: *Proc. CP98*, LNCS 1520, Springer Verlag, 1998, pp. 147–161.

9. Freuder E. Synthesizing Constraint Expression. *CACM* **21** (11) (1978), pp. 958–966.
10. Lhomme O. Consistency Tecniques for Numeric CSP's. In: *Proc. of the 13th IJCAI, R. Bajcsy (Ed)*. — IEEE Computer Society Press, 1993, pp. 232-238.
11. Macworth A.K. Consistency in Networks of Relations. *Artificial Intelligence* **8** (1) (1977), pp.99-118.
12. Mackworth A.K. The complexity of some polynomial network consistency algorithms for constraint satisfaction problems. *Artificial Intelligence* **8** (1997), pp. 99-118.
13. Mackworth A.K. and Freuder E.C.. Arc and path consistency revised. *Artificial Intelligence* **25** (1985), pp. 65-74.
14. Mohr R., Henderson T.C. Arc consistency for factorable relations. *Artificial Intelligence* **28** (1986), pp. 225-233.
15. Perlin M. A generic arc-consistency algorithm and its specializations. *Artificial Intelligence* **53** (1992), pp. 329-342
16. A. Semenov, D. Petunin, A. Kleymenov. GMACS — the general-purpose module architecture for building cooperative solvers. In: *Proceedings of the 2000 ERCIM/Compulog Net Workshop on Constraints*. Padova, Italy, June, 2000.
17. Tsang E. *Foundation of Constraint Satisfaction,* London: Academic Press Ltd., 1993.
18. Van Hentenryck P., Michel L. and Deville Y. *Numerica: a modeling language for global optimization,* Cambridge: MIT Press, 1997.
19. Van Hentenryck P., Deville Y, and Teng C.-M. Arc-consistency and arc-consistency again. *Artificial Intelligence* **57** (1992), pp. 291-321.
20. http://www.ilog.com/products/solver/

Интервальный алгоритм определения статического давления грунта на жёсткие круглые трубы в высокой насыпи в несколько ниток

М. Б. Бозоров

Навоийский государственный горный институт
Узбекистан, Навои
mamurjon@mail.ru

Аннотация. Предложена интервальная методика численного расчета по определению статического давления насыпи на круглые одиночные и многониточные трубы в высокой насыпи в несколько ниток. Для интервального расчета используется полная арифметика Каухера. Далее, дается описания комплекса программ, который предназначен для расчета труб, уложенных в среде в несколько ниток интервальным методом конечных элементов. Полученные численные результаты для двух и многониточных труб сравниваются с известными теоретическими и экспериментальными результатами.

В данной работе продолжены исследования, начатые в работе [1].

В последнее время, задачи оценки влияния различных неопределённых факторов: внешних возмущающих сил, неконтролируемых вариаций параметров, погрешностей в задании начальных условий на поведении решений рассматриваемых систем успешно решаются в рамках интервального анализа [2-4] - гарантированного (минимаксного) подхода. Применение интервального анализа позволяет заключить в интервалы решения задач, о входных данных которых известно лишь то, что они лежат в некоторых интервалах.

Данная работа отличается от [1] тем, что рассматриваемая задача решается в рамках интервального анализа. При этом используется минимаксная интервальная арифметика — полная арифметика Каухера.

При определении давления грунта на трубы необходимо учитывать следующие факторы [1]: количество ниток; рельеф насыпи; условия опирания труб и другие факторы, встречающиеся в проектной практике. Аналитические же решение, полученные в [1], позволяет учесть рельефы насыпи только для одиночной трубы, и дают приближенную теоретическую оценку для труб многониточной укладки.

В данной работе использован интервальный аналог метода конечных элементов (ИМКЭ), ориентированный на решении пространственной задачи линейной расчетной теории упругости. В качестве расчетной модели принимается весомая упругая среда, содержащая подкрепленные круговыми цилиндрами отверстия и другие включения (фундамент, неоднородности грунта и т.д). При этом для труб принимаем, что цилиндр спаян со средой (отсутствие проскальзывания грунта по поверхности трубы).

На внешнем контуре среды краевые условия имеет следующий вид :

- на вертикальных границах сдвигающие напряжения и горизонтальные перемещения, либо равны нулю, либо эти границы свободны;
- на нижней горизонтальной границе примыкающей к основанию насыпи, отсутствуют вертикальные и горизонтальные перемещения;
- верхняя поверхность либо свободно от внешних воздействий, либо загружена нагрузкой.

Размеры выбранной для расчета области должны быть оптимальными, либо от этого зависит время затрачиваемое на расчет по ИМКЭ и, следовательно эффективность работы алгоритма. Если грунт является изотропным материалом или рассматриваемая система трубы-грунт имеет ось симметрии, то возможно уменьшении расчетной области за счет приятия только симметричной её половины.

На основе предложенной методики составлен комплекс программ IMFE, который предназначен для расчета труб, уложенных в насыпи в несколько ниток.. Комплекс программ реализован на языке Turbo Pascal 7.0.

Вводимая информация содержит минимально необходимые данные: высота насыпи; размеры труб (диаметр и толщина стенки) их количество и расстояния между ними; тип опирания труб; характеристика материала трубы и грунта насыпи; формы и положения поверхностей нагрузки.

Для оценки достоверности получаемых результатов произведено тестирование IMFE на задачах линейной теории упругости, имеющих аналитические решения. Кроме того, производилось сравнение с экспериментальными данными полученными другими авторами для одиночной трубы. Наш анализ результатов численного моделирования показал, что использование методов интервального анализа, кроме классических точечных результатов моделирования даёт исследователю гарантированные верхние, нижние и усредненные значения искомых величин.

Список литературы

1. Бозоров М.Б., Ишмаматов М.Р. Алгоритм определения статического давления грунта на трубы методом конечных элементов // Вычислительные Технологии / Сб. научных трудов ИВТ СО РАН. — Т. 3, № 10. — Новосибирск, 1995. — С. 36–42.
2. Калмыков С.А., Шокин Ю.И, Юлдашев З.Х. Методы интервального анализа. — Новосибирск: Наука,1986.
3. Алефельд Г., Херцбергер Ю. Введение в интервальные вычисления. — М.: Мир, 1987.
4. Шарый С.П. Интервальные алгебраические задачи и их численное решение. Диссертация на соискание учёной степени доктора физико- математических наук. — Новосибирск : ИВТ СО РАН, 2000. — 327 с.

Возможности описания схемы соединения интервальной электрической цепи

Н.В. Киншт, Н.Н. Петрунько

Институт автоматики и процессов управления ДВО РАН,
Россия, 690041 г. Владивосток, ул. Радио, 5
kin@iacp.dvo.ru

Аннотация. Формальная структура уравнений задачи анализа электрических цепей (ЭЦ) связана со структурой ЭЦ, описываемой матрицами сечений и контуров ЭЦ, в выборе которых возможен значительный произвол, теоретически не влияющий на результат расчета. При интервальной постановке независимости результата анализа от структуры системы уравнений гарантировать нельзя. Рассматриваются возможности эвристического построения матриц основных сечений и контуров, ассоциируемых с деревьями графа ЭЦ специального вида, построенных с учетом интервальных характеристик режимов и параметров ветвей.

Попытки сформулировать общие принципы интервального анализа объектов, исследование которых может производиться в терминах теории электрических цепей (ЭЦ), предпринимались как отечественными исследователями, так и за рубежом. Не обращаясь к критическому обзору полученных результатов, будем опираться на методологию, основанную на постулатах теории электрических цепей и являющуюся органической ее частью [1]. Собственно теория ЭЦ заключена в учете особенностей цепи на уровне идеализированных элементов, обобщенных ветвей, подсхем, многополюсников, преобразований фрагментов ЭЦ с учетом топологии, аппроксимации схемных функций; использование всех этих теоретических построений приводит к повышению эффективности традиционных и созданию новых методов расчета. Опыт теории ЭЦ свидетельствует, что при всякой новой формализации задачи возникают как новые возможности, так и новые проблемы в схемном представлении ЭЦ идеализированными моделями в применении традиционных методов решения задач. Естественно стремление заложить идею систематичности, присущую теории ЭЦ в целом, в условие задачи исследования ЭЦ на возможно более раннем этапе, дающую возможность воспользоваться плодами многолетнего опыта теории ЭЦ.

Формальное описание задачи анализа сложной интервальной линейной ЭЦ постоянного тока производится полной системой уравнений по законам Кирхгофа и Ома:

$$D\,I_B = 0; \qquad (1)$$
$$B\,U_B = 0; \qquad (2)$$
$$U_B = E_B + R_B\,(J_B - I_B), \qquad (3)$$

(либо вместо (3) применяется эквивалентное выражение $I_B = J_B + G_B\,(E_B - U_B)$).

Здесь, как обычно,

D, B — матрицы основных сечений и контуров ЭЦ, соответственно, причем в частном случае в качестве матрицы D может применяться матрица инциденций A;

U_B, I_B, E_B, J_B – интервальные матрицы — столбцы напряжений, токов, источников э.д.с. и источников токов ветвей;

$R_B = G_B^{-1}$- диагональные интервальные матрицы сопротивлений и проводимостей ветвей; размеры упомянутых матриц определяются количеством ветвей ЭЦ *p* и узлов *q*.

В данной задаче все матрицы параметров ветвей ЭЦ E_B, J_B, R_B, G_B подразумеваются интервальными, и, соответственно U_B, I_B, как параметры режимов ветвей могут быть априори заданы своими интервальными ограничениями. Задача анализа интервальной ЭЦ, как правило, формулируется как построение наилучших интервальных оценок для параметров режима в пространстве $\{U_B,\ I_B\}$.

Первой особенностью этой системы, бросающейся в глаза, является ее разреженность. Теория ЭЦ в значительной степени заключена в использовании идей и способов планомерных преобразований, "сворачивания" задачи и ее декомпозиции для получения эффективного решения, и на этом основаны распространенные методы анализа ЭЦ – контурных токов, узловых потенциалов, напряжений ветвей дерева.

Так, метод узловых потенциалов (ориентирующийся на предпочтительное представление источников в виде источников тока), заключается в разрешении системы:

$$AG_BA^TU_У = -\ AJ_B, \qquad (4)$$

где $U_У$ — интервальный вектор узловых потенциалов размера $(q - 1)$, и вектор напряжений ветвей связан с ним через матрицу сечений: $U_B = A^T U_Y$. В конечном итоге формальное решение для векторов токов и напряжений может быть записано в виде:

$$U_B = A^T(AG_BA^T)^{-1}J_B$$
$$I_B = (1 - G_BA^T(AG_BA^T)^{-1}A)J_B$$

Аналогично, метод контурных токов (ориентирующийся на предпочтительное представление источников в виде источников э.д.с.) заключается в разрешении системы:

$$BR_BB^TI_k = — BE_B, \qquad (5)$$

где I_k — интервальный вектор контурных токов размера $(p — q +1)$, и вектор токов ветвей связан с ним через матрицу контуров: $I_B = B^T I_k$. Формальное решение для векторов токов и напряжений может быть записано в виде:

$$U_B = B^T(BR_BB^T)^{-1}E_B$$
$$I_B = (1 - R_BB^T(BR_BB^T)^{-1}B)\, E_B.$$

Кроме однородных координатных базисов (в приведенных случаях — U_Y и I_k) возможны и другие подстановки в полную систему уравнений ЭЦ.

Расчетные уравнения (1…3, 4 и 5) эквивалентны между собой с точки зрения "точной" математики, но очевидно, что с точки зрения интервальной постановки задачи они различны. Более того, в постановке этих задач имеется существенный произвол, заключающийся в выборе топологических матриц независимых контуров и независимых сечений B и D.

Наверное, можно было бы утверждать, что если вместо матриц независимых сечений и контуров ЭЦ D и B использовать матрицы всех (!) контуров и сечений D_0 и B_0, то можно было бы получить наилучшее интервальное решение для параметров режима U_B, I_B. Ясно, что этот путь совершенно неконструктивен ввиду "проклятия размерности". Утверждать же что-либо относительно качества решения задачи анализа ЭЦ при произвольно назначенных матрицах независимых сечений и контуров D и B в настоящее время пока не представляется возможным.

В работе [1] рассматривались вопросы решения интервальных уравнений ЭЦ, однако они были ориентированы, в основном, на анализ и аппроксимацию компонентных уравнений (3). Вместе с тем остается совершенно открытым вопрос об эффективном описании структуры – схемы соединения элементов интервальной ЭЦ.

В качестве одного из естественных эвристических принципов формирования системы интервальных уравнений можно принять идею о том, чтобы каждая интервальная переменная входила в систему уравнений минимальное число раз. Это в целом относится как к параметрам режима U_B, I_B, так и к пассивным параметрам ветвей R_B и G_B. При аналитических преобразованиях, необходимых для применения обычно применяемых методов, параметры ветвей неоднократно повторяются в различных уравнениях. (В общей теории интервальной постановки задачи такой возможности не предусмотрено, однако при описании интервальной ЭЦ это явление типично).

В теории ЭЦ активно эксплуатируется идея об оптимизации выбора системы основных контуров и сечений (графа) ЭЦ. В теоретических построениях и алгоритмах расчета, как правило, матрицы D и B ассоциируются с некоторыми деревьями специального вида. Этот выбор ориентируется на некоторые свойства элементов ветвей, а именно – линейность либо нелинейность элемента, его вид (резистор, индуктивность, емкость), или некоторые другие специфические параметры. Матрицы D и B, ассоциированные, соответственно, с деревьями T_D и T_B графа ЭЦ, обладают специальными свойствами. А именно – каждая ветвь дерева T_D входит в единственное сечение, порожденное ею, и каждая хорда дерева T_B входит в единственный контур, ею порожденный. Кроме того, матричные произведения типа (DG_BD^T) и (BR_BB^T) также обладают специальными свойствами. Это симметричные матрицы, на главной диагонали которых расположены суммы значений параметров элементов ветвей (G_B или R_B), входящих в соответствующее сечение (контур), а внедиагональные члены образованы элементами, одновременно инцидентными соответствующей паре сечений (контуров).

Очевидно, в задаче анализа интервальной ЭЦ основными параметрами, определяющими выбор системы основных сечений и контуров, следует считать свойства интервальности параметров элементов ЭЦ. В качестве численного критерия интервальности можно использовать величину интервала рассматриваемой переменной (или же относительную величину интервала, например, нормируя ее средним значением; последнее замечание в наибольшей степени применимо к параметрам, знак которых заведомо известен – например, к резистивным параметрам ветвей). Проиллюстрируем дальнейшее на примерах.

За основу возьмем граф ЭЦ, изображенный на рис.1. Упорядочим множество ветвей ЭЦ {1, 2, …12} в соответствии с различными интервальными характеристиками ветвей.

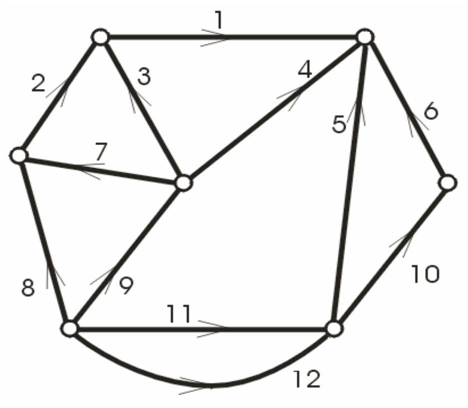


Рис.1. Граф ЭЦ для иллюстративного примера.

Так, пусть вначале ветви ЭЦ упорядочены в порядке убывания их "токовой" интервальной характеристики: {1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12}. В этом случае желательно, чтобы элементы с меньшими номерами входили в матрицу проводимостей единственный раз, и на основе этого критерия строим дерево T_D = {1, 2, 3, 5, 6, 8}. Ветви 4 и 7 не вошли в дерево, поскольку с предыдущими ветвями они образовывают контуры. Матрица основных сечений D предстанет в виде:

D = (6)

1	2	3	5	6	8	4	7	9	10	11	12	
1						1				1	1	1 –сечение
	1						-1	1		1	1	2 –сечение
		1				1	1	-1				3 –сечение
			1						1	-1	-1	5 –сечение
				1					-1			6 –сечение
					1			1		1	1	8 –сечение

Для дерева T_D = {1, 2, 3, 5, 6, 8} структура матрицы собственных и взаимных проводимостей дерева ($D\,G\,D^T$) получится в приводимом ниже виде. Здесь рядом с матрицей обозначены не номера строк и столбцов, а ассоциированные с ними номера ветвей дерева — сечений. Элементы матрицы также лишь обозначены своими номерами по следующему образцу: к примеру, для верхнего левого элемента вместо выражения (G_1+G_2+G_3) записан лишь перечень их нижних индексов (1, 2, 3). Кроме того, следует специально отметить, что отдельные члены внедиагональных элементов могут иметь как положительный, так и отрицательный знак, однако в данном контексте это несущественно. Пустым клеткам приведенных матриц соответствуют нулевые элементы.

$D\,G\,D^T$ = (7)

1	2	3	5	6	8	
1, 4,11,12	11,12	4	11,12		11,12	1
11,12	2, 7, 9, 11,12	7, 9	11,12		9,11,12	2
4	7, 9	3,4,7,9			9	3
11,12	11,12		5, 10, 11,12	10	11,12	5
			10	6, 10		6
11,12	9,11,12	9	11,12		8, 9, 11, 12	8

Как видно, в матрице (7) проводимости – параметры ветвей дерева входят лишь в диагональные элементы матрицы, причем однажды.

Продолжая анализ примера, допустим теперь, что ветви ЭЦ упорядочены по убыванию интервальных характеристик их напряжений: {1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12}. Это означает, что предпочтительным было бы, чтобы напряжения с наибольшими интервальными характеристиками входили в систему уравнений (2) по одному разу, образовывая ассоциированные с ними контуры, а, следовательно, дерево ЭЦ T_B следует строить, преимущественно используя ветви с максимальными номерами. Итак, дерево T_B примет вид: T_B ={3, 6, 8, 9, 10, 12}, а дополнение (множество хорд) образует независимые контуры, ассоциированные с номерами ветвей: {1, 2, 4, 5, 7, 11}.

B = (8)

1	2	4	5	7	11	3	6	8	9	10	12	
1						1	-1		1	-1	-1	1 –контур
	1					-1		1	-1			2 – контур
		1					-1		1	-1	-1	4 – контур
			1				-1			-1		5 – контур
				1				-1	1			7 – контур
					1						-1	11 – контур

Для рассмотренного дерева T_D структура матрицы собственных и взаимных контурных сопротивлений ($B\ R\ B^T$) получится в виде:

$B\ R\ B^T =$ (9)

1	2	4	5	7	11	
1,3,6,9,10,12	3,9	6,9,10,12	6,10	9	12	1–контур
3,9	2,3,9,8	9		8,9		2–контур
6,9,10,12	9	4,6,9,10,12	6,10	9	12	4–контур
6,10		6,10	5,6,10			5–контур
9	8,9	9		7,8,9		7–контур
12		12			11,12	11–контур

В заключение отметим, что если решение задачи анализа интервальной ЭЦ производится итерационным методом, то на различных шагах итераций интервальные характеристики элементов ветвей могут изменяться, коррелируя с текущими интервальными оценками напряжений, токов (и, быть может, пассивных параметров ветвей, если стоит задача их уточнения). Здесь, следовательно, напрашивается возможность итерационного изменения структуры расчетных уравнений ЭЦ, основанных на соответствующем выборе матриц контуров и сечений графа ЭЦ.

Список литературы

1. Киншт Н.В., Кац М.А. Интервальный анализ в задачах теории электрических цепей // *Электричество.* – 1999. – № 10. – С. 45–57.

Программный комплекс для графического представления процесса и результатов работы интервальных алгоритмов

Н. В. Панов, В.В. Колдаков

Новосибирский Центр Информационных Технологий "УниПро",
Россия, 630090 г. Новосибирск, просп. Лаврентьева, 6/1
rupo@ngs.ru kold@nbsp.nsk.su

Аннотация. Предметом нашей работы является наглядное представления процесса выполнения и результатов работы различных алгоритмов глобальной интервальной оптимизации функции двух переменных.

1 Постановка задачи

Данная работа посвящена разработке и реализации пакета программ, позволяющих ставить, решать задачу глобальной оптимизации функции двух переменных интервальными методами и наглядно демонстрировать процесс и результаты работы различных алгоритмов.

Реализованный программный комплекс основан на клиент-серверной архитектуре. Включает в себя мультиплатформенную клиентскую часть со встроенным собственным трехмерным ядром и серверную часть, объединяющую собственную программу-сервер и решатель, в свою очередь состоящий из интервального ядра, библиотеки алгоритмов и символьного интерпретатора.

Эта система призвана способствовать популяризации интервальных методов, в частности, методов глобальной оптимизации и создавалась с рекламно-демонстрационной целью, но может рассматриваться и как полноценный инструмент для решения двумерных задач глобальной оптимизации.

2 Специфика задачи

Математических пакетов с поддержкой интервальной математики все еще мало, общедоступные не очень удобны, имеется ряд существенных недостатков. При проектировании нашей системы, мы старались избежать основных недостатков существующих решений, мешающих программам такого рода завоевать массового пользователя:

1) “ресурсоемкость” (требуют от пользователя большой мощности процессора, много места на диске и оперативной памяти)
2) платформеннозависимость.

Использование клиент-серверной архитектуры позволяет разгрузить клиентскую машину, “унеся” все тяжелые вычисления на серверную сторону. Это не только практически снимает ограничение на мощность пользовательской машины, но и позволяет повысить скорость и точность результатов, за счет использования мощных серверных станций. Выбор Java в качестве языка реализации визуализационной части гарантирует мультиплатформенность. Математическая часть выполнена на языке C++, более подходящего для этих целей.

Одно из основных условий - система должна накладывать как можно меньше ограничений на потенциального пользователя и его рабочее место.

Вся работа осуществляется через сеть Internet при помощи любого Веб-браузера с поддержкой языка Java (Internet Explorer, Netscape Navigator, Mozilla, Opera), но без дополнительных модулей (например, для поддержки трехмерной графики).

Так как системе предстоит работать с трехмерной графикой, но, возможно, у пользователя не установлена система DirectX или OpenGL, а требовать от пользователя установить их мы не можем, встала необходимость создания собственного трехмерного ядра для клиентской части. Оно реализовано в виде внутреннего класса. В его задачи входит не только отрисовка трехмерных объектов средствами двумерной графики, но и немалая математическая часть:

1) вся рутина по вычислению проекций;
2) пересчет координат при повороте и прочие матричные преобразования;
3) расчет затененных областей;
4) удаления невидимых граней.

Далее взаимодействие с внутренним трехмерным ядром остальной программой осуществляется через предоставляемый набор стандартных интерфейсов, как если бы она пользовалась услугами, скажем DirectX.

В задачи сервера входит

1) непосредственно отслеживание и поддержание соединений,
2) аутентификация пользователей и разделение прав,
3) перекодировка и передача информации между клиентом и системой решения,
4) запуск решателя и постановка задачи для него.

Особенность серверной части в том, что она должна сочетать в себе модули, написанные на языках Java и C++. Серверная часть реализована в двух вариантах — в виде сервлета (Java Servlet) и в виде самостоятельного сервера. В силу большой специфичности круга задач, решаемых сервером и особенностей архитектуры, программа-сервер создавалась "с нуля", без использования готовых основ.

Математический решатель (Solver) состоит из интервального ядра, библиотеки алгоритмов, модуля обеспечения и символьного анализатора.

Интервальное ядро реализует представление типа данных «интервал» и набор базовых функций. Библиотека функций — расширяемый набор интервальных алгоритмов. Модуль обеспечения — отвечает за ввод-вывод, синхронизацию, целостность системы и взаимодействие с сервером (пользователем).

3 Краткий перечень требований, предъявляемый к создаваемому программному обеспечению

Разрабатываемое программное обеспечение должно позволять пользователю ставить задачу глобальной оптимизации. Поставленная задача должна быть решена за конечное время. Решение должно быть гарантированным, пропуск какого-либо решения недопустим. Создаваемое ПО должно обеспечивать наглядную демонстрацию работы интервальных алгоритмов решения задачи глобальной оптимизации. Результаты работы должны быть также представлены в наглядном виде. Система должна накладывать как можно меньше ограничений на потенциального пользователя и его рабочее место.

4 Описание алгоритмов трехмерной графики — приводится здесь в весьма сокращенном виде, порой, не затрагивая целые классы проблем, без решения которых трехмерное графическое ядро невозможно. Заинтересованного читателя мы отсылаем к работам [1, 3], подробно излагающим данный вопрос.

4.1 Матричные преобразования

В трехмерном пространстве любое движение, согласно теореме Шаля, может быть представлено в виде суперпозиции поворота и параллельного переноса. И именно поэтому нам достаточно знать, как сделать два преобразования — *перенос* и *поворот*.

4.1.1. Перенос точки. Точки будут рассматриваться как вектора с началом в начале координат и концом в собственно точке. Перенос — это сложение вектора-точки и вектора-переноса.

4.1.2. Поворот. Это занятие уже более интересное. Но тоже простое. Рассмотрим для примера поворот точки (x,y,z) относительно оси z. В этом случае z не меняется, а (x,y) меняются так же, как и при 2D повороте относительно начала координат. При этом координаты точки A' — результат поворота A(x,y) на угол можно узнать по формулам поворота в плоскости, для других осей поворота — аналогично.

$$x' = x*\cos(\alpha)-y*\sin(\alpha)$$
$$y' = x*\sin(\alpha)+y*\cos(\alpha)$$
$$z' = z$$

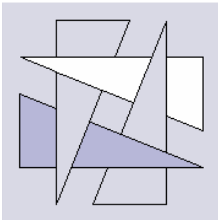
Можно заметить, что $\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) & 0 \\ \sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$

Но умножение матрицы на вектор требует больше операций, чем расчет x' и y' по формулам. Удобство матриц заключается как раз в свойстве A*(B*C) = (A*B)*C. Если мы делаем несколько поворотов подряд, например, пять (столько надо для поворота относительно произвольной оси), то, посчитав один раз матрицу преобразования, являющуюся комбинацией пяти поворотов, можно в дальнейшем без проблем делать сложное преобразование из пяти поворотов с помощью всего одного умножения матрицы на вектор. Таким образом, можно задать любой поворот матрицей, и любая комбинация поворотов также будет задаваться матрицей, которую легко посчитать.

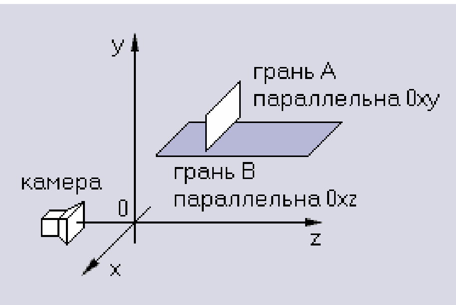
4.2 Удаление невидимых граней

4.2.1. Отброс нелицевых граней. Пусть у нас есть объект, внутри которого камера заведомо не окажется. Обычно такие объекты составляют большую часть или всю сцену. Тогда для каждой грани мы можем увидеть только одну ее сторону — лицевую, ее видно только из одного полупространства, из того, в которое "смотрит" нормаль к этой грани, направленная "из" объекта. Проверив, в какое полупространство попадает камера, можно сразу определить, является ли грань лицевой и надо ли ее рисовать.

Отдельный вопрос - как считать нормали к граням. Точнее, как выбрать одну из двух нормалей, которая будет смотреть из объекта. Обычно эта проблема решается еще на этапе построения 3D моделей - например, пакет 3D Studio записывает вершины граней в порядке A-B-C так, чтобы векторное произведение BA*CA и было нормалью. Еще один способ - выбрать внутреннюю точку для объекта (либо вручную, либо взять его центр тяжести, либо еще как-нибудь) и использовать ее: если для этой точки функция видимости положительна, то есть грань якобы видна, то надо поменять знак nx, ny и nz. Для выпуклых объектов этот метод полностью решает задачу об удалении невидимых частей.



4.2.2 Алгоритм художника. Пусть имеется некий набор граней (т.е. сцена), который требуется нарисовать. Отсортируем грани по удаленности от камеры и отрисуем все грани, начиная с самых удаленных. Довольно распространенная характеристика удаленности для грани ABC - это среднее значение z, mid_z = (A.z+B.z+C.z)/3. Вот и весь алгоритм. Просто, и обычно достаточно быстро. Существует, правда, несколько проблем. Во-первых, при некотором расположении граней этот алгоритм вообще не может дать правильного результата - в каком порядке грани не рисуй, получится неправильно. Стандартный пример справа. Во-вторых, при некотором расположении граней и использовании среднего значения z как характеристики удаленности алгоритм тоже дает неправильный результат. Пример слева. В этом случае горизонтальную грань надо рисовать второй, но по среднему значению z она лежит дальше и таким образом получается, что ее надо отрисовывать первой. Возможные пути решения этой проблемы - какие-то изменения характеристики удаленности, либо моделирование, не вызывающее таких ситуаций. И наконец, при использовании этого алгоритма отрисовываются вообще все грани сцены, и при большом количестве загораживающих друг друга граней мы будем тратить большую часть времени на рисование невидимых в конечном итоге частей. То есть совершенно впустую.



4.2.3. Z-буфер. Заведем буфер (собственно z-буфер) размером с экран, и забьем его каким-то большим числом. Для каждой рисуемой точки считаем значение z; если оно больше, чем значение в z-буфере (точка закрыта какой-то другой точкой), или меньше, чем -dist (точка находится за камерой), то переходим к следующей точке. Если меньше, то рисуем точку, а в z-буфер записываем текущее значение z. Вот и все. Это самый простой метод удаления невидимых частей, причем всегда дающий полностью правильные результаты.

5 Интервальные алгоритмы глобальной оптимизации функций нескольких переменных

Для решения задачи глобальной оптимизации применяются интервальные методы, основанные на адаптивном дроблении области определения в сочетании с оцениванием области значений по получающимся подобластям. Методы этого типа хорошо работают для функций сложного рельефа, надежно находя гарантированные оценки как для глобального оптимума, так и для доставляющих его значений аргументов.

5.1 Бисекция

Классический метод, на каждой итерации выбирает брус, доставляющий наихудшую оценку и делит его пополам.

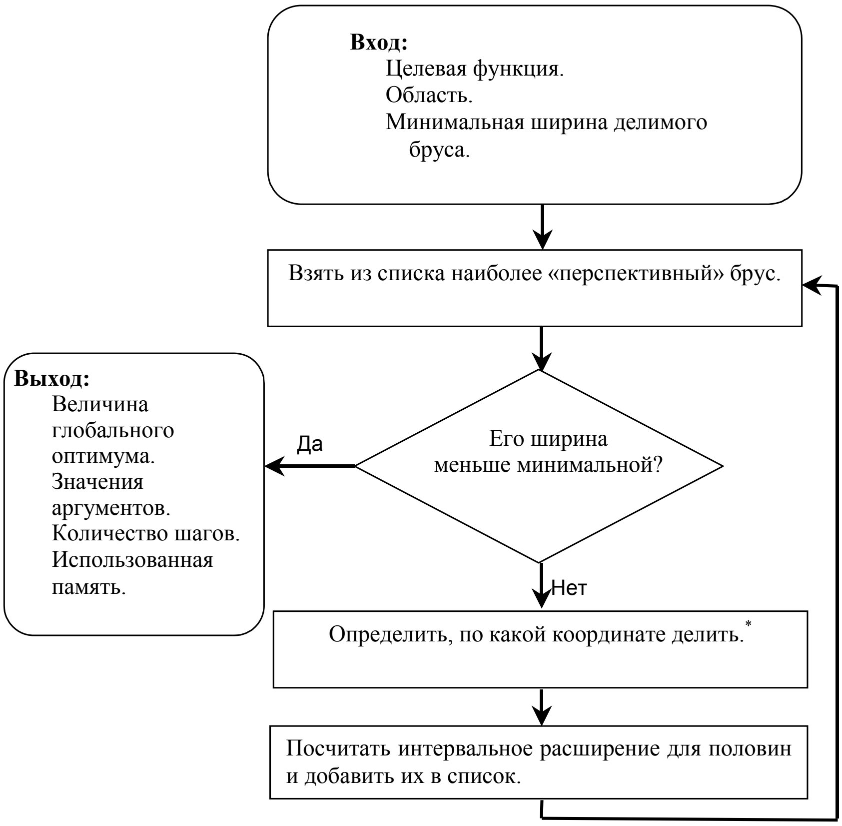
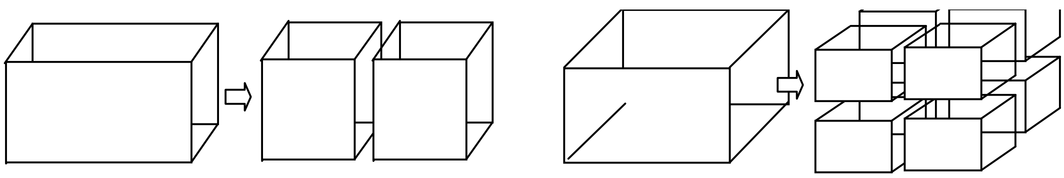


Рис 1. Блок-схема алгоритма глобальной оптимизации с использованием «Бисекции»

*) Деление может осуществляться несколькими способами.

– Естественное дробление (Дробление большой стороны).
  За одну итерацию алгоритма выполняется одно деление. Для деления выбирается более широкая сторона бруса. Это связано с тем, что точность интервальной оценки зависит от ширины интервала, в котором меняются аргументы, и от зависимости целевой функции по каждому из аргументов. Не делая никаких предположений о конфигурации функции, алгоритм последовательно уточняет интервальное расширение, деля наибольшую, а значит, вероятно, дающую наименее точную оценку сторону. Например, брус такого вида (рис. 2а) будет поделен таким (рис. 2б) образом.



а) б) в) г)

Рис 2. а), б) — деление по большей стороне. в), г) — тотальное деление

– Тотальное дробление (Дробление всех сторон).
  За одну итерацию алгоритма брус делится на восемь частей. Пример дробления по всем сторонам на рис. 3

– Дробление с памятью (Эвристическое дробление).

За одну итерацию алгоритма выполняется одно деление. Сторона для деления выбирается исходя из предыдущих итераций, используется вместе с алгоритмом естественного деления (деление большой стороны). Каждый раз, производя деление очередного бруса, этот алгоритм анализирует, насколько улучшилась интервальная оценка при делении именно по этой координате. В существующем варианте надежд не оправдал.

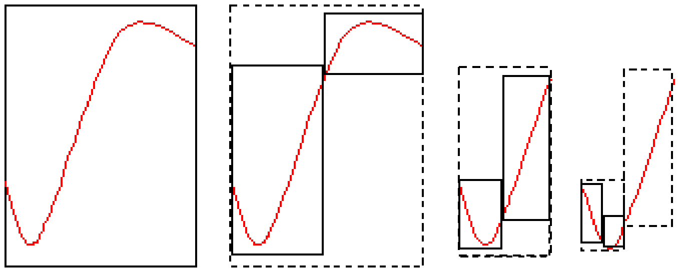


Рис 3. Поиск минимума методом «бисекции»

Вот пример, иллюстрирующий работу бисекции. Ищем минимум.

Бисекция — классический алгоритм, его достоинства — простота реализации и неприхотливость. В рамках данной работы был разработан ряд алгоритмов, также основанных на адаптивном дроблении.

5.2 Бисекция с перекрытием (Трисекция)

Алгоритм, основанный на бисекции. Деление происходит не пополам, а на две пересекающиеся части, по две третьих исходного бруса каждая.

Если при сравнении интервальных расширений получившихся брусов, рекордную оценку доставляет левый брус, делается предположение, что экстремум находится в первой трети, если правый брус рекорден, то наиболее вероятное положение экстремума для этого этапа — в последней трети. Таким образом, для каждой следующий итерации алгоритма будет получаться область в три раза меньше текущей. Но случай, когда оценки равновелики (Рис.3 *б*), требует отдельного рассмотрения. В этом случае необходимо проверить, не имеем ли мы дело с двумя равновеликими экстремумами в первой и третьей третях. Если это действительно так, алгоритм прекрасно распараллеливается, и каждый экстремум в дальнейшем обрабатывается в отдельном потоке. Если это не так, алгоритм предполагает, что экстремум находится во второй трети.

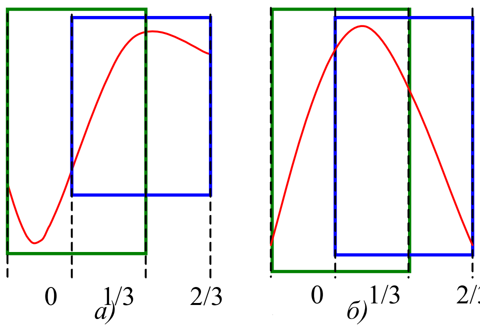


Рис 4. Пример разбиения исходной области перекрытием. (Метод «трисекции»)

Применение этого алгоритма для относительно хороших функций с единственным экстремумом на области дает выигрыш в полтора раза.

5.3 Алгоритмы с неравномерным дроблением

Эти алгоритмы отличаются от описанных выше тем, что деление очередного бруса происходит на неравные части, в зависимости от предыдущего дробления.

На рис.7 рассмотрено поведение подобных методов на основе алгоритма Неравномерной бисекции.

5.4 Вероятностный алгоритм

Другим популярным подходом к решению задач глобальной оптимизации является так называемый алгоритм *«симулированного отжига»* (известный еще как *алгоритм Метрополиса*). Это вероятностный метод, моделирующий одноименный физический процесс, и его имеет смысл применять в тех ситуациях, где доминирующими требованиями является желание получить в сложной задаче хоть какие-то результаты за заданное

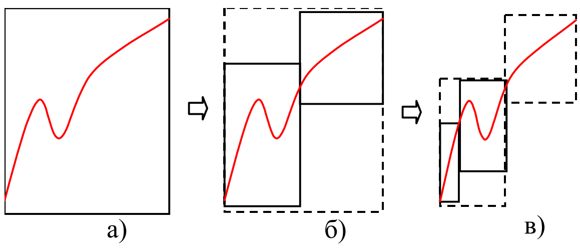


Рис 5. Пример работы алгоритма «Неравномерная бисекция»

ограниченное время. Разработанная С.П.Шарым и впервые реализованная в рамках данной работы интервальная версия алгоритма «симулированного отжига» для глобальной оптимизации функций сочетает в себе гарантированность интервальных методов с неприхотливостью вероятностного подхода.

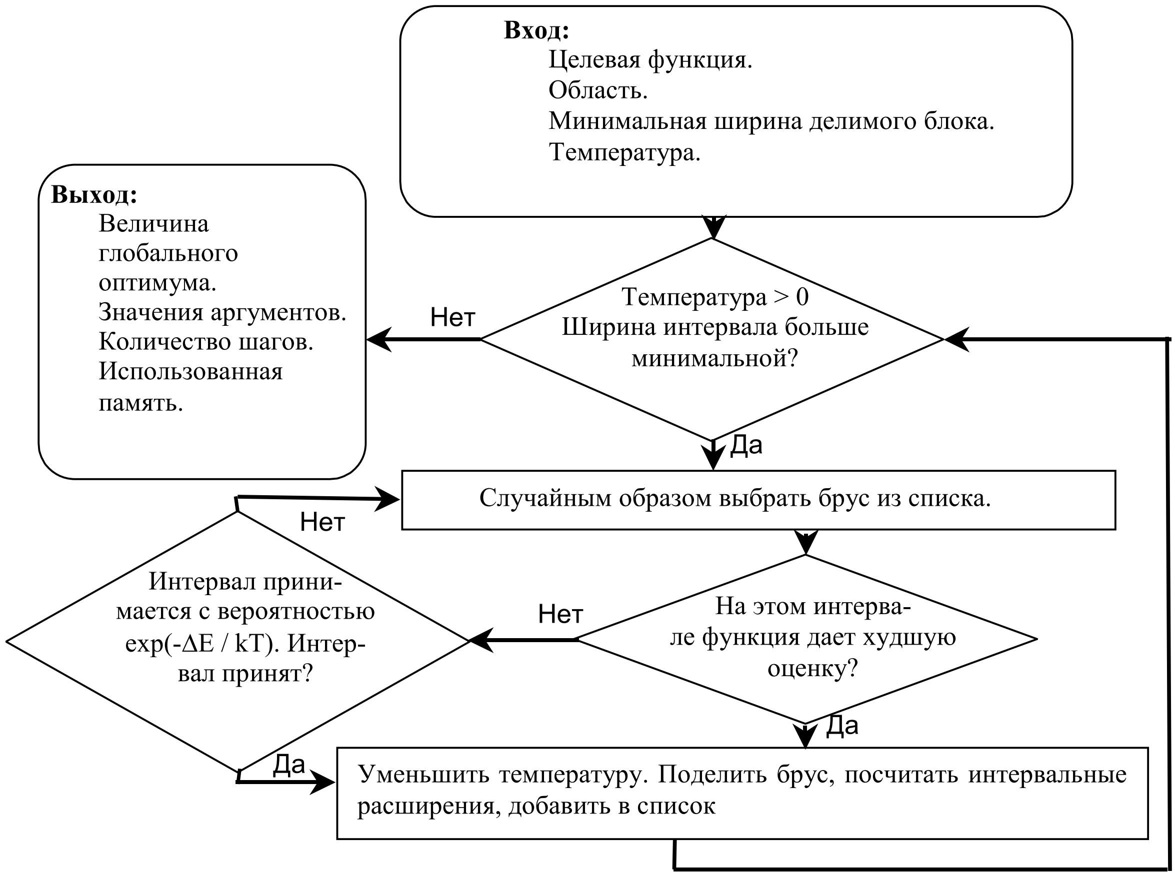


Рис 6. Блок-схема алгоритма интервального симулированного отжига для глобальной оптимизации.

6 Реализация

6.1 Решатель

Решатель реализован на языке С++. Платформеннозависим. Существующая реализация позволяет осуществлять переход между операционными системами Windows (поддерживается вся линейка, начиная с Windows 95) и клонами Linux без изменения кода. Поддержки графического интерфейса пользователя нет. Самодостаточен.

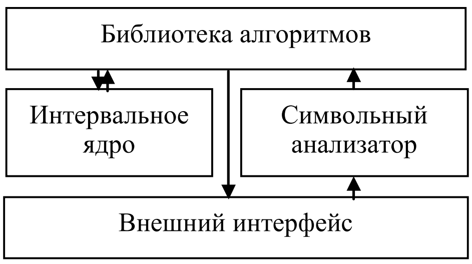


Рис 7. Структурная схема решателя.

Интервальное ядро реализует представление типа интервал и базовые функции. Может выступать в роли самостоятельного продукта. В библиотеке алгоритмов реализованны функции высокого уровня, алгоритмы глобальной оптимизации, методы построения графиков функций (поверхностей). Символьный анализатор позволяет ставить задачу, используя командную строку или файл инициализации. Позволяет определять константы, функции, переменные, понимает данные интервального типа, допускается ввод дробных чисел и чисел в экспоненциальном формате.

6.2 Сервер

Сервер реализован на языке Java в двух вариантах. В виде java программы и с использованием технологии Servlet.

Сервлет не является самостоятельной программой и должен исполняться в рамках какого-либо HTTP-сервера, поддерживающего работу сервлетов. Кроме того, он немного тяжелее в исполнении, хотя при повторных обращениях ситуацию несколько улучшает кэш сервера, пытаясь при этом подсунуть старые данные. При обращении к сервлету вызывается функция doGet. В ней, в зависимости от строки запроса либо ставится задача решателю, либо вызывается соответствующая native-функция для получения и обработки требуемых данных. Затем они посылаются клиенту, либо ему сообщается, что данные не готовы.

Такую же функциональность реализует TCP-IP сервер, для его работы не требуются какие-либо дополнительные программы. Поддерживает множество одновременных соединений.

Сервер и Решатель общаются через Посредника. Посредник кодирует и форматирует данные, решает проблемы разделения доступа и требований безопасности Java машины. Посредник реализован в виде динамически подключаемой библиотеки на языке C++.

6.3 Реализация клиентской части

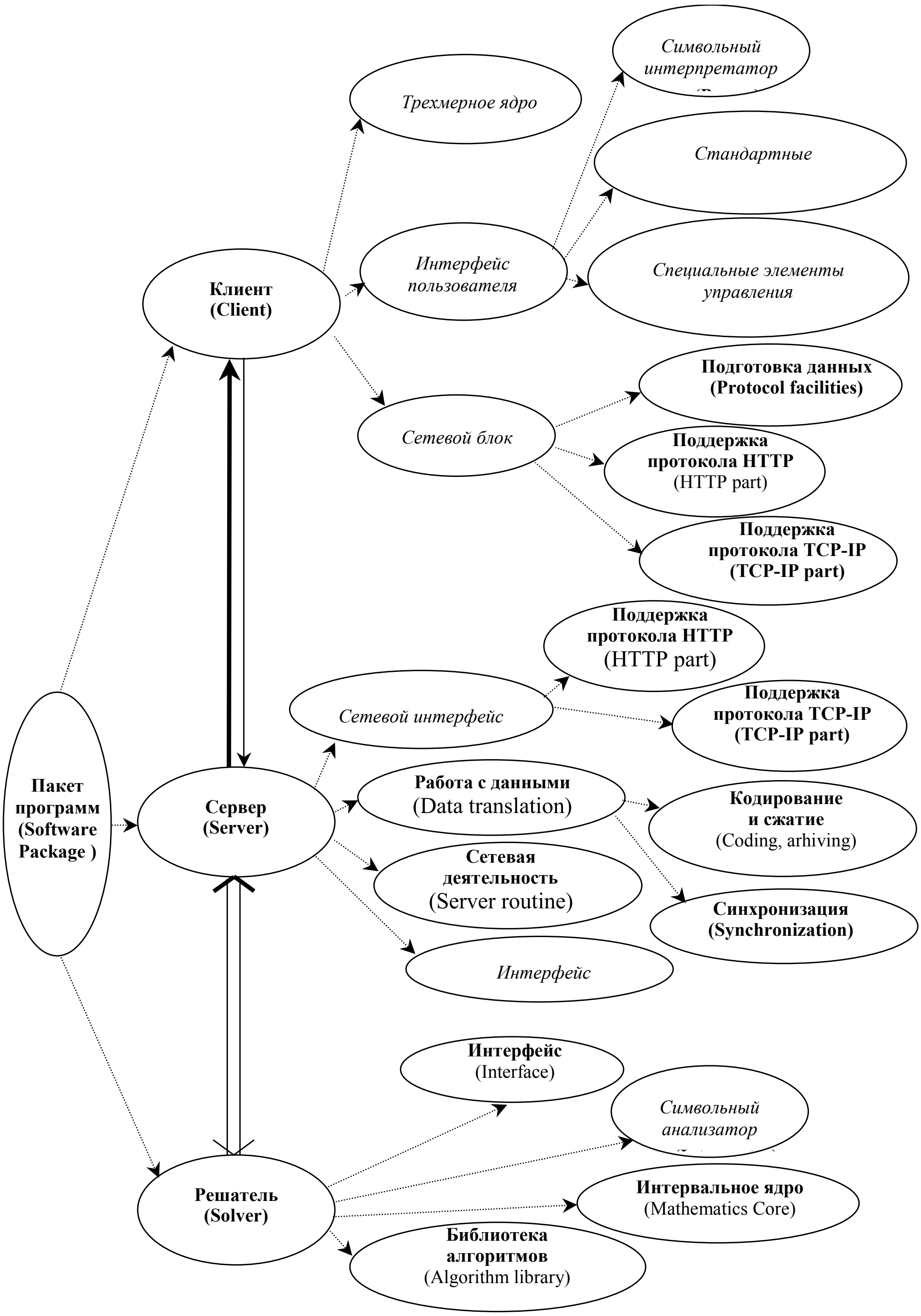
Клиентская часть реализована на языке Java в виде аплета, предоставляющего графический интерфейс пользователя со встроенным трехмерным ядром, собственной системой меню, диалогов и командной строкой.

Список литературы

[1] Калмыков С.А., Шокин Ю.И., Юлдашев З.Х. *Методы интервального анализа*. — Новосибирск, Наука, 1986.
[2] Фоли Дж., Дэм А.. *Машинная графика.* — 2т., Москва, Мир, 1988.
[3] Шарый С.П. *Интервальные алгебраические задачи и их численное решение.* — Диссертация на соискание ученой степени доктора физико-математических наук, Новосибирск, ИВТ СО РАН, 2000 г. — 327 с. Текст в Internet доступен по адресу http://www.ict.nsc.ru/interval/Library/InteDiss
[4] Hansen E.R. *Global optimization using interval analysis*. — Marcel Dekker, 1992.
[5] Kearfott R.B. *Rigorous Global Search: Continuous Problems.* — Dordrecht, Kluwer, 1996.
[6] http://www.enlight.ru/faq3d/main.html — *demo.design 3D. Programming FAQ.*

Приложение

Структура комплекса



О спектральной задаче Орра-Зомерфельда

Проскурин А.В., Сагалаков А.М.

Алтайский Государственный Университет
Россия, г. Барнаул
proskurin.574@phys.dcn-asu.ru

Аннотация. В работе рассмотрена проблема автоматизации трудоёмких вычислений спектров в задачах устойчивости параллельных течений. Использовался принцип аргумента. Кроме высокой надёжности вычислений, этот метод даёт дополнительную информацию о важных свойствах задачи.

Проблема устойчивости параллельных течений по-прежнему представляет значительный интерес. В простейшем случае течения в плоском канале задача сводится к уравнению Орра-Зоммерфельда

$$\omega^{(4)} - 2\alpha^2\omega'' + \alpha^4\omega = i\alpha\,\mathrm{Re}\left[(U - C)\left(\omega'' - \alpha^2\omega\right) - U''\omega\right],$$

(1)

$$\omega = \omega' = 0 \quad \text{при} \quad y = \pm 1,$$

где α, Re — параметры, ω — комплексная амплитуда малого возмущения, $U = 1 - y^2$ — невозмущённый профиль скорости, $C = X + iY$ — спектральный параметр, определяемый при решении задачи на собственные значения (1).

В связи с бурным ростом быстодействия компьютеров требования к вычислительным алгоритмам несколько изменились. Сейчас важным свойством численного метода является его надёжность, которая открывает широкие возможности для автоматизации.

Уравнение (1) решается методом прогонки, после чего формулируется дисперсионное соотношение. Интересным способом его исследования является так называемый принцип аргумента

(2)
$$N - P = \frac{1}{2\pi}\Delta_\gamma\left(\arg z\right),$$

где N и P — число нулей и полюсов в области, ограниченной контуром γ, с учётом их кратности.

В настоящее время проводится испытание этого метода на более трудной задаче исследования устойчивости течений проводящей жидкости в продольном магнитном поле.

Список литературы

[1] Гольдштик М.А., Штерн В.Н. *Гидродинамическая устойчивость и турбулентность.* – Новосибирск: Наука, 1977.

Интервальная техника в задаче параметрического оценивания

С.А. Назин, Б.Т. Поляк*

Институт проблем управления РАН
Москва 117997, ул. Профсоюзная, 65
{snazin,boris}@ipu.rssi.ru

Аннотация Рассматривается задача параметрического оценивания для линейной модели при наличии интервальной неопределенности. Строится алгоритм оптимального оценивания на основе интервального решения линейной интервальной системы алгебраических уравнений.

1 Введение и постановка задачи

Детерминированный подход к описанию неопределенностей в динамических системах хорошо известен в теории управления, где ошибки в модели предполагаются неизвестными, но ограниченными [1]. Интервальные ограничения при этом являются наиболее естественными. В рамках интервального анализа рассматривается множество различных задач с таким типом неопределенности [2,3,4]. Однако интервальные ограничения часто приводят к серьезным трудностям при решении, в частности, различных задач оценивания. Тем не менее, дальнейший анализ и развитие интервальной техники является здесь весьма перспективным.

В данной работе рассматривается проблема нахождения интервального решения линейной интервальной системы алгебраических уравнений применительно к задаче параметрического оценивания в линейных динамических системах при наличии интервальной неопределенности. Пусть задана линейная модель объекта

$$y = (A + \Delta A)\, x + w, \tag{1}$$

где $x \in \mathbb{R}^n$ – вектор параметров объекта, $y \in \mathbb{R}^m$ – известный вектор наблюдений, $A \in \mathbb{R}^{m \times n}$ – номинальная матрица системы, ΔA и w – интервальная матричная неопределенность в описании модели и вектор интервальных ошибок измерений соответственно. Будем полагать, что эти возмущения в системе удовлетворяют простым интервальным ограничениям:

$$\|\Delta A\|_\infty \le \varepsilon, \quad \|w\|_\infty \le \delta. \tag{2}$$

Здесь $\|.\|_\infty$ — ∞-норма матрицы или вектора, равная максимальному модулю значения элементов. Тогда все вектора $x \in \mathbb{R}^n$, удовлетворяющие (1) при данных ограничениях на неопределенность, образуют множество

$$X_m = \{\, x \mid \;\; y = (A + \Delta A)\, x + w,\; \|\Delta A\|_\infty \le \varepsilon,\; \|w\|_\infty \le \delta \,\}\,. \tag{3}$$

Индекс m означает количество измерений. Ставится задача оценить вектор параметров объекта по наблюдениям, то есть найти минимальный интервальный вектор $\mathbf{X} = \{(\mathbf{x}_1, ..., \mathbf{x}_n)^T \mid \forall i = 1, ..., n \;\underline{x_i} \le x_i \le \overline{x_i}\,\}$, содержащий X_m. Можно предполагать также, что априорно $x \in \mathbf{X}_0$, и по полученным измерениям пытаться уточнить эту начальную оценку.

Заметим, что в задачах оценивания число параметров обычно невелико ($n < 15$), в то время как количество измерений может быть большим ($m \gg n$).

* Работа выполнялась при частичной финансовой поддержке РФФИ (грант RFBR–02-01-00127) и INTAS (грант YSF 2002-181).

2 Случай скалярных измерений

Рассмотрим случай наличия одного измерения в системе $y = (a + \Delta a, x) + w$. Тогда множество X_m при $m = 1$ перепишется в явном виде

$$X_1 = \{x \mid \ |(a,x) - y| \le \varepsilon \|x\|_1 + \delta\}\,, \tag{4}$$

где $\|x\|_1 = \sum |x_i|$. Это множество представляет собой "расширенную" полосу в $\mathbb{R}^n$, которая в каждом ортанте есть выпуклое множество. Пусть также априорно известно, что $x \in \mathbf{X}_0$. В этом случае $x \in X_1 \bigcap \mathbf{X}_0$, и задача состоит в нахождении минимального $\mathbf{X}$, содержащего это пересечение.

Решение может быть получено путем линейного программирования для каждого ортанта. Действительно, фиксируя j-й ортант E^j в $\mathbb{R}^n$, мы однозначно определим вектор $\nu = \operatorname{sign} x$; его компоненты ν_i равны ± 1. Поэтому множество

$$X_1 \bigcap E^j = \left\{x \mid |(a,x) - y| \le \varepsilon \sum x_i \nu_i + \delta \right\}$$

является выпуклым. Если e_k — k-ый орт в $\mathbb{R}^n$, то решая задачу минимизации/максимизации линейной функции при линейных ограничениях,

$$\underline{x_k^j} = \arg \min_{x \in \mathbf{X}_0 \bigcap X_1 \bigcap E^j} (e_k, x), \qquad \overline{x_k^j} = \arg \max_{x \in \mathbf{X}_0 \bigcap X_1 \bigcap E^j} (e_k, x), \tag{5}$$

получим k-ю нижнюю/верхнюю интервальную границу. Таким образом, для каждого из 2^n ортантов отдельно находится интервальная оценка $\mathbf{X}^j$, $j = 1, ..., 2^n$, содержащая $\mathbf{X}_0 \bigcap X_1 \bigcap E^j$. Окончательно, интервальный вектор $\mathbf{X}$, содержащий объединение $\mathbf{X}^j$, находится тривиально путем выбора соответствующих минимальных и максимальных компонент интервальных векторов.

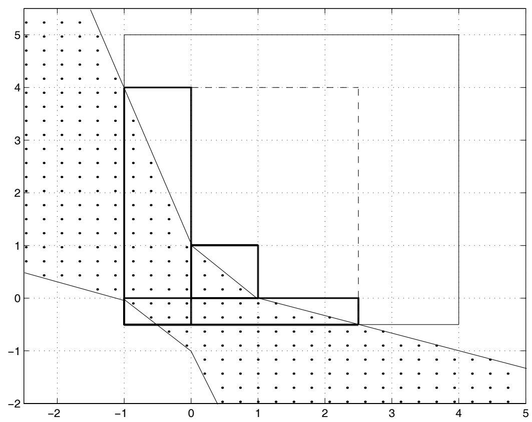


Рис. 1. Иллюстрация к Примеру 1.

При наличии нескольких измерений можно последовательно рассматривать скалярные уравнения в (1). Однако описанный алгоритм довольно громоздкий с вычислительной точки зрения, и его рекуррентное применение целесообразно только при малом количестве измерений. В противном случае необходимо использовать более экономичные способы, например, вычисляя интервальное решение линейной интервальной системы алгебраических уравнений.

Пример 1: Пусть

$$\mathbf{X}_0 = \begin{pmatrix} [-1,\ 4] \\ [-0.5,\ 5] \end{pmatrix},$$

а также $y = 0$, $A = (1, 1)$, $\varepsilon = \delta = 0.5$. Множество X_1 показано на рисунке штриховкой. Для каждого ортанта найдены интервальные оценки $\mathbf{X}^j$ (жирная линия). Тогда

$$\mathbf{X} = \begin{pmatrix} [-1,\ 2.5] \\ [-0.5,\ 4] \end{pmatrix}.$$

3 Решение линейной интервальной системы

Очевидным образом уравнение (1) при ограничениях (2) можно переписать в виде линейной интервальной системы

$$(A + \varDelta A)\, x = b + \varDelta b, \tag{6}$$

в которой $b = y$, а $\varDelta b = -w$ ($\|\varDelta b\|_\infty \le \delta$), и к решению которой свелась исходная задача. На данном этапе положим $m = n$, то есть число уравнений в (6) равно числу неизвестных. Пусть также матрица A является невырожденной.

Множество X всех допустимых векторов x из (6) называется множеством решений интервальной системы. Его детальное описание было приведено в [5]. Множество решений интервальной системы есть многогранник (в общем случае невыпуклый)

$$X = \{\, x \mid \ \ \|Ax - b\|_\infty \le \varepsilon \|x\|_1 + \delta \,\}\,. \tag{7}$$

Он является ограниченным, если выполнено условие регулярности (см. [6]) $\varepsilon < \vartheta(A) = 1/\|A^{-1}\|_{\infty,1}$, то есть интервальные матричные возмущения в системе (6) меньше, чем радиус невырожденности для интервальной матричной неопределенности. Здесь матричная норма

$$\|A\|_{\infty,1} = \max_{\|x\|_\infty \le 1} \sum |a_i^T x|,$$

где a_i — i-я строка матрицы A.

Интервальным же решением системы (6) будет любой интервальный вектор $\mathbf{X}_n \supseteq X_n$. Будем интересоваться только наименьшим интервальным решением $\mathbf{X}_n$. Однако его поиск является NP-сложной задачей (см. [7]). Но при этом удается обойтись без численного решения задач линейного программирования, а решать каждый раз только одно скалярное уравнение (см. [8,9]), что существенно уменьшает вычислительную трудность алгоритма. Приведем далее из [9] описание алгоритма нахождения интервального решения системы (6).

Обозначим $S = \{\ s \mid \forall i\ |s_i| = 1\ \}$. Рассмотрим скалярное уравнение относительно τ

$$\tau = \varepsilon \|A^{-1}(\tau s + b)\|_1 + \delta, \tag{8}$$

где вектор $s \in S$. Если $\varepsilon < \vartheta(A)$, то для фиксированного $s \in S$ это уравнение имеет единственное решение на интервале $(0, +\infty)$.

Лемма 1 (см. [9]): Пусть множество X_n ограничено ($\varepsilon < \vartheta(A)$) и Conv X_n — его выпуклая оболочка. Тогда Conv X_n имеет 2^n вершин, каждая из которых может быть найдена решением уравнения для своего фиксированного вектора $s \in S$. Если τ^* — решение (8), то

$$x^* = A^{-1}(\tau^* s + b)$$

будет вершиной Conv X_n.

Располагая всем набором таких вершин, очевидным образом можно найти интервальное решение $\mathbf{X}_n$ системы (6) посредством минимизации и максимизации соответствующих векторных компонент.

4 Алгоритм

В случае большого числа уравнений ($m > n$) в исходной модели (1) для нахождения интервальной оценки предлагается применить следующую схему. Полагаем изначально, что все уравнения в невозмущенной системе (то есть в системе $y = Ax$) линейно независимы. Пусть известно также, что $x \in \mathbf{X}_0$. Тогда

- Берем первые n уравнений, и находим интервальный вектор $\mathbf{X}^1$, который содержит $\mathbf{X}_0 \bigcap X_n^1$, используя алгоритм пункта 3.
- Если $m \ge 2n$, то берем следующие n уравнений, и находим $\mathbf{X}^2 \supseteq \mathbf{X}^1 \bigcap X_n^2$, и т.д.
- Если на k-м шаге получилось $m < kn$, то оставшиеся уравнения учитываем каждое по-отдельности по алгоритму пункта 2.

В итоге очевидно, что будет найдена оптимальная интервальная оценка вектора параметров x исходной системы (1).

5 Заключение

В работе рассмотрена задача параметрического оценивания при наличии неопределенности в описании объекта, когда ошибки стеснены интервальными ограничениями. Приведен алгоритм нахождения минимальной интервальной оценки для вектора параметров $x \in \mathbb{R}^n$ системы, основанный на получении интервального решения линейной интервальной системы алгебраических уравнений с квадратной матрицей. При этом каждое такое интервальное решение получается путем решения 2^n скалярных уравнений. Был проведен счет, который подтверждает эффективную работу алгоритма при небольших размерностях вектора параметров системы n ($n < 15$) и большом числе измерений ($m \sim 100$).

Список литературы

1. Schweppe F. *Uncertain dynamic systems.* – Englewood Cliffs, N.J.: Prentice Hall, 1973.
2. Калмыков С.А., Шокин Ю.И., Юлдашев З.Х. *Методы интервального анализа.* – Новосибирск: Наука, 1986.
3. Neumaier A. *Interval Methods for Systems of Equations.* – Cambridge: Cambridge University Press, 1990.
4. Шарый С.П. Внешнее оценивание обобщенных множеств решений интервальных линейных систем // *Вычислительные Технологии.* – 1999. – Т. 4, №4. – P. 82 – 110.
5. Oettli W., Prager W. Compatibility of approximate solution of linear equations with given error bounds for coefficients and right-hand sides // *Numerische Mathematik.* – 1964. — Vol. 6. – P. 405 – 409.
6. Polyak B.T. Robust linear algebra and robust aperiodicity // *Directions in Mathematical System Theory and Optimization*, A. Rantzer, C. Byrnes, eds. – Stockholm: Springer, 2003. – P. 249 – 260.
7. Kreinovich V., Lakeev A.V., Noskov S.I. Optimal solution of interval linear systems is intractable (NP-hard) // *Interval Computationa.* – 1993. – №1. – P. 6 – 14.
8. Rohn J. Systems of linear interval equations // *Linear Algebra and Its Applications.* – 1989. – Vol. 126. – P. 39 – 78.
9. Polyak B.T., Nazin S.A. Interval solutions for interval algebraic equations // Proc. 4th IMACS Symp. on Math. Modelling. Vienna, Austria, 2003. – P. 973 – 980.

Точечная аппроксимация множества решений интервальной системы линейных алгебраических уравнений

С.И. Носков

Иркутский государственный университет путей сообщения
Россия, 664074 г. Иркутск, ул. Чернышевского, 15
noskov_s@iriit.irk.ru

Рассмотрим интервальную систему линейных алгебраических уравнений (ИСЛАУ)

$$\mathrm{A}x = \mathrm{B}\,, \qquad (1)$$

где A – интервальная вещественная матрица размерности (n×m), B – n-мерный интервальный вещественный вектор, элементами которых являются соответственно интервалы $[a_{ki}^-, a_{ki}^+]$ и $[b_k^-, b_k^+]$, $k = \overline{1,n},\ i = \overline{1,m},\ x \in R^m$.

Введем естественные обозначения

$$A^- = \left\| a_{ki}^- \right\|,\ A^+ = \left\| a_{ki}^+ \right\|,\ k = \overline{1,n},\ i = \overline{1,m}, \quad b^- = (b_1^-, \ldots, b_n^-)^T,\ b^+ = (b_1^+, \ldots, b_n^+)^T.$$

Множество решений системы (1) может быть определено различными способами в зависимости от того, какими кванторами связаны коэффициенты матрицы и правой части [1]. Наиболее часто в литературе встречаются следующие множества решений ИСЛАУ:

$$\Re_1 = \{x \in R^m \mid \exists C \in \mathrm{A}\ \exists c \in \mathrm{B} \quad Cx = c\},$$

$$\Re_2 = \{x \in R^m \mid \forall C \in \mathrm{A}\ \exists c \in \mathrm{B} \quad Cx = c\},$$

$$\Re_3 = \{x \in R^m \mid \forall c \in \mathrm{B}\ \exists C \in \mathrm{A} \quad Cx = c\},$$

$$\Re_4 = \{x \in R^m \mid (\forall C \in \mathrm{A}\ \exists c \in \mathrm{A} \quad Cx = c)\ \&\ (\forall d \in \mathrm{B}\ \exists D \in \mathrm{A} \quad Dx = d)\}.$$

Известны (см., например, [2]), описания этих множеств в виде систем равенств и неравенств:

$$\Re_1 = \{x^1 - x^2 \mid x^1, x^2 \in R_+^m,\ (x^1, x^2) = 0,\ A^- x^1 - A^+ x^2 \le b^+,\ A^+ x^1 - A^- x^2 \ge b^-\},$$

$$\Re_2 = \{x^1 - x^2 \mid x^1, x^2 \in R_+^m,\ \ A^- x^1 - A^+ x^2 \ge b^-,\ A^+ x^1 - A^- x^2 \le b^+\},$$

$$\Re_3 = \{x^1 - x^2 \mid x^1, x^2 \in R_+^m,\ (x^1, x^2) = 0,\ A^- x^1 - A^+ x^2 \le b^-,\ A^+ x^1 - A^- x^2 \ge b^+\},$$

$$\Re_4 = \{x^1 - x^2 \mid x^1, x^2 \in R_+^m,\ (x^1, x^2) = 0,\ A^- x^1 - A^+ x^2 = b^-,\ A^+ x^1 - A^- x^2 = b^+\}.$$

Таким образом, лишь множество $\Re_2$ задается линейными ограничениями, в описании же множеств $\Re_1, \Re_3, \Re_4$ участвует нелинейное условие $(x^1, x^2) = 0$, что делает любую задачу оперирования ими NP-полной.

В случае, когда требуемое множество $\Re_i,\ i \in \{1,2,3,4\}$ оказывается пустым, вместо него можно использовать соответствующее квазирешение (описания всех таких квазирешений даны в [2]).

В реальных задачах не всегда удобно пользоваться множеством $\Re_i$, например, при идентификации параметров математической модели регрессионного типа с использованием интервальной выборки. Часто вполне достаточно выделения некоторого представителя $\Re_i$, обладающего специальными, следующими из обоснованных содержательных соображений, свойствами.

Таким представителем (назовем его **сильным** $\Re_i$ **- решением**) может быть точка x^*, обеспечивающая не просто выполнение линейных неравенств в описании $\Re_i$, но и гарантирующая максимальную разрешающую способность этих неравенств.

Так, для множества $\Re_2$ задача поиска сильного $\Re_2$ - решения сведется к задаче линейного программирования:

$$A^-x^1 - A^+x^2 - u \geq b^-$$
$$A^+x^1 - A^-x^2 + v \leq b^+$$
$$x^1 \geq 0, x^2 \geq 0, u \geq 0, v \geq 0$$
$$\sum_{i=1}^{n}(u_i + v_i) \to \max$$

После ее решения сильное $\Re_2$ -решение находится по правилу $x^* = x^1 - x^2$.

Линейные ограничения задач поиска сильных $\Re_i$ -решений, $i = 1, 3, 4$ дополнятся нелинейным условием $(x^1, x^2) = 0$. Эти задачи могут быть сведены к задачам частично булевого линейного программирования [2]. При этом переход от канонической формы записи ограничений для $\Re_4$ к нормальной можно легко осуществить с помощью стандартных приемов линейного программирования.

Список литературы

1. Ватолин А.А. О задачах линейного программирования с интервальными коэффициентами. // *Журн. вычисл. математ. и матем. физики.* – 1984. – Т. 24, №11. – С.1629–1637.
2. Носков С.И. *Технология моделирования объектов с нестабильным функционированием и неопределенностью в данных.* – Иркутск: Облинформпечать, 1996.

О проблеме реализации в пространстве состояний для интервальных динамических систем

С.Г. Пушков, С.Ю. Кривошапко

Бийский технологический институт
Алтайского государственного технического университета
кафедра информационных и управляющих систем
Россия, 659305 г. Бийск Алтайского края, ул. Трофимова, 27
psg@bti.secna.ru

Аннотация Рассматривается проблема реализации в пространстве состояний для интервальных динамических с дискретным временем. Анализируются возможные формулировки этой проблемы и пути ее решения. Сформулирован и доказан достаточный критерий алгебраической реализуемости импульной последовательности интервальных матриц. Предложен метод нахождения алгебраиеских реализаций для полностью неотрицательных и полностью неположительных систем. Приведен числовой пример.

1 Введение

Модели с пространством состояний играют чрезвычайно важную роль во многих областях применения. Модели такого рода являются естестевенной формой представления для динамических систем теории управления, в особенности теории автоматического управления. Однако метод пространства состояний оказывается применимым и во многих других случаях. Предметом рассмотрения данной работы является проблема построения модели в простанстве состояний на основе данных о поведении вход-выход линейной стационарной динамической системы с дискретным временем. В теории систем эта проблема известна как задача реализации.

Для класса линейных стационарных динамических систем с дискретным временем задача построения конечномерной реализации успешно решается в рамках формального алгебраического подхода к проблеме [1]. Однако при исследовании систем, подверженных различным искажениям, а также систем, фукционирующих в условиях неопределенности, приходится привлекать дополнительный математический аппарат. Чаще всего эти аппаратом являются методы математической статистики. В последнее время для анализа систем с неопределенностями и неоднозначностями в данных все чаще используются новые подходы, например, такие, как теория нечетких множеств и интервальный анализ.

Стандартным методологическим приёмом исследования объектов, функционирующих в условиях неопределённости, является рассмотрение семейства объектов, которое определяется принадлежностью параметров этого объекта некоторым множествам. В случае, когда эти множества являются интервалами, мы будем иметь дело с интервальной неопределённостью.

В настоящее время интервальные методы используются как для анализа статических систем (см., например, [2]), так и для решения проблем управления динамическими системами (см., например, [3-5]). Многие задачи математической теории управления допускают естественную "интервализацию"путем замены вещественных параметров и/или переменных на соответствующие интервальные. Большинство этих интервализованных задач оказываются вполне адекватными и интерпретируемыми с точки зрения практических приложений. Не исключеснием является и рассматриваемая в данной работе проблема реализации.

Далее в данной работе мы будем использовать следующие обозначения:

$\mathbb{R}$ — поле действительных чисел;
$\mathbb{IR}$ — классическая интервальная арифметика;
$\mathbb{KR}$ — полная интервальная арифметика (арифметика Каухера).

2 Классическая проблема реализации

Классическая проблема реализации состоит в определении модели пространства состояний для динамической системы, заданной своим поведением вход-выход. Поведение вход-выход линейной стационарной многомерной управляемой системы может быть охарактеризовано импульсной последовательностью матриц размера $p \times m$ (m — число входов, p — число выходов системы):

$$\{A_1, A_2, \ldots\}. \tag{1}$$

В этом случае для заданной последовательности векторов управлений (входной последовательности) $u(0), u(1), \ldots$ выходная последовательность векторов $y(0), y(1), \ldots$ определяется соотношениями

$$y(t) = \sum_{i=1}^{t} A_i u(t-i) \qquad (t = 1, 2, \ldots).$$

Задача реализации для данного класса систем состоит в определении математической модели этой системы в пространстве состояний, которая описывается разностными уравнениями

$$x(t+1) = Fx(t) + Gu(t),$$

$$y(t) = Hx(t) \quad (t = 0, 1, \ldots),$$

где $x(t)$ и $x(t+1)$ - вектора состояний в моменты времени t и $t+1$, соответственно. Определению подлежат матрицы F, G и H вместе с их размерностями. Хорошо известно [1], что задача реализации в этом случае сводится к нахождению тройки матриц (F, G, H) таких, что

$$A_i = HF^{i-1}G \quad (i = 1, 2, \ldots).$$

Известно также, что для систем над полями (и даже над коммутативными кольцами) реализуемость последовательности (1) эквивалентна ее рекуррентности [6]. Это означает, что если для последовательности (1) существует конечномерная реализация, то существет такое целое $r > 0$ и коэффициенты β_1, β_2, $\ldots$, β_r из поля (или кольца), над которым определена система, такие, что

$$A_{r+j+1} = \sum_{i=1}^{r} \beta_i A_{i+j} \quad (j = 0, 1, 2, \ldots).$$

Существует достаточное количество эффективных вычислительных процедур для построения конечномерных реализаций в случае систем над полями [7, 8] и, в частности, над полем действительных чисел $\mathbb{R}$. Данная проблема может быть обобщена также на случай систем над кольцами [6], и даже для систем с "шумом". В последнем случае речь идет о так называемой проблеме приближенной реализации [9].

3 Интервальные динамические системы

Математическая модель многомерного интервально заданного объекта управления обычно представляется в виде системы уравнений с интервальными параметрами и понимается как семейство математических моделей многомерных динамических объектов, параметры которых принадлежат заданным интервальным. Следуя этому подходу, введем следующее определение.

Определение 1. *Линейной стационарной динамической системой с дискретным временем (с m входами, n состояниями и p выходами) с интервальными параметрами будем называть объект $[\Sigma] = (\mathbf{F}, \mathbf{G}, \mathbf{H})$, динамическое поведение которого описывается уравнениями*

$$x(t+1) = \mathbf{F}x(t) + \mathbf{G}u(t), \quad y(t) = \mathbf{H}x(t), \tag{2}$$

где $u(t) \in \mathbb{R}^m$, $x(t) \in \mathbb{R}^n$, $y(t) \in \mathbb{R}^p$, и понимать как семейство математических моделей

$$x(t+1) = Fx(t) + Gu(t), \quad y(t) = Hx(t),$$

матрицы (F, G, H) которых принадлежат заданным интервальным матрицам $(\mathbf{F}, \mathbf{G}, \mathbf{H})$, т.е. $F \in \mathbf{F} \in \mathbb{IR}^{n\times n}$, $G \in \mathbf{G} \in \mathbb{IR}^{n\times m}$, $H \in \mathbf{H} \in \mathbb{IR}^{p\times n}$.

Следует заметить, что данный подход к определению интервально заданной линейной системы не является единственно возможным, не исчерпывает всего многообразия поведения объектов с интервальной неопределенностью и поэтому не может служить в качестве общего определения линейной системы с интервальной неопределенностью. Условно, называя определенный выше тип линейной стационарной динамической системой с дискретным временем с *интервальными параметрами*, мы далее приведем определения еще двух типов линейных динамических систем с интервальной неопределенностью.

Определение 2. *Линейной стационарной динамической системой с дискретным временем с интервальными объектами будем называть объект* $\mathbf{\Sigma} = (F, G, H)$*, динамическое поведение которого описывается уравнениями*

$$\mathbf{x}(t+1) = F\mathbf{x}(t) + G\mathbf{u}(t), \quad \mathbf{y}(t) = H\mathbf{x}(t), \tag{3}$$

где

$$\mathbf{x}(t),\ \mathbf{x}(t+1) \in \bar{X} \cong \mathbb{IR}^n, \qquad \mathbf{u}(t) \in \bar{U} \cong \mathbb{IR}^m, \qquad \mathbf{y}(t) \in \bar{Y} \cong \mathbb{IR}^p,$$
$$F \in \mathbb{R}^{n \times n},\ G \in \mathbb{R}^{n \times m},\ H \in \mathbb{R}^{p \times n}.$$

К анализу систем такого типа не применимы методы классической реализации, поскольку $\bar{X}$, $\bar{U}$ и $\bar{Y}$ мы не можем рассматривать как $\mathbb{R}$-модули.

Определение 3. *Линейной стационарной динамической системой с дискретным временем с интервальными объектами и интервальными параметрами будем называть объект* $[\mathbf{\Sigma}] = (\mathbf{F}, \mathbf{G}, \mathbf{H})$*, динамическое поведение которого описывается уравнениями*

$$\mathbf{x}(t+1) = \mathbf{F}\mathbf{x}(t) + \mathbf{G}\mathbf{u}(t), \quad \mathbf{y}(t) = \mathbf{H}\mathbf{x}(t), \tag{4}$$

где

$$\mathbf{x}(t), \mathbf{x}(t+1) \in \bar{X} \cong \mathbb{IR}^n, \mathbf{u}(t) \in \bar{U} \cong \mathbb{IR}^m, \quad \mathbf{y}(t) \in \bar{Y} \cong \mathbb{IR}^p,$$
$$\mathbf{F} \in \mathbb{IR}^{n \times n}, \mathbf{G} \in \mathbb{IR}^{n \times m}, \mathbf{H} \in \mathbb{IR}^{p \times n}.$$

Представленные типы интервальных динамических систем являются примерами обобщения обычных линейных стационарных динамических систем с дискретным временем. Эти типы далеко не исчерпывают ни всех возможных типов обобщений, ни видов интервальной неопределенности. Например, имеет право на существование множество типов интервальных динамических систем с интервальной неопределенностью поведения, когда равенства в (2)-(4) заменяются на включения. Можно рассматривать промежуточные типы интервальных динамических систем.

4 Проблема реализации для интервальных систем

С интервальными динамическими системами первого и третьего перечисленных выше типов можно связать импульсную последовательность матриц

$$\mathbf{A}_i = \mathbf{H}\mathbf{F}^{i-1}\mathbf{G}, \quad i = 1, 2, \ldots, \tag{5}$$

где матричные произведения выполняются справо налево, т.е. сначала вычисляется произведение FG, затем $F(FG)$ и т.д. Этой последовательности матриц можно поставить в соответствие отображение вход-выход, под которым в зависимости от того, с системой какого типа мы имеем дело, можно понимать в общем случае разные объекты.

Для системы с интервальными параметрами под отображением вход-выход следует понимать семейство отображений $f_\alpha : U^{\mathbb{Z}_+} \to Y^{\mathbb{Z}_+}$, порождаемых соотношениями

$$y(t) = \sum_{i=1}^{t} A_i^\alpha u(t-i) \qquad (t = 1, 2, \ldots). \tag{6}$$

для

$$A_i^\alpha \in \mathbf{A}_i \in \mathbb{IR}^{p \times m}, \qquad i = 1, 2, \ldots. \tag{7}$$

Для системы с интервальными объектами и интервальными параметрами под отображением вход-вход целесообразно понимать отображение $f : \bar{U}^{\mathbb{Z}_+} \to \bar{Y}^{\mathbb{Z}_+}$, порождаемое интервальными соотношениями

$$\mathbf{y}(t) = \sum_{i=1}^{t} \mathbf{A}_i \mathbf{u}(t-i) \qquad (t = 1, 2, \ldots). \tag{8}$$

Для заданного отображения вход-выход интервальной системы, представленного импульсной последовательностью интервальных матриц можно поставить задачу построения динамического поведения (2) (или (4)), т.е. задачу реализации. Одной из возможных формулировок задачи реализации для интервальных систем может быть такая:

для заданной последовательности интервальных матриц размера $p \times m$

$$\{\mathbf{A}_1, \mathbf{A}_2, \ldots\}, \qquad \mathbf{A}_i \in \mathbb{IR}^{p\times m}, \qquad i = 1, 2, \ldots. \tag{9}$$

определить размерность n и тройку интервальных матриц $(\mathbf{F}, \mathbf{G}, \mathbf{H})$ таких, что выполняются интервальные уравнения (5), где $\mathbf{F} \in \mathbb{IR}^{n\times n}$, $\mathbf{G} \in \mathbb{IR}^{n\times m}$, $\mathbf{H} \in \mathbb{IR}^{p\times n}$.

Данную задачу мы далее будем называть *задачей алгебраической реализации.* Поставленная задача, которая по своей сути является задачей точной реализации, в общем случае является трудно разрешимой. Более того, в силу "плохих"свойств интервальной арифметики $\mathbb{IR}$ она вообще может не иметь решения. Проблема остается таковой и в случае рассмотрения систем над $\mathbb{KR}$ вместо $\mathbb{IR}$. Поэтому в данном случае необходимы другие формулировки задачи реализации, в большей степени отвечающие духу систем с интервальными неопределенностями.

Для заданной последовательности интервальных матриц (9) будем понимать уравнение (5) как семейство троек матриц (F, G, H) над $\mathbb{R}$ таких, что

$$A_i = HF^{i-1}G, \quad i = 1, 2, \ldots. \tag{10}$$

для $A_i \in \mathbf{A_i}$ $(i = 1, 2, \ldots)$. Тогда можно поставить целый ряд задач приближенной реализации, среди которых выделим следующие:

1) определить мощность множества алгебраических реализаций (сколько их с точностью до изоморфизма);
2) определить алгебраические реализации минимальной размерности;
3) для подходящего n (которое также нужно определить) определение внешней оценки для (F, G, H), т.е. определение интервальных матриц $\mathbf{F}$, $\mathbf{G}$, $\mathbf{H}$, содержащих все матрицы F, G, H, удовлетворяющие (10) (при определенном значении n);
4) для подходящего n (которое также нужно определить) определение внутренней оценки для (F, G, H), т.е. определение интервальных матриц $\mathbf{F}$, $\mathbf{G}$, $\mathbf{H}$, содержащих все матрицы F, G, H, которые содержатся во множестве решений уравнений (5) (при определенном значении n).
5) описать структуру множества конечномерных реализаций (или хотя бы найти способ такого описания).

Очевидно, решение поставленных задач неединственно даже для фиксированной размерности системы. Наибольший интерес представляют в определенном смысле оптимальные оценки. Очевидным критерием качества интервальных решений в задачах оценивания является степень близости (в том или ином смысле) полученной интервальной оценки к точному множеству решений. Для задач внешнего оценивания такая оптимальность обычно понимается в смысле минимальности по включению, а для задач внутреннего оценивания - максимальности по включению. Причем, нужно быть готовыми к ситуации, что даже такие оптимальные оценки окажутся неединственными.

Положение еще более усложняется, если ставится задача нахождения реализации минимальной размерности. Здесь мы имеем дело с ситуацией, когда минимизация размерности реализации находится в противоречии с оптимизацией оценок реализации. В связи с этим имеет право на существование целый ряд постановок задач приближенной реализации, являющихся способом разрешения этого противоречия или компромиссного решения.

5 Достаточное условие алгебраической реализуемости

Как мы видели в разделе 1, для классического (неинтервального) случая рекуррентность заданной импульсной последовательности матриц является необходимым и достаточным условием ее реализуемости. Для интервальных систем рекуррентность остается достаточным условием реализумости.

Определение 4. *Будем говорить, что последовательность интервальных матриц (9) интервально рекуррентна, если существет такое целое $r > 0$ и коэффициенты $\beta_1, \beta_2, \ldots, \beta_r \in \mathbb{IR}$ такие, что*

$$\mathbf{A}_{r+j+1} = \sum_{i=1}^{r} \beta_i \mathbf{A}_{i+j} \qquad (j = 0, 1, 2, \ldots). \tag{11}$$

Для интервальных динамических систем, поведение вход-выход которых описывается импульсной последовательностью интервальных матриц, имеет место следующий результат.

Предложение 1. *Если последовательность интервальных матриц интервально рекуррентна, то для нее существует алгебраическая интервальная реализация.*

Доказательство. Пусть исходная последовательность интервальных матриц (9) интервально рекуррентна с соотношением рекуррентности(11). Рассмотрим интервальную систему $(\mathbf{F}, \mathbf{G}, \mathbf{H})$, у которой

$$\mathbf{F} = \begin{pmatrix} \mathbf{O} & \mathbf{I} & \mathbf{O} & \ldots & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & \mathbf{I} & \ldots & \mathbf{O} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \mathbf{O} & \mathbf{O} & \mathbf{O} & \ldots & \mathbf{I} \\ \beta_1 \mathbf{I} & \beta_2 \mathbf{I} & \beta_3 \mathbf{I} & \ldots & \beta_r \mathbf{I} \end{pmatrix}, \mathbf{G} = \begin{pmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \\ \ldots \\ \mathbf{A}_{r-1} \\ \mathbf{A}_r \end{pmatrix}, \mathbf{H} = \begin{pmatrix} \mathbf{I} \, \mathbf{O} \ldots \mathbf{O} \end{pmatrix},$$

где

$\mathbf{O}$ — нулевая интервальная матрица размера $p \times p$, т.е. матрица, все элементы которой имеют вид $[0, 0]$;
$\mathbf{I}$ — единичная интервальная матрица размера $p \times p$, т.е. матрица, на главной диагонали которой стоят элементы $[1, 1]$, а остальные имеют вид $[0, 0]$;
$\beta_1, \beta_2, \ldots, \beta_r$ — интервальные коэффиценты рекуррентности из соотношения (11).

Легко проверить, что полученная таким образом система является реализацией исходной последовательности матриц. □

6 Граничные реализации

Рассмотрим проблему вычисления точной конечномерной реализации для интервальной последовательности матриц (9). В некоторых частных случаях эта проблема может быть успешно решена. В частности, для полностью неотрицательных интервальных систем можно легко получить одно из возможных решений.

С импульсной последовательностью интервальных матриц можно связать две обычные (вещественные) импульсные последовательности, определяемые верхними и нижними границами интервальных матриц.

Определение 5. *Для последовательности интервальных матриц*

$$\{\mathbf{A}_1, \mathbf{A}_2, \ldots\} = \{\, [\underline{A}_1, \overline{A}_1], [\underline{A}_2, \overline{A}_2], \ldots \,\} \tag{12}$$

реализации последовательности

$$\{\, \underline{A}_1, \underline{A}_2, \ldots \,\} \tag{13}$$

будем называть нижними граничными реализациями последовательности (12), а реализации последовательности

$$\{\, \overline{A}_1, \overline{A}_2, \ldots \,\} \tag{14}$$

будем называть верхними граничными реализациями последовательности (12).

Как уже отмечалось ранее, в рамках интервальной арифметики (как классической, так и полной) не выполняется закон дистрибутивности операций сложения и умножения интервалов. Однако этот закон имеет место, когда мы имеем дело с неотрицательными интервалами, т.е. интервалами, обе границы которых неотрицательны. Это позволяет нам сформулировать и доказать следующий результат, касающейся граничных реализаций.

Предложение 2. *Если для нижней и верхней граничных реализаций одинаковой размерности* $(\underline{F}, \underline{G}, \underline{H})$ *и* $(\overline{F}, \overline{G}, \overline{H})$ *некоторой последовательности интервальных матриц выполняется*

1) $\underline{F}, \underline{G}, \underline{H}, \overline{F}, \overline{G}, \overline{H}$ *— неотрицательные;*
2) $\underline{F} \leq \overline{F}, \underline{G} \leq \overline{G}, \underline{H} \leq \overline{H}$*,*

то интервальная система $([\underline{F}, \overline{F}], [\underline{G}, \overline{G}], [\underline{H}, \overline{H}]$ *является интервальной точной (алгебраической) реализацией этой последовательности.*

Доказательство. Легко заметить, что для неотрицательных интервальных матриц (т.е. матриц, все элементы которых являются неотрицательными интервалами) результат произведения двух неотрицательных матриц будет также неотрицательной матрицей. Для неотрицательных интервальных матриц $\mathbf{B} = [\underline{B}, \overline{B}]$ и $\mathbf{C} = [\underline{C}, \overline{C}]$ имеет место

$$\mathbf{BC} = [\underline{B}, \overline{B}][\underline{C}, \overline{C}] = [\underline{BC}, \overline{BC}].$$

Кроме того, из-за дистрибутивности операций сложения и умножения неотрицательных интервалов выполняется закон ассоциативности для умножения неотрицательных интервальных матриц.

Пусть $(\underline{F}, \underline{G}, \underline{H})$ и $(\overline{F}, \overline{G}, \overline{H})$ — две граничные реализации одинаковой размерности для последовательности интервальных матриц (12). А именно, $(\underline{F}, \underline{G}, \underline{H})$ - реализация последовательности (13), $(\overline{F}, \overline{G}, \overline{H})$ - реализация последовательности (14). Тогда, используя перечисленные выше свойства операции умножения неотрицательных интервальных матриц, получаем

$$[\underline{H}, \overline{H}][\underline{G}, \overline{G}] = [\underline{HG}, \overline{HG}] = [\underline{A_1}, \overline{A_1}] = \mathbf{A_1},$$

$$[\underline{H}, \overline{H}][\underline{F}, \overline{F}][\underline{G}, \overline{G}] = [\underline{HFG}, \overline{HFG}] = [\underline{A_2}, \overline{A_2}] = \mathbf{A_2},$$

$$\cdots\cdots\cdots$$

$$\cdots\cdots\cdots$$

$$[\underline{H}, \overline{H}][\underline{F}, \overline{F}]^{i-1}[\underline{G}, \overline{G}] = [\underline{HF}^{i-1}\underline{G}, \overline{HF}^{i-1}\overline{G}] = [\underline{A_i}, \overline{A_i}] = \mathbf{A_i},$$

и т.д. Следовательно, интервальная система $([\underline{F}, \overline{F}], [\underline{G}, \overline{G}], [\underline{H}, \overline{H}])$ является реализацией последовательности (12), что и требовалось доказать. □

Однако очень часто оказывается, что найденные граничные реализации не удовлетворяют условиям предложения 2. Но и в этих случаях ситуация не является безнадежной. Дело в том, что подходящим выбором базисов пространств состояний граничных реализаций часто можно добиться того, чтобы граничные реализации удовлетворяли условиям предложения 2. Легко доказывается следующее утверждение, являющееся следствием предложения 2.

Предложение 3. *Если для граничных реализаций одинаковой размерности* $(\underline{F}, \underline{G}, \underline{H})$ *и* $(\overline{F}, \overline{G}, \overline{H})$ *некоторой последовательности интервальных матриц найдутся такие матрицы* T_1 *и* T_2*, что выполняются неравенства*

$$\underline{\hat{F}} \leq \overline{\hat{F}}, \underline{\hat{G}} \leq \overline{\hat{G}}, \underline{\hat{H}} \leq \overline{\hat{H}},$$

где

$$\underline{\hat{F}} = T_1 \underline{F} T_1^{-1}, \underline{\hat{G}} = T_1 \underline{G}, \underline{\hat{H}} = \underline{F} T_1^{-1}, \tag{15}$$

$$\overline{\hat{F}} = T_1 \overline{F} T_1^{-1}, \overline{\hat{G}} = T_1 \overline{G}, \overline{\hat{H}} = \overline{F} T_1^{-1}, \tag{16}$$

$\underline{\hat{F}}, \underline{\hat{G}}, \underline{\hat{H}}, \overline{\hat{F}}, \overline{\hat{G}}, \overline{\hat{H}}$ *— неотрицательные, то интервальная система* $([\underline{\hat{F}}, \overline{\hat{F}}], [\underline{\hat{G}}, \overline{\hat{G}}], [\underline{\hat{H}}, \overline{\hat{H}}]$ *является интервальной алгебраической реализацией этой последовательности.*

Таким образом, предложение 3 переводит проблему вычисление интервальной алгебраичесой реализации в задачу линейной матричной алгебры. С использованием предложений 2 и 3 можно строить конечномерные реализации интервальных импульсных последовательностей во многих частных случаях. Алгоритм действий здесь следующий. Сначала ищутся граничные реализации одинаковой размерности. Если они удовлетворяют условиям предложения, то соответствующая интервальная реализация будет искомой реализацией заданной импульсной последовательности. В противном случае с помощью преобразований подобия (15)-(16) нужно попытаться найти эквивалентные (с точностью до изоморфизма) граничные реализации, для которых условия предложения 2 выполняются. Во многих случаях такими реализациями являются системы в наблюдаемой или управляемой канонической форме.

Пример. Рассмотрим интервальную импульсную последовательность для системы с одним входом и одним выходом:

$$\mathbf{A_1} = [0, 0.1], \qquad \mathbf{A_2} = [0.81, 1.22], \qquad \mathbf{A_3} = [0.729, 1.574], \qquad \mathbf{A_4} = [1.3122, 3.2308]. \tag{17}$$

Граничные реализации, вычисленные с помощью метода, представленного в [10], имеют вид

$$\underline{F} = \begin{pmatrix} 0.9 & 1 \\ 0.81 & 0 \end{pmatrix}, \underline{G} = \begin{pmatrix} 0 \\ 0.81 \end{pmatrix}, \underline{H} = \begin{pmatrix} 1 & 0 \end{pmatrix},$$

$$\overline{F} = \begin{pmatrix} 12.2 & 1 \\ -133.1 & -11 \end{pmatrix}, \overline{G} = \begin{pmatrix} 0.1 \\ 0 \end{pmatrix}, \overline{H} = \begin{pmatrix} 1 & 0 \end{pmatrix}.$$

Очевидно, эта пара не удовлетворяет условиям предложения 2. Применив к этим реализациям преобразования подобия

$$T_1 = \begin{pmatrix} 1 & 0 \\ 0.9 & 1 \end{pmatrix}, T_2 = \begin{pmatrix} 1 & 0 \\ 12.2 & 1 \end{pmatrix},$$

мы получаем другую, эквивалентную исходной, пару реализаций в наблюдаемой канонической форме:

$$\underline{\hat{F}} = \begin{pmatrix} 0 & 1 \\ 0.81 & 0.9 \end{pmatrix}, \underline{\hat{G}} = \begin{pmatrix} 0 \\ 0.81 \end{pmatrix}, \underline{\hat{H}} = \begin{pmatrix} 1 & 0 \end{pmatrix},$$

$$\overline{\hat{F}} = \begin{pmatrix} 0 & 1 \\ 1.1 & 1.2 \end{pmatrix}, \overline{\hat{G}} = \begin{pmatrix} 0.1 \\ 1.22 \end{pmatrix}, \overline{\hat{H}} = \begin{pmatrix} 1 & 0 \end{pmatrix}.$$

Легко проверяется, что интервальная система, составленная на основе этих граничных реализаций, т.е. система с матрицами

$$\hat{\mathbf{F}} = \begin{pmatrix} [0,0] & [1,1] \\ [0.81, 1.1] & [0.9, 1.2] \end{pmatrix}, \hat{\mathbf{G}} = \begin{pmatrix} [0, 0.1] \\ [0.81, 1.22] \end{pmatrix}, \hat{\mathbf{H}} = \begin{pmatrix} [1,1] & [0,0] \end{pmatrix}.$$

является интервальной реализацией интервальной последовательности(17).

Следует заметить, что граничные реализации последовательности (17), полученные с помощью методов, представленных работах [7, 8] сразу удовлетворяют условиям предложения 2.

Данный подход к вычислению алгебраических реализаций можно применять также для импульсных последовательностей полностью неположительных матриц. Действительно, легко показать, что если $(\mathbf{F}, \mathbf{G}, \mathbf{H})$ — граничная реализация для последовательности (11), то для последовательности

$$\{\, [-\overline{A}_1, -\underline{A}_1], [-\overline{A}_2, -\underline{A}_2], \ldots \,\}$$

реализациями будут системы $(\mathbf{F}, -\mathbf{G}, \mathbf{H})$ и $(\mathbf{F}, \mathbf{G}, -\mathbf{H})$.

Хотя метод граничных реализаций является обоснованным для полностью неотрицательных и полностью неположительных систем, работа с многочисленными системами показала, что он оказывается применимым и для других классов систем. Однако пока не удалось очертить достаточно точные границы применимости этого метода.

7 Заключительные замечания

Последовательное изложение теории и методов реализации интервальных систем потребует развития методов интервальной алгебры. Потребуется пересмотр и переопределение таких привычных понятий линейной алгебры, как невырожденная матрица, ранг матрицы и т.п. Без решения этих вопросов не удастся получить не только эффективные методы реализации, но и решить вопрос реализуемости заданной импульсной последовательности интервальных матриц. По-видимому, наибольшего прогресса можно достигнуть путем рассмотрения их в $\mathbb{KR}$ и погружения в линейное пространство [11]. Для погружений множеств входных сигналов, состояний и выходных сигналов интервальной системы в линейные пространства удвоенной размерности имеет место следующая коммутативная диаграмма множеств и отображений

$$
\begin{array}{ccccccc}
\mathbb{KR}^m & \underset{\mathbf{G}}{\longrightarrow} & \mathbb{KR}^n & \underset{\mathbf{F}}{\longrightarrow} & \mathbb{KR}^n & \underset{\mathbf{H}}{\longrightarrow} & \mathbb{KR}^p \\
\Omega \downarrow\uparrow \Omega^{-1} & & \Pi \downarrow\uparrow \Pi^{-1} & & \Pi \downarrow\uparrow \Pi^{-1} & & \Gamma \downarrow\uparrow \Gamma^{-1} \\
\mathbb{R}^{2m} & \underset{\hat{G}}{\longrightarrow} & \mathbb{R}^{2n} & \underset{\hat{F}}{\longrightarrow} & \mathbb{R}^{2n} & \underset{\hat{H}}{\longrightarrow} & \mathbb{R}^{2p} \\
\updownarrow id & & \uparrow \varphi & & \uparrow \varphi & & \updownarrow id \\
\mathbb{R}^{2m} & \underset{\tilde{G}}{\longrightarrow} & \mathbb{R}^{2n} & \underset{\tilde{F}}{\longrightarrow} & \mathbb{R}^{2n} & \underset{\tilde{H}}{\longrightarrow} & \mathbb{R}^{2p}
\end{array}
$$

В общем случае индуцированные погружениями отображения $\hat{F}$, $\hat{G}$ и $\hat{H}$, соответствующие $\mathbf{F}$, $\mathbf{G}$ и $\mathbf{H}$, являются нелинейными. Однако в ряде частных случаев можно добиться того, чтобы они были линейными. Матрицы $\hat{F}$, $\hat{G}$ и $\hat{H}$ должны иметь специальный вид. Этого можно добиться с помощью преобразования подобия φ для реализации $(\tilde{F}, \tilde{G}, \tilde{H})$, полученной в линейном пространстве.

Список литературы

1. Калман Р., Фалб П., Арбиб М. *Очерки по математической теории систем.* - М.: Мир, 1971.
2. Шарый С.П. Алгебраический подход к анализу линейных статических систем с интервальной неопределенностью // *Известия РАН. Теория и системы управления.* – 1997. – №3. – С. 51 – 61
3. Смагина Е.М., Моисеев А.Н. Слежение за полиномиальным сигналом в интервальной динамической системе // *Вычислительные технологии.* – 1998. – Т. 3, №1. – С. 67 – 74.
4. Ивлев Р.С., Соколова С.П. Построение векторного управления многомерным интервально заданным объектом // *Вычислительные технологии.* – 1999. – Т. 4, №4. – С. 3 – 13.
5. Богомолов А.С., Сперанский Д.В. Об одной разновидности задачи стабилизации линейной дискретной системы // *Автоматика и телемеханика.* – 2002. – №9. – С. 111 – 124.
6. Sontag E.D. Linear systems over commutative rings: A survey // *Ricerche di Automatica.* – 1976. – Vol. 7. – P. 1 – 34.
7. Пушков С.Г. Об алгоритме конечномерной реализации // *Автоматика и телемеханика.* – 1991. – №10. – С. 56 – 63.
8. Пушков С.Г. Конечномерные реализации импульсной характеристики, основанные на псевдообращении ганкелевой матрицы // *Известия РАН. Теория и системы управления.* – 2002. – №3. – С. 5 – 11.
9. Пушков С.Г. Об одном подходе к описанию наблюдаемого процесса линейной динамической системой // *Известия РАН. Теория и системы управления.* – 2001. - №1. – С. 39 – 44.
10. Пушков С.Г. О вычислении конечномерной реализации // *Кибернетика и системный анализ.* – 1991. – №6. – С. 107 – 112.
11. Шарый С.П. Алгебраический подход во "внешней задаче"для интервальных линейных систем // *Вычислительные технологии.* – 1998. – Т. 3, №2. – С. 67 – 114.

Интервальный подход к решению оптимизационной задачи удовлетворения ограничений

Е.В. Бревнов

Институт систем информатики имени А.П.Ершова,
Россия, 630090 г. Новосибирск, просп. Лавретньева, 6
bev@sib3.ru

Ключевые слова: Глобальная оптимизация, интервальная арифметика, метод распространения ограничений, метод ветвей и границ.

Введение

Данная работа посвящена проблеме решения задач глобальной оптимизации с дополнительными ограничениями, определенными на множестве переменных задачи. Основной упор сделан на решение целочисленных задач оптимизации. В работе, с успехом, используются идеи и методы интервальной математики.

Во многих научных исследованиях по вещественной и целочисленной глобальной оптимизации просматриваются идеи использования классического метода ветвей и границ [9], [10], интервальной арифметики и ее методов [12]. Ряд статей посвящены различным способам увеличения эффективности работы интервального алгоритма ветвей и границ. Так, в статьях [6] и [7] рассматриваются способы выбора области для дальнейшего поиска экстремума, а также доказывается оптимальность такого выбора. Различные стратегии деления множества значений переменных подробно рассматриваются в [5] и [11]. Результатом этих исследований становится вариативный способ деления допустимых значений переменных на разное количество частей. Один из способов комбинирования стратегий поиска локального и глобального экстремумов для нахождения глобального экстремума рассматривается в статье [4]. Результаты статей и наработанная база методов интервальной математики оказали решающее воздействие на дальнейшую работу.

Глава 1 посвящена описанию метода распространения ограничений для достижения локальной совместности. Интервальный метод оптимизации представлен в главе 2. Глава 3 содержит анализ и сравнение интервального метода с оптимизационным методом решателя SibCalc [8].

1 Интервальный метод распространения ограничений

Интервальная арифметика — это арифметика, определённая на множестве интервалов. Существует ряд работ, достаточно подробно описывающих методы и свойства интервальной математике (например [1]), поэтому мы их рассматривать не будем. Однако, эти свойства будут нами использоваться в ходе дальнейшего изложения интервального метода распространения ограничений, который фактически основан на интервальной математике.

Одним из методов достижения локальной совместности является интервальный метод распространения ограничений (**CP**). Рассмотрим более детально данный метод. Пусть математическая модель задана системой из m алгебраических уравнений от n переменных:

$$F(x) = 0, \tag{1.1}$$

где $F = (f_1(x), f_2(x), \ldots, f_m(x))$ — вектор-функция и $x = (x_1, \ldots, x_n)$ — вектор переменных, причем m может быть не равно n. Множество всех решений системы (1.1) обозначается через $P = \{x \in A^n \mid F(x) = 0\}$, где P принадлежит ограниченному параллелепипеду $D = \{x \in A^n \mid l_i \le x_i \le r_i : l_i, r_i \in A,\ i = 1, 2, \ldots, n\}$. Так как в общем случае i-е уравнение системы содержит не все переменные, то его можно записать в следующем виде:

$$f^{(i)}(x_{i_1}, \ldots, x_{i_{n_i}}) = 0, \tag{1.2}$$

где $n_i \le n$. Предполагается, что из этого уравнения можно выразить каждую переменную через другие, тогда получится n_i следующих уравнений:

$$x_{i_1} = f_1^{(i)}(x_{i_2}, \ldots, x_{i_{n_i}}),$$

$$\ldots$$

$$x_{i_j} = f_j^{(i)}(x_{i_1}, x_{i_2}, \ldots, x_{i_{j-1}}, x_{i_{j+1}}, \ldots, x_{i_{n_i}}), \quad (1.3)$$

. . .

$$x_{i_{n_i}} = f_{n_i}^{(i)}(x_{i_1}, x_{i_2}, \ldots, x_{i_{n_i-1}}).$$

Если записать в таком виде все уравнения системы (1.1), то будет получено $\mu = n_1 + n_2 + \ldots + n_m$ равенств вышеуказанного вида, определяющих в общем случае для каждой i-й переменной системы (1.1) m_i $(1 \le m_i \le m)$ функций от разных наборов переменных. Правые части равенств (1.1.3) называются *функциями интерпретации*.

Считается, что каждое значение переменной x_i недоопределено и может принадлежать интервалу $[l_i, r_i]$. **СР** сужает интервалы для каждой переменной x_i по описанному ниже алгоритму.

Не ограничивая общности, равенства (1.3) можно переписать для интервальных типов данных с операциями интервальной математики:

$$^*x_j = {^*f}_j^{(i)}(^*x_1, {^*x}_2, \ldots, {^*x}_{j-1}, {^*x}_{j+1}, \ldots, {^*x}_{n_i}) \quad (1.4)$$

Итерационный процесс уточнения недоопределенных значений переменных строится на основе представления (1.4) и состоит из следующих шагов:

Для k = 0 вычисляются все ($\mu = n_1 + n_2 + \ldots + n_m$) функции интерпретации (1.1.4), полагая $^*x_i^0 = [l_i, r_i]$.

- Для всех переменных, значения которых изменились после вычисления соответствующих функций интерпретации, помечаются функции, имеющие эти переменные в качестве аргумента. Множество всех помеченных функций называются *активными функциями*. Если множество активных функций пусто, то процесс вычислений заканчивается.
- Вычисляются значения активных функций для k+1 шага по формулам:

$$^*x_j^{k+1} = {^*f}_j^{(i)}(^*x_1^{k+1}, {^*x}_2^{k+1}, \ldots, {^*x}_{j-1}^{k+1}, {^*x}_{j+1}^{k}, \ldots, {^*x}_{n_i}^{k}) \cap {^*x}_j^k \quad (1.5)$$

- Переход на шаг 2.

Таким образом, значения переменных в процессе итерационных вычислений представляются последовательностью не расширяющихся интервалов $^*x_j^k$. Такая последовательность стабилизируется за конечное число шагов, причем порядок выполнения функций интерпретации не влияет на окончательный результат вычислений. В результате применения итерационного процесса (1.5) будет получен многомерный параллелепипед $^*D$, являющийся декартовым произведением интервальных значений $^*x_j = \lim\limits_k {^*x}_j^k,\ j = \overline{1,n}$. Если $^*D \ne \varnothing$, то он гарантированно содержит все корни системы (1.1), в противном случае ($^*D = \varnothing$) — система несовместна. Более подробное описание работа метода можно найти в соответствующей литературе [2].

2 Интервальный метод ветвей и границ

Впервые метод ветвей и границ был предложен Лендом и Дойгом [9] в 1960 для решения общей задачи целочисленного линейного программирования. Интерес к этому методу и фактически его “второе рождение” связано с работой Литтла, Мурти, Суини и Кэрела [10], посвященной задаче коммивояжера [3].

Начиная с этого момента, появилось большое число работ, посвященных методу ветвей и границ и различным его модификациям. Столь большой успех объясняется тем, что авторы первыми обратили внимание на широту возможностей метода, отметили важность использования специфики задачи и сами воспользовались спецификой задачи коммивояжера. Отметим, что эффективность работы алгоритма определяется двумя факторами: качеством оценочной функции и возможностью получать “хорошие” значения целевой функции на ранних стадиях.

В основе интервального метода **IBB** лежат идеи интервальной арифметики, интервального метода распространения ограничений и классического метода ветвей и границ. Задача **ЗУО** есть подзадача **ОЗУО**, так как решениями задачи **ОЗУО** являются те решения задачи **ЗУО**, на которых достигается экстремум функции цели. Следовательно, для решения задачи **ОЗУО** необходимо сначала научиться решать задачу **ЗУО**. Одной из наиболее эффективных стратегий решения задач **ЗУО**, является стратегия «Forward Checking». Основное звено стратегии «Forward Checking» составляют методы достижения локальных совместностей. Интервальный метод распространения ограничений помогает устанавливать локальную совместность на алгебраических системах, что делает возможным эффективное использование этого метода при решении более сложной задачи **ОЗУО**. С одной стороны, задача **ОЗУО** очевидно сложнее, чем задача **ЗУО**, с другой, использование специфики задачи **ОЗУО** позволяет облегчить поиск решения. Одним из методов, ориентированных на решение задач оптимизации, является классический метод ветвей и границ. Использование знания о значении целевой функции дает возможность уменьшить пространство для поиска корня. Совместное использование метода ветвей и границ и интервальной арифметики позволяет достичь хорошей эффективности при решении задачи **ОЗУО**.

Для простоты дальнейшего изложения материала, не ограничивая общности, будем считать, что задача **ОЗУО** есть задача максимизации. На вход методу подается задача **P = (X, D, C, F)**. На выходе ожидается вектор-решение для задачи **P**. На первом шаге происходит инициализация *интервального рекорда* минимальным допустимым числом. Затем исходная область **D** подается на вход методу **CP** для достижения **2B** совместности. На выходе получаем область $\mathbf{D}' \subseteq \mathbf{D}$, причем количество решений в областях **D'** и **D** совпадает. Далее, область **D'** делится на подобласти $\mathbf{D_1}$ и $\mathbf{D_2}$ так, что текущая переменная деления принимает максимальное допустимое значение в области $\mathbf{D_1}$ и все остальные значения в области $\mathbf{D_2}$. Области $\mathbf{D_1}$ и $\mathbf{D_2}$ проходят этап установления локальной совместности. В случае невозможности достижения локальной совместности, область исключается из дальнейшей обработки ввиду отсутствия решений в **ЗУО** задаче. Иначе вычисляется интервальное расширение целевой функции, которое сравнивается с *интервальным рекордом*, и, при необходимости, происходит обновление *интервального рекорда*. Весь процесс повторяется до тех пор, пока не обработаются все области. Область-вектор на котором достигается значение равное значению *интервального рекорда,* является вектором-решением задачи **P**. К сожалению, данный алгоритм не лишен недостатков. Одним из недостатков, является существенная зависимость эффективности работы данного алгоритма от близости решения к краю области поиска. Удачный выбор значения переменной деления быстро приводит к хорошим значениям *интервального рекорда*, что позволяет эффективно сократить пространство поиска экстремума. Напротив, неудачный выбор значения для переменной деления ведет к полному перебору. Также, способ деления путем означивания переменных делает затруднительным и мало эффективным применение алгоритма в случае вещественных переменных. Наличие вышеперечисленных недостатков можно устранить при помощи деления переменной на две или три равные части, и выбора наиболее перспективной области для дальнейшего поиска решения. Перспективность области определяется по верхней границе интервальной оценки значения целевой функции. Для дальнейшего деления выбирается область с наибольшей такой границей. Псевдокод модифицированного алгоритма приводится в таблице 2.1. Рассмотрим работу модифицированного интервального метода ветвей и границ (**IBBM**) на примере:

Пример 2.1: Определим целочисленные переменные x1 и x2:

x1 = [-1, 2]; x2 = [-1, 10];

Система ограничений задается в следующем виде:

$x1^2 — x2 > 0$; $2{*}x2 + x1 <= 1$;

Целевая функция есть: $F = x1^2 — x1 — x2$;

Задача заключается в поиске таких значений переменных, при которых удовлетворяются все ограничения, и достигается максимум целевой функции. Поиск корня происходит слева направо сверху вниз в соответствии с рисунком 2.1. Овалы обозначают области, полученные в результате деления по фиксированной переменной. Прямоугольники содержат локально совместные области. Прежде чем начать поиск корня, исходная область подается на вход методу **CP** для установления локальной совместности, в результате чего происходит уменьшение пространства поиска более чем в 2 раза. Суженая область подвергается первому делению по переменной x2. Применение метода **CP** для подобластей и вычисление интервального расширения целевой функции приводит к отсечению крайней правой ветки из-за невозможности достижения локальной совместности. Далее, поиск корня происходит вдоль крайней левой ветки, ввиду наибольшего значения верхней границы интервального расширения целевой функции. После полной обработки левой ветки происходит возврат на центральную ветвь. К этому моменту, значение *интервального рекорда* равно 3. Верхняя граница интервального расширения целевой функции оказывается меньше текущего значения *интервального рекорда*, следовательно, поиск корня вдоль центральной ветки не имеет смысла. Итак, в результате найдено два корня с максимальным значением целевой функции равным 3. Рисунок 2.1 свидетельствует, что благодаря применению техники метода **IBBM** пространство поиска составило меньше трети от возможного. Из примера 2.1 видно, что в ряде случаев удается избежать полного перебора за счет отсечения несовместных областей и областей, на которых *интервальное значение* целевой функции больше *интервального рекорда*. Для задач больших размерностей и с большим числом ограничений данное свойство алгоритма может существенно облегчить поиск экстремума.

```
IBBM(X, D, C, F): Maximum
best =I — infinity;
D = CP(X, D, C);
Queue = {D};
Pick domain D from Queue
    Split any Dx ∈ D into D1, D2, D3
        D1 = CP(X, D1, C);
        D2 = CP(X, D2, C);
        D3 = CP(X, D3, C);
        Fmax = max{F(D1), F(D2), F(D3)};
        D = {Di | F(Di) = Fmax, i = 1, 2, 3};
        Put domain Dj (i ≠ j, j = 1, 2, 3) into Queue;
        If (F(D) <I best) then
            Pick next domain;
        End If;
        If (Is empty(X)) then
            If (F(D) >I best) then
                best =I F(D);
            End If;
        End If;
    End;
End;
Return best;
End.
```

Таблица 2.1

3 Анализ интервального метода ветвей и границ.

Теперь дадим качественную оценку методу **IBBM**. Прежде всего, опишем класс задач, покрываемый методом **IBBM**. Этот класс составляют задачи целочисленной, вещественной, и смешанной глобальной оптимизации с ограничениями. Система ограничений есть произвольная система алгебраических уравнений. Целевая функция выражается в виде допустимой комбинации параметров, целочисленных и вещественных переменных. Кроме того, в уравнениях системы и целевой функции могут присутствовать интервальные параметры. Теперь остановимся на свойствах данного алгоритма:

- **глобальность** (найденное решение является глобальным экстремумом)
- **полнота** (все глобальные решения содержатся среди найденных)
- **существование субоптимум**а (в любой момент времени существует приближение к глобальному экстремуму)
- **универсальность**

Все эти свойства, несомненно, являются положительными сторонами алгоритма **IBBM**. Большой диапазон задач, которые могут быть решены при помощи данного алгоритма, ставит вопрос об эффективности их решения. Наивно ожидать, что все задачи будут решаться одинаково хорошо. Эффективность решения может существенно зависеть от вида ограничений и целевой функции, хотя некоторые теоретические оценки все же можно сделать. Из свойств интервальной арифметики следует, что в случае однократного вхождения каждой переменной в целевую функцию получаем точную нижнюю и верхнюю оценки множества значений целевой функции. Например, в случае линейной целевой функции мы не будем иметь эффекта расширения ее интервального значения. Точность оценки изменения значений целевой функции увеличивает возможность отсечения областей, выходящих за границы *интервального рекорда*. Не последнюю роль играет структура системы ограничений. В случае, когда система ограничений сильно нелинейная, интервальный метод распространения ограничений существенно сужает допустимые области изменения значений переменных, что значительно сокращает пространство поиска экстремума. Напротив, сложно ожидать хороших результатов, когда переменные многократно входят в целевую функцию, а ограничения линейны. Наихудший случай работы алгоритма — это полный перебор. Поэтому наихудшее время работы алгоритма может быть оценено как $T = O(e^n)$, где n — мощность множества переменных. Память, необходимая в этом случае оценивается как $V = O(n^2*d)$, где n — мощность множества переменных, d — максимальная мощность областей значения переменных. Дать более точную оценку эффек-

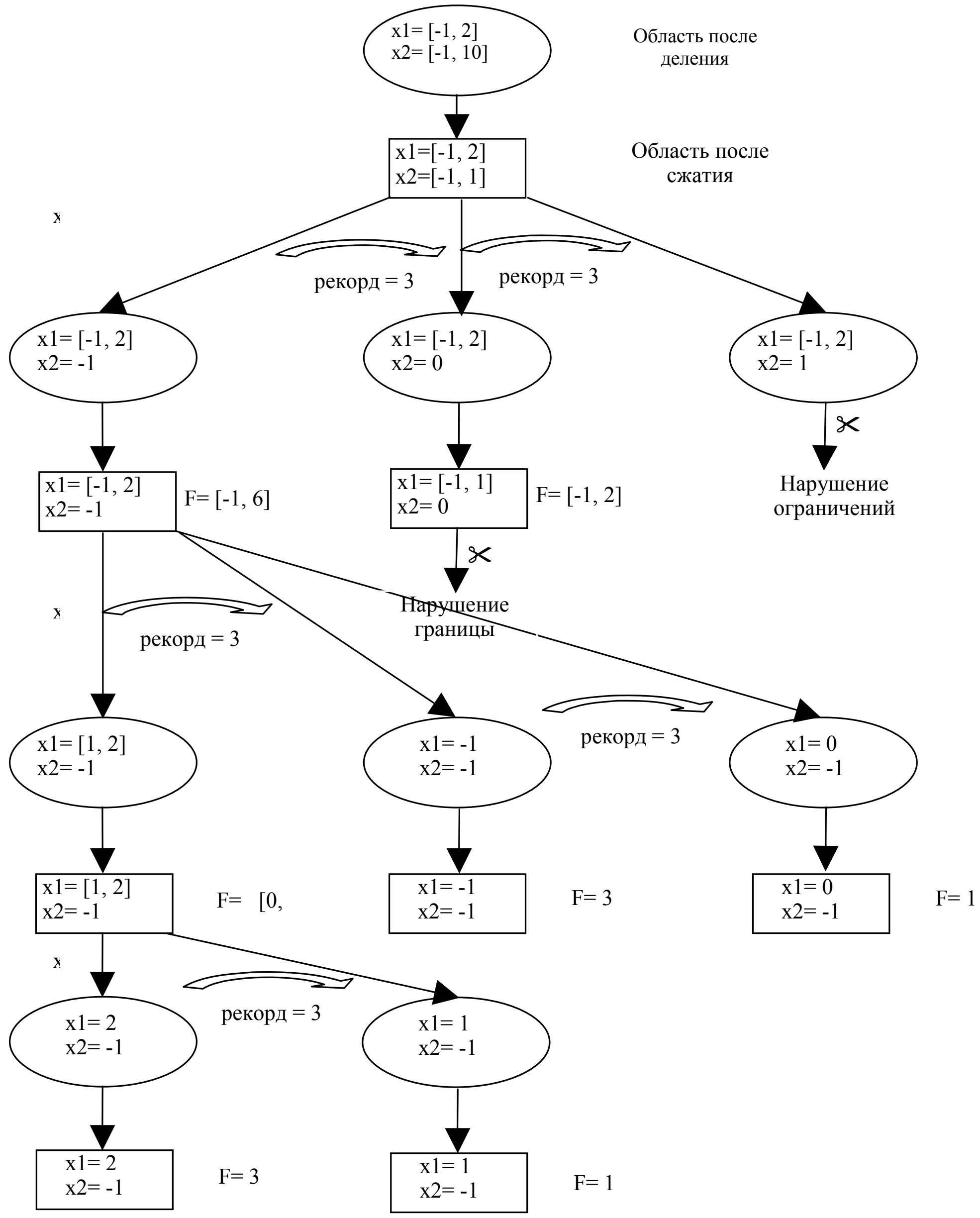


Рисунок 2.1

тивности алгоритма весьма затруднительно, что подчеркивает важность проведения практических экспериментов.

Как и в любом алгоритме поиска, очень важным моментом оказывается порядок переменных, по которым ведется поиск. Без дополнительных алгоритмов анализа структуры задачи наиболее успешной стратегией оказывается динамический выбор последовательности переменных поиска. В случае целочисленных переменных правомерно ожидать, что дерево поиска окажется минимальным при выборе самой узкой переменной для поиска. Например, пусть переменная x1 имеет ширину 2, а переменная x2 имеет ширину 3. В случае поиска по переменной x1, а затем по переменной x2, количество узлов в дереве поиска при полном переборе составит 8, а при противоположном порядке переменных 9. С вещественными переменными дело обстоит иначе. Выбор самой широкой переменной для поиска увеличивает шансы сильного сжатия области поиска с помощью алгоритма **CP**.

Кроме выбора последовательности переменных для поиска, существенную роль играет способ деления множества допустимых значений переменной. Принимая во внимание исследования других авторов в этой области (см. [5], [11]), был выбран метод трисекции. Деление не на две, а на три части оказывается более эффективным за счет уменьшения вероятности, что произойдет пересечение множества значений целевой функции на всех трех областях. Если в области с большим значением целевой функции будет найдено решение, то область с меньшим значением может быть исключена из дальнейшего рассмотрения без обработки. Также, в результате деления на большее количество частей повышается точность интервальной оценки целевой функции, что имеет положительный эффект в работе алгоритма.

Ряд изученных работ [6], [7] посвящены построению критерия выбора области для дальнейшего поиска из множества доступных областей. Несомненно, «хороший» критерий помогает существенно ускорить работу метода. Ввиду использования интервальной математики для оценки множества значений изменения целевой функции, поиск корня ведется в области с наибольшим значением верхней границы целевой функции. Данный критерий позволяет ускорить нахождение «хороших» *интервальных рекордов*, что помогает при дальнейшем поиске.

Интересным моментом в работе алгоритма является способ хранения текущих экстремумов. Все найденные экстремумы хранятся в обычном линейном списке. При нахождении новой области, которая удовлетворяет всем ограничениям, необходимо либо добавить ее в список, либо заменить весь список областью, либо отбросить ее.

Name	***CP + Bisect***		***IBB***	
	Time	**Roots**	**Time**	**Roots**
Группа 1				
Reg_Bag_500_5	***57.6***	1	***0.7***	1
Reg_Bag_1000_15	***3074.3***	1	***5.6***	1
Reg_Bag_50_50	***218.6***	1	***8.5***	1
Reg_Bag_50_800	***> 1 day***	0	***233.8***	1
1Rand_Bag_20_40000	***1024.2***	1	***0.17***	1
2Rand_Bag_20_40000	***3136.2***	1	***0.15***	1
3Rand_Bag_20_40000	***> 2 hours***	0	***0.41***	1
1Rand_Bag_50_20000	***> 15 hours***	0	***8.5***	1
2Rand_Bag_50_20000	***> 15 hours***	0	***3.7***	1
3Rand_Bag_50_20000	***> 15 hours***	0	***0.8***	1
Группа 2				
1Rand_Com_5	***2329.8***	1	***0.12***	1
2Rand_Com_5	***2087.3***	1	***0.14***	1
3Rand_Com_5	***2152.2***	1	***0.12***	1
1Rand_Com_7	***> 15 hours***	0	***12.3***	1
2Rand_Com_7	***> 15 hours***	0	***11.9***	1
3Rand_Com_7	***> 15 hours***	0	***8.6***	1
1Rand_Com_10	***> 1 day***	0	***2343.7***	1
2Rand_Com_10	***> 1 day***	0	***1402.6***	1
3Rand_Com_10	***> 1day***	0	***2408.9***	1
Группа 3				
Golomb_5	***0.1***	1	***0.09***	4
Golomb_6	***1.7***	1	***2***	8

Таблица 3.1

Если интервальное значение целевой функции на области строго меньше наибольшего значения целевой функции из списка, то происходит отбрасывание данной области. Иначе происходит добавление области в список и отбрасывание всех областей с меньшими значениями целевой функции. Введем дополнительный интервал, нижняя граница которого определяется как максимум нижних границ значений целевой функции из списка, верхняя как максимум верхних границ. Введение такого интервала позволяет отказаться от сравнения значения целевой функции на области со всеми значениями из списка. Достаточно лишь сравнения нового значения с введенным интервалом и однократной фильтрацией списка в конце работы алгоритма. Такой подход хранения множества текущих экстремумов в отличие от предлагаемых в литературе приводит к существенному сокращению числа машинных операций на этапе обновления этого множества.

Разработанный метод является частью мощного вычислительного инструмента SibCalc. При помощи данного метода были решены некоторые сложные практические задачи. В таблице 3.1 проводится сравнительный анализ метода **IBB** с существующим методом оптимизации универсального решателя SibCalc. Таблица разбита на три группы по типам решаемых задач. Первая группа — задачи о ранце для различных весов, ценностей и количества предметов. Вторая группа — задачи коммивояжера для различного числа городов и расстояний между ними. Третья группа — задача Голомба размерности 5 и 6. Результаты решения задач первой и второй группы показывают существенное преимущество метода **IBB** по сравнению с методом оптимизации решателя **CP+Bisection.** Группа номер три отражает свойство полноты множества найденных решений. Результаты сравнения двух вышеупомянутых методов и метода **IBBM** приводятся в таблице 3.2. Двадцать тестовых задач вошли в таблицу и образовали две группы. Первая группа задач одинаково хорошо решена всеми тремя методами. Вторая группа задач показывает эффективность применения стратегии трисекции и используемого критерия выбора области для деления. Результаты решения задачи H72 свидетельствуют об увеличении времени выполнения одной итерации. Несмотря на это обстоятельство, метод **IBBM** за одно и тоже время сумел сильнее приблизиться к оптимальному решению, чем метод **IBB**, что свидетельствует об оптимальности пути поиска экстремума.

Name	***CP + Bisection***			***IBB***			***IBBM***		
	Time	**Iterations**	**Roots**	**Time**	Iterations	**Roots**	**Time**	**Iterations**	**Roots**
Группа 1									
Eq1	0.02	368	1	0.01	3	1	0.04	31	1
Flaudas3	0.01	418	1	0.03	3	1	0.02	7	1
H104	0.01	Inconsistent		0.01	Inconsistent		0.01	Inconsistent	
H77	0.1	Inconsistent		0.1	Inconsistent		0.1	Inconsistent	
H78	0.05	2259	1	0.05	16	6	0.05	15	6
Hock7	0.02	189	1	0.01	2	2	0.01	4	2
Ineq1	0.03	944	1	0.2	58	1	0.23	70	1
Группа 2									
Flaudas4	545.0	36077810	1	0.02	5	1	0.02	6	1
H93	> 2000		---	12.85	Inconsistent		10.4	Inconsistent	
H76	> 2000		---	0.1	21	1	0.1	28	1
H79	0.6	50234	1	1.35	546	1	0.3	71	1
H110	> 2000		---	63.87	660	1	5.1	42	1
H13	8.5	753597	1	102.76	16701	2	20.34	2327	2
H73	> 2000		---	15.42	3011	1	8.28	1103	1
Combag	> 2000		---	252.3	11747	1	261.1	13648	1
O32	979.0	572036397	1	40.9	7561	20	5.7	570	20
H100	> 2000		---	89.14	21619	1	75.27	12515	1
H106	> 2000		---	> 2000		---	39.11	Inconsistent	
H111	> 2000		---	631.4	6747	1	8.4	74	1
H72	Possible global minimum = 212, Time limit = 180 sec								
	> 2000		---	180	62517	5 (7143)	180	44653	75 (673)

Таблица 3.2

Заключение

В данной работе был представлен интервальный метод нахождения глобального экстремума для оптимизационной задачи удовлетворения ограничений. Данный метод является универсальным, и может быть использован для нахождения глобального экстремума для целочисленных, вещественных, и смешанных задач оптимизации. Применение интервальных операций при машинных вычислениях позволяет гарантировать нахождение всех глобальных оптимумов. Отсутствие решений по завершению работы метода гарантирует невозможность удовлетворения ограничений, и как следствие, отсутствие глобального оптимума. Ограничение времени работы метода позволяет находить приближение к глобальному оптимуму. Оптимальность такого приближения может быть оценена при помощи интервальной техники.

Результаты сравнения интервального метода ветвей и границ с внутренним методом оптимизации универсального решателя SibCalc свидетельствуют об эффективности комбинирования классических методов поиска экстремума с интервальными техниками. Применение метода распространения ограничений в процессе поиска решений позволяет сузить области поиска и уточнить интервальную оценку множества значений целевой функции. Метод, как часть системы SibCalc, может быть использован при решении широкого круга оптимизационных задач, возникающих в практической деятельности.

Предполагается ряд дальнейших работ по развития алгоритма таких, как применение дополнительных критериев отсечения областей, особенно в случае вещественных переменных, повышение точности внешней оценки для множества допустимых значений целевой функции, привлечения стратегий поиска локального экстремума для увеличения эффективности нахождения рекордов.

Список литературы

1. Алефельд Г., Херцбергер Ю. Введение в интервальные вычисления, М.: Мир, 1987.
2. Кашеварова Т.П., Семенов А.Л. Некоторые вопросы сходимости метода недоопределенных вычислений. В: Проблемы представления и обработки не полностью определенных знаний. — Москва–Новосибирск: РосНИИ ИИ, 1996. — С. 31-37.
3. Корбут А.А., Финкельштейн Ю.Ю. Дискретное программирование М. Наука. Гл. ред. физ.-мат. лит. (1969).
4. Casado L.G., Garcia I. New load balancing criterion for parallel global optimization algorithm (1998).
5. Casado L.G., Garcia I., Csendes T. A new multisection technique in interval methods of global optimization (1998 — 1999).
6. Csendes T. Ratz D. Subdivision direction selection in interval methods for global optimization (1997).
7. Kreinovich V. Csendes T. Theoretical justification of a heuristic sub-box selection criterion for interval global optimization (2001).
8. Kleymenov A., Petunin D., Semenov A., Vazhev I. A Model of Cooperative Solvers for Computational Problems. Proceed. of the 4th International Conference PPAM 2001, Poland, September. LNCS, v 2328, pp.797-802.
9. Land A.H., Doig A.G. An automatic method of solving discrete programming problems. Econometrica. v28 (1960), pp 497-520.
10. Little J.D.C., Murty K.G., Sweeney D.W., and Karel C. An algorithm for the traveling salesman problem. Operations Research. v11 (1963), pp. 972-989.
11. Markot M. Csendes T. Csallner A. Multisection in interval Branch-and-Bound methods for global optimization 2 (2000).
12. Van Hentenryck P. McAllester D. Kapur D. Solving polynomial systems using a branch and prune approach (1994).

О решениях систем нелинейных уравнений

Т. П. Кашеварова, А. Л. Семенов

Институт систем информатики им. А.П. Ершова СО РАН
Новосибирск, пр.ак. Лаврентьва, 6
toma@iis.nsk.su, semenov@iis.nsk.su

Аннотация. Данная работа посвящена исследованию решений нелинейных систем уравнений с использованием интервальных методов. Разработанные вычислительные алгоритмы эффективны для решения нелинейных систем средних размерностей и применимы для случая так называемых ползущих решений. Предлагается подход, основанный на сочетании интервальных методов и методов распространения ограничений для задач с интервальными типами данных.

Необходимость решения систем алгебраических уравнений с недоопределенностью возникает практически в любой задаче моделирования или проведения инженерных расчетов. Данная работа посвящена численному исследованию решений нелинейных систем уравнений и разработке различных интервальных вычислительных методов нахождения решений нелинейных систем небольших и средних размерностей, применимых для случая так называемых ползущих решений [8].

В первой части доклада приводятся определения решений, получаемых численными методами, вводится определение ε - решения и описываются условия, когда ε - решение является окрестностью точного решения. Далее проводится анализ существующих интервальных методов, позволяющих гарантированно находить все ε - решения нелинейных систем уравнений. Так, интервальный метод Ньютона [6] позволяет гарантированно находить ε - решения и имеет квадратичную скорость сходимости в случае, если корень отделен и производная на рассматриваемом интервале отлична от нуля. Интервальные методы Хансена [2, 3] и Кравчика [5] способны отделить изолированные решения, так как они используют расширенную интервальную арифметику.

Метод Кравчика является модификацией интервального метода Ньютона, в котором вычисление обратной интервальной матрицы заменяется ее оценкой. В исходной формулировке метод недостаточно вычислительно эффективен и его улучшением является алгоритм Хансена-Сенгупты [4], позволяющий сравнительно быстро найти решение на небольшом начальном интервале или определить его отсутствие. В случае большого начального интервала этот алгоритм может не сойтись к решению. Позднее появляются различные улучшения и модификации вышеупомянутых алгоритмов [7, 9] связанные в основном с более точным оцениванием обратной интервальной матрицы.

Другими подходами к решению систем нелинейных уравнений являются методы распространения ограничений (СР) [1], активно развиваемые в области искусственного интеллекта. Методы СР применимы для произвольных систем уравнений и неравенств, то есть в систему могут одновременно входить линейные и нелинейные уравнения, различные неравенства и логические выражения, причем число уравнений может быть не равно числу неизвестных. Как плата за универсальность, время нахождения решения может оказаться значительным. В вычислительном плане метод СР может быть проинтерпретирован как недетерминированный итерационный процесс, управление которым осуществляется по данным, простейшим частным случаем которого можно рассматривать интервальный метод Зейделя .

Метод СР решает внешнюю интервальную задачу, т.е. находит многомерный брус, содержащий все решения системы уравнений. Для локализации отдельных решений применяется бисекция в сочетании с методом СР. В соответствии с алгоритмом выбирается переменная, по которой начинается поиск решения и направление поиска по этой переменной – слева или справа. Выбором последовательности переменных в поиске можно управлять. По умолчанию выбирается переменная, имеющая наименьшую ширину интервала. После того как интервал по выбранной переменной поделен пополам или в другой пропорции, вся система запускается на счет сначала на левом подынтервале, если выбрано левое направление поиска, и на правом – в случае правого направления, т.е. направление поиска определяет очередность просмотра подынтервалов.

Если система на одном из подынтервалов совместна, то происходит его следующее деление. Если же система на обоих подынтервалах несовместна, то они не содержат решения, и в дальнейшем весь объединенный интервал не рассматривается. Процесс деления по выбранной переменной происходит до тех пор, пока ширина интервала по этой переменной ни станет меньше заданной точности ε. Далее просматриваются интервалы всех остальных переменных. Если среди них есть интервалы с шириной больше ε, то разбиение повторяется по одной из этих переменных. Корень считается найденным, если максимальная ширина интервалов по всем переменным не превышает заданную точность и ему соответствует вся найденная ε-окрестность.

От выбора последовательности переменных, по которым осуществляется бисекция, существенно зависит расчетное время нахождения решения. В данной работе рассматривается стратегия выбора последовательности переменных, которая обеспечивает разумный компромисс между скоростью вычислений и сложностью алгоритма выбора. Данная стратегия основана на анализе системы на связность и оценке скорости распространения ограничений по переменным системы. Для оценки скорости распространения ограничений используется скорость изменения интервального решения по отдельным переменным. Вычислительные эксперименты демонстрируют достаточно высокую эффективность выбранного алгоритма. Зачастую вышеописанный алгоритм по качеству решения и скорости его нахождения превосходит аналоги интервальных методов Ньютона.

Отдельно рассматривается поиск решения для систем нелинейных уравнений малой размерности. Решение таких систем уравнений часто встречается в геометрическом моделировании (различные задачи на плоскости, задачи пересечения параметрических кривых в пространстве). Из всех алгоритмов практический интерес представляют те, в которых требуется минимальное время для нахождения всех решений уравнений. Для случая кратных решений характерно то, что условие малости невязки выполняется в брусе, длина ребра которого может быть много больше ε. Такие ситуации называются ползущим решением. Проблема оптимального нахождения этого бруса является важной задачей, особенно в вычислительных системах реального времени. Предложенный алгоритм поиска позволяет с высокой степенью эффективности гарантированно находить все области, в которых невязка мала.

Идея алгоритма состоит в следующем. Начальная область решения оценивается с помощью метода СР. С грубой точностью, порядка одной десятой максимальной ширины интервалов искомых переменных, запускается алгоритм бисекции в сочетании с методом СР. Это позволит быстро отсечь некоторые брусы, не содержащие решения В каждой полученной области решения ищутся не последовательно, а находятся первые решения по разным направлениям, которые составляют брус. Для этого бруса анализируется интервальное значения якобиана системы. Если якобиан знакопостоянен на всем брусе, то это значит, что отображение взаимно однозначно и в этом брусе возможен только один корень, который находится любым интервальным методом, например методом Кравчика. В другом случае, когда интервальное значение якобиана содержит нуль, для рассматриваемого бруса оценивается интегральная мера невязки по квадратурной формуле Гаусса. Если интегральная мера невязки не превышает заданной точности, то весь брус является решением, в противном случае проводится дополнительная его бисекция.

Разработанные вычислительные алгоритмы позволяют эффективно решать произвольные системы уравнений с ограничениями и интервальными данными, что позволяет использовать предложенную методику для решения прикладных задач.

Список литературы

1. Collavizza H., Delobel F., Rueher M. A note on partial consistencies over continuous domains // Proc. CP98, Pisa, Italy, October 26-30, 1998. – Maher M., Puget. J.-F. (Ed.). – P 147-161
2. Hansen E.A. A globally convergent interval method for computing and bounding real roots // BIT. – 1978. – Vol. 18. – P. 415-424.
3. Hansen E. Global Optimization Using Interval Analysis. – New York: Marcel, Dekker, 1992
4. Hansen E.A., Sengupta S. Bounding solutions of systems of equations using interval analysis // BIT. – 1981. – . – Vol. 21. – P. 203-211.
5. Krawczyk R. A class of interval Newton operators // Computing. – 1986. – Vol. 37. – P. 179-183
6. Moore R.E. Interval analysis. – Englewood Cliffs, N.-J.: Prentice Hall, 1966.
7. Neumaier A. Interval methods for systems of equation. – Cambridge University Press, 1990
8. Semenov, T. Kashevarova, A. Leshchenko, D. Petunin. Combining Various Techniques with the Algorithm of Subdefinite Calculations // Proc. of the 3rd Intern. Conf. on the Practical Application of Constraint Technology (PACT'97), London, England, April 1997. – P. 287-306
9. Wolfe M.A. A modification of Krawczyk's algorithm // SIAM Journal on Numerical Analysis. – 1980. – Vol. 17. – P. 376-379

Технология построения кооперативных решателей для решения сложных вычислительных задач

А. Е. Клейменов

Институт систем информатики СО РАН
Россия, 630090 г. Новосибирск, просп. Лаврентьева, 6
san@sib3.su

Аннотация. В статье описывается применение кооперативного подхода к решению задач математического моделирования имеющих сложную формулировку. Предлагаются разработанная технология для построения кооперативных решателей и программная среда обеспечивающая их создание. После описания архитектуры среды и ее текущей реализации приволятся примеры построения действующих кооперативных решателей эффективных при решении конкретного класса задач.

Введение

В настоящее время для решения прикладных вычислительных задач все чаще применяются решатели, основанные на интервальных методах распространения ограничений. Применение этих методов позволяет упрощать формулировки задач, решать задачи в нестандартных постановках, с неточными данными, объединять переменные разных типов и т.д. Однако эти методы все же не являются универсальными и существуют классы задач, для которых эти методы либо неприменимы вообще, либо имеют низкую эффективность (например, системы линейных алгебраических уравнений). Вместе с тем, объединение методов распространения ограничений со специализированными методами позволяет значительно увеличить эффективность и мощность каждого из методов. Поэтому построение кооперативных решателей, объединяющих различные методы и использующих современные вычислительные технологии (параллельность, дистрибутивность и т.д.), является очень перспективным и важным направлением в решении сложных вычислительных проблем.

Существует достаточно большое количество работ в области построения кооперативных решателей [6], [7]. Большинство из них либо рассматривают, какие методы следует использовать при построении кооперативных решателей, либо описывают некоторый конкретный кооперативный решатель [1], [9]. На наш взгляд, многие из уже созданных решателей и архитектур имеют весьма ограниченные возможности для их расширения другими методами либо используют очень жесткую схему вычислений. Поэтому в нашей работе основной целью является не разработка некоторого кооперативного решателя, а создание среды для разработки семейства кооперативных решателей и средств их построения. Эти средства не только не противоречат уже разработанным методологическим подходам, а предоставляют удобную и универсальную базу для их реализации. На базе этого подхода были созданы прототипы кооперативных решателей, которые показывают высокую перспективность данной разработки.

Статья организована следующим образом: во второй части даются основные определения и описываются принципы построения кооперативных решателей являющиеся базой разработанной технологии. В третьей части рассматривается програмная среда для построения кооперативных решателей: ее архитектура, возможности и описывается технология построения кооперативных решателей. В четвертой части приводятся примеры готовых кооперативных решателей. В заключении описываются некоторые реализационные моменты и направления дальнейших работ.

1 Основные определения и принципы построения кооперативных решателей

Для того, чтобы перейти к формальному определению понятий, связанных с кооперативными решателями, нам необходимо ввести ряд дополнительных определений, включая понятие модели, метода, кооперативности и т.д. Так как наше определение модели основано на терминологии задач удовлетворения ограничений, то необходимо также ввести основные определения, используемые в этой области искусственного интеллекта [10], [2].

1.1 Задачи удовлетворения ограничений

Ограничение — это предикат $C(x_1,\dots,x_n)$, на некотором множестве переменных, принимающих значения из заданных областей определения. При этом переменные, входящие в предикат, могут быть разных типов.

Таким образом, ограничения описывают множества допустимых значений переменных, и дают некоторую частичную информацию об интересующих нас переменных.

Задача удовлетворения ограничений — это тройка $P=(X, D, C)$, где

$X=\{x_1, \dots, x_n\}$ - набор переменных,

$D=\{D_1, \dots, D_n\}$ - области определения переменных,

$C=\{C_1, \dots, C_m\}$ - множество ограничений на этих переменных.

В качестве областей D_i используются множества целых чисел или других объектов, а также вещественные интервалы. В последнем случае значениями переменной являются подынтервалы исходного интервала и для работы с ними используются алгоритмы интервального анализа.

Вектор $(a_1, \dots, a_n)\in D$ называется **решением задачи удовлетворения ограничений**, если для всех ограничений $C_j(x_{j1}, \dots, x_{jk})\in C$ значения $C_j(a_{j1}, \dots, a_{jk})$ истинны (в этом случае также говорят, что кортеж $(a_{j1}, \dots, a_{jk})$ удовлетворяет ограничению C_j). Если такого вектора не существует, то решение задачи считается равным пустому множеству, а задача удовлетворения ограничений называется **несовместной**.

1.2 Модели и методы

Математическую модель мы отождествляем с понятием задачи удовлетворения ограничений, а ее **решением** называем построение по входной модели одной или нескольких новых моделей, которые удовлетворяют некоторым, заранее определенным требованиям. Такими требованиями могут быть, например, получение множества значений переменных D' в виде интервалов с заданной максимальной шириной и удовлетворяющих ограничениям C исходной модели; разделение входной модели на линейную и нелинейную части и т.д. Таким образом, решение модели есть преобразование

$M \to (M_1, M_2, \dots, M_n)$.

При этом, если соответствующая модели задача удовлетворения ограничений несовместна, то результирующие модели также считаются **несовместными**, т.е. их решением являются пустые множества.

Решение модели может производиться с помощью метода или решателя. **Метод** есть наименьшая неделимая сущность, обладающая свойством получения по данной модели одной или нескольких других моделей. Метод может работать только с входной моделью. Обобщением метода является **решатель**, который для выполнения преобразований может использовать один или несколько методов и других решателей. В простейшем случае решатель может просто вызывать один метод. Если решатель использует более одного метода, то он называется **кооперативным решателем**. В дальнейшем мы не будем делать различий между методом и решателем, рассматривая и тот и другой как некоторый преобразователь MS, на вход которого подается исходная модель и на выходе получается одна или несколько новых, то есть

$MS(M) = (M_1, M_2, \dots, M_n)$

Метод представляет собой сложный объект, который включает в себя непосредственно вычислительный алгоритм, каналы обмена данными (определяемые ниже), механизмы управления работой вычислительного алгоритма, а также ряд других компонент.

1.3 Кооперативность

Под **кооперативным взаимодействием** мы понимаем совместное решение частей (в общем случае пересекающихся) исходной задачи различными методами, где каждый из методов предоставляет результаты своих вычислений другим методам.

Кооперативное взаимодействие методов определяется **схемой вычислений** - некоторым ориентированным графом, узлами которого являются методы, а дугами - связывающие их информационные потоки. Схема вычислений полностью определяет ход решения задачи. Она содержит информацию о наборе используемых методов, порядке их запуска, способе и направлении обмена данными между методами. Фактически, схема вычислений описывает, откуда берется входная информация, как она преобразуется, и что должно получиться на выходе. Задание вычислительной схемы определяет вид кооперативного взаимодействия методов.

Передача данных между методами осуществляется посредством каналов. **Каналом** мы назовем некоторую сущность, имеющую один или несколько входов и выходов и способную получать данные на входах и передавать их на выходы. Каналы могут инкапсулировать некоторый сетевой транспортный протокол, что автоматически позволяет переносить решатель на распределенную архитектуру. Использование такой компоненты позволяет естественным образом осуществлять коммуникацию с методом: прием/передачу данных задачи и управляющих команд.

Из приведенных выше определений видно, что возможности кооперативного решателя напрямую связаны с возможностями составляющих его методов. Кроме того, с реализационной точки зрения, эффективность работы решателя зависит от средств, используемых для коммуникации методов, - позволяют ли данные средства проводить распределенные или параллельные вычисления, какова скорость передачи данных и т.д.

При разработке средств для построения кооперативных решателей большое значение имеет способ описания вычислительной схемы, лежащей в основе создаваемого решателя. Для этих целей существуют два подхода:

- Использование специальных языковых средств, синтаксис которых удобно и интуитивно понятно отражает как конфигурацию кооперативного взаимодействия методов, так и стратегию ее изменения в ходе решения задачи [8].
- Использование графических средств описания вычислительной схемы [5], [4].

Кроме описательных средств, для разработки кооперативных решателей очень важной является возможность пополнения набора используемых при построении решателя методов. Здесь можно выделить следующие два подхода:

- Дополнение набора используемых методов некоторыми уже существующими приложениями. При этом подходе должен быть создан механизм встраивания ранее разработанных вычислительных библиотек и/или вычислительных приложений. На данный момент существует большое количество готовых и очень эффективных в своих областях применения вычислительных библиотек и приложений, которые можно использовать для построения кооперативных решателей. Для обеспечения данного подхода, механизмы взаимодействия методов должны поддерживать работу встроенных приложений в различных операционных системах.
- Реализация нового метода на основе методов, имеющихся в библиотеке кооперативного решателя. При этом подходе используется некоторый внутренний язык реализации методов, при помощи которого можно либо реализовать метод "с нуля", либо использовать уже имеющиеся методы для создания нового.

Из всего вышесказанного следует, что средства для построения кооперативных решателей должны обеспечивать следующие возможности:

- возможность взаимодействия методов при решении задачи,
- интуитивно понятного описания вычислительной схемы,
- распараллеливания/распределения вычислений,
- контроля процесса вычислений,
- динамического изменения вычислительной схемы,
- пополнения библиотеки используемых методов,
- использования в вычислениях гетерогенных кластеров.

2 Среда для построения кооперативных решателей

Здесь мы опишем результат, полученный в ходе разработки нашего подхода к кооперативному решению сложных вычислительных задач. Этим результатом является создание библиотеки средств, объединяющей вычислительные компоненты и средства построения кооперативных решателей, которые удовлетворяют всем перечисленным в предыдущем разделе требованиям. Данная среда позволяет достаточно легко строить кооперативные решатели для конкретных классов задач или моделей.

2.1 Компоненты среды

В основе созданной нами среды лежат следующие основные части:

- Вычислительное ядро
- Модуль описания математической модели
- Модуль создания и исполнения кооперативных решателей

Вычислительное ядро среды представляет собой библиотеку методов и приложений, используемых для решения задач, языка описания моделей и средств перевода с этого языка в некоторые структуры, обеспечивающие решение задачи требуемыми методами. Все методы, встроенные в библиотеку, имеют общий интерфейс, что обеспечивает общий механизм организации их взаимодействия и запуска.

В настоящее время вычислительное ядро представлено решателем SibCalc [3]. Библиотека методов SibCalc включает в себя:

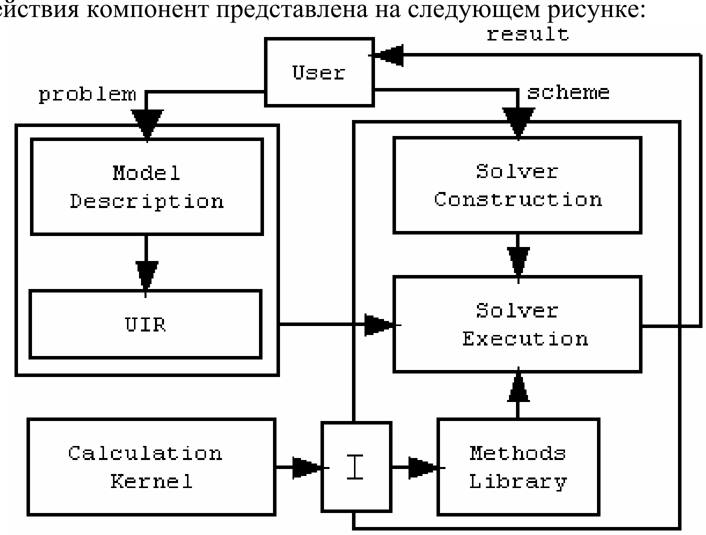
- алгоритмы распространения ограничений над конечными областями (AC-4, AC-5)4
- алгоритмы распространения ограничений над непрерывными областями (интервальный AC-3);
- интервальный метод Ньютона;
- методы решения систем интервальных алгебраических линейных уравнений;
- линейного программирования (интервальное расширение метода внутренних точек);
- большой набор методов поиска над непрерывными и конечными областями;
- автоматическое дифференцирование;
- ряд специализированных методов различных назначений.

Модуль описания математической модели предоставляет возможность описания математической модели в форме, максимально приближенной к общепринятой математической нотации. При помощи специального механизма данное декларативное представление модели трансформируется в универсальное внутреннее представление (UIR – Universal Internal Representation) задачи, которое используется всеми методами, участвующими в расчетах в качестве входных и выходных данных.

Модуль создания и исполнения кооперативных решателей предоставляет следующие средства:

- Единый интерфейс, который должен быть реализован методом, участвующем в кооперативном взаимодействии. Методы, реализующие данный интерфейс, составляют библиотеку методов системы.
- Языковые средства реализации новых методов - как с нуля, так и на основе имеющихся в библиотеке.
- Каналы, посредством которых производится взаимодействие методов и управление ими.
- Механизм, позволяющий задавать схему взаимодействия методов. Такая схема представляет собой декларативное описание кооперативного решателя и может рассматриваться как новый метод для использования наряду с другими методами при построении нового кооперативного решателя. Данный механизм позволяет создавать иерархические кооперативные решатели.
- Механизм, позволяющий распределять методы на гетерогенном вычислительном кластере, производить запуск кооперативного решателя и осуществлять контроль за вычислениями. В ходе работы кооперативного решателя набор участвующих методов и направления действующих информационных потоков могут динамически изменяться, что позволяет интерактивно влиять на процесс вычисления. В отличие от декларативного описания схемы вычислений, данный механизм представляет процедурное описание схемы поведения кооперативного решателя.

Общая схема взаимодействия компонент представлена на следующем рисунке:



2.2 Технология построения кооперативных решателей

После рассмотрения основных компонент среды, рассмотрим общую схему построения кооперативного решателя.

При возникновении практической необходимости в расчете некоторой математической модели, пользователь среды сначала формулирует ее на высокоуровневом специализированном языке.

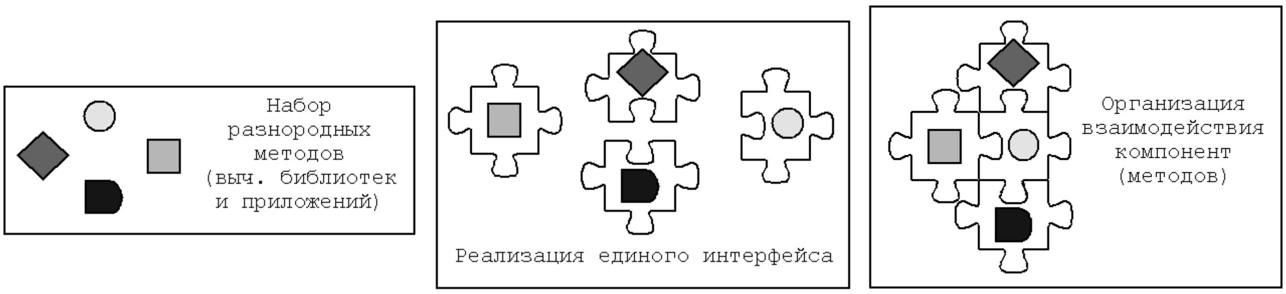
Далее разрабатывается стратегия расчета полученной математической модели, включающая в себя следующие этапы:

- декомпозиция математической модели на составные части, которые могут быть решены отдельными методами библиотеки и/или другими приложениями.
- поиск наиболее эффективных вычислительных средств, способных решать каждую из полученных частей задачи.
- схематичная разработка кооперативного взаимодействия вычислительных средств (создание вычислительной схемы кооперативного взаимодействия).

После того, как стратегия решения задачи выбрана, выбирается набор методов, которые будут использованы для кооперативного решения задачи. При этом, для каждого вычислительного средства, необходимого для решения задачи и не входящего в текущую библиотеку методов среды, реализуется единый интерфейс методов библиотеки. На основе выбранных методов и разработанной схемы их кооперативного взаимодействия при помощи механизмов, предоставляемых модулем создания и исполнения кооперативных решателей, конструируется решатель. Общая схема данного этапа представлена на следующем рисунке:

Следующим этапом является отладка кооперативного решателя. На этом этапе совершенствуется выбранная стратегия расчета математической модели с целью достижения максимальной эффективности при ее решении.

Затем производится отторжение готового решателя и его дальнейшее самостоятельное использование.



3 Примеры кооперативных решателей

В данном разделе мы проиллюстрируем общую схему создания кооперативного решателя на конкретных действующих примерах.

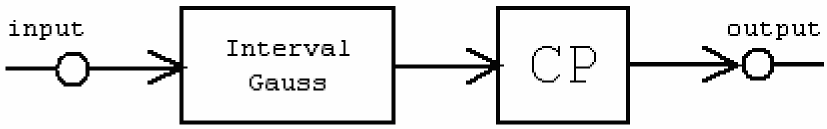
3.1 Комбинирование методов для решения нелинейной задачи с большими линейными подсистемами

Рассмотрим пример решателя, производящего расчет нелинейных математических моделей с большими линейными подсистемами.

На встроенном языке математического моделирования сформулируем задачу отыскания всех решений линейной системы, лежащих в брусе, ограниченным шаром заданного радиуса:

```
const int N, M;
real A[N,M], b[N], x[M];
A*x = b;
sum i in <1..M> {x[i]*x[i]} < R^2;
```

Данная модель параметризуется размерами линейной задачи и радиусом шара. Хорошо работающий на нелинейных задачах метод интервального распространения ограничений (*CP*) имеет на моделях данного типа крайне низкую эффективность. Это служит поводом к созданию кооперативного взаимодействия специализированного метода для решения линейных систем – интервального метода Гаусса (*Interval Gauss*), с методом *CP*. Вычислительная схема такого взаимодействия может быть представлена следующим образом:



Вышеприведенное схематическое представления решателя имеет следующий аналог описания на внутреннем языке системы (в качестве которого используется скриптовый язык Python [11]):

```
from scf.methods import IGauss, CP

gauss = IGauss()
cp = CP()
gauss.output.connect(cp.input)
```

Инициализация решателя исходными данными, запуск и получение результата производится следующим образом:

```
model = Model("example_1.uir");

gauss.input(model)
gauss.run()
cp.run()

print "Solution:", cp.output()
```

Это пример простейшего последовательного кооперативного взаимодействия, называемого `sequential evaluation`. В нижеследующей таблице приведены времена расчета тестовых задач различной размерности при помощи одного метода *CP* и его кооперативного взаимодействия с методом *Interval Gauss*.

Размерность	*CP*	*CP+Interval Gauss*
10x10	2.04	0.012
20x20	17.39	0.320
50x50	1240.64	3.2

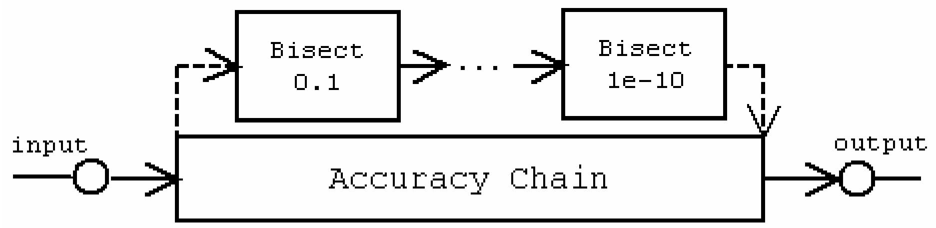
3.2 Решатель, реализующий последовательную кооперативность с изменяющейся точностью вычислений

Здесь мы ищем глобальный оптимум вещественной функции, параметры которой связаны набором дополнительных ограничений выраженных равенствами и неравенствами:

```
/* minimize objf */
real x1, x2, x3, x4, objf;

objf = x1^2 + 2.*x2^2 + 3.*x3^2 + 4.*x4^2;
x1^2 + 3.*x2^2 - 4.*x3*x4 - 1. = 0.;

x4 = max(x4, 1.e-4);
sqrt(5.*abs(x1/x4)) + sin(x2) - 1. = 0.;

(x1 - x2 - x3 + x4)^2 - 1.=0;
x2*x3 + (x1 - sin(x4))^2 - 3.=0;

-5 <= x1; x1 <= 3;
-3 <= x2; x2 <= 3;
-5 <= x3; x3 <= 4;
-6 <= x4; x4 <= 2;
```

Мы ищем минимум функции `objf` при помощи разделения области ее значений слева направо и применяя метод *CP* на соответствующих областях определения. Данный алгоритм реализуется методом *Bi-*

`sect`. Нам необходимо найти значение минимума с точностью 10^-10. Если сразу решать задачу с требуемой точностью, используя только бисекцию, то ее решение находится за 36 секунд. Но если решать задачу с постепенным увеличением точности и последующей бисекцией, так как показано на следующей вычислительной схеме:



то для ее решения требуется всего 2.6 секунды.

Приведенная выше вычислительная схемы соответствует следующему методу, описанному на внутреннем языке:

```
from scf.methods import Scheme, Bisect

class AccuracyChain(Scheme):

    def __init__(self, acc, target, *args):
        self.accuracy = acc
        self.target = target
        Scheme.__init__(*args)

    def staff(self):
        accuracy = 0.1
        channel = self.input
        while accuracy >= self.accuracy:
            bisect = Bisect()
            bisect.target(self.target, Bisect.MIN|Bisect.LEFT)
            bisect.accuracy(accuracy)
            channel.connect(bisect.input)
            channel = bisect.output
            self.append(bisect)
            accuracy *= 0.1
        channel.connect(self.output)
```

Данное описание демонстрирует создание кооперативного решателя, допускающего дальнейшее переиспользование. Создание экземпляра решателя, его инициализация, запуск и получение результата производится следующим образом:

```
from scf.methods import AccuracyChain

model = Model("example_2.uir")
chain = AccuracyChain(1e-10, "objf")
chain.input(model)
chain.run()
print "Solution:", chain.output()
```

5 Заключение

В статье была представлена разработанная нами среда для построения кооперативных решателей. Эта среда функционирует на различных платформах и используется нами для создания прототипов кооперативных решателей различного назначения. Другим применением рассматриваемой среды является исследование эффективности применения получаемых решателей для решения практических задач, имеющих сложную математическую формулировку.

При разработке среды был использован Python - многофункциональный, развивающийся скриптовый язык

[11]. Мощь скриптовых языков при переиспользовании готовых компонент позволяет легко встраивать в систему новые вычислительные библиотеки. Так, при встраивании решателя SibCalc в качестве вычислительного ядра, для всей библиотеки решателя был реализован Python интерфейс, что обеспечило возможность ее использования. Подобным образом среду можно расширить другими приложениями.

В процессе реализации коммуникационных каналов было испытано много различных средств: CORBA, RPC, MPI, XML-RPC, ACE, имеющих как преимущества, так и недостатки [12], [13], [14]. В конечном итоге, используя стандартные средства, предоставляемые системой Python, мы реализовали собственную, не зависящую от платформы высокопроизводительную библиотеку.

Конечным результатом работы системы является набор скриптов Python, которые при помощи программ типа py2exe трансформируются в исполняемые модули, не требующие инсталяции Python environment на вычислительных станциях пользователя.

Данная работа не является законченным продуктом, и мы продолжаем работу над ней. В частности, мы занимаемся повышением вычислительной эффективности создаваемых решателей, повышением удобства работы пользователя со средой и с конечными решателями. В рамках этих разработок мы планируем следующие действия по развитию среды.

1. Создание средств отладки работы вычислительной схемы.
2. Дальнейшая разработка визуального представления вычислительной схемы.
3. Разработка механизма идентификации методов для создания интеллектуальных построителей вычислительных схем.
4. Пополнение библиотеки методов и поиск наиболее эффективных их комбинаций.

Список литературы

1. F. Arbab and E. Monfroy. Using coordination for cooperative constraint solving. In Proceedings of the 1998 ACM Symposium on Applied Computing; Special Track on Coordination Models, Languages and Applications, Atlanta, Georgia, February-March 1998. ACM.
2. Bartak, R. Constraint Programming: In Pursuit of the Holy Grail. In Proceedings of the Week of Doctoral Students (WDS99), Part IV, MatFyzPress, Prague, June 1999, pp. 555-564.
3. Kleymenov, D. Petunin, A. Semenov, I.Vazhev. A Model of Cooperative Solvers for Computational Problems// Proceedings of the 4th International Conference PPAM 2001, Poland, September, LNCS, 2328. – P.797-802.
4. Kleymenov, D. Petunin, A. Semenov. GMACS: the general-puprose module architecture for building cooperative solvers. In Proceedings of the 1999 ACM Symposium on Applied Computing, ACM Press (1999) 19-24
5. Kurt Lichtner, Paulo Alencar, and Donald Cowan. Using View-Based Models to Formalize Architecture Description. Third International Software Architecture Workshop, ISAW-3, Oct 1998.
6. Monfroy, E.: Solver Collaboration for Constraint Logic Programming. PhD Thesis. Centre de Recherche en Informatique de Nancy INRIA-Lorraine 1996.
7. Monfroy, E.: From Solver Collaboration Expressions to Communicating and Coordinated Agents. In Proceedings of the International Workshop for Object-oriented and Constraint Programming for Time Critical Applications, COTIC'99, Lisbon, Portugal, 1999.
8. G.A. Papadopoulos and F. Arbab. Coordination Models and Languages. In M. Zelkowitz, editor, Advances in Computers, The Engineering of Large Systems, volume 46. Academic Press, August 1998.
9. M. Rueher. An Architecture for Cooperating Constraint Solvers on Reals In Podelski A., editor, Constraint Programming: Basics and Trends, vol. 910 of LNCS, pages 231-250, Springer, 1995.
10. E. Tsang. Foundations of Constraint Satisfaction. Academic Press, Essex, 1993.
11. Official website for the Python language. http://www.python.org
12. Home page of ACE (The ADAPTIVE Communication Environment) http://www.cs.wustl.edu/~schmidt/ACE.html
13. XML-RPC Home Page. http://www.xmlrpc.com/
14. MPICH-A Portable Implementation of MPI. http://www-unix.mcs.anl.gov/mpi/mpich/

К вопросу сравнения по эффективности интервальных алгоритмов решения систем нелинейных уравнений

М. Б. Бозоров, Б.Х.Бердиев

Навоийский государственный горный институт
Узбекистан, Навои
mamurjon@mail.ru

Обсуждаются общие положения, лежащие на основе сравнения эффективности интервального решения систем нелинейных уравнений. Выделяются параметры, определяющие особенности каждой системы уравнений; кратко рассмотрены некоторые интервальные методы решения систем нелинейных уравнений и критерии сравнения. Даются рекомендации, касающихся выбора метода решения основных типов проблем.

Введение

Построению и исследованию интервальных методов решения систем нелинейных алгебраических уравнений вида $f(x) = 0$ посвящено большое число работ(см., например,[1]–[3] и приведенную в них библиографию). Среди работ последних лет отметим исследований [4]–[7]. К сожалению, в настоящее время для решения таких систем нет такой же общей теории, какой является теория интервальных методов для решения систем линейных алгебраических уравнений. Эквивалентность задач решения системы нелинейных уравнений и безусловной оптимизации некоторых функционалов позволяет использовать для решения нелинейных систем большую группу методов численной оптимизации. Все это создает значительные трудности как при выборе алгоритма для решения конкретной системы нелинейных уравнений, так и при разработке и исследованию новых методов.

Создание программных средств для интервального решения систем нелинейных уравнений предполагает разработку комплекса программ, реализующих различные интервальные алгоритмы. В связи с этим возникает необходимость систематизации имеющихся методов, оценки их эффективности на классах задач и разработки способов описания алгоритмов с единой точки зрения.

В данной работе делается попытка обсудить общие положения, которые могут, лежат на основе сравнения по эффективности интервального решения систем нелинейных уравнений. Выделяются параметры, определяющие особенности каждой системы уравнений; кратко рассмотрены некоторые интервальные методы решения систем нелинейных уравнений и критерии сравнения. В данном докладе мы ограничимся лишь рассмотрением интервальных методов типа Ньютона и методов Кравчука. Тщательное сравнение этих методов с другими интервальными методами могут стать предметом отдельной большой работы.

Общие положения

В последнее время опубликован ряд работ, посвященных разработке и реализации интервальных методов решения систем нелинейных алгебраических уравнений. Однако, в большинстве работ при описании нового численного метода в лучшем случае проводится сравнение эффективности программ решения рассматриваемой задачи на двух–трёх примерах (иногда простейших, иногда весьма специального вида) с одним из традиционных методов. С другой стороны, основные результаты этих исследований, вообще говоря, не применимы к ЭВМ используемых в странах СНГ, поскольку компьютеры разного типа различаются по точности представления чисел, скорости, различной структуры команд и, кроме того, быстрота решения задачи в существенной мере зависит от применяемого языка программирования и типа транслятора с этого языка.

Имеется ещё три очевидные причины, по которым необходимо проводить тестирования методов. Прежде всего, перед решением практической задачи надо быть твердо быть уверенным, что в программе отсутствуют ошибки. Другая причина заключается в определении степени надёжности программы, т.е. в её способности давать верные(или хотя бы разумные) результаты при правильно заданных параметрах решения и информацию о невозможности решения в противном случае. Третья причина — определения степени эффективности программы для задач данного класса. Причем, когда говорится о таких свойствах программ как надёжность и эффективность, то подразумевается обычно решение различных типов задач с несколькими значениями заданной точности.

В основе сравнения по эффективности решения систем нелинейных алгебраических уравнений лежат следующие три «составные части»: а) задачи, подлежащие сравнению (вид конкретных систем и начальные параметры); б) тестируемые методы (информация о методе решения и реализующем его алгоритме) и в) критерии сравнения эффективности. Все эти составные части являются важными, а именно они определяют результат исследования и от них зависят о целесообразности применения алгоритма (и реализующей его программы) к решению практических задач.

Необходимо также отметить, что существуют определённая «дистанция» между понятием «метод решения» и его реализаций в виде программы. Это различие существует потому, что каждый метод допускает много различных способов своей программной реализация, зависящей от языка программирования, квалификации программиста и т.п. Поэтому сравнение различных методов решения происходит не в «чистом виде», а опосредованное через используемые программы, некоторые из них могут быть далеки от совершенство.

Параметры задачи

Перед рассмотрением типов решаемых задач следует выделить параметры, которые в достаточно полной мере определяют специфические особенности каждой системы уравнений и предъявляют определенные требования к методу решения. В общем виде можно записать так (**p** — от английского слово **«Problem»**):

$$p = (F, X^{(0)}, \partial F/\partial X)$$

Первые два параметра определяют как бы математическую часть задачи: $F = (F_1, F_2, ..., F_n)$ — вид интервального расширения левой часты системы; $X^{(0)}$ — начальный интервальный вектор из области определения f; наконец последний параметр — $\partial F/\partial X$ также является важной характеристикой задачи (интервальная матрица-Якобиан), поскольку он определяет степень сложности задачи, кроме того, многие алгоритмы требуют вычислений $\partial F/\partial X$ в ряде интервальных значениях.

При описании тестовых примеров необходимо указывать причины рассмотрения именно этих примеров. Нам кажется, что должны быть подвергнуты тестированию следующие типы задач : одна из модельных систем, характерных для определенной области науки и техники, одна из относительно «трудных» задач (реальный пример из практики), а также тестируемые ранее задачи другими авторами. Нами в данной работе рассмотрены 20 модельных систем нелинейных алгебраических уравнений, взятых из различных областей науки и техники, которые для различных значений X_0 решаются различными методами.

Некоторые интервальные методы решения систем нелинейных алгебраических уравнений

Интервальный аналог метода типа Ньютона был предложен Р. Муром. Идея метода заключается в следующем.

Рассмотрим задачу нахождения решения системы

$$f_i\ (x_1, x_2, ..., x_n) = 0, i = \overline{1,n} \quad (1)$$

О функциях f_i предположим, что они непрерывно дифференцируемо при $x \in A = (A_1, A_2, ..., A_n)$ **,** причем $\partial f_i/\partial x_j$ обладают монотонные по включению интервальными расширениями $F_{ij}(x)$**,** определенными при $x \in A$**.** Через $J(X)$ обозначим матрицу с элементами $F_{ij}(x)$**.**

Предложенный Р.Муром интервальный аналог метода Ньютона для системы (1) определяется равенством

$$X^{(k+1)} = N(X^k) \cap X^k, k = 0,1,..., n \quad (2),$$

где оператор $N(X^k) = m(X^k) - B(X^k) \cdot f(m(X^k)$, $f = (f_1, f_{2,}, ..., f_n)$, а матрица $B(X^{(k)})$ содержит все обратные к матрицам из $J(X^{(k)})$. Если $\xi = (\xi_1, \xi_2, ..., \xi_n)$ удовлетворяет системы (1), то из включение $\xi \in X^k$ следует, что $\xi \in N(X^{(k)})$ тем самим $\xi \in X^{(л+1)}$. Таким образом, если $X^{(0)}$ такой, что $\xi \in X^{(0)}$, то $\xi \in X^{(k)}$ при всех k.

В дальнейшем, другими учеными, были исследованы различные модификаций интервального аналога метода Ньютона. Кроме того, в работах Р. Мура и его учеников показано, как применить интервальные методы для проверки существование и единственность решение нелинейных систем в заданной области пространства R^n, а также для проверки сходимости соответствующих итерационных процессов.

Среди других методов следует особо выделить интервальный метод Р. Кравчука [8]. Рассмотрим оператора K предложенный Р.Кравчуком и задаваемый равенством

$$K(X, x, y) = x - yf(x) + (e - yF^{'}(X))(X - x), \qquad (3),$$

$X \subset A,\ x, f \in R^{(n)}, y \in R^{n \times n}, F^{'}(X) \in I(R^{n \times n})$. На практике в качестве $x^{(k)}$ часто берут $m(X^{(k)})$, а $y^{(k)} \approx (m(F^{'}(X^{(0)})))^{-1}$ или $y^{(k)} \approx (m(F^{'}(X^{(k)})))^{-1}$. Отметим, что в отличие от интервального метода Ньютона метод Кравчука не требует обращения интервальных матриц.

Р. Кравчук определяет последовательность интервалов $X^{(k)}$ равенством

$$X^{(k+1)} = K(X^{(k)}, x^{(k)}, y^{(k)}) \bigcap X^{(k)},$$

где $x^{(k)} \in X^{(k)}$. Тогда $\xi \in X^{(k)}$ при всех k=1,2,….

В последние годы Р. Кравчуком и другими авторами были предложены и исследованы различные модификации оператора Кравчука. Например, в работах [9]-[12] , были предложены следующие операторы (использованные обозначения авторами были приняты нами):

$$K(X) = \check{x} - a(X)G(\check{x}) + r(X)(EradX), \qquad (4)$$

$$N(X) = \check{x} - a(X)G(\check{x}) + q(X)(E(a(X)G(\check{x}))), \qquad (5)$$

$$NV(X) = \check{x} - [e - q(X), e + q(X)](a(X)G(\check{x})) \qquad (6)$$

Далее, если L(X) — интервальная матрица Липшица положительно — определена, то оптимальный оператор - O(X) определяется равенством

$$O(X) = [\underline{x} - \bar{l}^{-1}(X)\bar{g}(\underline{x}), \bar{x} - \bar{l}^{-1}(X)\underline{g}(\bar{x})] \qquad (7),$$

где приняты следующие обозначения: $X = [\underline{x}, \bar{x}] \in IR^{(n)}$, $\check{x} = mid(X) = (\underline{x} + \bar{x})/2$,
$rad(X) = (\bar{x} - \underline{x})/2$, $a(X) = (midL(X))^{-1}$, $r(X) = |a(X)| \cdot radL(X)$
$q(X) = (e - r(X))^{-1} \cdot r(X)$, $E = [-e, e]$, e — единичная матрица, $G(x) = [\underline{g}(x), \bar{g}(x)] : D \in IR^{(n)} \to IR^{(n)}$.

Все эти операторы дают последовательности вложенных интервалов $X^{(0)} \supseteq X^{(1)} \supseteq X^{(2)} \supseteq ......$, содержащих решения системы (1).

В данной версии работы из большего количества различных методов и алгоритмов выбраны шесть алгоритмов, которые дали наиболее удовлетворительные результаты.

Критерии сравнения

Чем больше критериев сравнения, тем труднее сделать выводы — какой метод (или программа) лучше для решения задач данного типа. Вид задачи и её интервальное расширение определяет очень многое, однако и алгоритм решения и его программная реализация играют не последнюю роль. Нам кажется, что основными критериями сравнения по эффективности работы различных интервальных алгоритмов решения систем нелинейных уравнений следует считать: число вычислений правых частей и время счета, количество итераций, полуширина и середина получаемых интервальных решений. Мы не включили параметр количества вычислений $(\partial F/ \partial X)^{-1}$ в число основных критериев эффективности потому, что для вычисления этого параметра существуют и разными авторами реализуются различные способы обращения матриц(которые, кстати, также нуждаются в отдельном тестирования на различных задачах в диапазоне заданной точности). Кроме того, в некоторых методах нет необходимости вычислять $(\partial F/ \partial X)^{-1}$, а в тех алгоритмах, где нужно — число необходимых вычислений сильно зависит как от метода решения, так и от самой задачи; тогда как время обращения $(\partial F/ \partial X)^{-1}$ входит в суммарное время решения.

Список литературы

1. Moore R. Methods and Applications of Interval Analysis.- SIAM, Philaladelphia,1979.
2. Калмыков С.А., Шокин Ю.И, Юлдашев З.Х. Методы интервального анализа.- Новосибирск: Наука,1986.
3. Алефельд Г., Херцбергер Ю. Введение в интервальные вычисления. — М.: Мир, 1987.
4. Neumaier A. Interval Methods for systems of Equations.- Cambridge: Cambridge Universite Press,1990.
5. Kearfott R. Rigorous Global Search: Continuous Problems.- Dordrecht:Kluwer,1996.
6. Шарый С.П. Интервальные алгебраические задачи и их численное решение. Диссертация на соискание учёной степени доктора физико- математических наук.- Новосибирск: ИВТ СО РАН, 2000.- 332 с.
7. Krawczyk R. Newton-Algorithmen zur Bestimmung von Nullstellen mit Fehlerschranken// Computing.-1969.-vol. 4.- N 3.- p.187-201.
8. Krawczyk R. Properties of interval operators// Computing.-1986.-vol. 37,-p.227-245.
9. Krawczyk R. On interval operators obtained by Splitting the Lipschitz Matrix// Freiburger Interval-Berichete.-1987.-87/1.
10. Wolf M.A. A modification of Krawczyk’s algorithms // SIAM J. Numer. Anal.-1980.-17,376.
11. Rump S.M. Solution of linear and nonlinear algebraic problems with sharp guaranteed bounds// Comp. Suppl.-1984.-Vol.5.-p.147-168.

Задачи практической (интервальной) устойчивости с заданной областью предельных отклонений

А.Н. Рогалев

Институт вычислительного моделирования СО РАН
Россия, 660036 г. Красноярск, Академгородок
ran@krsk.info

Аннотация В докладе описано применение гарантированных оценок множеств решений в задачах практической устойчивости. Получаемые в результате работы алгоритма числовые значения гарантированных границ множеств решений учитывают влияние на решения постоянно действующих возмущений, представляющих многозначные функции. Знание таких границ позволяет формулировать и обосновывать математически строгие результаты, касающиеся практической устойчивости.

1 Постановки задач практической устойчивости

В классической постановке задачи об устойчивости движения предполагается, что возмущающие силы вносятся только в начальный момент времени, далее возмущенные движения осуществляются при действии тех же сил, учтенных при описании невозмущенного движения. Однако свойства устойчивости по Ляпунову и даже асимптотической устойчивости недостаточны для практической устойчивости [4], названной также технической устойчивостью или устойчивостью при постоянно действующих возмущениях [9], [3], [4], [8]. Все эти определения описывают свойство ограниченности решений на конечном интервале. Длина этого интервала времени выбирается в зависимости от смысла поставленной задачи.

Положение равновесия может быть математически неустойчиво, тем не менее, система может совершать колебания в достаточной близости от положения равновесия, так что ее режим является вполне приемлемым. Многие авиационные и технические установки ведут себя именно таким образом. Подобные системы функционируют в течение конечного промежутка времени. При этом представляет интерес не только факт устойчивости или неустойчивости, но и количественные оценки поведения этих систем, а также приемлемость этих оценок в реальных условиях. Первые постановки задач об устойчивости на конечном промежутке времени принадлежат Четаеву Н.Г. [21], Моисееву Н.Д. [9], Каменкову Г.В. [3]. Эти задачи получили дальнейшее развитие как теория практической или технической устойчивости в монографиях Ла Салля Ж.П. – Лефшеца С., Мартынюка А.А., Матросова В.М.

В данной работе оцениваются области решений при конечных постоянно действующих возмущениях при помощи гарантированных методов, основанных на аппроксимации оператора сдвига вдоль траектории.

Пусть рассматривается система обыкновенных дифференциальных уравнений

$$\frac{dy}{dt} = F(t, y),\ t \geq 0. \tag{1}$$

Положение равновесия расположено в начале координат $F(t, y) \equiv 0$ при всех $t \geq 0$. Возмущенная система

$$\frac{dy}{dt} = F(t, y) + p(t, y),\ t \geq 0. \tag{2}$$

Можно привести несколько типов постоянно действующих возмущений, встречающихся в теории устойчивости движения.

- Первый вид функций возмущения можно назвать классическими постоянно действующими возмущениями. Функции $p(t, y)$ практически неизвестны. Предполагается, что они удовлетворяют общим условиям существования решений системы (2) и определенным условиям малости и при этом не обращаются в нуль при $y = 0$. Такие возмущения легко моделировать с помощью интервальных величин. Возмущения $p(t, y)$ могут быть малы в среднем.
- Возмущения $p(t, y)$ — это известные функции переменных t, y, быть может более высокого порядка малости по сравнению с функциями $f(t, y)$, обращающиеся в нуль при $y = 0$ и удовлетворяющие условиям ограниченности.

– Возмущения $p(t, y)$ — это функции, такие, что их явные выражения практически неизвестны, являются малыми величинами, и $p(t, y) = 0$ для любого $t \geq t^0$.

Наряду с системой (2) рассмотрим систему (1). Пусть невозмущенное движение $y = 0$ этой системы обладает некоторым свойством $\mathcal{P}$. Требуется выяснить ограничения, налагаемые на систему (1) и возмущения $p(t, y)$, при которых свойство $\mathcal{P}$ "деформируется", оставляя движение $y = 0$ системы устойчивым. Очевидно, что система (1) с "сильно устойчивым"к разрушению невозмущенным движением $y = 0$ допускает более широкие ограничения на функции постоянно действующих возмущений $p(t, y)$.

Чтобы сформулировать определение практической устойчивости, надо предположить, что область начальных отклонений характеризуется множеством $S_0(t^0) \subset R^n$, изменяющимся при выборе $t^0 \in T$. Область последующих отклонений оценивается множеством $S(t)$, для которого примем следующие предположения:

$$S_0(t^0) \subset S(t),\ \forall t \in T,\ \partial S_0 \cap \partial S = \oslash,\ \forall t \in T.$$

Для этих задач постоянно действующие возмущения оцениваются допустимым множеством $P(t)$. Воспользуемся понятием практической устойчивости согласно [2], [5].

Определение 1 . *При заданных оценках областей $S_0(t^0), S(t), P(t)$ невозмущенное движение $y(t)$ системы (1) называется устойчивым при малых постоянно действующих возмущениях (далее ПДВ), если для любых $t^0 > 0$, $y(t^0) \in S_0$ и возмущениях $p(t) \in P(t)$ решение $y(t)$ остается внутри области $S(t)$, то есть $y(t) \in S(t)$ при всех $t \in T$.*

Условие малости постоянно действующих возмущений может быть заменено на непрерывную зависимость двух множеств, включающих возмущения, что расширит определение 1. Пусть выбраны число $\delta > 0$ и два множества Q и Q_0 в пространстве $\mathbb{R}^n$. Считаем, что Q — замкнутое ограниченное множество, содержащее начало координат, а Q_0 — подмножество множества Q. Пусть y^p — решение возмущенной системы (2), удовлетворяющей начальному условию $y^p(t^0) = y^0$. Далее P — множество всех возмущений p, удовлетворяющих условию $\| p(t, y) \| < \delta$ для всех $t \geq 0$ и всех y.

Определение 2 . *Невозмущенное движение $y = 0$ системы (1) назовем практически (интервально) устойчивым при постоянно действующих возмущениях (ПДВ), если для каждого возмущения $p \in P$, для любых $t^0 > 0$ и интервала начальных данных $\mathbf{Y}^0 \in Q_0$ гарантированная оценка $\mathbf{Y}$ множества решений $\{y^p(t)|_{y^0} \in \mathbf{Y}^0\} \subset \mathbf{Y}$ остается в множестве Q при всех $t \geq 0$.*

Понятие практической устойчивости включает в себя число δ и два множества Q и Q_0. Множество Q является множеством допустимых состояний системы, описываемой ОДУ. Если решение y^p системы (2) в момент времени t находится внутри Q, то система работает в это момент надежно. Подмножество Q^0 является совокупностью начальных положений.

Интервальная (практическая) устойчивость означает равномерную ограниченность решений относительно множества начальных значений Q_0 и совокупности P возмущающих воздействий. Однако для практической устойчивости требуется не только существование ограничивающей постоянной для решений, но и чтобы эта постоянная имела значения, такие, что решения, начинающиеся в множестве Q_0, все время оставались в Q.

Система ОДУ с постоянно действующими возмущениями (2) может быть сведена к задаче следующего вида

$$\frac{dy}{dt} = g(t, y),\ t \geq 0, \tag{3}$$

где $y(t^0) \in \mathbf{Y}^0,\ g(t, y) \in \mathbf{G}(t, \mathbf{Y})$. Обозначим множество решений этой задачи как некоторую многозначную функцию

$$\mathcal{Y}(t) \supseteq \cup_{y^0 \in \mathbf{Y}^0} y(t).$$

Во многих случаях вычисление значений множества решений эффективнее и удобнее проводить, если функция заменена на ее интервальное расширение $\mathbf{Y}(t) \supseteq \mathcal{Y}(t)$.

Предложенные для количественного оценивания решений и исследования устойчивости алгоритмы [11], [17], [31] опираются на возможность в любой точке t^k построить символьную формулу решения и вычислять согласно этой формуле включения множеств решений, появляющихся при возмущении правых частей и начальных данных.

Определение 3 . *Символьный метод (аналитический метод) – запись метода преобразования символьной информации (символьных формул) на языке математического анализа. В дальнейшем при записи символьных формул, аппроксимирующих оператор сдвига вдоль траектории, допускается включение в них числовых констант с отложенным выполнением арифметических действий над ними.*

Выполнение гарантированных методов, основанных на аппроксимации оператора сдвига вдоль траектории, разделено на два этапа.

На первом этапе, происходит построение (запись) символьных формул приближенных решений как векторных функций

$$\mathcal{S}^n(\mathcal{Y}^0) \circ \mathcal{S}^{n-1}(\mathcal{Y}^0) \circ \ldots \mathcal{S}^1(\mathcal{Y}^0),$$

где вектор $\mathcal{Y}^0(t) = (\mathcal{Y}_1^0, \mathcal{Y}_2^0, \ldots, \mathcal{Y}_n^0)$ – вектор начальных значений, рассматриваемых как символьные величины. В работе [17] введено следующее определение.

Определение 4 . *Пусть $E_{i,h}$ – последовательность нормированных пространств, возможно зависящих от параметра h, принадлежащего множеству P; $\mathcal{F}^i(i = 1, \ldots, n-1)$ – последовательность символьных формул непрерывных отображений F^i, таких что соответствующие функции (числовые оригиналы символьных формул) F^i определены на прямом произведении $E_1 \times E_2 \times \ldots \times E_i$, отображают это произведение в пространство E_{i+1} и задают зависимость между значениями решения в каждой точке области определения и начальными значениями. Тогда результат последовательного исполнения преобразований формул*

$$\begin{aligned}
&\mathcal{Y}^1 = \mathcal{F}^1(t^0, t^1, \mathcal{Y}', \mathcal{Y}^1, b) = \mathcal{S}^1(\mathcal{Y}^0),\\
&\mathcal{Y}^2 = \mathcal{F}^2(t^0, t^1, t^2, \mathcal{Y}^0, \mathcal{Y}^1, \mathcal{Y}^2, b) = \mathcal{S}^2(\mathcal{Y}^0) \circ \mathcal{S}^1(\mathcal{Y}^0, b),\\
&\ldots\\
&\mathcal{Y}^i = \mathcal{F}^i(t^0, \ldots, t^i, \mathcal{Y}^0, \mathcal{Y}^1, \ldots, \mathcal{Y}^i, b) = \mathcal{S}^i(\mathcal{Y}^0, b) \circ \ldots \circ \mathcal{S}^1(\mathcal{Y}^0, b),\\
&\ldots\\
&\mathcal{Y}^m = \mathcal{F}^m(t^0, \ldots, t^m, \mathcal{Y}^0, \mathcal{Y}^1, \ldots, \mathcal{Y}^m, b) = \mathcal{S}^m(\mathcal{Y}^0, b) \circ \ldots \circ \mathcal{S}^1(\mathcal{Y}^0, b)
\end{aligned} \tag{4}$$

будет называться символьной формулой метода сдвига вдоль траектории решения системы ОДУ.

Очевидно, что для большинства методов $E_i = R^i, i = 1, 2, 3, \ldots, m-1$, и $E_m = R^m$, где R^m есть m-мерное евклидово пространство, и $\mathcal{F}_i$ – есть формула отображения, основанного на элементарных арифметических операциях.

На втором этапе определяется гарантированная оценка глобальной ошибки приближенного решения и вычисляется числовое значение границ множества приближенных решений, осуществляя некоторый алгоритм вычисления интервального расширения для формулы $\mathbf{S}_h(t, \mathbf{Y}^0) = \mathrm{val}\,(\mathcal{S}_h(t, h, \mathcal{Y}^0))$. Оценка глобальной ошибки получается как включение решения уравнения для глобальной ошибки, выписанного в некоторой промежуточной точке интервала $[y(t_i), y(t_{i+1})]$. Для нахождения этой оценки предварительно строится "коридор"точных решений [18], а затем определяется интервальная оценка, как формула решения этого уравнения. Таким образом, в методе не накапливается сумма интервалов, включающих глобальную ошибку, то есть устраняется влияние так называемого "wrapping"эффекта, проявляющегося практически для всех известных интервальных и двусторонних методов решения ОДУ с неточно заданными данными [25], [26], [24], [27],[28] [29],[30],[23] .

В завершение работы алгоритма к множеству (значению многозначной или интервальной функции) $\mathbf{S}_h(t, \mathbf{Y}^0)$ добавляется величина, задающая гарантированные границы глобальной ошибки. В итоге будет получена требуемая гарантированная оценка множества точных решений $\mathcal{Y}(t) \subseteq \mathbf{Y}(t)$.

Будем обозначать гарантированный метод решения задачи (3), основанный на аппроксимации оператора сдвига вдоль траектории и нахождению его символьной формулы, аббревиатурой $\mathcal{IMSHT}$.

В качестве символьной формулы $\mathcal{F}_i$ метода сдвига вдоль траектории можно выбрать формулы, числовые оригиналы которых соответствуют приближенным методам решения систем обыкновенных дифференциальных уравнений, например линейным одношаговым и многошаговым методам, коллокационным методам. Эти символьные формулы (последовательности имен переменных и действий) должны быть такими, чтобы хранение формул в памяти машины и обработка проводилась за минимальное время и с разумными затратами памяти. После подстановки в эти формулы вместо символьных переменных числовых значений определяется включение всего множества точных решений, учитывающее гарантированные оценки глобальных ошибок каждого из приближенных решений этой совокупности. Применение гарантированного метода $\mathcal{IMSHT}$ позволяет получать оценку, для которой выполняется включение

$$\mathcal{Y}(t) \subseteq \mathbf{Y}(t), \tag{5}$$

при этом для проекции на координатные оси каждой компоненты множества точных решений $\mathcal{Y}(t)$ при стремлении шага разбиения $\delta(t^0)$ к нулю выполняется оценка

$$\| \mathcal{Y}_k - \mathbf{Y}_k \| \leq h^p \| \mathbf{G}_h \|,$$

где p – порядок символьной формулы $\mathcal{S}$, $\| \mathbf{G}_h \|$ – норма гарантированной интервальной оценки глобальной ошибки [17].

Теорема 1 *Пусть выполнены все условия, обеспечивающие разрешимость задачи (3) и предположения относительно множеств P^0, P. Если в результате вычислений в методе $\mathcal{IMSHT}$ были получены включения $\mathbf{Y}(t) \in Q$ для любого $t \in T$, то решение задачи (3) интервально (практически) устойчиво.*

Доказательство этой теоремы является прямым следствием оценок, полученных при обосновании метода $\mathcal{IMSHT}$ в [16], [17], [18].

2 Результаты применения гарантированных методов оценки практической устойчивости

Исследуем вопросы практической устойчивости системы ОДУ

$$\begin{aligned} \frac{dy_1}{dt} &= -y_1 + 2y_1^2 y_2, \\ \frac{dy_2}{dt} &= -y_2. \end{aligned} \tag{6}$$

Система имеет единственное положение равновесия, асимптотически устойчивое по Ляпунову. Граница области асимптотической устойчивости задается уравнением $y_1 y_2 = 1$ [2]. Эта система часто использовалась в качестве тестового примера для методов, оценивающих границы области притяжения. Расчеты, проведенные гарантированным методом, показали существование области притяжения решений и позволили установить границы области. Увеличение размеров многозначной функции (постоянно действующего возмущения) демонстрируют потерю устойчивости решениями системы. Рис. 1, Рис. 2 соответствуют случаю, когда выбранные границы области ПДВ обеспечивают устойчивость решений.

При увеличении области возмущений система (6) теряет устойчивость, что показано на Рис. 3.

В качестве следующего примера применения гарантированного метода оценивания множеств решений в задачах практической устойчивости рассмотрим систему Ван-дер-Поля. Известно, что генератор Ван-дер-Поля имеет “мягкий” режим возникновения автоколебаний, не требующий толчка. В безразмерных величинах уравнение Ван-дер-Поля запишется в виде системы

$$\begin{aligned} \frac{dy_1}{dt} &= \frac{dy_2}{dt}, \\ \frac{dy_2}{dt} &= \frac{1}{\mu}(1 - y_2^2)y_2 - y_1. \end{aligned} \tag{7}$$

Графики гарантированных оценок решений (7) показаны на Рис. 4, Рис. 5, Рис. 6.

Рассмотрим применение предложенных гарантированных методов оценки практической устойчивости на примере стабилизации в поле центральной силы кругового движения материальной точки, управляемой реактивной силы [19]. Запишем уравнения возмущенного движения в виде

$$\begin{aligned} \frac{dy_1}{dt} &= y_2, \\ \frac{dy_2}{dt} &= -\frac{\mu_0}{(r_0 + y_1)^2} + \frac{((\mu_0 r_0)^{1/2} + y_3)^2}{(r_0 + y_1)^3} - 2by_2 - (\frac{2}{r_0 b + y_1} - 2\lambda \cdot y_3) b \frac{r_0 + y_1}{r_0}, \end{aligned}$$

$$\frac{dy_3}{dt} = \frac{r_0 + y_1}{r_0} - 2by_2 - \Big(\frac{2}{r_0 b + y_1} - 2\lambda \cdot y_3\Big) b\frac{r_0 + y_1}{r_0}, \tag{8}$$
$$u = r_0 c_\phi \frac{d\ln m}{dt},$$
$$b = \frac{c_t}{c_\phi r_0}.$$

Здесь $y_1 = r - r_0$, $y_2 = \frac{dr}{dt}$, $y_3 = r^2\frac{d\varphi}{dt} - (\mu_0 r_0)^{1/2}$, $\mu_0 = \gamma M$, r, ϕ — полярные координаты точки, r_0 — радиус невозмущенной круговой орбиты, γ — постоянная всемирного тяготения, M — масса притягивающего центра, c_r, c_ϕ — проекции относительной скорости отделяющейся частицы на направление радиуса и поперечное направление, m — масса точки. Компонента правой части $-2by_2 - (\frac{2}{r_0 b + y_1} - 2\lambda \cdot y_3) b\frac{r_0+y_1}{r_0}$ — стабилизирующее управление, предложенное в [19].

Для расчетов выбраны следующие параметры возмущенного движения материальной точки $\mu = 1$, $r_0 = 150$, $b = 0.8$. Интервальные оценки множества решений возмущенной системы просчитаны при разных значениях возмущенного движения. Были найдены границы области практической устойчивости. В качестве примера приведены графики таких областей, соответствующие значениям начальных данных. $y_1^0 \in [-0.999, 7.151], y12 \in [0.01069412, 14.7791137], y13 \in [-0.124432692, 0.317325192]$. Стабилизация движений системы с ПДВ была отмечена при $t \in [0, 50], t \in [0, 100], t \in [0, 200]$. Однако движение является практически устойчивым только относительно координат вектора решений y_2, y_3. Относительно координаты вектора решений y_1 наблюдается стабилизация на конечном интервале или локальная ограниченность. Эти результаты являются строгими.

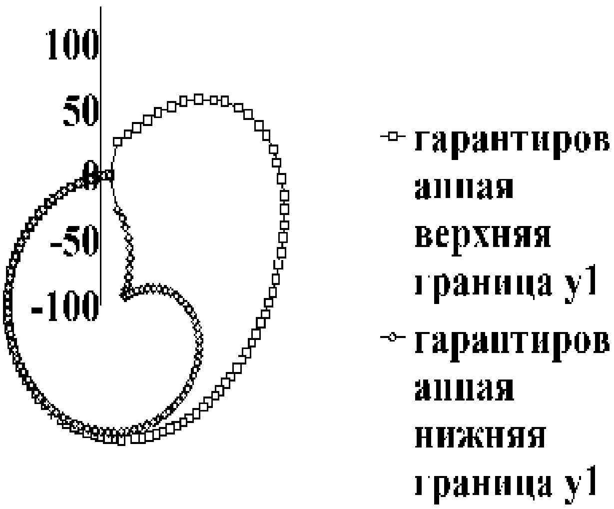


Рис. 1. Проекция на оси t, $y1$ гарантированной оценки множества решений возмущенной системы (6). Показаны верхняя и нижняя границы. График имеет вид области в полярной системе координат, где ось времени совпадает с угловой осью φ. Постоянные значения, которые принимает возмущенное решение отложены вдоль окружностей с центром в начале координат и радиуса, совпадающего с этим значением.

Рис. 2. Проекция на оси t, y_1 гарантированной оценки множества решений возмущенной системы (6). Показаны верхняя и нижняя границы. График изображает область решений в декартовой системе координат.

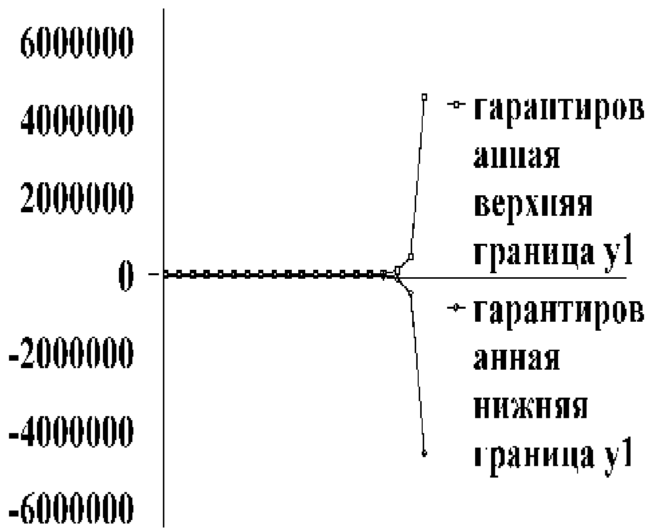


Рис. 3. Проекция на оси t, $y1$ гарантированной оценки множества решений возмущенной системы (6). Показаны верхняя и нижняя границы. График изображает область в декартовой системе координат. Виден неограниченный рост границ множеств решений.

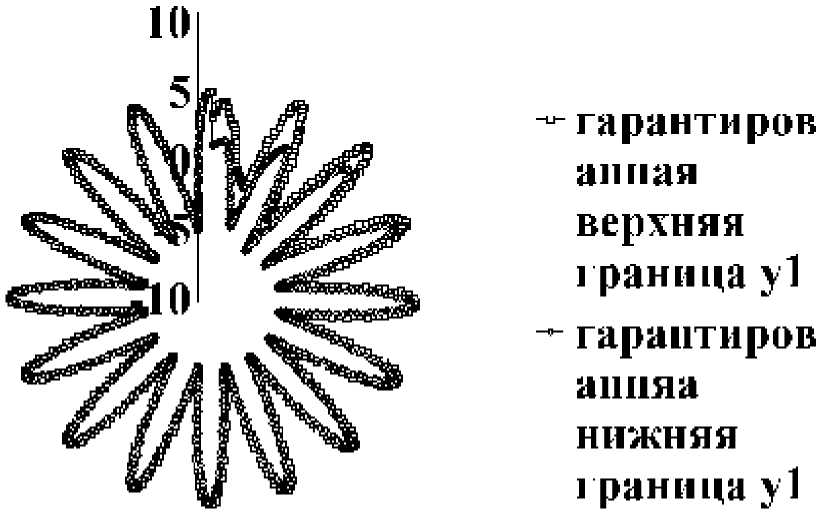


Рис. 4. Проекция на оси t, y_1 гарантированной оценки множества решений возмущенного уравнения Ван-дер-Поля. Показаны верхняя и нижняя границы. График имеет вид области в полярной системе координат, где ось времени совпадает с угловой осью φ. Постоянные значения, которые принимает возмущенное решение отложены вдоль окружностей с центром в начале координат и радиуса, совпадающего с этим значением.

Рис. 5. Проекция на оси t, y_2 гарантированной оценки множества решений возмущенного уравнения Ван-дер-Поля. Показаны верхняя и нижняя границы. График имеет вид области в декартовой системе координат.

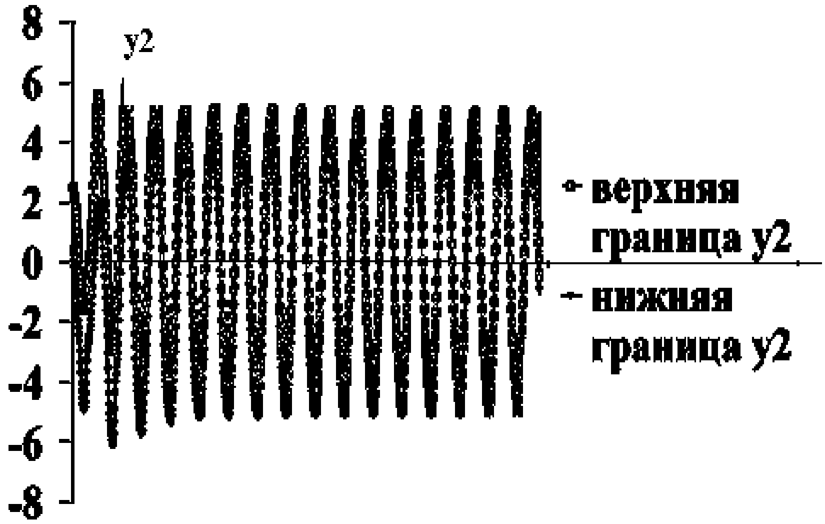


Рис. 6. Проекция на оси t, y_1 гарантированной оценки множества решений возмущенного уравнения Ван-дер-Поля. Показаны верхняя и нижняя границы. График имеет вид области в декартовой системе координат.

Рис. 7. Проекция на оси t, y_1 гарантированной оценки множества решений возмущенной системы кругового движения материальной точки, управляемой реактивной силой. Показаны верхняя и нижняя границы.

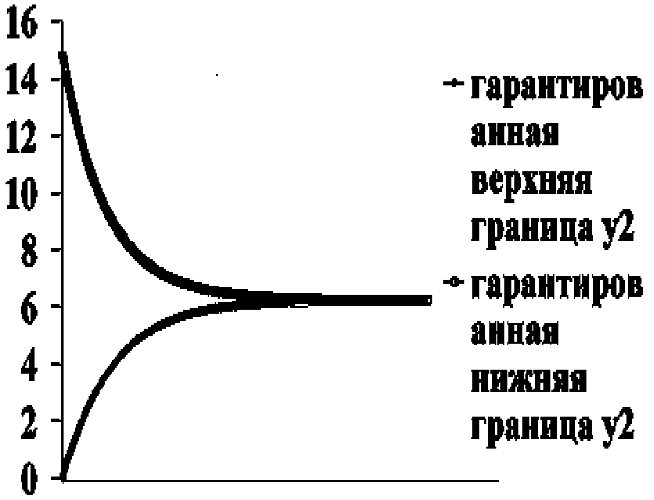


Рис. 8. Проекция на оси t, y_2 гарантированной оценки множества решений возмущенной системы кругового движения материальной точки, управляемой реактивной силой. Показаны верхняя и нижняя границы.

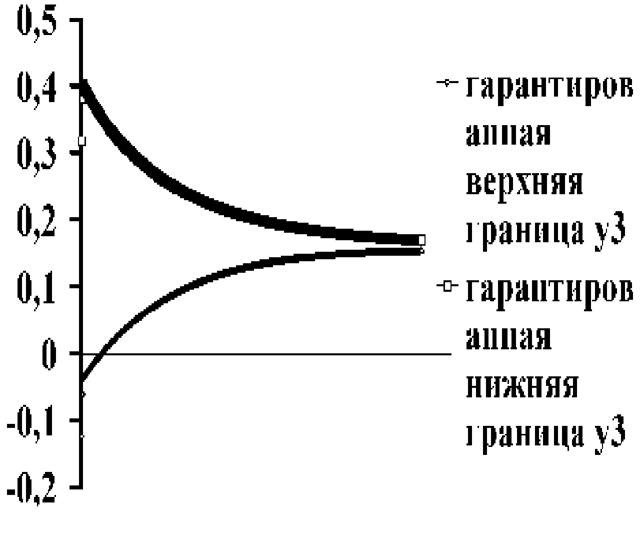


Рис. 9. Проекция на оси t, y_3 гарантированной оценки множества решений возмущенной системы кругового движения материальной точки, управляемой реактивной силой. Показаны верхняя и нижняя границы.

Список литературы

1. Абгарян К.А. *Введение в теорию устойчивости движения на конечном интервале времени.* – М.: Наука, 1992. –
2. Зубов В.И. *Устойчивость движения.* – М.: Высшая школа, 1973.
3. Каменков Г.В. Об устойчивости движения на конечном интервале времени // *Прикладная математика и механика.* – 1953. – Т. 17, №5. – С. 529 – 540.
4. Ла-Салль Ж., Лефшец С. *Исследование устойчивости прямым методом Ляпунова.* – М.: Мир, 1964.
5. Мартынюк А.А., Гутовски Р. *Интегральные неравенства и устойчивость движения.* – Киев: Наукова Думка, 1979.
6. Мартынюк А.А. *Практическая устойчивость движения.* – Киев: Наукова Думка, 1983.
7. Матросов В.М., Анапольский Л.Ю., Васильев С.Н. *Метод сравнения в математической теории систем.* – Новосибирск: Наука, 1980.
8. Матросов В.М. Метод векторных функций Ляпунова: анализ динамических свойств нелинейных систем. – М.: Физматлит, 2001.
9. Моисеев Н.Д. О некоторых методах теории технической устойчивости // *Труды Военно-воздушной Академии им. Жуковского.* – 1945. – Вып. 135. – С. 3 – 17.
10. Рогалев А.Н. Нахождение оптимальных гарантированных оценок множеств решений систем ОДУ с интервальными данными // *Вычислительные Технологии.* – 1995. – Т. 4, №13. – С. 58 – 64.
11. Рогалев А.Н., Шокин Ю.И. Исследование и оценка решений обыкновенных дифференциальных уравнений интервально-символьными методам // *Вычислительные Технологии.* – 1999. - Т. 4, №4. – С. 51 – 76.
12. Рогалев А.Н. Поле аппроксимаций обыкновенных дифференциальных уравнений с интервальными коэффициентами // *Вопросы математического анализа.* – Красноярск: КГТУ, 2001. – Вып. 4. – С. 152 – 170.
13. Рогалев А.Н. Динамические разностные схемы для построения интервальных методов решения ОДУ с неточно заданными начальными данными // *Вычислительные технологии.* – 2001. – Т. 6, ч. 2. – С. 510 – 518.
14. Рогалев А.Н. Динамика множеств решений обыкновенных дифференциальных уравнений и интервальные решения // *Вопросы математического анализа.* – Красноярск: КГТУ, 2001. – Вып. 5.
15. Рогалев А.Н. Использование границ глобальной ошибки в гарантированных оценках решений обыкновенных дифференциальных уравнений // *Вычислительные технологии.* – 2002. – Т. 7, ч. 4. – С. 88 – 95.
16. Рогалев А.Н. Исследование практической устойчивости при постоянно действующих возмущениях // *Вычислительные технологии.* – 2002. – Т. 7, ч. 5. – С. 148 – 150.
17. Рогалев А.Н. Гарантированные методы решения систем обыкновенных дифференциальных уравнений на основе преобразования символьных формул // *Вычислительные технологии.* – 2003. – представлена в журнал.
18. Рогалев А.Н. Включение множеств решений дифференциальных уравнений и гарантированные оценки глобальной ошибки // *Вычислительные технологии.* – 2003. – представлена в журнал.
19. Румянцев В.В., Озиранер А.С. *Устойчивость и стабилизация движения по отношению части переменных.* – М.: Наука, 1987.
20. Румянцев В.В. Об оптимальной стабилизации управляемых систем // *Прикладная математика и механика.* – 1970. – Т. 34, вып. 3. – С. 440 – 456.
21. Четаев Н.Г. Устойчивость движения. – М.: Наука, 1965.
22. Adams B., Adams E., Spreuer H. Practical stability and stochastic point processes // IMACS Ann. Comput. and Appl. Math., 1989. – Vol. 1, No. 1–4. – P. 81 – 89.
23. Anguelov R., Markov S. Wrapping effect and wrapping function // *Reliable Computing.* – 1998. – Vol. 4, No. 4. – P. 311 – 330.
24. Berz M., Makino K. Verified integration of ODE's and flows using diifferential algebraic methods on high-order Taylor models // *Reliable Computing.* – 1998. – Vol. 4, No. 4. – P. 361 – 369.
25. Lohner R. On the ubiquity of the wrapping effect in the computation of error bound // SCAN'2000 – Interval'2000, GAMM/IMACS International Symposium on Scientific Computing, Computer Arithmetic and Validated Numerics, Sept. 18–22, 2000. – Universität Karlsrue, Institut für Angewandte Mathematic (Germany). – P. 36.
26. Nedialkov N., Jackson K., Corliss G. Validated solutions of initial value problems for ordinary differential equations // *Applied Mathematics and Computations.* – 1999. - Vol. 105, No. 1. – P. 21 – 68.
27. Nedialkov N., Jackson K. An interval Hermite-Obreschkoff method for computing rogorous bounds on the solution of an initial value problem for an ordinary differential equation // *Reliable Computing.* – 1999. – Vol. 5, No. 5. – P. 289 – 310.
28. Nedialkov N., Jackson K. The design and implementations of an object-orientd validated ODE solver // Numerical Analysis Technical Report, Department of Computer Science, University of Toronto, Canada. – Toronto, 2001. – 32 p.
    Электронная версия доступна в Интернете на `http://www.cs.toronto.edu/NA/reports.html#ned.software.01`
29. Nedialkov N., Jackson K., Pryce J. An effective high-order interval method for validating existence and uniqueness of the solution of an IVP for an ODE // *Reliable Computing.* – 2001. – Vol. 7, №6. – P. 449–465.
30. Nedialkov N., Jackson K. A New Perspective on the Wrapping Effect. // Interval Methods for Initial Value Problems for Ordinary Differential Equations. Perspectives on Enclosure Methods. Kulisch U., Lohner R., Facius A., eds. – Berlin: Springer-Verlag, 2001. – P. 219 – 264.
    Электронная версия доступна в Интернете на `http://www.cs.toronto.edu/NA/reports.html`

31. Rogalev A. N. The interval methods for ordinary differential equations and shift along trajectories // IMACS-GAMM International Symposium on Scientific Computing, Computer Arithmetic and Validated Numerics, September 10–12, 1997. École Normale Superieure de Lyon (France). – P. X9–X13.

Экспоненциальная устойчивость одного класса нелинейных интервальных динамических систем

Р. С. Ивлев

Институт проблем информатики и управления МОН РК
Казахстан, 480100 г. Алма-Ата, ул. Пушкина, 125
ivlevruslan@newmail.ru

Аннотация Рассматривается класс нелинейных интервальных динамических систем с нелинейностью секторного типа, математическая модель которой задана в пространстве состояний. На основе прямого метода Ляпунова получены достаточные условия экспоненциальной устойчивости нулевого положения равновесия рассматриваемой нелинейной динамической системы.

1 Введение

Изучение свойств нелинейных динамических систем в условиях параметрической неопределенности интервального типа представляет большой научный интерес. Многие из вопросов, касающиеся исследования устойчивости нелинейных интервальных динамических систем, заданных в пространстве состояний, до сих пор остаются открытыми. В настоящей работе рассматривается нелинейность секторного типа. Задачи исследования динамических систем с нелинейностью секторного типа, математические модели которых точно известны, восходят к работам А.И.Лурье [1] и В.М.Попова [2] и составляют два взаимосвязанных направления современной теории абсолютной устойчивости. Наличие интервальной неопределенности обусловило появление нового витка актуальности задач исследования нелинейных систем А.И. Лурье в условиях параметрической неопределенности. Так, например, в работе [3] получены робастные модификации частотных критериев абсолютной устойчивости при неопределенности в линейной части системы. В отличии от указанной работы, в которой линейная часть задана в виде семейства полиномов, наибольший интерес в этой области представляет исследование нелинейных систем, заданных в пространстве состояний. В работе [4], используя функционалы Ляпунова-Красовского, получены достаточные условия абсолютной устойчивости нелинейной интервальной системы с запаздыванием в состоянии и нелинейностью секторного типа. В настоящей работе на основе прямого метода Ляпунова [5] и методов интервального анализа [6] решается задача исследования экспоненциальной устойчивости нелинейной интервальной динамической системы, математическая модель которой задана в пространстве состояний.

2 Обозначения и постановка задачи

Придерживаясь схожей с [4] нотации, условимся полужирным шрифтом обозначать интервальные величины, обычным шрифтом — неинтервальные. Символом нижнего и верхнего подчеркивания будут обозначаться соответственно нижняя и верхняя границы интервала. Операции взятия нижней и верхней границ интервалов применительно к матрицам и векторам будут пониматься в поэлементном смысле.

Возмущенное движение исследуемой нелинейной интервальной системы задано в пространстве состояний в виде следующего соотношения

$$\dot{x}(t) \in \mathbf{A}x(t) + \mathbf{b}\varphi(\sigma), \qquad x(t_0) = x_0, \qquad t \in [t_0, \infty), \tag{1}$$

где $x(t) = (x_i(t))$ — вектор состояний, компонентами которого являются непрерывные на $[t_0, \infty)$ функции $x_i(t)$, т.е. $x_i(t) \in C[t_0, \infty)$, $1 \le i \le n$, $\mathbf{A} \in \mathbb{IR}^{n\times n}$ — постоянная интервальная матрица размерности $n \times n$, $A = (\mathbf{a}_{ij})$, $\mathbf{a}_{ij} = [\underline{\mathbf{a}}_{ij}, \overline{\mathbf{a}}_{ij}]$, $1 \le i, j \le n$, $\mathbb{IR}$ — множество всех вещественных интервалов [6], $\mathbf{b} \in \mathbb{IR}^n$ — интервальный вектор размерности $n \times 1$, $\mathbf{b} = (\mathbf{b}_i)$, $\mathbf{b}_i = [\underline{\mathbf{b}}_i, \overline{\mathbf{b}}_i]$, $1 \le i \le n$, $\varphi(\sigma)$ — непрерывно дифференцируемая функция, $\varphi : \mathbb{R} \to \mathbb{R}$, такая, что справедливыми являются ограничения секторного типа (график функции $\varphi(\sigma)$ расположен в секторе между прямыми $\varphi = 0$ и $\varphi = \mu\sigma$, $\mu \in \mathbb{R}^+$, $\mu > 0$), которые в формальном виде могут быть записаны

$$0 \le \varphi/\sigma \le \mu, \tag{2}$$

где $\sigma \neq 0$, и при $\sigma = 0$ необходимо $\varphi(0) = 0$. Величина $\sigma \in \mathbb{R}$ определяется согласно выражению

$$\sigma = r^T x(t), \tag{3}$$

где $r \in \mathbb{R}^n$ — вектор размерности $n \times 1$.

Математическая модель (1) понимается как семейство математических моделей вида

$$\dot{x}(t) = Ax(t) + b\varphi(\sigma), \qquad x(t_0) = x_0,\ t \in [t_0, \infty), \qquad A \in \mathbf{A},\ b \in \mathbf{b}. \tag{4}$$

Двойное неравенство (2) можно переписать в виде одного неравенства. Учитывая выражение для σ, имеем

$$\varphi(\mu\sigma - \varphi) = \mu\varphi r^T x - \varphi^2 = F(x, \varphi) \geq 0. \tag{5}$$

Левая часть неравенства (5) является квадратичной формой переменных x и φ, которая определяет нелинейное ограничение на x и φ.

Будем предполагать, что пара интервальных матриц $(\mathbf{A}, \mathbf{b})$ полностью управляема, т.е. для любых $A \in \mathbf{A}$ и $b \in \mathbf{b}$ полностью управляемой является пара (A, b).

Определение 1 *Под решением задачи* (1)–(3) *будем понимать всякую абсолютно непрерывную функцию $x(t, t_0, x_0)$, удовлетворяющую при некоторых значениях $A \in \mathbf{A}$ и $b \in \mathbf{b}$ следующей нелинейной системе дифференциальных уравнений*

$$\begin{cases} \dot{x}(t) = Ax(t) + b\varphi(\sigma) \\ u(t) = \varphi(\sigma),\ \sigma = r^T x(t), \end{cases} \qquad x(t_0) = x_0,\ t \in [t_0, \infty). \tag{6}$$

В силу $\varphi(0) = 0$ существует тривиальное решение $x(t, t_0, x_0) = x(t) \equiv 0$ системы.

Определение 2 *Тривиальное ршение $x(t) \equiv 0$ системы* (1)–(3) *называется экспоненциально устойчивым при $t \to \infty$, если для любых значений $A \in \mathbf{A}$ и $b \in \mathbf{b}$ и всякого решения $x(t, t_0, x_0)$ этой системы справедливо неравенство*

$$\|x(t, t_0, x_0)\| \leq N \, \|x(t_0)\| \, \exp\bigl(-\alpha(t - t_0)\bigr),$$

где $\|\cdot\|$ — Евклидова норма, N и α — положительные постоянные, не зависящие от выбора $x(t, t_0, x_0)$.

Задача: *определить условия наличия свойства экспоненциальной устойчивости положения равновесия $x(t) \equiv 0$ интервальной динамической системы* (1)–(3) *с нелинейностью секторного типа* (2) *в смысле Определения 2.*

3 Основной результат

Решение поставленной задачи будет осуществлено на основе прямого метода Ляпунова посредством выбора функции Ляпунова в виде квадратичной формы

$$V(x) = x^T H x, \tag{7}$$

где $H \in \mathbb{R}^{n\times n}$, $H = H^T$ — некоторая симметрическая положительно-определенная матрица.

Используя арифметические операции классической интервальной арифметики, построим интервальную матрицу $\mathbf{A}_1 = (\mathbf{a}_{1_{ij}})_{ij=1}^n \in \mathbb{IR}^{n\times n}$

$$\mathbf{A}_1 = \mathbf{A} + \frac{\mu}{2}\mathbf{b}r^T$$

и введем в расмотрение множество угловых (крайних) матриц

$$\operatorname{vert} \mathbf{A}_1 = \left\{ \hat{A}_1 \in \mathbb{R}^{n\times n} \mid \hat{A}_1 = (\hat{a}_{1_{ij}})_{ij=1}^n,\ \hat{a}_{1_{ij}} \in \{\underline{\mathbf{a}}_{1_{ij}}, \overline{\mathbf{a}}_{1_{ij}}\} \right\}.$$

Условия экспоненциальной устойчивости исследуемой системы дает следующая теорема.

Теорема 1 *Пусть для заданных интервальной матрицы $\mathbf{A} \in \mathbb{IR}^{n\times n}$, интервального вектора $\mathbf{b} \in \mathbb{IR}^n$, вектора $r \in \mathbb{R}^n$ и скаляра $\mu \in \mathbb{R}^+$ выполнены следующие условия:*

– *пара интервальных матриц $(\mathbf{A}, \mathbf{b})$ полностью управляма;*

– *существует такая положительно-определенная симметрическая матрица* $H \in \mathbb{R}^{n\times n}$, *что следующие матрицы*

$$\hat{A}_1^T H + H\hat{A}_1 + HH + \frac{\mu^2}{4} rr^T \underline{\mathbf{b}^T\mathbf{b}}, \qquad \hat{A}_1^T H + H\hat{A}_1 + HH + \frac{\mu^2}{4} rr^T \overline{\mathbf{b}^T\mathbf{b}}, \qquad \hat{A}_1 \in \mathrm{vert}\ \mathbf{A}_1$$

являются отрицательно-определенными.

Тогда тривиальное решение $x(t) \equiv 0$ *нелинейной интервальной системы* (1)–(3) *экспоненциально устойчиво для заданного класса нелинейностей.*

Доказательство. В области пространства $\mathbb{R}^n \times \mathbb{R}$, выделяемой ограничениями секторного типа (5), вычислим производную функции Ляпунова (7) по времени на траекториях движения системы (4) для произвольных, но фиксированных значений $A \in \mathbf{A}$ и $b \in \mathbf{b}$. Для этого, применяя S-процедуру [7], составим S-форму

$$S(x,\varphi) = \dot{V}(x) + \tau F(x,\varphi),$$

где $\tau \in \mathbb{R}^+$ — некоторое положительное число.

$$\begin{aligned}
S(x,\varphi) &= \dot{x}^T Hx + x^T H\dot{x} + \tau\mu\varphi r^T x - \tau\varphi^2 = \\
&= (Ax + b\varphi)^T Hx + x^T H(Ax + b\varphi) + \tau\mu\varphi r^T x - \tau\varphi^2 = \\
&= x^T(A^T H + HA)x + \varphi b^T Hx + x^T Hb\varphi + \tau\mu\varphi r^T x - \tau\varphi^2 = \\
&= x^T(A^T H + HA)x + \varphi(b^T H + \frac{\tau\mu}{2} r^T)x + x^T(Hb + \frac{\tau\mu}{2} r)\varphi - \tau\varphi^2.
\end{aligned}$$

Выбирая $\tau = b^T b > 0$, поскольку пара $(\mathbf{A}, \mathbf{b})$ является полностью управляемой, и, выделяя полный квадрат, имеем

$$\begin{aligned}
&S(x,\varphi) = x^T(A^T H + HA)x + \varphi(b^T H + b^T b r^T \frac{\mu}{2})x + x^T(Hb + \frac{\mu}{2} r b^T b)\varphi - b^T b\varphi^2 = \\
&= x^T(A^T H + HA)x + \varphi b^T(H + br^T\frac{\mu}{2})x + x^T(H + \frac{\mu}{2} rb^T)b\varphi - b^T b\varphi^2 = \\
&= x^T(A^T H + HA)x + \varphi b^T Gx + x^T G^T b\varphi - b^T b\varphi^2 = \\
&= x^T(A^T H + HA)x + \varphi b^T Gx + x^T G^T b\varphi - b^T b\varphi^2 - x^T G^T Gx + x^T G^T Gx = \\
&= x^T(A^T H + HA)x - (Gx - b\varphi)^T(Gx - b\varphi) + x^T(H + br^T\frac{\mu}{2})^T(H + br^T\frac{\mu}{2})x = \\
&= x^T(A^T H + HA)x - (Gx - b\varphi)^T(Gx - b\varphi) + \\
&+ x^T\left(HH + Hbr^T\frac{\mu}{2} + \frac{\mu}{2} rb^T H + \left(\frac{\mu}{2}\right)^2 rb^T br^T\right)x = \\
&= x^T\left((A + \frac{\mu}{2} br^T)^T H + H(A + \frac{\mu}{2} br^T) + HH + \left(\frac{\mu}{2}\right)^2 rb^T br^T\right)x - \\
&-(Gx - b\varphi)^T(Gx - b\varphi) \le \\
&\le x^T\left((A + \frac{\mu}{2} br^T)^T H + H(A + \frac{\mu}{2} br^T) + HH + \left(\frac{\mu}{2}\right)^2 rb^T br^T\right)x = \\
&= x^T(A_1^T H + HA_1 + HH + C)x = -x^T Qx,
\end{aligned}$$

где $G = H + \frac{\mu}{2} br^T$, $A_1 = A + \frac{\mu}{2} br^T$, $C = (\frac{\mu}{2})^2 rb^T br^T = \frac{\mu^2}{4} rr^T b^T b$ и $-Q = A_1^T H + HA_1 + HH + C$.

Поскольку $A_1 \in \mathbf{A}_1$ при $A \in \mathbf{A}$ и $b \in \mathbf{b}$, а также $C \in \frac{\mu^2}{4} rr^T[\underline{\mathbf{b}^T\mathbf{b}}, \overline{\mathbf{b}^T\mathbf{b}}]$ при $b \in \mathbf{b}$, следовательно

$$A_1 = \sum_{\hat{A}_1 \in \mathrm{vert}\ \mathbf{A}_1} \hat{A}_1 \eta_{\hat{A}_1}, \qquad \sum_{\hat{A}_1 \in \mathrm{vert}\ \mathbf{A}_1} \eta_{\hat{A}_1} = 1, \qquad \eta_{\hat{A}_1} \ge 0$$

и

$$C = \frac{\mu^2}{4} rr^T\left(\eta_C \underline{\mathbf{b}^T\mathbf{b}} + (1 - \eta_C)\overline{\mathbf{b}^T\mathbf{b}}\right), \qquad 0 \le \eta_C \le 1.$$

Подставляя последние два соотношения в выражение для $-Q$, имеем

$$-Q = \left(\sum_{\hat{A}_1 \in \text{vert } \mathbf{A}_1} \hat{A}_1 \eta_{\hat{A}_1} \right)^T H + H \sum_{\hat{A}_1 \in \text{vert } \mathbf{A}_1} \hat{A}_1 \eta_{\hat{A}_1} + HH+$$
$$+\frac{\mu^2}{4} rr^T \left(\eta_C \underline{\mathbf{b}^T \mathbf{b}} + (1-\eta_C) \overline{\mathbf{b}^T \mathbf{b}} \right) =$$
$$= \sum_{\hat{A}_1 \in \text{vert } \mathbf{A}_1} \left(\hat{A}_1^T H + H\hat{A}_1 + HH + \frac{\mu^2}{4} rr^T \left(\eta_C \underline{\mathbf{b}^T \mathbf{b}} + (1-\eta_C) \overline{\mathbf{b}^T \mathbf{b}} \right) \right) \eta_{\hat{A}_1} =$$
$$= \sum_{\hat{A}_1 \in \text{vert } \mathbf{A}_1} \left((\eta_C + (1-\eta_C)) \left(\hat{A}_1^T H + H\hat{A}_1 + HH \right) + \right.$$
$$\left. +\frac{\mu^2}{4} rr^T \left(\eta_C \underline{\mathbf{b}^T \mathbf{b}} + (1-\eta_C) \overline{\mathbf{b}^T \mathbf{b}} \right) \right) \eta_{\hat{A}_1} =$$
$$= \sum_{\hat{A}_1 \in \text{vert } \mathbf{A}_1} \left(\hat{A}_1^T H + H\hat{A}_1 + HH + \frac{\mu^2}{4} rr^T \underline{\mathbf{b}^T \mathbf{b}} \right) \eta_C \eta_{\hat{A}_1} +$$
$$+ \sum_{\hat{A}_1 \in \text{vert } \mathbf{A}_1} \left(\hat{A}_1^T H + H\hat{A}_1 + HH + \frac{\mu^2}{4} rr^T \overline{\mathbf{b}^T \mathbf{b}} \right) (1-\eta_C) \eta_{\hat{A}_1}.$$

Матрицы

$$\hat{A}_1^T H + H\hat{A}_1 + HH + \frac{\mu^2}{4} rr^T \underline{\mathbf{b}^T \mathbf{b}}, \qquad \hat{A}_1^T H + H\hat{A}_1 + HH + \frac{\mu^2}{4} rr^T \overline{\mathbf{b}^T \mathbf{b}}, \qquad \hat{A}_1 \in \text{vert } \mathbf{A}_1$$

являются симметрическими. По условию теоремы эти матрицы также отрицательно определены, а так как $\eta_C \eta_{\hat{A}_1} \geq 0$ и $(1-\eta_C)\eta_{\hat{A}_1} \geq 0$, то матрица $-Q$ является отрицательно-определенной для любых $A \in \mathbf{A}$ и $b \in \mathbf{b}$. Поэтому в области пространства $\mathbb{R}^n \times \mathbb{R}$, выделяемую ограничениями секторного типа (5), для любых значений $A \in \mathbf{A}$ и $b \in \mathbf{b}$ имеет место соотношение

$$\dot{V}(x) \leq -x^T Q x < 0, \qquad \forall x \in \mathbb{R}^n \backslash \{0\}.$$

Следовательно, тривиальное решение $x(t) \equiv 0$ экспоненциально устойчиво. Теорема доказана.

Для постоянных N и α справедливы следующие соотношения

$$N = \sqrt{\max_p \lambda_p(H) / \min_p \lambda_p(H)},$$
$$\alpha = \min_{A \in \mathbf{A}, b \in \mathbf{b}} \min_p \lambda_p(Q) / \left(2 \max_p \lambda_p(H) \right),$$

где $\lambda_p(\cdot)$ — собственные значения матриц. В силу симметричности матриц H и Q собственные значения этих матриц вещественны.

Список литературы

1. Лурье А.И. *Некоторые нелинейные задачи теории автоматического регулирования.* – М.: Гостехиздат, 1951.
2. Попов В.М. *Гиперустойчивость автоматических систем.* – М.: Наука, 1970.
3. Джури Э.И., Премаратне К., Эканайаке М.М. Робастная абсолютная устойчивость дискретных систем // *Автоматика и Телемеханика.* – 1999. – №3. – С. 97–118.
4. Ивлев Р.С. Абсолютная устойчивость нелинейных динамических систем с параметрической неопределенностью интервального типа и запаздывающим аргументом // Материалы Международной конференции "Вычислительные технологии и математическое моделирование в науке, технике и образовании" – ВТММ-2002, Новосибирск-Алматы, 2002. – С. 27–34.
5. Демидович Б.П.*Лекции по математической теории устойчивости.* – М.: Наука, 1967.
6. Алефельд Г., Херцбергер Ю. *Введение в интервальные вычисления.* – М.: Мир, 1987.
7. Гелиг А.Х., Леонов Г.А., Якубович В.А. *Устойчивость нелинейных систем с неединственным состоянием равновесия.* – М.: Наука, 1978.

Построение областей притяжения устойчивых решений дифференциальных уравнений с помощью монотонного метода и М–матриц

Н.В. Перцев*

Омский госуниверситет, кафедра математического моделирования
Россия, 644077, г. Омск, проспект Мира, 55а, ОмГУ
pertsev@homlab.omsk.ru

Аннотация Работа посвящена задаче построения областей притяжения для устойчивых положений равновесия определенного класса систем дифференциальных уравнений запаздывающего типа. Предложен способ построения таких областей, опирающийся на метод двусторонних приближений и свойства невырожденных М-матриц. Для заданного положения равновесия x^* область притяжения строится в форме многомерного параллелепипеда с центром в точке x^*. Границы параллелепипеда $x^* - z^1 \le x \le x^* + z^2$ задаются с помощью векторов z^1, z^2, которые находятся как начальные точки некоторого итерационного процесса.

Рассматривается система функционально-дифференциальных уравнений запаздывающего типа с заданным начальным условием

$$\dot{x}(t) = f(x_t) - \lambda\, x(t), \quad t \ge 0, \tag{1}$$

$$x(t) = \psi(t), \quad t \in I = [-\omega, 0]. \tag{2}$$

Здесь

$x(t) = \mathrm{col}(x_1(t), x_2(t), ..., x_m(t))$ – искомая вектор-функция,
$\dot{x}(t)$ — ее правосторонняя производная,
$\psi(t) = \mathrm{col}(\psi_1(t), \psi_2(t), ..., \psi_m(t))$ — заданная начальная функция,
$0 < \omega < \infty$ – фиксированное число,
$f(x_t) = \mathrm{col}(f_1(x_t), f_2(x_t), ..., f_m(x_t))$ — некоторое отображение,
$\lambda = \mathrm{diag}(\lambda_1, \lambda_2, ..., \lambda_m)$ — диагональная матрица с элементами $\lambda_i > 0$, $1 \le i \le m$.

Полагаем, что $x_t(\theta) = x(t+\theta)$, $\theta \in I$, $f(z) : C(I, D) \to \mathbb{R}^m$, $C(I, D)$ – множество непрерывных функций $z : I \to D$ с нормой $||z||_C = \max_{\theta \in I} |z(\theta)|$, где D – выпуклое, открытое подмножество $\mathbb{R}^m$. Предполагается, что ψ является непрерывной функцией, а f – непрерывным отображением. Считаем, что неравенства между векторами из $\mathbb{R}^m$ понимаются как неравенства между их компонентами. Запись $\xi \in \mathbb{R}^m$, $\xi > 0$, эквивалентна тому, что все компоненты вектора ξ положительны. Если $x^1, x^2 \in C(I, D)$, то неравенство $x^1 \le x^2$ $(x^2 \ge x^1)$, означает, что для всех $\theta \in I$ верно $x^1(\theta) \le x^2(\theta)$. Дополнительно предполагаем, что для любых $z \in C(I, D)$ имеет место представление

$$f(z) = g(z, z), \tag{3}$$

где непрерывное отображение $g(x, y) : C(I, D) \times C(I, D) \to \mathbb{R}^m$ изотонно по x и антитонно по y, то есть

$$g(x^1, y^1) \le g(x^2, y^2)$$

для любых пар $(x^i, y^i) \in C(I, D) \times C(I, D)$, $i = 1, 2$, таких, что $x^1 \le x^2$, $y^1 \ge y^2$.

Уравнения (1)–(2) возникают при построении математических моделей сложных процессов и систем, регулируемых с помощью положительных и отрицательных обратных связей. Особенностью некоторых моделей является высокая размерность, наличие различных запаздываний, большое число параметров, что вызывает трудности при аналитическом исследовании свойств их решений. Другая особенность многих моделей проявляется в том, что правые части уравнений содержат монотонные функции или операторы вида (3), отражающие определенные закономерности изучаемых процессов.

* Работа выполнена при финансовой поддержке РФФИ (грант №02-01-01166).

Одной из важнейших проблем является изучение условий существования и устойчивости положений равновесия x^* системы (1). Другая важная проблема состоит в построении такого множества $\Psi_* \subset D$, что для любой начальной функции $\psi(t) \in \Psi_*$, $t \in I$, задача (1), (2) имеет решение $x(t) \to x^*$ при $t \to +\infty$. Для решения этих проблем можно использовать различные подходы. Один из них опирается на монотонный метод и свойства невырожденных М-матриц. Определим функцию

$$h(u, w) = \lambda^{-1} g(u, w),\ u,\ w \in D$$

и рассмотрим ряд свойств решений задачи (1), (2).

Лемма *Предположим, что существует пара $(u^0, w^0) \in D \times D$, удовлетворяющая неравенствам*

$$u^0 \leq w^0, \quad u^0 \leq h(u^0, w^0), \quad w^0 \geq h(w^0, u^0). \tag{4}$$

Рассмотрим систему уравнений

$$u = h(u, w), \quad w = h(w, u), \quad u^0 \leq u \leq w^0, \quad u^0 \leq w \leq w^0 \tag{5}$$

и последовательности векторов

$$u^n = h(u^{n-1}, w^{n-1}), \quad w^n = h(w^{n-1}, u^{n-1}), \quad n = 1, 2, ..., . \tag{6}$$

Справедливы следующие утверждения:

1) *любая пара (u^n, w^n) удовлетворяет неравенствам (4);*
2) *существуют $u^* = \lim_{n\to\infty} u^n$, $w^* = \lim_{n\to\infty} w^n$ такие, что $u^* \leq w^*$ и пара (u^*, w^*) является решением системы (5);*
3) *решение (u^*, w^*) системы (5) единственно тогда и только тогда, когда $u^* = w^*$;*
4) *система (1) имеет положение равновесия x^* такое, что $u^0 \leq x^* \leq w^0$;*
5) *при $u^* = w^*$ система (1) не имеет других положений равновесия $\bar{x}$, принадлежащих $u^0 \leq u \leq w^0$, кроме $x^* = u^* = w^*$.*

Теорема *Пусть $x^* \in D$ – положение равновесия системы (1), для которого выполнены следующие условия:*

1) *функция $h(u, w)$ имеет непрерывные частные производные первого порядка в некоторой окрестности пары (x^*, x^*);*
2) *$(I - A_*)$ является невырожденной М-матрицей, где матрица $A_* = (a_{ik})$ задается элементами*

$$a_{ik} = \partial h_i(x^*, x^*)/\partial u_k - \partial h_i(x^*, x^*)/\partial w_k \geq 0, \quad 1 \leq i,\ k \leq m.$$

Тогда существует вектор $\xi^* \in \mathbb{R}^m$, $\xi^* > 0$, $(I - A_*)\xi^* > 0$, такой, что для любой начальной функции

$$\psi(t) \in \Psi_* = \{\psi \in \mathbb{R}^m : x^* - \xi^* \leq \psi \leq x^* + \xi^*\}, \qquad t \in I,$$

задача (1)–(2) имеет решение $x(t)$, $t \in [0, \infty)$ и $x(t) \to x^*$ при $t \to +\infty$.

Практическое применение этих результатов к исследованию конкретной задачи (1)–(2) предполагает несколько этапов. На первом этапе необходимо построить отображение (3), удовлетворяющее приведенному выше условию монотонности. Укажем, что для конкретной системы такое представление может быть неединственным.

На втором этапе требуется получить условия, при которых $Q_* = (I - A_*)$ является невырожденной М-матрицей. Здесь, в частности, можно использовать следующие эквивалентные критерии:

а) все главные миноры Q_* положительны,
б) существует такой $u \in \mathbb{R}^m$, $u > 0$, что $Q_* u > 0$,
в) матрица Q_*^{-1} существует и имеет неотрицательные элементы,
г) все собственные числа матрицы Q_* имеют положительные действительные части (иначе, матрица $-Q_*$ устойчива).

Третий этап связан с построением вектора ξ^*, задающего границы множества Ψ_*. Для этого следует решить систему неравенств

$$Q_* \xi > 0, \quad \xi > 0, \qquad \xi \in \mathbb{R}^m.$$

Зададим далее $u^0 = x^* - q\,\xi$, $w^0 = x^* + q\,\xi$, где $q > 0$ — достаточно малое число. Построим последовательности векторов (6), которые монотонно сходятся к векторам u^* и w^*. Если окажется, что $u^* = w^*$, то тогда $x(t) \to x^* = u^* = w^*$ при $t \to +\infty$ и множество

$$\Psi_q = \{\psi \in \mathbb{R}^m : x^* - q\,\xi \le \psi \le x^* + q\,\xi\}$$

служит оценкой границ изменения начальных функций ψ, для которых указанный предел существует. В этом случае параметр $q > 0$ можно увеличить. Если же при выбранном $q > 0$ окажется, что $u^* \ne w^*$, то этот параметр следует уменьшить. В обоих случаях вычисления повторяются заново, что дает возможность получить границы множества Ψ_*.

Предложенный подход использовался для изучения свойств решений нескольких моделей (1)–(2), которые не «поддавались» стандартным способам исследования.

В завершение отметим, что эффективная реализация описанного подхода требует привлечения методов интервального анализа и соответствующей программной поддержки.

Список литературы

1. Ортега Дж., Рейнболт В. *Итерационные методы решения нелинейных систем уравнений со многими неизвестными.* – М.: Мир, 1975.
2. Berman A., Plemmons R.J. *Nonnegative matrices in the mathematical sciences.* – New York: Academic Press, 1979.
3. Курпель Н.С., Шувар Б.А. *Двусторонние операторные неравенства и их применения.* – Киев: Наукова Думка, 1980.
4. Дьери И., Перцев Н.В. Об устойчивости положений равновесия функционально-дифференциальных уравнений запаздывающего типа, обладающих свойством смешанной монотонности // *ДАН СССР.* – 1987. – Т. 297, №1. – С. 23 – 25.
5. Gyori I. Global attractivity in delay differential equations using mixed monotone technique // *Journal of Math. Anal. and Appl.* – 1990. – Vol. 152, No. 1. – P. 131 – 155.
6. Перцев Н.В. Применение монотонного метода и М-матриц к анализу поведения решений некоторых моделей биологических процессов // *Сибирский журнал индустриальной математики.* – 2002. – Т. 5, №4(12). – С. 110 – 122.

Об интервальных методах для дифференциальных задач с неопределенностями в виде выпуклых множеств

Г. Ш. Утюбаев*

Сибирский государственный аэрокосмический университет
Россия, 660014, г. Красноярск, а/я 486
prmath@saa.ru

Аннотация Работа состоит из пяти частей: свойств нестандартной арифметики, общих свойств квазилинейных пространств, свойств интервальных функций, сходимости интервальных методов, приложений. Введены определения следующих понятий:полного квазилинейного пространства, середины и ширины элемента этого пространства, квазилинейного подпространства, внешней (внутренней) аппроксимации, устойчивости, внешней (внутренней) сходимости одношаговых интервальных методов для задачи Коши с интервальными параметрами. Получена оценка решения интервального аппроксимирующего метода для задачи Коши обыкновеных дифференциальных уравнений (ОДУ) с неопределенностями в начальных данных и в параметрах. Доказана теорема о внешней и внутренней сходимости для динамическихи систем с аналогичными неопределенностями. Приведены приближенный метод в пространстве n-мерных эллипсоидов для линейных систем ОДУ внешне сходящийся к точному решению с вторым порядком и модифицированный метод параллелепипедов.

1 Свойства нестандартной интервальной арифметики

Применение стандартных интервальных операций [1-3] приводит к увеличению ширины результирующего интервала , т.е.

$$\text{wid}\ \sum_{k=1}^{n} \boxed{\text{i}} A_k = \sum_{k=1}^{n} \text{wid}\ A_k,$$

где $\boxed{\text{i}}$ — знак суммы или разности интервальной стандартной операции.

Для дальнейшего введем следующие обозначения для любого интервала. В частности для интервала A получаем:

$$A = \{a \mid ia \le a \le sa\} \in \mathbb{IR}.$$

Определение 1. [4] Для любых A, B определим нестандартную операцию вычитания:

$$A - B = [\min\{ia - ib, sa - sb\}, \max\{ia - ib, sa - sb\}]. \tag{1}$$

Замечание 1. Операция (1) является частным случаем обобщенной операции вычитания в конечнопорожденном квазилинейном пространстве (см. ниже (4)).

Замечание 2. Нестандартная операция вычитания применяется при определениях производной интервальной функции [4] и производной в квазилинейном пространстве [6].

Операция (1) не монотонна по включениям , т.е. в общем случае

$$a - b \notin A - B, \quad a \in A, b \in B.$$

В настоящей части работы докажем условия сохранения монотонности и выполнения точных равенств в нестандартной арифметике.

Свойство 1. Для любых A, B, C верно включение

* Работа выполнена при финансовой поддержке фонда фундаментальных исследований Министерства образования Российской Федерации, грант Т00-14.2-127.

$(A+B)-C \subseteq A+(B-C)$. Равенство выполнено , тогда и только тогда, когда

$$(\text{wid }(B)-\text{wid }(C))\text{wid }(A) \geq 0.$$

Свойство 2. Для любых A, B, C верно $A-(B-C) \subseteq (A-B)+C$.

Равенство имеет место , тогда и только тогда, когда выполнено одно из условий:

$$1.\text{wid }(A) \geq \text{wid }(B) \geq \text{wid }(C), \qquad 2.(\text{wid }(A)-\text{wid }(B))\text{wid }(C)=0.$$

Свойство 3. Пусть $A \subseteq B$. Кроме того wid $(C) \geq (B)$ или wid $(C) \leq \min\{sa-ib, sb-ia\}$, тогда $A-C \subseteq B-C$.

Свойство 4. Субдистрибутивность: $AB \ - \ AC \subseteq A(B-C), \ \ \forall A,B,C \in \mathbb{IR}$.

Здесь и далее во всей работе стандартная интервальная операция умножения.

Свойство 5. $AA-BB \subseteq A(A-B)+B(A-B)$.

Свойство 6. $(A-B)-(C-D)=(A-C)-(B-D)$.

Замечание 3. Все вышеприведенные свойства верны в конечно-порожденном квазилинейном пространстве с базисом (см. определение 10) и аксиомой сокращения.

2 Основные свойства квазилинейных пространств

2.1 Используемые определения

Пусть задано поле P с порядком и стабильное по операции сложения.

Определение 2. Множество Q с двумя операциями (сложением элементов множества и умножением элемента поля на элемент множества) называется квазилинейным пространством над полем P , если выполнено:

I. $< Q, + >$ - коммутативный моноид с нейтральным элементом Θ,

II. Для второй операции верны следующие аксиомы:

$$\forall A, \quad 1A = A,$$

$$\forall A, \alpha, \beta \quad (\alpha\beta)A = \alpha(\beta A),$$

$$\forall A, B, \alpha, \ \ \alpha(A+B) = \alpha A + \alpha B,$$

$$\forall A, \alpha, \beta : \alpha\beta \geq 0, \ \ (\alpha+\beta)A = \alpha A + \beta A.$$

Записывать это определение мы будем аналогично определению линейных пространств в виде $< Q, +, \cdot, P >$ или просто Q.

Определение 3. [6] Элемент $A \in Q$ называется групповым , если существует противоположный элемент $A' \in Q$.

Определение 4. [7] Элемент A пространства Q называется дистрибутивным, если

$$\forall \alpha, \beta \in P \quad (\alpha \ + \ \beta)A \ = \ \alpha A \ + \ \beta A.$$

Определение 5. [7] Элемент A называется симметричным, если $A \ = \ (-1)A$.

Обозначим соответственно через $Q_G, \ \ Q_D, \ \ Q_I, \ \ Q_S$ — множество всех групповых, дистрибутивных, идемпотентных, симметричных элементов из Q.

Любое линейное (векторное) пространство является квазилинейным дистрибутивным пространством. Обратное утверждение неверно.

Определение 6. [8] Q_1 называется квазилинейным подпространством пространства Q, если $Q_1 \subseteq Q$ и Q_1 является квазилинейным пространством с операциями из Q.

Пусть заданы $< Q_1, +, \cdot, P >$ и $< Q_2, \oplus, \odot, P >$.

Определение 7. [8,23] Отображение F из Q_1 в Q_2 называется гомоморфизмом, если

$$\forall A, \ B \ \in \ Q_1, \quad F(A+B) \ = \ F(A) \ \oplus \ F(B),$$

$$\forall \alpha \ \in P, \ \ \forall A \ \in Q_1, \quad F(\alpha \cdot A) \ = \ \alpha \odot F(A).$$

Определение 8. [8,23] Квазилинейное пространство Q называется полным, если

$$\forall A \in Q, \exists B \in Q: \quad (-1)A \ = \ A \ + \ B.$$

Определение 9. [8,23] Пусть $A \ = \ B \ + \ C, \quad B \in Q_D, C \in Q_S$, тогда $B, \ C$ назовем соответственно серединой и шириной элемента А.

Определение 10. [6] Множество элементов $B_1, B_2, ..., B_n$ называется базисом квазилинейного пространства Q, если

1). любой элемент квазилинейного пространства является линейной комбинацией этого множества.

2).$\forall i, B_i$ не является линейной комбинацией остальных B_j.

2.2 Свойства квазилинейных пространств [8-10]

Приведем свойство квазилинейного пространства с аксиомой сокращения, верное и в случае линейного пространства.

Свойство 7. $(\alpha = 0) \vee (X = \Theta) \iff \alpha \cdot X = \Theta$.

Следствие 1. $Q_G \subseteq Q_D$.

Следствие 2. В квазилинейном пространстве с аксиомой сокращения единственным идемпотентом является нейтральный элемент.

Свойство 8. $\alpha X = \beta X \iff (|\alpha| = |\beta|) \vee (X = \Theta)$.

Следствие 3. В квазилинейном пространстве с аксиомой сокращения единственным идемпотентом является нейтральный элемент.

Замечание 4. В квазилинейном пространстве без аксиомы сокращения верны следующие свойства:

$$1. \quad \alpha \cdot \Theta = \Theta, \ \forall \alpha \in P, \quad 2. \quad 0 \cdot X \ = \ \Theta, \text{ если } \forall X : 0 \cdot X \in Q_G,$$

$$3. \quad 0 \cdot X \in Q_I, \quad 0 \cdot X \notin Q_G \iff X \neq \Theta,$$

$$4. \quad \forall X, \ \ 0 \cdot X \in \ Q_S \cap Q_D.$$

Свойство 9. Квазилинейное уравнение

$$\gamma X = \alpha X + \beta X \tag{2}$$

для различных значений параметров α, β, γ имеет следующие решения:

1. При любых $\alpha, \ \beta, \ \gamma \quad X = \theta$ является решением .

2. При $\gamma \ = \ 0$ и $\alpha \ = \ -\beta, \quad \forall X \ \in \ Q_D$ верно (2).

Замечание 5. В дальнейшем в работе будем опускать выражения типа "верно (2)"для различных вариаций параметров.

3.$\gamma \neq 0$ 3.1 $\alpha\beta \ \geq \ 0$ 3.1.1 $\gamma \ = \ \alpha \ + \ \beta \ , \forall X \in Q$.

3.1.2 $-\gamma \ = \ \alpha \ + \ \beta \ , \forall X \in Q_S$.

3.2 $\alpha\beta < 0$ 3.2.1 $|\gamma| > max\{|\alpha|, |\beta|\}, (\gamma - \alpha = -\beta) \vee (\gamma - \beta = -\alpha), \forall X \in Q_S$.

3.2.2 $|\gamma| \ \leq max\{|\alpha|, |\beta|\}, \alpha - \ \gamma = \ \beta) \vee (\beta - \gamma = - \ \alpha), \quad \forall X \in Q_D$.

Следствие 4. Квазилинейное уравнение

$$X + (-1)X = \alpha X \tag{3}$$

имеет решение только при $(\alpha = 0) \vee (\alpha = 2)$.

Получены решения для уравнения $X = A + \beta X$.

Свойство 10. $< Q_1, +, \cdot, P >$ является квазилинейным подпространством пространства Q, если

$$A + B \in Q_1, \quad \forall A, B \in Q_1,$$

$$\alpha \cdot A \in Q_1, \quad \forall A \in Q_1, \quad \forall \alpha \in P.$$

Определим оператор: $H(A) = \alpha(A + (-1)A)$.

Из свойств 1 и 2 для $\forall \alpha \neq 0$: $Q_D \ = \ KerH(\cdot), \quad Q_S \ = \ ImH(\cdot)$.

При этом $Q_D(Q_S)$ являются максимальным линейным (симметричным) подпространством пространства Q.

Теорема 1. *Если Q — полное квазилинейное пространство с аксимой сокращения, тогда Q является прямой суммой линейного пространства множества дистрибутивных элементов и квазилинейного пространства симметричных элементов.*

Замечание 6. В произвольном квазилинейном пространстве получаем:

$$Q = Q_D \oplus Q_S \oplus Q_N,$$

таким образом $\forall X \in Q$ имеем

$$X = X_D + X_S + X_N,$$

где X_D, X_S — решение (3) при $\alpha = 0$ и $\alpha = 2$ соответственно, X_N — если (3) не имеет решения.

Следствие 5. Если Q является конечно-порожденным квазилинейным пространством, тогда

$$Q \cong \mathbb{R}^n \oplus \mathbb{R}_+^m \oplus \mathbb{R}_+^{2k}, \quad A = \mathbb{R}^{n\times n} \oplus \mathbb{R}_+^{m\times m} \oplus \mathbb{R}_+^{2k\times 2k},$$

где A — матрица линейного преобразования $F(\cdot)$.

Таким образом $X \in Q$, где Q — полное пространство, то $X = X_D + X_S$. Естественно, ввести определения середины и ширины элемента X: $X_D = \mathrm{mid}\ (X)$, $X_S = \mathrm{wid}\ (X)$.

Замечание 7. Теорема 1 верна и в более общем случае, если $H(\cdot)$ — гомоморфное отображение Q, такое что $KerH(\cdot)$ является квазилинейным групповым пространством.

Из следствия 5 естественно определим операцию вычитания в полном квазилинейном пространстве Q:

$$X - Y = (X_{D1} - Y_{D1}, \ldots, X_{Dn} - Y_{Dn}, |X_{S1} - Y_{S1}|, \ldots, |X_{Sm} - Y_{Sm}|). \tag{4}$$

Рассмотрим квазилинейную систему

$$FL(X) + A = FR(X) + B,$$

где $FL(\cdot), FR(\cdot)$ — гомоморфизмы пространства Q. Если $A = B$, тогда квазилинейная система в квазилинейном пространстве с аксиомой сокращения называется однородной, если $A \neq B$, тогда будем называть неоднородной.

Свойство 11. [10] Решение однородной квазилинейной системы является квазилинейным подпространством пространства Q, (см. свойство 3.)

Свойство 12. [10] Решение неоднородной квазилинейной системы есть прямая сумма частного решения неоднородной квазилинейной системы и решения однородной квазилинейной системы.

2.3 Построение квазилинейного пространства при помощи линейного

Пусть $< L, +, \cdot, P >$ — линейное пространство, $S(L)$ — множество подмножеств этого постранства L. Для любых $X, Y \in S(L)$ и $\forall \alpha \in P$ известны операции Минковского:

$$X \oplus\ Y = \ \{z \mid\ z\ = x + y, x \in X, y \in Y\} \tag{5}$$

$$\alpha \odot X = \ \{z \mid z = \alpha x, \ \ x \in\ X\}. \tag{6}$$

Таким образом стандартные интервальные операции являются частными случаями (5),(6).

Пусть $V(L)$ — множества всех выпуклых подмножеств .

Теорема 2. $< V(L), \oplus, \odot, P >$ *является максимальным квазилинейным пространством, т.е. любое квазилинейное пространство, порожденное операциями (5),(6) является квазилинейным подпространством пространства $V(L)$.*

Построены различные примеры квазилинейных пространств.

3 Исследования задачи Коши для ОДУ и интервальных дифференциальных уравнений

В этом параграфе мы исследуем связи решений вышеописанных задач. Введено определение производной интервальной функции в нормированном квазилинейном пространстве [11].

3.1 Сходимость одношаговых интервальных методов [22-27]

Эффект обертывания (wrapping effect) [1,2,12-21] для систем дифференциальных уравнений с интеральными данными усложняет построение приближенных методов даже в случае линейных систем. Проблема эффекта обертывания с алгебраической точки зрения заключается в выполнении общего закона субдистрибутивности (включения) для любых бинарных операций $*, +$:

$$A * (B + C) \subseteq A * B + A * C.$$

Пусть задана динамическая система

$$\dot{x} = f(x,p), \qquad t \in [0, t_f], \qquad x \in \mathbb{R}^n. \tag{7}$$

Начальные данные из некоторого выпуклого множества X^0:

$$x(0) \in X^0, \qquad X^0 \subseteq \mathbb{R}^n. \tag{8}$$

Для параметра задано условие:

$$p \in P, \qquad P \in \mathbb{IR}^l \text{ — множество интервальных векторов.} \tag{9}$$

Стандартно выпишем множество достижимости задачи (7)–(9):

$$X^*(t, X^0, P) = \{\, x(\cdot) \mid \dot{x}(t) = f(x(t), x_0, p), x_0 \in X^0, p \in P \,\}.$$

Таким образом, имеем:

$$X^*(t, X^0, P) = \bigcup_{x_0 \in X^0} \bigcup_{p \in P} x(t, x_0, p).$$

В дальнейшем множество $X^*(t, X^0, P)$ будем записывать как $X(t)$. Для задачи Коши (7)–(9) из предположения, что для любого $x_0 \in X^0$ и любого $p \in P$ существует единственное решение $x(t, x_0, p)$, выпишем точное решение этой задачи для $t \geq \tau$ в виде:

$$X(t) = C_1(t, \tau, X(\tau)) = C(t, \tau) \circ X(\tau).$$

Введем дискретизацию на отрезке $[0, t_f]$:

$$t_0 = 0, \qquad t_i = t_{i-1} + h_i, \qquad i = \overline{1, r}, \qquad t_r = t_f, \tag{10}$$

h_i — шаг i—го приближенного интегрирования.

В операторном виде верно:

$$X(t_i) = C_i \circ X(t_{i-1}), \quad \text{где } \ C_i = C(t_i, t_{i-1}). \tag{11}$$

Таким образом, для задачи (7)–(9): $C_i : 2^{\mathbb{R}^n} \to 2^{\mathbb{R}^n}, \quad i = \overline{1, r}$.

Выпишем общую формулу одношагового интервального метода в квазилинейном пространстве $Q \equiv V(\mathbb{R}^n)$ для задачи (7)–(9). Здесь $V(\mathbb{R}^n)$ — множество всех выпуклых подмножеств линейного пространства $\mathbb{R}^n$:

$$X_i = S_i(X_{i-1}), \quad S_i : V(\mathbb{R}^n) \to V(\mathbb{R}^n), \quad i = \overline{1, r}, \quad X_0 = X^0. \tag{12}$$

Определение 11. Интервальный метод (12) внешне аппроксимирует задачу (7)–(9) с порядком m, если выполняются следующие условия для $i = \overline{1, r}$:

$$q(C_i X, S_i X)) \ \leq \ k_1 h^{m+1}, \quad \forall\, X \in V(R_n), \tag{13}$$

$$C_i X \subseteq S_i X, \quad \forall X \in 2^{\mathbb{R}^n}, \tag{14}$$

здесь и далее в работе $k_j (j \geq 1)$ — константы не зависящие от i и h_i, но зависящие от $q(\cdot,\cdot)$ — метрики квазилинейного пространства Q. Одной из таких метрик будет расстояние по Хаусдорфу для произвольных подмножеств A, B линейного пространства $\mathbb{R}^n$ с евклидовой метрикой $\rho(\cdot,\cdot)$:

$$q(A,B) = \max\{\max_{a\in A}\min_{b\in B}\rho(a,b), \max_{b\in B}\min_{a\in A}\rho(a,b)\}.$$

Замечание 8. Естественным образом можно ввести определение внутренней аппроксимации интервального метода (12) изменив операцию включения в (14).

Замечание 9. Если $X \notin V(\mathbb{R}^n)$ в (12), тогда вместо X нужно взять со X — выпуклую оболочку X :

$$\text{co}\, X = \bigcap_{X\subseteq Y, Y\in V(\mathbb{R}^n)} Y.$$

Замечание 10. Если в Q выполнены два условия:

$$1. \quad X(t_i) + D_i(h_i, X_0, P) = S_i X(t_{i-1}), \forall i, \quad 2. \quad \Theta \subseteq D_i(\cdot,\cdot,\cdot), \forall i.$$

Тогда очевидно выполнены условия внешней аппроксимации, здесь Θ — нейтральный элемент квазилинейного пространства Q.

Замечание 11. Внешнюю(внутреннюю) аппроксимацию естественным образом вводится для произвольных динамических систем, только в качестве аппроксимирующей динамической системы вместо метода (12) необходимо рассмотреть другую динамическую систему с начальным условием типа (8).

Для краткости введем обозначения: $S_{ij} = S_i \circ S_{i-1} \circ \cdots \circ S_j, \quad i \geq j.$

Будем предполагать, что операторы S_i из (10) монотонны по включениям и удовлетворяют следующему условию Липшица: $q(S_{ij}X, S_{ij}Y) \leq k_{ij} q(X,Y), \quad \forall X, Y \in Q.$

Определение 12. Метод (12) называется устойчивым, если выполнено неравенство:

$$k_{ij} \leq k_4, \quad \forall i,j \in 1,\ldots N, \quad i \geq j.$$

Определение 13. Интервальный метод $Y_i = D_i(Y_{i-1})$, $Y_0 = X^0$ называется подчиненным методу (12), если $Y_i \subseteq X_i \ \forall i$ и для любой дискретизации.

Определение 14. Интервальный метод (12) является внешне сходящимся к решению задачи (7)–(9) с порядком m, если выполнены следующие два условия:

$$1.\ q(X_i, X(t_i)) \leq k_5 h^m, \quad \forall i, \qquad 2.\ X(t_i) \subseteq X_i, \quad \forall i.$$

Замечание 12. Очевидно, если изменить включения в (14), тогда получим внутреннюю сходимость метода (12).

Теорема 3. *Пусть для $x_0 \in X^0$ и для любого $p \in P$ решение задачи (7)–(9) единственно, тогда $X^*(t, X^0, A)$ является связным множеством.*

Теорема 4. *Пусть в (1) $n = 1, a \in \mathbb{R}^l$. Тогда для метода (12), аппроксимирующего задачу (7), (8), имеем оценку:*

$$\text{wid}\ X_i \leq k_2 \text{wid}\ X_0 + k_3 h^m, \quad i = \overline{1,r}, \tag{15}$$

где $h_i = \max_{1\leq i\leq r}(h_i)$, $X = [i(X), s(X)]$, wid $(X) = s(X) - i(X)$.

Замечание 13. Оценка (15) не гарантирует сходимость методов, которые аппроксимируют задачу теоремы 4. Действительно, построим интервальный метод:

$$X_i = [\underline{s}_i i(X_{i-1}), \overline{s}_i s(X_{i-1})], \quad \underline{s}_i(\cdot) \leq c_i(\cdot) \leq \overline{s}_i(\cdot), i = \overline{1,r},$$

где $\underline{s}_i, \overline{s}_i$ — аппроксимирующие точечные методы , но не с сходящиеся. Тогда будем иметь интервальную аппроксимацию, но не иметь сходимости.

Теорема 5. *Если метод (12) внешне (внутренне) аппроксимирует задачу (7)–(9) с порядком m и метод (12) устойчив, тогда метод (12) является внешне (внутренне) сходящимся к решению задачи (7)–(9) с порядком m.*

4 Приложения

4.1 О методе эллипсоидов

Пусть поставлена задача Коши

$$\dot{x} = a(t)x, \qquad a(\cdot) \in \mathbb{R}^{n\otimes n}, \quad x \in \mathbb{R}^n, \quad t \in [0, t_f] \tag{16}$$

$$x(0) \in y(0) + X(0), \quad X(0) \in E(\mathbb{R}^n). \tag{17}$$

Здесь $E(\mathbb{R}^n)$ — множество эллипсоидов в $\mathbb{R}^n$, т.е. $X \subset \mathbb{R}^n$:

$$X = \{x \in \mathbb{R}^n \mid (d_X^{-1}x, x) \leq 1\},$$

d_X — строго положительная матрица ,

$a(t)$ — гладкая (бесконечно-дифференцируемая) матричная функция , например, элементы $a(t)$ — многочлены от аргумента t.

Выпишем интервальный пошаговый метод для задачи в n-мерном пространстве эллипсоидов $E(\mathbb{R}^n)$:

$$d_i = s_i \cdot d_{i-1} \cdot s_i^*, \quad i = \overline{1, r}, \qquad X_i = \{x \mid (d_i^{-1}x, x) \leq 1\}. \tag{18}$$

Теорема 6. [23,24] *Если метод (18) внешне (внутренне) аппроксимирует решение задачи Коши (16), (17) (с порядком m) и устойчив, тогда интервальный метод (18) внешне (внутренне) сходится к решению этой задачи (с порядком m).*

Теорема 7. [23,24] *Пусть в методе (18)*

$$s_i = e + h_i(a(t_{i-1}) + a(t_{i-1} + 2h/3))(3e + h_i a(t_{i-1}) + 2h_i a(t_{i-1)} + h_i a(t_i)))/4 \tag{19}$$

и система (16) такова, что

$$\forall x \in \mathbb{R}^n, \qquad \forall \theta \in [0, t_f], \qquad (a^3(\theta)x, x) < 0, \tag{20}$$

тогда метод эллипсоидов (18) с оператором (19) внешне сходится к решению задачи (16)–(17), удовлетворяющих условию (20), с порядком 2.

4.2 О методе параллелепипедов

Рассмотрим линейную систему (16) с начальным условием

$$x(0) \in X(0) \in \mathbb{IR}^n. \tag{21}$$

Выпишем формулу модифицированного метода Никеля в пространстве n –мерных параллелепипедов $P(\mathbb{R}^n)$:

$$X_i = (\sum_{j=0}^{m} h_i^j \cdot \psi_j(t_i)/j!) \cdot X_{i-1} + \text{пр}_{\alpha_j}((h_i^{m+1} \cdot \Psi_{m+1}(T_i) \cdot Z_i)/(m+1)!), \tag{22}$$

где α_j — вектор соответствующей диагонали параллелепипеда, вычисляемой по формуле :

$$\alpha_j = a(t_i)(x_j - \sum_{j=0}^{n} x_j/n).$$

Теорема 8. [26,27] *Метод (22) с постоянным шагом h, выбранный в специальном соответствии с ошибкой округления ε_L, численно сходится в метрике Хаусдорфа к решению задачи Коши (16), (21), если $a(t)$ — полиномиальная матричная функция с целыми коффициентами.*

Рассмотрим следующую интервальную задачу Коши с постоянными коэффициентами

$$\dot{x} = ax + b, \qquad x(0) \in X(0), a \in A, b \in B, A \in \mathbb{IR}^{n\otimes n}, B \in \mathbb{IR}^n. \tag{23}$$

Для оценки стационарных точек системы вводится определение [21,28,29] числовой характеристики $\gamma(A)$ интервальной матрицы A:

$$\gamma(A) = \sup_{a \in A} \|e - Ya\|,$$

где $\|\cdot\|$ — произвольная матричная норма, e — единичная матрица, $Y = \text{mid } (A)^{-1}$. Доказаны свойства введенной характеристики. В частности, она является обобщением понятия относительной полуширины.

Для множества достижимости X^* стационарной линейной системы с интервальными коэффициентами имеем следующую оценку [21,28,29]:

$$\text{diam } X^* \leq \|Y \text{ wid } B\| + 2\gamma(A)\|B\|/(1 - \gamma(A)).$$

Исследованы внешние оценки для движения КА [30] в кеплеровой постановке задачи с возмущениями в начальных данных. Получен оптимальный радиус в зависимоти от начального возмущения в случае круговой орбиты. Рассмотрены аналогичные задачи для движения КА с малыми возмущениями в тяге.

Проведены численные расчеты: для классических пошаговых методов, в пространствах эллипсоидов второго, четвертого порядков сходимости и в пространстве параллелепипедов 4–6 порядков.

Список литературы

1. Moore R.E. *Interval analysis.* – N. J.: Prentice Hall, 1966.
2. Калмыков С.А., Шокин Ю.И., Юлдашев З.Х. *Методы интервального анализа.* — Новосибирск: Наука, 1986.
3. Алефельд Г., Херцбергер Ю. *Введение в интервальные методы.* – М. Мир, 1987.
4. Markov S.M. A non-standard substraction of intervals // *Serdica.* – 1997. – P. 359 – 370.
5. Mayer O. Algebraische und metrische Strukturen in der Intervallrechnung und einige Anwendungen // *Computing.* – 1970. – Vol. 5. – P. 144 – 162.
6. Ratchek H. Nichtnumerishe Aspekte der Intervallarithmetik // *Interval Mathematics 1975*, K. Nickel, ed. – Berlin: Springer, 1975. – P. 48 – 49. – (*Lecture Notes in Computer Science*; vol. 29).
7. Schroder G., Charakterisierung des quasilinearen Raumes I(R) und Klassifizierung der quasilinearen Raume der dimension 1 und 2 // *Computing.* – 1972. – Vol. 10. – P. 111 – 120.
8. Утюбаев Г.Ш. Гомоморфизмы квазилинейных пространств и их свойства // *Вычислительные технологии.* – 2001. – Т. 6, ч. 2. – С. 632 – 638.
9. Утюбаев Г.Ш. Гомоморфизмы квазилинейных пространств с аксиомой сокращения и их свойства // Труды Межд. конференции “Симметрия и дифференциальные уравнения”. – Красноярск, 2000. – С. 229 – 231.
10. Утюбаев Г.Ш. О решении квазилинейных систем // Тезисы Межд. сем. по теор. групп. Екатеринбург: УрГУ, 2001. – С .228 – 230.
11. Утюбаев Г.Ш. Производная в квазилинейном пространстве // *Вопросы математического анализа.* Сборник трудов. – Красноярск: КГТУ, 1997. – С. 129 – 131.
12. Черноусько Ф.Л. *Оценивание фазового состояния динамических систем. Метод эллипсоидов.* – М.: Наука, 1988.
13. Nickel K.L.E. Using interval methods for numerical solution of ODEs // *ZAMM.* – 1986. – Bd. 66, No. 11. – P. 513 – 523.
14. Вербицкий В.И., Горбань А.Н., Утюбаев Г.Ш., Шокин Ю.И. Эффект Мура в интервальных пространствах // *ДАН СССР.* – 1989. – Т. 304, №1. – С. 17 – 22.
15. Рогалев А.Н., Шокин Ю.И. Исследование и оценка решений обыкновенных дифференциальных уравнений интервально-символьными методами // *Вычислительные технологии.* – 1999. – Т. 4, №4. – С. 51 – 75.
16. Neumaeir A. The wrapping effect, ellipsoids arithmetic, stability and confidence regions // *Computing Supplement 9.* – 1993. – P. 175 – 190.
17. Neumaier A. Global, rigorous and realistic bounds for the solution of dissipative differential equations // *Computing.* – 1994. – Vol. 52. – P. 315 – 336.
18. Kuehn W. Rigorously computed orbits of dinamical systems without the wrapping effect // *Computing.* – 1998. – Vol. 61. – P. 47 – 67.
19. *Nedialkov N.S., Jackson K.R., Corliss G.F.* Validated solutions of initial value problems for ordinary differential equations // *Applied Mathematics and Computation.* – 1999. – Vol. 105. – P. 21 – 68.
20. Добронец Б.С., Шайдуров В.В. *Двусторонние численные методы.* – Новосибирск: Наука, 1990.
21. Новиков В.А., Рогалев А.Н. Построение сходящихся верхних и нижних оценок решений систем обыкновенных дифференциальных уравнений // *ЖВМиМФ.* – 1993. – Т. 33, №2. – С. 219 – 231.
22. Новиков В.А., Утюбаев Г.Ш. О сходимости интервальных методов решения задачи Коши для обыкновенных дифференциальных уравнений // Красноярск, 1988. – С. 33–36. – (Препринт ВЦ СО АН СССР; №6).
23. Утюбаев Г.Ш. Построение и анализ сходящихся интервальных методов решения задачи Коши для обыкновенных дифференциальных уравнений. – Автореф. диссерт. на соиск ... к.ф.-м.н. – Красноярск, 1992 – 18 с.

24. Утюбаев Г.Ш. О внешней (внутренней) сходимости одношаговых интервальных методов // *Вычислительные технологии* (Спец. выпуск). – 2002. – Т. 7 (4). – С. 211 – 218.
25. Утюбаев Г.Ш. О методе эллипсоидов для линейной системы обыкновенных дифференциальных уравнений. – Красноярск, 1990. – (Препринт ВЦ СО АН СССР; №16). – С. 29 – 32.
26. Утюбаев Г.Ш. Модификация метода параллелепипедов для линейных динамических систем с начальными данными из интервальных множеств // *Вестник СибГАУ*. – Красноярск, 2002. – С. 32 – 37.
27. Утюбаев Г.Ш. Интервальные методы для динамических систем с неопределенностями в параметрах // *Вестник САА*, сборник научных трудов. – Красноярск, 2001. – Вып. 2. – С. 29 – 38.
28. Утюбаев Г.Ш. Числовая характеристика интервальной матрицы // Труды Межд. конф. по алгебре и лин. оптим., ред. Еремин И.И. – Екатеринбург: ИММ УрО РАН, 2002. – С. 206 – 212.
29. Утюбаев Г.Ш., Шокин Ю.И. О числовой характеристике интервальной матрицы // *Вычислительные технологии* (Спец. выпуск). – 2002. – Т.7 (4). – С. 219 – 224.
30. Utyubaev G.Sh. On interval methods for dinamical systems with uncertainty in parameters // Book of Abstracts of Int. Symp. on Sci. Comp., Rene Alt, ed. – Paris, 2002. – P. 138 – 139.

Интервально-аффинная арифметика с управляемой точностью

Р.Р. Ахмеров

Алтайский госуниверситет,
Россия, г. Барнаул
arr@ctta.ru

Аннотация. Работа посвящена интервально-аффинной арифметике с управляемой точностью – новому способу построения внешних оценок областей значений функций. Рассмотрены проблемы классического подхода и способы их преодоления. Показаны преимущества нового способа оценивания и сложности, которые при этом возникают. Предложен метод управления точностью получаемых оценок. Приводятся результаты вычислительных экспериментов

Основной областью применения классической *интервальной арифметики* [1] является внешнее оценивание областей значений функций при известных интервалах изменения аргументов. Любой алгоритм, использующий на каком-либо этапе интервальную арифметику, за редким исключением, эксплуатирует ее именно в этих целях. В этом смысле интервальная арифметика обладает рядом достоинств и недостатков.

Достоинства интервальной арифметики:

1. Простота реализации и использования.
2. Интервальные операции могут быть реализованы эффективно и выполняться лишь немного медленнее, чем соответствующие вещественные операции.
3. Возможность применять для широкого класса задач.
4. Гарантированность получаемых результатов.

Главным недостатком применения интервальной арифметики для задач оценивания является малая точность поставляемых ей оценок. Часто оценки оказываются малоинформативными и практически бесполезными. Это препятствует широкому распространению интервального подхода на практике.

С целью повышения качества оценок был разработан ряд методов. Значительная часть этих методов основана на анализе функций и на нижнем уровне, так или иначе, использует интервальную арифметику. Такие методы мы назовем интервально-аналитическими. К методам этого типа можно отнести: среднезначную форму, разложения в ряд Тейлора, технику наклонов (slopes), интерполяционные формы и т.д. [1, 2, 3]

Все без исключения интервально-аналитические методы не дают гарантии повышения точности оценок. Для широких входных интервалов они обычно дают даже худшие оценки, чем простые интервальные расширения. Достоинство интервально-аналитических методов – более высокий порядок сходимости приближенных оценок к точным при стремлении ширин входных интервалов к нулю. Это подразумевают некоторую процедуру дробления области изменения аргументов, что сопряжено со значительными вычислительными затратами даже для задач сравнительно небольшой размерности, порядка нескольких десятков или сотен.

Другая проблема связана с анализом функций заданных алгоритмически, таких как метод Гаусса для решения систем линейных уравнений. В этом случае применение интервально-аналитического подхода сопряжено со значительными трудностями, часто непреодолимыми. Надо заметить, что наклоны (slopes) в принципе могут быть использованы для решения подобного рода задач, если для расчета функции наклона $g(x, y)$ использовать процедуру, описанную, например, в [3]. В этом случае стоит ожидать некоторое повышение точности оценок.

Альтернативой интервально-аналитическому подходу является использование арифметик другого вида. Самые известные из них – обобщенная арифметика Хансена [4] и аффинная арифметика [5]. Их действие базируется на хранении большего количества информации о величине, чем это делает интервальная арифметика. Это дает возможность получать более точные оценки на каждом шаге алгоритма. Повышение точности оценок достигается ценой увеличения времени выполнения операций. Тем самым повышается сложность основанных на таких арифметиках алгоритмов. Например, метод Гаусса, основанный на интервальной арифметике, имеет сложность $O(n^3)$. Тот же метод, основанный на арифметике Хансена или на аффинной арифметике в ее оригинальной реализации, имеет сложность $O(n^5)$.

В данной работе мы представляем арифметику нового вида – *интервально-аффинную арифметику с* *управляемой точностью*. Рассмотрены основания для построения такой арифметики. Показано, что интервально-аффинная арифметика всегда точнее, чем интервальная. Приводится метод управления точностью оценок, поставляемых этой моделью, что дает возможность изменять сложность алгоритмов. Рассмотрены нюансы реализации арифметики на языке программирования C++. В конце работы приводятся результаты вычислительных экспериментов с интервально-аффинной арифметикой.

Список литературы

[1] Moore R.E. *Methods and Applications of Interval Analysis*. – Philadelphia: SIAM, 1979.

[2] Neumaier A. *Interval Methods for Systems of Equations*. – Cambridge: Cambridge University Press, 1990.

[3] Hansen E. *Global Optimization Using Interval Analysis*. – New York: Marcel Dekker, 1992.

[4] Hansen E. A generalized interval arithmetic // Interval Mathematics, K. Nickel, ed. – Berlin: Springer Verlag, 1975. – P. 7–8. – (Lecture Notes in Computer Science; vol. 29).

[5] Stolfi J. and de Figueiredo L.H. *Self-Validated Numerical Methods and Applications* – Notes of 21st Brazilian Mathematics Colloquium, 1997.

Гарантированно субоптимальные решения задач линейной оптимизации

А.Г. Ершов

Компания SIB3
Россия, 630090 г. Новосибирск
просп. Лаврентьева, 6

Аннотация. В данной статье рассматривается вопрос о возможности эффективного вычисления на ЭВМ субоптимальных решений задач линейной оптимизации, гарантированно удовлетворяющих условиям допустимости; вводится класс задач, для которых проблема поиска таких решений является в некотором смысле корректной — класс так называемых устойчиво допустимых задач линейной оптимизации. Представляется новый алгоритм поиска субоптимального гарантированно допустимого решения задачи линейной оптимизации, основанный на методе внутренних точек (Interior Point) и использующий технику малого возмущения решаемой задачи. Указывается способ построения с помощью такого алгоритма субоптимальной внешней и субмаксимальной внутренней интервальных оценок множества субоптимальных решений задачи линейной оптимизации.

Введение

Данная работа выполнялась в рамках экспериментальной системы — расширяемого вычислительного сервера, основанного на математическом аппарате программирования в ограничениях (рабочее название Solver). Программирование в ограничениях — это универсальный подход к решению систем ограничений, позволяющий работать с вычислительными моделями, содержащими переменные различных типов, связанные разнородными ограничениями (см. [5]). Помимо универсальности, такой подход имеет еще два важных преимущества: гарантированность вычислений и возможность работы с недоопределенными данными. Во время работы методов программирования в ограничениях происходит постепенное уточнение информации об области возможных значений переменных модели, - такой подход позволяет работать с недоопределенными данными. В системе Solver информация об области возможных значений вещественных переменных модели представляется в виде внешней интервальной оценки, то есть интервала, который содержит все возможные значения переменной. Известно, что алгоритмы программирования в ограничениях имеют экспоненциальную сложность вычислений, поэтому применение программирования в ограничениях оправдано для решения сложных систем алгебраических уравнений, содержащих трансцендентные функции, или сложных смешанных задач, анализ которых затруднен, но не слишком адекватно для решения хорошо изученных в вычислительной математике задач. К таким задачам относятся, например, системы линейных уравнений и задачи линейной оптимизации (также называемых задачами линейного программирования, далее ЗЛП), для которых существуют специальные полиномиальные методы решения. Мы будем исследовать вопрос о возможности получения гарантированных интервальных оценок на оптимальное значение целевой функции и множество решений ЗЛП с помощью полиномиальных алгоритмов.

В данной статье будем рассматривать задачи оптимизации — как классические, так и интервальные. Классическими постановками задач линейной оптимизации являются общая, каноническая и стандартная задачи линейной оптимизации

$$\begin{cases} \min \leftarrow c^T * x \\ A_1 * x = b_1 \\ A_2 * x \le b_2 \\ A_3 * x \ge b_3 \\ x \in R_+^n \end{cases} \quad (1о) \qquad \begin{cases} \min \leftarrow c^T * x \\ A * x = b \\ x \in R_+^n \end{cases} \quad (1к) \qquad \begin{cases} \min \leftarrow c^T * x \\ A * x \le b \\ x \in R_+^n \end{cases} \quad (1с)$$

Заметим, что эти постановки эквивалентны друг другу - задачу любого из трех видов можно преобразовать к любому другому виду.

Определение. Означивание переменных ЗЛП называется допустимым решением, если оно удовлетворяет всем ограничениям ЗЛП.

Определение. Означивание переменных задачи линейной минимизации (максимизации) называется оптимальным решением, если

1) Оно является допустимым решением.
2) Значение целевой функции на нем не больше (не меньше), чем на любом другом допустимом решении. Такое значение называется оптимальным значением целевой функции (далее ОЗЦФ).

Рассмотрим следующий пример.

Пример 1. $\begin{cases} \max \leftarrow x+y+z \\ 3*x+2*y=3 \\ 2*x+3*y=3 \\ x+z \le 1 \\ x,y,z \ge 0 \end{cases}$. Все коэффициенты в модели — представимые в ЭВМ целые, однако оптимальное решение $\left(\frac{3}{5},\frac{3}{5},\frac{2}{5}\right)$ не представимо в ЭВМ с обычной бинарной разрядной сеткой. Более того, множество допустимых решений $\left\{\left(\frac{3}{5},\frac{3}{5},t\right), t\in[0,\frac{2}{5}]\right\}$ не содержит представимых в ЭВМ векторов.

Известно, что можно гарантировать существование у ЗЛП допустимого решения, если ЗЛП имеет так называемое "ε-допустимое решение с достаточно малым параметром ε" (см. [3]), однако оценка такого "достаточно малого ε" включает в себя неизвестные априори величины, вычисление которых затруднено, потому качественный анализ существования допустимого решения ЗЛП на практике не проводится. Таким образом, получение точного оптимального (или хотя бы точного допустимого) решения произвольной ЗЛП на ЭВМ, равно как и гарантирование его существования, либо невозможно, либо сложно в вычислительном плане. Возникают следующие вопросы:

1) Для какого класса задач линейной оптимизации существуют методы, позволяющие численно получать машинно-представимые оптимальные решения?
2) Каким образом можно ослабить условия оптимальности решения так, чтобы класс задач, для которых численный поиск машинно-представимых решений такого вида имел бы смысл, был достаточно широк?

Сначала мы займемся исследованием второго вопроса, изучив свойства множеств субоптимальных ("близких к оптимальным") решений задач линейного программирования.

Субоптимальные решения

Сформулируем условия субоптимальности решения ЗЛП:

1) Решение должно быть допустимым — то есть удовлетворять всем ограничениям ЗЛП.
2) Решение должно быть близко к оптимальному. Характеристики оптимальности могут быть разными, так, мы будем считать решение x субминимальным (субмаксимальным) с параметром ε в абсолютном смысле, если для любого допустимого решения y выполняется $c^T*y \ge c^T*x-\varepsilon$ (выполняется $c^T*y \le c^T*x+\varepsilon$) (2.1); субминимальным (субмаксимальным) с параметром ε в относительном смысле, если для любого допустимого решения y выполняется $c^T*y \ge (c^T*x)*(1-\varepsilon*sign(c^T*x))$ (соответственно выполняется $c^T*y \le (c^T*x)*(1+\varepsilon*sign(c^T*x))$) (2.2). Мы не будем далее рассматривать субоптимальные решения в относительном смысле для задач с нулевым ОЗЦФ.

Так как условия субоптимальности решения (2.1), (2.2) при $\varepsilon=0$ совпадают с условиями оптимальности, множество всех оптимальных решений (далее МОР) является частным случаем множества всех субоптимальных решений (далее МСР). Кроме того, МСР задачи (1о) совпадает с множеством всех допустимых решений (далее МДР) той же самой задачи с одним дополнительным ограничением $c^T*x \le l$ (2.3), где l - константа, зависящая от ε и типа субоптимальности; из этого следует выпуклость МСР. Раз все оптимальные решения являются субоптимальными, то МСР ЗЛП содержит МОР и содержится в МДР. Легко видеть, что МСР с параметром ε_1 содержит любое МСР с параметром ε_2 при $0 \le \varepsilon_2 \le \varepsilon_1$. Все перечисленные выше качественные свойства МСР не зависят от типа оптимальности (максимальность или минимальность) и типа субоптимальности (абсолютная или отрицательная).

Теперь мы докажем четыре предложения об общих свойствах МСР и МДР (предложения 1 и 2), МСР и МОР (предложения 3 и 4).

Предложение 1. Для любого типа субоптимальности, и для любого значения параметра субоптимальности $\varepsilon > 0$, мощность МСР совпадает с мощностью МДР.

Доказательство. Как известно, МДР может иметь мощность континуума, 1 или 0. Если допустимое решение единственно, то оно оптимально, а значит, и субоптимально. Если допустимых решений нет, то, так как любое субоптимальное решение допустимо, нет и субоптимальных решений. Докажем, что континуальность МДР ЗЛП влечет континуальность МОР той же самой ЗЛП для любого значения параметра субоптимальности, большего 0. Действительно, рассмотрим отрезок, соединяющий некоторое оптимальное решение с некоторым другим допустимым решением - по свойству выпуклости МДР, отрезок между ними целиком состоит из допустимых решений, а в силу линейности целевой функции, целый интервал ненулевой длины состоит из решений, удовлетворяющих условию субоптимальности.

В то же время мощность МОР (то есть МСР с параметром $\varepsilon = 0$) не всегда равна мощности МДР.

Предложение 2. МСР ЗЛП в абсолютной мере с параметром субоптимальности $\varepsilon > 0$ имеет внутреннюю точку тогда и только тогда, когда МДР той же ЗЛП имеет внутреннюю точку.

Доказательство. Достаточность. Пусть в МДР содержится некоторый шар X. Пусть y — некоторое оптимальное решение. Тогда множество $\{\gamma * x + (1-\gamma) * y \mid x \in X\}$ имеет внутреннюю точку для любого $\gamma \in (0,1]$, и состоит полностью из субоптимальных решений для всех достаточно малых γ. Необходимость очевидна, так как любое МСР содержится в МДР.

Аналогичный факт имеет место и для МСР в относительной мере для значения параметра субоптимальности $\varepsilon > 0$ при условии неравенства ОЗЦФ нулю. Заметим, что существование внутренней точки у МДР и МСР не влечет существование внутренней точки у МОР.

Предложение 3. МСР ЗЛП является ограниченным тогда и только тогда, когда МОР той же ЗЛП является ограниченным.

Доказательство. Как уже было замечено, МОР и МСР ЗЛП (1o) являются МДР ЗЛП, полученными из ЗЛП (1o) добавлениями ограничений вида $c^T * x \le l_o$ (для МОР) и $c^T * x \le l_c$ (для МСР). Ограниченность МДР задачи с добавленным ограничением эквивалентна ограниченности ЗЦФ всех задач вида $\begin{cases} \min \leftarrow x_i \\ c^T * x \le l_o \\ \dots \end{cases}$ и $\begin{cases} \max \leftarrow x_i \\ c^T * x \le l_o \\ \dots \end{cases}$ (для МОР) и $\begin{cases} \min \leftarrow x_i \\ c^T * x \le l_c \\ \dots \end{cases}$ и $\begin{cases} \max \leftarrow x_i \\ c^T * x \le l_c \\ \dots \end{cases}$ (для МСР). Если мы рассмотрим пары задач для МОР и МСР, отличающиеся лишь коэффициентом правой части в добавленном ограничении, то заметим, что двойственные к ним задачи будут отличаться только одной координатой вектора целевой функции, следовательно, МДР таких задач совпадают. Так как непустота МДР двойственной задачи определяет ограниченность ЗЦФ прямой задачи, мы получаем эквивалентность ограниченности МОР и МСР.

Очевидно, что МСР и МДР ЗЛП могут не являться одновременно ограниченными.

Определение. Внешней (внутренней) интервальной оценкой множества в R^n называют содержащий его (содержащийся в нем) n-мерный брус $[x_l, x_r]$ с гранями, параллельными координатным плоскостям (см. [4]).

Определение. Наименьшей внешней (наибольшей внутренней) интервальной оценкой называют такую оценку, что любая другая оценка содержит ее (содержится в ней). Такие оценки называют еще оптимальными.

Определение. Минимальной внешней (максимальной внутренней) интервальной оценкой называют такую оценку, что никакая другая оценка не содержится в ней (не содержит ее).

Предложение 4. Оптимальные внешние интервальные оценки ограниченного непустого МСР ЗЛП непрерывно зависят от параметра субоптимальности ε.

Доказательство. Оптимальная интервальная внешняя оценка МСР ЗЛП по координате i имеет границы, равные ОЗЦФ ЗЛП $\begin{cases} \min \leftarrow x_i \\ c^T * x \le l \\ \dots \end{cases}$ и $\begin{cases} \max \leftarrow x_i \\ c^T * x \le l \\ \dots \end{cases}$, полученных из исходной ЗЛП добавлением одного ограничения и изменением целевой функции, при этом l является величиной, непрерывно зависящей от ε. Для определенности рассмотрим ЗЛП для определения нижней границы. Ее МДР есть МСР исходной задачи. Двойственной к ней будет ЗЛП $\begin{cases} \max \leftarrow b^T * y + l * w \\ \dots \end{cases}$, МДР которой никак не зависит от параметра ε. ОЗЦФ этой двойственной ЗЛП (равный искомой нижней границе внешней оценки по координате i) будет непрерывно зависеть от l, а значит, и от ε, что и доказывает предложение.

Устойчиво допустимые задачи линейного программирования.

Теперь мы вернемся к первому поставленному во введении вопросу, — для какого класса задач линейной оптимизации возможен поиск оптимальных, и для какого — субоптимальных решений. Вспомним, что МДР, МОР и МСР ЗЛП являются выпуклыми множествами, и что имеет место следующее предложение.

Предложение. Выпуклое множество в евклидовом пространстве имеет внутреннюю точку тогда и только тогда, когда оно не содержится ни в какой гиперплоскости пространства.

Если множество содержится в некоторой гиперплоскости, то вряд ли оно содержит машинно-представимые элементы. Действительно, практическая вероятность того, что гиперплоскость в евклидовом пространстве имеет общие точки с некоторым фиксированным конечным множеством точек, крайне мала (она равна нулю, например, для случая равномерного распределения всех коэффициентов уравнения гиперплоскости в произвольных конечных интервалах ненулевой длины). Таким образом, можно считать поиск машинно-представимых элементов выпуклых множеств, не имеющих внутренних точек, некорректным. В то же время, если некоторое выпуклое множество имеет внутреннюю точку, то, в силу возможности аппроксимации любого вектора из R^n векторами из множества $\{k/2^n; k, n \in Z\}$, на ЭВМ с достаточно большой бинарной разрядной сеткой можно представить некоторые элементы такого выпуклого множества. Итак, мы будем считать проблему представления элементов выпуклых множеств на ЭВМ корректной для выпуклых множеств, имеющих внутреннюю точку.

Как уже было замечено, из существования внутренней точки у МДР следует существование внутренней точки у МСР, но не следует существование внутренней точки у МОР. Потому нам придется ввести два класса "хороших" ЗЛП.

Определение. ЗЛП называется устойчиво допустимой, если у нее существует допустимое решение, целая окрестность которого состоит из допустимых решений (то есть множество допустимых решений имеет внутреннюю точку).

Определение. ЗЛП называется устойчиво оптимизируемой, если у нее существует оптимальное решение, целая окрестность которого состоит из оптимальных решений.

Как уже было доказано, МДР и МСР одновременно имеют или не имеют внутренние точки, следовательно, устойчиво допустимые задачи являются также и "устойчиво субоптимизируемыми". К сожалению, рассматриваемые на практике задачи не являются устойчиво оптимизируемыми.

Предложение. Если вектор целевой функции не равен нулю, задача не является устойчиво оптимизируемой.

Доказательство. Пусть существует устойчиво оптимизируемая задача с ненулевым целевым вектором. Возьмем ее оптимальное решение x, целая окрестность которого состоит из оптимальных решений. Легко видеть, что значение целевой функции на x будет отличаться от значения целевой функции на другом оптимальном решении, полученным малым сдвигом из x по ненулевой координате целевого вектора, что противоречит условию оптимальности одного из этих решений.

Итак, класс устойчиво оптимизируемых задач является, в практическом смысле этого слова, вырожденным, что же касается класса устойчиво допустимых задач, то в него входят некоторые практически важные задачи ЗЛП.

Пример 2. $$\begin{cases} \max \leftarrow \sum\limits_j c_j * x_j \\ \sum\limits_{j=1}^{m} a_{i,j} * x_j \le b_i, i \in \overline{1,n} \\ x_j \ge 0, j \in \overline{1,m} \\ b_i, c_j, a_{i,j} \in R_+^n \end{cases} \tag{3.1}$$

Задача оптимизации выпуска продукции. Предприятие занимается выпуском продукции m типов, производящейся из сырья n типов, на выпуск единицы продукции типа j требуется $a_{i,j}$ единиц сырья типа i. Запасы сырья типа i равны b_i. Выпуск единицы продукции типа j приносит прибыль c_j. Надо максимизировать прибыль от выпуска продукции. Легко видеть, что если все коэффициенты b_i строго положительны, задача является устойчиво допустимой. Если же какие-то коэффициенты b_i равны нулю, то после удаления ограничений, их содержащих, и множества переменных $\{x_j \mid \exists i, a_{i,j} > 0, b_i = 0\}$ задача становиться устойчиво допустимой.

Пример 3. $$\begin{cases} \min \leftarrow \sum\limits_j c_{i,j,k} * x_{i,j,k} \\ \sum\limits_{j=1}^{n} \sum\limits_{k=1}^{t} v_k * x_{i,j,k} \le s_i, i \in \overline{1,n} \\ \sum\limits_{i=1}^{n} x_{i,j,k} \ge d_{j,k}, j \in \overline{1,n}, k \in \overline{1,t} \\ x_{i,j,k} \in R_+, i \in \overline{1,n}, j \in \overline{1,n}, k \in \overline{1,t} \end{cases} \tag{3.2}$$

Задача об оптимальном транспортном плане. Есть n городов, на каждом есть склад объема s_i и магазин, нуждающийся в m типах товаров. На хранение единицы товара типа k требуется пространство v_k. Необходимо составить генеральный план перевозок $x_{i,j,k}$ из складов в магазины так, чтобы обеспечить наличие в магазине j товара типа k в количестве не менее $d_{j,k}$, и не превысить объема склада s_i. Эта задача является разрешимой при $\sum\limits_{k=1}^{t}(\sum\limits_{j=1}^{n} d_{j,k}) * v_k \le \sum\limits_{i=1}^{n} s_i$, устойчиво допустимой при $\sum\limits_{k=1}^{t}(\sum\limits_{j=1}^{n} d_{j,k}) * v_k < \sum\limits_{i=1}^{n} s_i$.

Пример 4. $$\begin{cases} \min \leftarrow \sum\limits_j v_j * x_j \\ \sum\limits_{j=1}^{n} (\sum\limits_{s=1}^{t} d_{s,j}) * x_j \ge \sum\limits_{s=1}^{t} p_t, t \in \overline{1,T} \\ x \in R_+^n \end{cases} \tag{3.3}$$

Задача о минимальном инвестировании. Есть n видов ценных бумаг, цена одной бумаги вида j равна v_j, а доход от нее за период времени $[t, t+1]$ равен $d_{t,j}$. Необходимо минимизировать общую цену покупаемых бумаг так, чтобы обеспечить покрытие трат p_t за периоды времени $[t, t+1]$. Эта задача всегда является устойчиво допустимой.

Естественной формой записи устойчиво допустимой ЗЛП является запись без равенств. Действительно, если в задаче есть хотя бы одно ограничение-равенство, то множество ее допустимых решений содержится в гиперплоскости, которая задается этим равенством, откуда следует, что задача не является устойчиво допустимой. Потому следует, если это возможно, формулировать ЗЛП в стандартном виде, не вводя дополнительные переменные для приведения ее к каноническому виду.

Получение гарантированных интервальных оценок.

Рассмотрим проблему поиска гарантированно субоптимального решения устойчиво допустимых задач линейной оптимизации. Для решения ЗЛП применим метод Interior Point. Методом используется возмущенная система Карош-Кунн-Такера ЗЛП в стандартной постановке

$$\begin{cases} A*x-t=b \\ A^T*y+z=c \\ X*Z=\mu*e \\ T*Y=\mu*e \\ t\in R_+^m, z\in R_+^n, x\in R_+^n, y\in R_+^m \end{cases} \quad (4),$$

последовательное решение с $\mu\to 0$ ее ньютоновских приближений позволяет решить прямую и двойственную задачу за полиномиальное время (см. [1], [2]). Теоретически, последовательные приближения лежат строго внутри МДР, но фактически они быстро прижимаются к его границе, не позволяя верифицировать их допустимость прямым счетом с направленными округлениями. Предлагается решать равномерно возмущенную задачу

$$\begin{cases} A_j*x-t_j=b_j+\varepsilon*\|A_j\| \\ (A^i)^T*y+z_i=c_i-\varepsilon*\|A^i\| \\ X*Z=\mu*e \\ T*Y=\mu*e \\ t\in R_+^m, z\in R_+^n, x\in R_+^n, y\in R_+^m \end{cases} \quad (5),$$

где $\varepsilon\to 0, \mu\to 0$. Заметим, что малое возмущение такого рода устойчиво допустимых задач дает устойчиво допустимые задачи. Такое возмущение позволит верифицировать допустимость промежуточных решений прямым счетом с направленными округлениями.

Кроме поиска гарантированного субоптимального решения (под гарантированным субоптимальным решением мы будем понимать субоптимальное решение, допустимость которого можно верифицировать на ЭВМ с помощью вычислений с направленными округлениями), представляет интерес оценка параметра субоптимальности такого решения. Из теории двойственности ясно, что гарантированной внешней интервальной оценкой (далее ГВИО) ОЗЦФ задачи линейной минимизации (максимизации) может служить любой интервал $[b^T y, c^T x]$ (соответственно, $[c^T x, b^T y]$), где x — произвольное допустимое решение прямой, а y - произвольное допустимое решение двойственной задачи. Получив такой интервал можно с легкостью оценить сверху величину параметра субоптимальности как в абсолютном, так и в относительном (если оценочный интервал не содержит нуля) смысле. В процессе работы метода, последовательно решающего возмущенные задачи с $\varepsilon\to 0$, мы будем, таким образом, получать на каждом шаге все более точные ГВИО ОЗЦФ (алгоритмы получения интервальных оценок такого рода называются последовательно гарантирующими — см. [6]).

При успешном интервальном оценивании ОЗЦФ далее можно проводить интервальное внешнее оценивание МСР (а значит - МОР), решая задачи покоординатной оптимизации

$$\begin{cases} \min\leftarrow x_i \\ c^T*x\le l_+ \\ A\ *x\le b \\ x\in R_+^n \end{cases}, \text{ и } \begin{cases} \max\leftarrow x_i \\ c^T*x\le l_+ \\ A\ *x\le b \\ x\in R_+^n \end{cases},$$

где l_+ — гарантированная верхняя оценка на ОЗЦФ.

Полученная таким образом внешняя оценка МСР будет субоптимальной с известной ГВИО на меру субоптимальности по каждой координате и по каждой границе, а значит — субоптимальной внешней оценкой МОР. Заметим, что при условии несовпадения l_+ с ОЗЦФ задачи (1о), эти задачи будут устойчиво допустимыми тогда и только тогда, когда будет устойчиво допустимой исходная задача (1о). В любом случае, можно в качестве l_+ вместо полученной гарантированной верхней оценки целевой функции задачи (1о) подставить в задачи покоординатной оптимизации увеличенное на малое $\varepsilon>0$ ее значение, гарантируя тем самым сохранение устойчивой допустимости.

Для получения интервальных внутренних оценок МСР можно использовать алгоритм улучшения внутренней интервальной оценки, начиная с гарантированно субоптимального решения (которое является точечной внутренней интервальной оценкой МСР). Имея внутреннюю интервальную оценку $X_{i-1}=[X_{i-1}^-, X_{i-1}^+]$, мы можем получить оценку $X_i = [X_{i-1}^- - \alpha_i * d_i^- * e_i, X_{i-1}^+ + \alpha_i * d_i^+ * e_i]$, где коэффициент $\alpha_i \in [0,1]$, e_i — i-тый базисный орт, а величины d_i^-, d_i^+ равны:

$$d_i^+ = \min_j \begin{cases} \left\lfloor \left\lceil (b_j - \min\limits_{x \in X_{i-1}} (a_j^T * x)) \right\rceil / a_{i,j} \right\rfloor & \textit{для отношения вида } a_j^T * x \ge b_j, \textit{где } a_{i,j} < 0 \\ \left\lfloor \left\lfloor (b_j - \max\limits_{x \in X_{i-1}} (a_j^T * x)) \right\rfloor / a_{i,j} \right\rfloor & \textit{для отношения вида } a_j^T * x \le b_j, \textit{где } a_{i,j} > 0 \\ 0, \textit{ для отношения вида } a_j^T * x = b & \\ \infty, \textit{ во всех других случаях} & \end{cases},$$

$$d_i^- = \min_j \begin{cases} \left\lfloor \left\lfloor (\min\limits_{x \in X_{i-1}} (a_j^T * x) - b_j) \right\rfloor / a_{i,j} \right\rfloor & \textit{для отношений вида } a_j^T * x \ge b_j, \textit{где } a_{i,j} > 0 \\ \left\lfloor \left\lceil (\max\limits_{x \in X_{i-1}} (a_j^T * x) - b_j) \right\rceil / a_{i,j} \right\rfloor & \textit{для отношений вида } a_j^T * x \le b_j, \textit{где } a_{i,j} < 0 \\ 0, \textit{ для отношений вида } a_j^T * x = b & \\ \infty, \textit{ во всех других случаях} & \end{cases},$$

а $\min\limits_{x \in X_{i-1}} (a_j^T * x) = \left\lfloor \sum_k \left\lfloor a_{k,j} * \begin{cases} (X_{i-1}^-)_k^-, a_{k,j} \ge 0 \\ (X_{i-1}^+)_k^+, a_{k,j} < 0 \end{cases} \right\rfloor \right\rfloor$, $\max\limits_{x \in X_{i-1}} (a_j^T * x) = \left\lceil \sum_k \left\lceil a_{k,j} * \begin{cases} (X_{i-1})_k^+, a_{k,j} \ge 0 \\ (X_{i-1})_k^-, a_{k,j} < 0 \end{cases} \right\rceil \right\rceil$.

Если последовательно по каждой координате таким образом улучшать внутреннюю оценку, полагая $\alpha_i = 1$, мы получим субмаксимальную внутреннюю интервальную оценку (если не принимать в расчет ошибки округлений, то она является максимальной). Временные затраты получения внутренней оценки равны $O(n^2 * m)$, что позволяет считать ее вычисление эффективным.

Таким образом, мы получаем численный, корректно реализуемый на ЭВМ способ построения гарантированных субоптимальных решений и гарантированных субоптимальных внешних и субмаксимальных внутренних интервальных оценок множества всех субоптимальных решений устойчиво допустимой задачи линейной оптимизации.

Интервальные данные в задачах линейной оптимизации.

Входные данные ЗЛП — это матрица ограничений, вектор правой части и вектор коэффициентов целевой функции. Рассмотрим задачу $\begin{cases} \min \leftarrow (c^T * x) \\ [A_1^-, A_1^+] * x = [b_1^-, b_1^+] \\ [A_2^-, A_2^+] * x \le [b_2^-, b_2^+] \\ [A_3^-, A_3^+] * x \ge [b_3^-, b_3^+] \\ x \in R_+^n \end{cases}$ (6.1), возникающую из задачи (1о) при интервализации матрицы ограничений и вектора правой части. Легко видеть, что в силу условия неотрицательности переменных, возникающие линейные интервальные ограничения эквивалентны линейным скалярным ограничениям:

$$\sum_j [a_{i,j}^-, a_{i,j}^+] * x_j = [b_i^-, b_i^+] \Leftrightarrow \begin{cases} \sum\limits_j a_{i,j}^- * x_j \le b_i^+ \\ \sum\limits_j a_{i,j}^+ * x_j \ge b_i^- \end{cases} \quad (6.2)$$

$$\sum_j [a_{i,j}^-, a_{i,j}^+] * x_j \le [b_i^-, b_i^+] \Leftrightarrow \sum_j a_{i,j}^- * x_j \le b_i^+ \quad (6.3)$$

$$\sum_j [a_{i,j}^-, a_{i,j}^+] * x_j \ge [b_i^-, b_i^+] \Leftrightarrow \sum_j a_{i,j}^+ * x_j \ge b_i^- \quad (6.4)$$

Таким образом, ЗЛП (1о) с интервальной матрицей ограничений и вектором правой части эквивалентна обычной ЗЛП вида (1с). Если на переменные модели не накладывать условия неотрицательности, то задача с интервальными коэффициентами в ограничениях $\begin{cases} \min \leftarrow (c^T * x) \\ [A_1^-, A_1^+] * x = [b_1^-, b_1^+] \\ [A_2^-, A_2^+] * x \le [b_2^-, b_2^+] \\ [A_3^-, A_3^+] * x \ge [b_3^-, b_3^+] \\ x \in R^n \end{cases}$ (6.5) в общем случае уже не будет иметь эквивалентной ЗЛП. Действительно, в общем случае множество решений интервальной системы линейных уравнений, которое является множеством допустимых решений такой интервальной задачи, не является выпуклым, можно утверждать лишь выпуклость его пересечения с каждым из пространственных ортантов (см. [4]). Таким образом, невозможно применить к (6.5) впрямую методы линейного программирования. Одним из способов решения (6.5) может быть решение исходной задачи в каждом из ортантов, то есть решение задачи вида (6.1), и выбор среди полученных решений оптимального, однако такой подход очевидно неэффективен в силу экспоненциального роста временных затрат с увеличением числа переменных.

Интервальность коэффициентов целевой функции приводит к задаче: $\begin{cases} \min \leftarrow \sum_{i=1}^{n} [c_i^-, c_i^+] * x_i \\ A_1 * x = b_1 \\ A_2 * x \le b_2 \\ A_3 * x \ge b_3 \\ x \in R_+^n \end{cases}$ (6.6). Легко видеть, что все свойства допустимых решений и их множеств (например, выпуклость) задачи (1о) справедливы и для задачи (6.6). Тем не менее, МОР такой задачи может быть невыпуклым.

Пример 5. $\begin{cases} \min \leftarrow [-1,-1] * x + [-2,2] * y \\ x + y \le 1 \\ -x + y \le 1 \\ x - y \le 1 \\ -x - y \le 1 \end{cases}$. Множеством всех допустимых решений является множество $\{y = -x-1, x \in [-1,0]\} \cup \{y = x+1, x \in [-1,0]\}$, легко видеть, что оно не является выпуклым, то же касается и МСР с достаточно малым параметром субоптимальности. Следовательно, не существует задачи вида (1о), эквивалентной данной.

Экспериментальные результаты

Сравнение проводилось с демо-версией специализированного оптимизационного пакета OPL Studio, использующего решатель ILOG CPLEX. К сожалению, из-за встроенного в нее лимита на число переменных и ограничений задачи не удалось сравнить производительность систем на достаточно больших задачах. Тем не менее, результаты позволяют говорить о достаточно высокой эффективности реализованного в Solver метода Interior Point.

Задача	Число переменных	Число ограничений	OPL Studio CPLEX	Solver Interior Point Method
Trans256	256	16	0.04	0.04
Gas	300	40	0.28	0.22
Trans4096	4096	64	X	3.62
Bonds	2500	20	X	6.95

X — решения не найдено.

Кроме того, в результате работы Solver мы имеем ГВИО ОЗЦФ, гарантию допустимости полученного решения и гарантию того, что оно отличается от оптимального не более чем на ширину ГВИО ОЗЦФ.

Задача	Гарантированная интервальная оценка оптимального значения
Trans256	[410.64898,410.64915]
Gas	[-4634621.786,-4634621.751]
Trans4096	[32367.361, 32367.367]
Bonds	[15365.3033,15365.3038]

Задачи Trans256, Trans4096 — это задачи (3.2) с данными разных объемов, задача Gas — это задача о планировании производства (3.1), задача Bonds — задача о минимальном инвестировании (3.3).

Список литературы

1. J. Gondzio, T.Terlaky. “A Computational View of Interior-Point Methods for Linear Programming”, 1994.
2. E.D.Andersen, J.Gondzio, C.Meszaros, Xiaojie Xu. “Implementation of Interior Point Method for Large Scale Linear Programming” // Interior Point Methods of Mathematical Programming”, Chapter 6. — Dordrecht: Kluwer Academic Publisher, 1996.
3. Хачиян Л. Г. “Сложность задач линейного программирования”. — М.: Знание, 1987.
4. Шарый С. П. “Интервальные алгебраические задачи и их численное решение”. Диссертация на соискание ученой степени доктора физико-математических наук. —Новосибирск: ИВТ СО РАН, 2000.
5. Кашеварова Т.П. “Применение метода недоопределенных вычислений в математическом моделировании”. Диссертация на соискание ученой степени кандидата физико-математических наук. — Новосибирск: ИСИ СО РАН, 1999.
6. Шокин Ю.И. Об интервальных задачах, интервальных алгоритмах и их трудоемкости // *Вычислительные технологии.* — 1996. — Т. 1, № 1.

Символьно-интервальная эвристика для минимизации при краевых ограничениях

Е.С. Петров*

Ecole des Mines de Nantes
Chantrerie, 4 Alfred Kastler, BP 20722
44307 Nantes Cedex 3, France
evgueni.petrov@emn.fr

Аннотация Глобальная оптимизация при краевых ограничениях дает ответ на многие практические вопросы в химии, молекулярной биологии, экономике. Большинство алгоритмов для решения задач глобальной оптимизации является сочетанием интервальных методов и полного перебора. Эффективность таких алгоритмов характеризуется их способностью обнаруживать и исключать области субоптимальных допустимых решений. Эта их способность повышается наличием хорошей оценки сверху для глобального минимума. В этой статье мы представляем символьно-интервальный алгоритм для вычисления оценок сверху в задачах глобальной минимизации при краевых ограничениях и сообщаем результаты нескольких экспериментов с этим алгоритмом.

1 Введение

Глобальная оптимизация при краевых ограничениях дает ответ на многие практические вопросы в химии, молекулярной биологии, экономике. Большинство алгоритмов для решения задач глобальной оптимизации является сочетанием интервальных методов и полного перебора [3,5,9]. Эффективность таких алгоритмов характеризуется их способностью обнаруживать и исключать области субоптимальных допустимых решений. Эта их способность повышается наличием хорошей оценки сверху для глобального минимума.

Нетривиальная оценка сверху для минимума любой рациональной функции может быть получена при помощи модальных интервалов или арифметики Каухера [8,1]. Что же касается нелинейных функций вообще, то в настоящее время подобных средств не имеется и приходится пользоваться сочетаниями *ad hoc* локального поиска и выборки значений. Достоинство такого подхода в том, что он применим к любой функции, заданной “черным ящиком”, преобразующим аргументы в значения функции. Его наиболее существенный недостаток в том, что этот подход игнорирует символьное представление минимизируемой функции, которое несет полезную информацию о её глобальном поведении.

В этой статье мы представляем символьно-интервальный алгоритм для вычисления оценок сверху в задачах глобальной минимизации при краевых ограничениях и сообщаем результаты нескольких экспериментов с этим алгоритмом.

Статья организована следующим образом. В разделе 2 вводятся основные определения. В разделе 3 представлен сам алгоритм. Раздел 4 содержит данные экспериментов с функциями Dixon-Szegö и другими классическими функциями. Раздел 5 подводит итоги статьи. Приложение A содержит тестовые функции.

2 Основные определения и обозначения

Этот раздел содержит определения, связывающие вещественные и интервальные функции, их символьное представление и частные производные. Интервалами называются замкнутые выпуклые множества вещественных чисел, *брусами* называются векторы интервалов. Множество вещественных чисел и множество интервалов обозначаются $\mathbb{R}$, $\mathbb{I}$. Интервальная функция $[f]$, которая возвращает точный интервал значений вещественной функции f исходя из интервалов значений ее аргументов, называется *интервальным расширением* f. Мы называем сложение, вычитание, умножение, деление, степенную функцию для целых

* При финансовой поддержке Франко-русского центра им. Ляпунова (проект 06–98) и проекта ESPRIT COCONUT IST–2000–26063.

показателей, синус, косинус, тангенс, экспоненту и обратные к ним функции *базовыми функциями.* Мы пишем "вектор", "функция" вместо "вещественный вектор", "вещественная функция".

Символы $\mathcal{B}$, обозначающие базовые функции, называются *функциональными символами.* Символы арифметических операций являются *бинарными*, остальные символы являются *унарными.* Символы, обозначающие вещественные числа, называются *символами констант.* Символы v_i-е, отличные от констант и функциональных символов, называются *переменными.*

Множество *термов* над переменными v_i-ми, построенных из символов $\mathcal{B}$ и символов констант, определяется обычным образом. Терм t' является *подтермом* терма t, $t' \sqsubseteq t$, если t' входит в t. Терм, построенный из символа $\alpha \in \mathcal{B}$ и термов t', t'', записывается в виде $\alpha(t', t'')$ (для бинарного α) или в виде $\alpha(t')$ (для унарного α). Множество переменных в терме t обозначается $\textsc{var}(t)$.

Рассмотрим следующие алгебры, имеющие одни и те же функциональные символы $\alpha \in \mathcal{B}$: термы $\mathcal{T}_n$ над переменными $v_1, \ldots, v_n$, где операция (обозначенная) α возвращает терм, построенный из аргументов этой операции и символа α; функции $\mathcal{R}_n$ n-мерных векторов, где α возвращает композицию аргументов и базовой функции, обозначенной α; интервальные функции $\mathcal{I}_n$ n-мерных брусов, где α возвращает композицию аргументов и интервального расширения базовой функции, обозначенной α.

Алгебра $\mathcal{T}_n$ называется алгеброй *вещественных термов.* Подалгебра алгебры $\mathcal{R}_n$, порожденная проекторами из $\mathbb{R}^n$ в $\mathbb{R}$ и константными функциями, называется алгеброй *элементарных функций.* Подалгебра алгебры $\mathcal{I}_n$, порожденная проекторами из $\mathbb{I}^n$ в $\mathbb{I}$ и функциями, возвращающими точечные интервалы, называется алгеброй *интервальных элементарных функций.*

Элементарные и интервальные элементарные функции являются образами вещественных термов под действием гомоморфизмов $(\cdot)^{\mathcal{R}_n}$ и $[\cdot]_n$, отображающих переменные и символы констант в порождающие элементы соответствующей алгебры. Мы говорим, что терм t *задает* элементарную функцию $t^{\mathcal{R}_n}$ и интервальную элементарную функцию $[t]_n$. Ниже n является некоторым фиксированным неотрицательным целым; мы пишем $(\cdot)^{\mathcal{R}}$, $[\cdot]$ вместо $(\cdot)^{\mathcal{R}_n}$, $[\cdot]_n$.

Терм α' задает производную от функции, заданной $\alpha(v_1)$ для унарных символов $\alpha \in \mathcal{B}$. Термы α'_1, α'_2 задают частные производные функции, заданной $\alpha(v_1, v_2)$ для бинарных символов $\alpha \in \mathcal{B}$. Терм, полученный из некоторого терма t одновременной подстановкой термов t_i-ых вместо переменных v_i-ых, обозначается через $t(t_1, \ldots, t_n)$. Когда $\textsc{var}(t)$ совпадает с $\{v_1, v_2\}$ или $\{v_1\}$ мы используем сокращения $t(t_1, t_2)$, $t(t_1)$. Функция $\partial t/\partial(\cdot)$, рекуррентно определенная для подтермов терма t следующим образом:

$$
\begin{aligned}
\partial t/\partial t &= 1 \\
\partial t/\partial t' &= \alpha'(t') \cdot \partial t/\partial\alpha(t') && \text{для унарных } \alpha \in \mathcal{B} \text{ и } \alpha(t') \sqsubseteq t \\
\partial t/\partial t' &= \alpha'_1(t', t'') \cdot \partial t/\partial\alpha(t', t'') && \text{для бинарных } \alpha \in \mathcal{B} \text{ и } \alpha(t', t'') \sqsubseteq t \\
\partial t/\partial t' &= \alpha'_2(t'', t') \cdot \partial t/\partial\alpha(t'', t') && \text{для бинарных } \alpha \in \mathcal{B} \text{ и } \alpha(t'', t') \sqsubseteq t
\end{aligned}
$$

называется *дифференцированием* t. Терм $\partial t/\partial t'$ называется *частной производной* t по $t' \sqsubseteq t$. Мы считаем, что символы $\cdot$ и 1 обозначают умножение и единицу. Предположим, что t содержит единственное вхождение t'. Пусть $t = t''(v_1, \ldots, v_n, t')$ и $\textsc{var}(t'') = \{v_1, \ldots, v_{n+1}\}$. Функция $(\partial t/\partial t')^{\mathcal{R}} : \mathbb{R}^n \to \mathbb{R}$ является композицией частной производной $D_{n+1}\left[(t'')^{\mathcal{R}_{n+1}}\right]$ функции $(t'')^{\mathcal{R}_{n+1}} : \mathbb{R}^{n+1} \to \mathbb{R}$ по последнему аргументу и вложения $\mathbb{R}^n$ в $\mathbb{R}^{n+1}$, переводящего вектор p в вектор $\left(p, (t')^{\mathcal{R}}(p)\right)$.

3 Символьно-интервальная эвристика

Глобальный минимум любой функции можно оценить сверху при помощи выборки значений и локальной минимизации. Применение локальной минимизации к функции, не являющейся *a priori* выпуклой, имеет два недостатка: риск медленной сходимости и риск нарушения краевых ограничений. Оба риска снижаются по мере приближения точки для начала локальной минимизации к глобальному минимуму. Целью для нашего алгоритма является нахождение такой точки, где выполняются краевые ограничения и значение целевой функции достаточно мало.

Эскиз алгоритма Пусть вещественный терм t и брус b задают целевую функцию и краевые ограничения. Наш алгоритм многократно использует следующее наблюдение, касающееся минимума целевой функции $t^{\mathcal{R}}$ в брусе b. Пусть для $t^{\mathcal{R}}$ и некоторых элементарных функций f и g верно $\forall p \in \mathbb{R}^n$ $t^{\mathcal{R}}(p) = f(p, g(p))$ и пусть a является значением g в некоторой точке из b. Тогда минимум целевой функции ограничен сверху $\min_{p \in b \cap g^{-1}(a)} f(p, a)$, где $g^{-1}(a)$ является прообразом a под действием g.

Выбор конкретного a определяется соображениями эффективности. Идеальным (и непрактичным) выбором было бы значение g в точке глобального минимума целевой функции. На практике мы требуем, чтобы f была монотонна по последнему аргументу и чтобы мини- и максимизация g была доступна по вычислительным затратам. При этих предположениях мы используем $a = \min_b g$ (если f возрастает) или $a = \max_b g$ (если f убывает). Эти предположения справедливы, если g задается подтермом $t' \sqsubseteq t$ таким, что (1) ни одна переменная не входит в t' дважды и (2) функция, заданная частной производной $\partial t/\partial t'$, сохраняет знак в b. Максимальные относительно $\sqsubseteq$ подтермы t, обладающие этими двумя свойствами, называются *хорошими подтермами*. Наш алгоритм показан в Таблице 1 и Таблице 2.

Таблица 1. Эвристика для минимизации при краевых ограничениях; вектор p является точкой для начала локальной оптимизации, множество ℓ является множеством хороших подтермов t, терм $t' \sqsubseteq t$ является одним из хороших подтермов, функция $\mathrm{PROJ}(\cdot,\cdot)$ определена в Таблице 2, функция $\mathrm{sign}(\cdot)$ получает терм, задающий знакопостоянную функцию в b, и возвращает ее знак (+1 или −1), функции $\mathrm{mid}(\cdot)$ и $\mathrm{rad}(\cdot)$ возвращают центр и радиус интервалов, интервал b_i является i-м элементом бруса b

```
p = (0, ..., 0)
DO WHILE ( VAR(t) ≠ ∅ )
    ℓ := множество хороших подтермов t
    IF  ℓ = ∅
        BREAK
    ELSE IF  VAR(t') = VAR(t) для всех t' ∈ ℓ  THEN
        t' := такой подтерм из ℓ, что t^ℛ(PROJ(t', mid[t'](b) − sign(∂t/∂t') · rad[t'](b)))
        наименьшее
    ELSE IF  мощность VAR(t') равна 1 для одного из t' ∈ ℓ  THEN
        t' := такой подтерм
    ELSE
    t' := произвольный подтерм из ℓ
    END IF
    p = p + PROJ(t', mid[t'](b) − sign(∂t/∂t') · rad[t'](b))
    заменить в t каждую v_i ∈ VAR(t') на p_i
    p_i := mid(b_i) для всех v_i ∈ VAR(t)
END DO
```

Пример Протрассируем алгоритм для функции Бранина, заданной термом

$$t = \left(v_2 - \frac{5.1}{4\pi^2}v_1^2 + \frac{5v_1}{\pi} - 6\right)^2 + 10\left(1 - \frac{1}{8\pi}\right)\cos v_1 + 10$$

в брусе $b = ([-5, 10], [0, 15])$ (мы записываем t в инфиксной форме). Мы имеем только один хороший подтерм $t' = 10\left(1 - \frac{1}{8\pi}\right)\cos v_1 \sqsubseteq t$ (можно убедиться, что $\partial t/\partial t' = 1 \cdot 1 \cdot 1$). Минимизация $(t')^{\mathcal{R}}$ в b даст v_1 значение $-\pi$ (или π, или 3π — в зависимости от фактической реализации $\mathrm{PROJ}(\cdot,\cdot)$). Поскольку $\mathrm{VAR}(t) = \{v_2\} \neq \emptyset$, тело цикла выполняется еще раз. Теперь весь терм t является своим собственным хорошим подтермом (ведь v_1 заменена на $-\pi$!). Минимизация $t^{\mathcal{R}}$ даст v_2 значение 12.275 и цикл завершится, поскольку $\mathrm{VAR}(t)$ теперь равняется $\emptyset$. Точкой p для начала локальной оптимизации является $(-\pi, 12.275)$. В действительности это один из глобальных минимумов функции Бранина в b.

Корректность Алгоритм Таблицы 1 и Таблицы 2 является корректным, т.е. он генерирует вектор p, который удовлетворяет краевым ограничениям. Корректность алгоритма следует из того факта, что для термов t', содержащих не более одного вхождения каждой переменной, интервальная функция $[t']$ является интервальным расширением функции $(t')^{\mathcal{R}}$. В частности, это означает, что границы $[t'](b)$ являются глобальными экстремумами $(t')^{\mathcal{R}}$ в b. Следовательно, i-й элемент вектора $\mathrm{PROJ}(t', a)$, где a — любая из границ $[t'](b)$, принадлежит интервалу b_i как только $v_i \in \mathrm{VAR}(t')$.

Таблица 2. Определение функции PROJ$(\cdot,\cdot)$; символ v_i является переменной, c, c', c'' являются вещественными числами, функция g является базовой функцией, которую обозначает $\alpha \in \mathcal{B}$

$$\textsc{proj}(v_i, c) = (0, \dots, 0, c, 0, \dots, 0), \qquad \text{где } c \text{ занимает } i\text{-е место}$$

$$\textsc{proj}(\alpha(t'), c) = \textsc{proj}(t', c'), \qquad \text{где } c = g(c'),\ c' \in [t'](b)$$

$$\textsc{proj}(\alpha(t', t''), c) = \textsc{proj}(t', c') + \textsc{proj}(t'', c''), \qquad \text{где } c = g(c', c''),\ c' \in [t'](b),\ c'' \in [t''](b)$$

Сложность в худшем случае В худшем случае число вычислений базовых функций и их интервальных расширений составляет $O(s^2/n)$, число логических операций составляет $O(n \cdot s)$, где s является размером терма, задающего целевую функцию, n является числом переменных.

Детали компьютерной реализации Компьютерная реализация алгоритма Таблицы 1 и Таблицы 2 должна корректно оперировать с вещественными числами, непредставимыми в формате с плавающей точкой. Вместо таких чисел мы используем тонкие интервалы, ограниченные числами с плавающей точкой. Таким образом, компьютерной реализации нашего алгоритма требуется 4 числа с плавающей точкой на интервал. Интервалы $[\partial t/\partial t'](b)$ и $[t'](b)$ вычисляются для всех подтермов $t' \sqsubseteq t$ интервальной версией автоматического дифференцирования [4].

4 Эксперименты со стандартными функциями

В этом разделе мы кратко описываем эксперименты с реализацией на C++ алгоритма, описанного в разделе 3. Функции для этих экспериментов были взяты из [2], [7] и 1st International Contest on Evolutionary Optimization (см. приложение A). Для каждой тестовой функции мы сообщаем число n переменных в соответствующем терме, глобальный минимум и оценку сверху, вычисленную по нашей эвристике (первые 4 цифры из их десятичной записи, см. Таблицу 1). Общее наблюдение, касающееся нашего алгоритма, заключается в том, что он имеет тенденцию давать лучшие результаты для сепарабельных целевых функций.

функции Dixon-Szegö			
функция	n	min	оценка
Shekel 5	4	–10.38	–10.38
Shekel 7	4	–10.48	–10.48
Shekel 10	4	–10.53	–10.53
Goldstein-Price	2	3.000	600.0
6 hump camel	2	–1.031	150.9
Hartmann 3	3	–3.862	–3.761
Hartmann 6	6	–3.322	–3.203
Shubert	2	–186.7	19.87
Branin	2	0.3978	0.3978
функции Jansson-Knüppel			
Rosenbrock	2–80	0.000	0.000
Levy 8	2–80	0.000	0.000

функции Jansson-Knüppel (продолж.)			
функция	n	min	оценка
Levy 8 bis	2–80	0.000	0.000
Levy 13	2–80	0.000	0.000
функции Int. Contest on Evol. Optim.			
ICEO 2	5	0.000	0.000
ICEO 2	10	0.000	0.000
ICEO 3	5	–10.40	–10.40
ICEO 3	10	–10.20	–10.20
ICEO 4	5	–4.687	–4.488
ICEO 4	10	–9.660	–8.048
ICEO 5	5	–1.499	–1.499
ICEO 5	10	–1.500	–1.500

Рис. 1. Оценки сверху, найденные нашим алгоритмом для тестовых функций Dixon-Szegö, Jansson-Knüppel и 1st International Constest on Evolutionary Optimization; каждая строка содержит имя функции из приложения A, использованные размерности, глобальный минимум, вычисленную оценку сверху

5 Заключение

В этой статье мы представили символьно-интервальную эвристику для минимизации при краевых ограничениях. Ключевая идея нашей эвристики состоит в упрощении задачи глобальной минимизации путем игнорирования зависимостей между некоторыми подвыражениями в символьном представлении целевой функции. Наши эксперименты показывают, что этот подход позволяет получать хорошие оценки сверху в ряде задач минимизации при краевых ограничениях.

Наша работа в будущем будет сфокусирована на теоретических свойствах представленной эвристики, а также на интеграции ее в современные алгоритмы глобальной оптимизации, подобные [6].

Список литературы

1. Armengol Llobet J. Application of modal interval analysis to the simulation of the behaviour of dynamic systems with uncertain parameters. – Ph.D. thesis: University of Girona, 1999.
2. Dixon L. C. W., Szegö G. P., eds. *Towards Global Optimization.* – North-Holland, Amsterdam, 1975–1978.
3. Fiacco A. V., McCormick G. P. *Nonlinear Programming. Sequential unconstrained minimization techniques.* – New York: Willey, 1968.
4. Griewank A. On automatic differentiation // *Mathematical Programming: Recent Developments and Applications*, M. Iri and K. Tanabe, eds. – Dordrecht: Kluwer Academic Publishers, 1989. – P. 83 – 108.
5. Hansen E. *Global optimization using interval analysis*, – New York: Marcel Dekker, 1992.
6. Huyer W., Neumaier A. Global optimization by multilevel coordinate search // *Journal of Global Optimization.* – 1999. – Vol. 14. – P. 331 – 355.
7. Jansson Ch., Knьppel O. A global minimization method: the multi-dimensional case // Technische Informatik-III, TU Hamburg-Harburg. – 1992.
8. Kaucher E. Interval analysis in the extended interval space $\mathbb{IR}$ // *Computing Supplement 2.* – 1980. – P. 33 – 49.
9. Nocedal J., Wright S.J. *Numerical Optimization.* – Berlin: Springer, 1999.

A Тестовые функции

По соображениям полноты изложения мы приводим термы и брусы, которые задают тестовые целевые функции и соответствующие краевые ограничения. Мы обозначаем n-элементный вектор $(x,\dots,x)$ через x^n. Термы записываются в инфиксной форме. Формальные суммы и произведения должны быть раскрыты в соответствующие цепочки лево-ассоциативных сложений и умножений.

Shekel m

$b = [0,10]^4$, $t = -\sum_{i=1}^{m} 1 \Big/ \Big(c_i + \sum_{j=1}^{4}(v_j - a_{ij})^2\Big)$

$$c = \big(1\ 2\ 2\ 4\ 4\ 6\ 3\ 7\ 5\ 5\big)/10, \quad a^{\mathrm{T}} = \begin{pmatrix} 4\ 1\ 8\ 6\ 3\ 2\ 5\ 8\ 6\ 7 \\ 4\ 1\ 8\ 6\ 7\ 9\ 5\ 1\ 2\ 3.6 \\ 4\ 1\ 8\ 6\ 3\ 2\ 3\ 8\ 6\ 7 \\ 4\ 1\ 8\ 6\ 7\ 9\ 3\ 1\ 2\ 3.6 \end{pmatrix}$$

Goldstein–Price

$b = [-2,2]^2$,

$$\begin{aligned} t = &\big(1 + (v_1 + v_2 + 1)^2 \cdot \big(19 - 14 \cdot v_1 + 3 \cdot v_1^2 - 14 \cdot v_2 + 6 \cdot v_1 \cdot v_2 + 3 \cdot v_2^2\big)\big) \times \\ &\big(30 + (2 \cdot v_1 - 3 \cdot v_2)^2 \cdot \big(18 - 32 \cdot v_1 + 12 \cdot v_1^2 + 48 \cdot v_2 - 36 \cdot v_1 \cdot v_2 + 27 \cdot v_2^2\big)\big) \end{aligned}$$

6 hump camel

$$b = ([-3,3],[-2,2]), \quad t = \big(4 - 2.1 \cdot v_1^2 + v_1^4/3\big) \cdot v_1^2 + v_1 \cdot v_2 + 4 \cdot \big(v_2^2 - 1\big) \cdot v_2^2$$

Hartmann 3

$b = [0,1]^3$, $t = -\sum_{j=1}^{4} c_j \cdot \exp\left(-\sum_{i=1}^{3} a_{ij} \cdot (v_i - p_{ij})^2\right)$,

$$a = \begin{pmatrix} 3 & 0.1 & 3 & 0.1 \\ 10 & 10 & 10 & 10 \\ 30 & 35 & 30 & 35 \end{pmatrix}, \quad p = \begin{pmatrix} 0.36890 & 0.46990 & 0.10910 & 0.03815 \\ 0.11700 & 0.43870 & 0.87320 & 0.57430 \\ 0.26730 & 0.74700 & 0.55470 & 0.88280 \end{pmatrix},$$
$$c = \begin{pmatrix} 1 & 1.2 & 3 & 3.2 \end{pmatrix}$$

Hartmann 6
$b = [0,1]^6$, $t = -\sum_{j=1}^{4} c_j \cdot \exp\left(-\sum_{i=1}^{6} a_{ij} \cdot (v_i - p_{ij})^2\right)$,

$$a^{\mathrm{T}} = \begin{pmatrix} 10 & 3 & 17 & 3.5 & 1.7 & 8 \\ 0.05 & 10 & 17 & 0.1 & 8 & 14 \\ 3 & 3.5 & 1.7 & 10 & 17 & 8 \\ 17 & 8 & 0.05 & 10 & 0.1 & 14 \end{pmatrix}, p^{\mathrm{T}} = \begin{pmatrix} 0.1312 & 0.1696 & 0.5569 & 0.0124 & 0.8283 & 0.5886 \\ 0.2329 & 0.4135 & 0.8307 & 0.3736 & 0.1004 & 0.9991 \\ 0.2348 & 0.1451 & 0.3522 & 0.2883 & 0.3047 & 0.6650 \\ 0.4047 & 0.8828 & 0.8732 & 0.5743 & 0.1091 & 0.0381 \end{pmatrix},$$
$$c = \begin{pmatrix} 1 & 1.2 & 3 & 3.2 \end{pmatrix}$$

Shubert
$b = [-10,10]^2$, $t = \prod_{j=1}^{2} \sum_{i=1}^{6} i \cdot \cos((i+1) \cdot v_j + i)$

Rosenbrock
$b = [-2.048, 2.048]^n$, $t = \sum_{i=1}^{n-1} 100 \cdot \left(v_i - v_{i+1}^2\right)^2 + (1 - v_{i+1})^2$

Levy 8
$b = [-10,10]^n$,

$$t = 10 \cdot \sin^2 \pi \cdot \frac{v_1 + 3}{4} + \left(\frac{v_n - 1}{4}\right)^2 + \sum_{i=1}^{n-1} \left(\frac{v_i - 1}{4}\right)^2 \cdot \left(1 + 10 \cdot \sin^2 \pi \cdot \frac{v_{i+1} + 3}{4}\right)$$

Levy 8 bis

$$b = [-10,10]^n, \quad t = 10 \cdot \sin^2 \pi \cdot v_1 + (v_n - 1)^2 + \sum_{i=1}^{n-1} (v_i - 1)^2 \cdot \left(1 + 10 \cdot \sin^2 \pi \cdot v_{i+1}\right)$$

Levy 13
$b = [-10,10]^n$,

$$t = \sin^2 3 \cdot \pi \cdot v_1 + (v_n - 1)^2 \cdot \left(1 + \sin^2 3 \cdot \pi \cdot v_n\right) + \sum_{i=1}^{n-1} (v_i - 1)^2 \cdot \left(1 + \sin^2 3 \cdot \pi \cdot v_{i+1}\right)$$

ICEO 2
$b = [-600,600]^n$, $t = \sum_{i=1}^{n} (v_i - 100)^2 \big/ 4000 - \prod_{i=1}^{n} \cos\left((v_i - 100) \big/ \sqrt{i}\right) + 1$

ICEO 3
$b = [0,10]^n$, $t = -\sum_{j=1}^{30} 1 \big/ \left(c_j + \sum_{i=1}^{n} (v_i - a_{ji})^2\right)$ (см. матрицу a и вектор c ниже)

ICEO 4
$b = [0,\pi]^n$, $t = -\sum_{i=1}^{n} \sin v_i \cdot \sin^{20}\left(i \cdot v_i^2 \big/ \pi\right)$

ICEO 5
См. матрицу a и вектор c ниже.

$$b = [0,10]^n, \quad t = -\sum_{j=1}^{30} c_j \cdot \exp\left(-\sum_{i=1}^{n} (a_{ji} - v_i)^2 \Big/ \pi\right) \cdot \cos\left(\pi \cdot \sum_{i=1}^{n} (a_{ji} - v_i)^2\right)$$

Коэффициенты для ICEO 3 и ICEO 5

$$a = \begin{pmatrix}
9.681 & 0.667 & 4.783 & 9.095 & 3.517 & 9.325 & 6.544 & 0.211 & 5.122 & 2.020 \\
9.400 & 2.041 & 3.788 & 7.931 & 2.882 & 2.672 & 3.568 & 1.284 & 7.033 & 7.374 \\
8.025 & 9.152 & 5.114 & 7.621 & 4.564 & 4.711 & 2.996 & 6.126 & 0.734 & 4.982 \\
2.196 & 0.415 & 5.649 & 6.979 & 9.510 & 9.166 & 6.304 & 6.054 & 9.377 & 1.426 \\
8.074 & 8.777 & 3.467 & 1.863 & 6.708 & 6.349 & 4.534 & 0.276 & 7.633 & 1.567 \\
7.650 & 5.658 & 0.720 & 2.764 & 3.278 & 5.283 & 7.474 & 6.274 & 1.409 & 8.208 \\
1.256 & 3.605 & 8.623 & 6.905 & 0.584 & 8.133 & 6.071 & 6.888 & 4.187 & 5.448 \\
8.314 & 2.261 & 4.224 & 1.781 & 4.124 & 0.932 & 8.129 & 8.658 & 1.208 & 5.762 \\
0.226 & 8.858 & 1.420 & 0.945 & 1.622 & 4.698 & 6.228 & 9.096 & 0.972 & 7.637 \\
7.305 & 2.228 & 1.242 & 5.928 & 9.133 & 1.826 & 4.060 & 5.204 & 8.713 & 8.247 \\
0.652 & 7.027 & 0.508 & 4.876 & 8.807 & 4.632 & 5.808 & 6.937 & 3.291 & 7.016 \\
2.699 & 3.516 & 5.874 & 4.119 & 4.461 & 7.496 & 8.817 & 0.690 & 6.593 & 9.789 \\
8.327 & 3.897 & 2.017 & 9.570 & 9.825 & 1.150 & 1.395 & 3.885 & 6.354 & 0.109 \\
2.132 & 7.006 & 7.136 & 2.641 & 1.882 & 5.943 & 7.273 & 7.691 & 2.880 & 0.564 \\
4.707 & 5.579 & 4.080 & 0.581 & 9.698 & 8.542 & 8.077 & 8.515 & 9.231 & 4.670 \\
8.304 & 7.559 & 8.567 & 0.322 & 7.128 & 8.392 & 1.472 & 8.524 & 2.277 & 7.826 \\
8.632 & 4.409 & 4.832 & 5.768 & 7.050 & 6.715 & 1.711 & 4.323 & 4.405 & 4.591 \\
4.887 & 9.112 & 0.170 & 8.967 & 9.693 & 9.867 & 7.508 & 7.770 & 8.382 & 6.740 \\
2.440 & 6.686 & 4.299 & 1.007 & 7.008 & 1.427 & 9.398 & 8.480 & 9.950 & 1.675 \\
6.306 & 8.583 & 6.084 & 1.138 & 4.350 & 3.134 & 7.853 & 6.061 & 7.457 & 2.258 \\
0.652 & 2.343 & 1.370 & 0.821 & 1.310 & 1.063 & 0.689 & 8.819 & 8.833 & 9.070 \\
5.558 & 1.272 & 5.756 & 9.857 & 2.279 & 2.764 & 1.284 & 1.677 & 1.244 & 1.234 \\
3.352 & 7.549 & 9.817 & 9.437 & 8.687 & 4.167 & 2.570 & 6.540 & 0.228 & 0.027 \\
8.798 & 0.880 & 2.370 & 0.168 & 1.701 & 3.680 & 1.231 & 2.390 & 2.499 & 0.064 \\
1.460 & 8.057 & 1.336 & 7.217 & 7.914 & 3.615 & 9.981 & 9.198 & 5.292 & 1.224 \\
0.432 & 8.645 & 8.774 & 0.249 & 8.081 & 7.461 & 4.416 & 0.652 & 4.002 & 4.644 \\
0.679 & 2.800 & 5.523 & 3.049 & 2.968 & 7.225 & 6.730 & 4.199 & 9.614 & 9.229 \\
4.263 & 1.074 & 7.286 & 5.599 & 8.291 & 5.200 & 9.214 & 8.272 & 4.398 & 4.506 \\
9.496 & 4.830 & 3.150 & 8.270 & 5.079 & 1.231 & 5.731 & 9.494 & 1.883 & 9.732 \\
4.138 & 2.562 & 2.532 & 9.661 & 5.611 & 5.500 & 6.886 & 2.341 & 9.699 & 6.500
\end{pmatrix}
\quad
c = \begin{pmatrix}
0.806 \\ 0.517 \\ 0.1 \\ 0.908 \\ 0.965 \\ 0.669 \\ 0.524 \\ 0.902 \\ 0.531 \\ 0.876 \\ 0.462 \\ 0.491 \\ 0.463 \\ 0.714 \\ 0.352 \\ 0.869 \\ 0.813 \\ 0.811 \\ 0.828 \\ 0.964 \\ 0.789 \\ 0.360 \\ 0.369 \\ 0.992 \\ 0.332 \\ 0.817 \\ 0.632 \\ 0.883 \\ 0.608 \\ 0.326
\end{pmatrix}$$

Линейный критерий для задачи оптимизации на графах с интервальными параметрами

Г.Л. Козина, Р. Кудерметов

Запорожский национальный технический университет
Украина, г. Запорожье
ains@comint.net

Мы предлагаем новый критерий для сравнения решений в дискретных задачах оптимизации. В этой работе рассматривается задача о минимальном остовном дереве ([1], [2]) на графах с весами ребер, изменяющимися в заданных интервалах.

Пусть дан граф $G = (V, E)$ с множеством вершин V и множеством ребер E. Каждому ребру $e_i \in E$ сопоставлен вес c_i, который может меняться в интервале $C_i = \left[\overline{C}_i, \underline{C}_i\right]$, $c_i \in C_i$. Пусть $x = (V, E_x)$, $E_x \subset E$, остовное дерево графа G. Вещественным весом остовного дерева x является сумма весов c_i ребер $e_i \in E_x$: $f(x,c) = \sum c_i$, где $c = (c_1, c_2, ..., c_N)$.

Задача о минимальном остовном дереве состоит в нахождении остовного дерева, имеющего минимальный вес.

Пусть вектор $c \in C$, где $C = (C_1, C_2, ..., C_N)$ и C_i, ($i = \overline{1,N}$), являются интервалами. Интервальным весом остовного дерева x является сумма интервальных весов C_i ребер $e_i \in E_x$: $f(x,C) = \sum C_i = [f_1(x),\ f_2(x)]$.

В случае интервальных весов становится трудно сравнивать различные деревья в соответствии с их весами.

Допустимое остовное дерево называется сильным, если оно минимально для всех реализаций параметров, изменяющихся в заданных интервалах ($\forall c \in C$). Допустимое остовное дерево называется слабым, если оно минимально для некоторой реализации параметров, изменяющихся в заданных интервалах ($\exists c \in C$).

Все слабые решения образуют множество слабых деревьев W. Мы принимаем это множество в качестве решения интервальной задачи об остовных деревьях [2].

Для принятия решения мы должны найти оценку качества слабого решения. Мы вводим понятие вероятности оптимальности для слабых решений. Такой вещественный критерий позволяет нам сравнивать решения и соответственно оценивать их качество.

Для каждого слабого решения имеется реализация, при которой это решение является оптимальным. Но интуитивно ясно, что для одного решения таких реализации будет «много», а для других «недостаточно», то есть одни решения более вероятны, чем другие.

Введем понятие вероятности оптимальности слабого решения. Интервальный вектор C можно рассматривать как набор всех реализации параметров. Очевидно, что $C = C_1 \times C_2 \times ... \times C_N \subseteq R^N$.

Обозначим через Q_k множество реализаций, при которых слабое решение x_k является оптимальным, $k = \overline{1,l}$. Очевидно, что $\bigcup\limits_{k=1}^{l} Q_k = C$.

Обозначим через $\widetilde{Q}_k$ множество внутренних точек множества Q_k. Справедлива следующая теорема.

Теорема. Если все веса ребер являются строгими интервалами, то есть $\overline{C_i} - \underline{C}_i > 0$, то множества Q_m и Q_k могут пересекаться только по границе для любых двух $m \neq k$, то есть $\widetilde{Q}_m \cap \widetilde{Q}_k = \varnothing$.

Введем меру $\mu(\bullet)$ на множестве подмножеств множества Q. Тогда вероятность оптимальности слабого решения x_k есть $P(x_k) = \dfrac{\mu(Q_k)}{\mu(C)}$.

Таким образом, каждому слабому решению соответствует некоторое число, который может отражать качество этого решения. Чем больше вероятность оптимальности данного решения, тем оно более предпочтительно.

Как находить вероятность оптимальности слабых решений?

Авторами написана программа IntGraph, которая позволяет получать вероятности оптимальности решения с помощью имитационного моделирования параметров задачи. Программа IntGraph реализована на языке С [5] и имеет графический интерфейс для представления результатов. При моделировании процесса программа генерирует веса ребер графа как случайные равномерно распределенные числа и находит оптимальные решения, которые по определению являются слабыми. Для каждого слабого решения подсчитывается число реализаций, при которых данное решение и находиттся оценка вероятности оптимальности слабого решения.

Такой подход может применяться к другим проблемам оптимизации на графах.

Список литературы

1. Kozina G.L., Perepelitsa V.A. Spanning Trees Problem: Solvability and Computational Complexity // *Interval Computations.* – 1994. – №1. – P. 42–50.
2. Yaman H., Karasan O., Mustafa O. Minimum Spanning The Problem with Interval Data. – *Techical Report 9909*, Bilkent University, Department of Industrial Engineering, July, 1999.

О выборе портфеля инвестиционных проектов

П.А. Кунташев

Белгородский филиал Московского государственного
университетастатистики и информатики (МЭСИ)
Россия, г. Белгород, ул. Коммунистическая, 64
pavelbelg@intbel.ru

Инвестиционный риск связан с тем, что в силу действия социально-экономических факторов, имеющих случайный и неопределенный характер, эффект от проекта может оказаться меньше ожидаемого или даже убыточным. Высокая степень случайности и неопределенности социально-экономических явлений связана с динамичностью обстановки, массовым характером производства, активной реакцией экономических субъектов, психологическими эффектами в закономерностях потребления , установлении рыночных цен с учетом ожиданий будущей ситуации и так далее.

Это приводит к тому, что в большинстве задач прогноза экономических явлений текущие и будущие параметры имеют принципиальную неопределенность. Для их описания наиболее адекватным математическим аппаратом является интервальный анализ. Риск проекта связан с тем, что рыночные цены на планируемый к выпуску товар и цены необходимых производственных ресурсов могут меняться в неблагоприятном направлении на периоде проекта, может изменяться темп инфляции и т.д. Поэтому возникает следующий подход к оценке инвестиционного риска. На основе интервальных экспертных оценок для исходных данных задачи требуется оценить интервал для эффекта, который позволит дать количественную оценку инвестиционного риска.

В данной работе используется классическая интервальная арифметика [1]. Для интервалов $A=[\underline{a};\ \overline{a}]$, $B=[\underline{b};\ \overline{b}]$ вводятся арифметические операции сложения, вычитания, умножения и деления.

(1) $A+B=[\underline{a}+\underline{b};\ \overline{a}+\overline{b}]$; $A-B=[\underline{a}-\overline{b};\ \overline{a}-\underline{b}]$; $A/B=[\underline{a};\ \overline{a}]*[1/\overline{b};\ 1/\underline{b}]$;
(2) $A*B=[\min\{\underline{a}\,\underline{b};\underline{a}\,\overline{b};\ \overline{a}\underline{b};\ \overline{a}\,\overline{b}\};\ \max\{\underline{a}\,\underline{b};\underline{a}\,\overline{b};\ \overline{a}\underline{b};\ \overline{a}\,\overline{b}\}]$.

Эффект от проекта, рассчитанного на n лет, измеряется чистой приведенной стоимостью, которая для простейшего случая инвестиций, сводящихся к первоначальным капитальным вложениям S, имеет вид [2]:

(3) $$NPV=\sum_{k=1}^{n}(CF_k)[(1+i_0)^k(1+i_1)...(1+i_k)]^{-1}\text{-}S.$$

Здесь i_k - прогнозные уровни инфляции для соответствующего года проекта, i_0-реальная ставка доходности. Денежный поток от производственной деятельности равен чистой прибыли плюс амортизация $C_{амор}$:

(4) $$CF_k=(1-i_{нал})[Q_k(P_k-\sum_i P_{ik}d_i)-C_k]+C_{амор},$$

где $i_{нал}$-ставка налога на прибыль, d_i-количество i-го ресурса ,необходимого для производства единицы товара.

Будем считать, что на основе экспертных оценок инфляции и маркетинговых исследований установлены интервалы для изменения следующих параметров в каждом k-м году проекта: Q_k — объем производства товара, P_k — цена товара, P_{ik} — цена i-го производственного ресурса, входящего в состав переменных затрат, C_k — постоянные затраты, прогнозные уровни инфляции и др

Интервал $[-A_j; B_j]$ возможного изменения для эффекта от проекта с номером j может быть получен путем подстановки в NPV (3),(4) вместо значений параметров проекта соответствующих им интервалов и выполнения на компьютере соответствующих интервальных арифметических операций по (1),(2). Это означает, что при неблагоприятном развитии экономической ситуации инвестиции в данный проект могут принести убытки A_j. В качестве измерителя риска проекта предлагается использовать эту величину максимальной суммы возможных убытков для данного проекта:

(5) $\varphi_j = A_j$.

Далее рассматривается задача о выборе структуры x_j портфеля делимых проектов[3]. Понятие делимого проекта подразумевает, что вместо исходного проекта может быть реализована его некоторая доля x_j. При этом все производственные мощности, объемы выпуска продукции и затраты пропорционально уменьшаются (соответствующим образом переписывается бизнес-план и технико-экономическое обоснование проекта). В силу очевидных линейных свойств чистой приведенной стоимости, эффект от доли проекта равен доле эффекта исходного проекта. Под x_j понимается доля исходного j-го проекта, которая принимается к финансированию. Требуется максимизировать критерий чистой приведенной стоимости портфеля $NPV_{порт}$, при ограничениях на заданный капитал портфеля $V_{порт}$ и заданных требуемых инвестициях V_j исходных проектов:

(6) $NPV_{порт} = \sum_j NPV_j * x_j \to \max; \sum_j V_j * x_j \le V_{порт}; 0 \le x_j \le 1.$

В данной работе эта формулировка задачи [3] расширена на случай интервальной неопределенности исходных данных, а так же учета риска портфеля проектов.

Вводятся критерии: f_k -среднего интервального эффекта проекта и f -среднего интервального эффекта портфеля проектов:

(7) $f_j = (B_j - A_j)/2; f = (B - A)/2; \quad A = \sum_j A_j * x_j; B = \sum_j B_j * x_j.$

Наряду с риском φ_j проекта (5), вводится риск портфеля делимых проектов, равный сумме максимальных убытков по портфелю проектов при неблагоприятном развитии экономической ситуации:

(8) $\varphi = A.$

Требуется выбрать структуру x_j портфеля делимых проектов так, чтобы максимизировать f — среднеинтервальный эффект портфеля при риске портфеля φ, не превышающем заданный инвестором уровень φ_0, при ограничениях на заданный капитал портфеля $V_{порт}$:

(9) $f = 1/2 * \sum_j (B_j - A_j) * x_j \to \max; \quad F = \varphi_0 - \sum_j A_j x_j \ge 0;$

$\Psi = V_{порт} - \sum_j V_j * x_j \ge 0; x_j \ge 0; (1 - x_j) \ge 0;$

Предлагаемая экономико-математическая модель (9) для задачи о выборе структуры портфеля делимых проектов с учетом риска портфеля относится к хорошо исследованному классу задач линейного программирования. С учетом взаимосвязи между чистой приведенной стоимостью и индексом доходности проекта [3], находим интервалы $[C_j; D_j]$ возможного изменения для индекса доходности проекта:

(10) $PI_j = (NPV_j)/(V_j) + 1; \quad C_j = 1 - A_j / V_j \ge 0; \quad D_j = 1 + B_j / V_j \ge 0;$

Переходя от искомых переменных долей проектов x_j к переменным капиталов $t_j = x_j * V_j$, вложенных в проекты, получаем экономико-математическую модель для задачи о выборе структуры портфеля делимых проектов с учетом риска портфеля в виде:

(11) $f + V_{порт} = 1/2 * \sum_j (D_j + C_j) * t_j \to \max; \quad F = \varphi_0 - \sum_j A_j t_j / V_j \ge 0;$

$\Psi = V_{порт} - \sum_j t_j \ge 0; t_j \ge 0; (V_j - t_j) \ge 0;$

Вводя соответствующие множители Лагранжа, получаем функцию Лагранжа:

(12) $L = (f + V_{порт}) + \lambda_{риск} * F + \lambda_{кап} * \Psi + \sum_j \lambda_j * t_j + \sum_j \alpha_j * (V_j - t_j);$

Условия оптимальности Куна-Такера [4], которые являются необходимыми, а в силу выпуклости задачи (11) и достаточными условиями оптимальности портфеля, сводятся к следующему. Если ограничения на риск портфеля и капитал портфеля являются критическими, то должны существовать положительные множители

Лагранжа $\lambda_{риск}$, $\lambda_{кап}$, такие, что:

1) для проектов $(x_j = 1)$, включенных в портфель в полном объеме, $h_j > \lambda_{кап}$;

2) для проектов $(x_j = 0)$,вообще не включенных в портфель, $h_j < \lambda_{кап}$;

3) для проекта $(0 < x_j < 1)$, включенного в портфель в частичном объеме (если такой найдется), $h_j = \lambda_{кап}$;

где введена характеристическая функция, зависящая от индекса доходности и риска проекта:

(13) $h_j = (D_j + C_j)/2 - \lambda_{риск} * A_j / V_j$.

Фигурирующие здесь величины имеют следующий экономический смысл:
$(D_j + C_j)/2$ — это среднеинтервальный индекс доходности проекта, который показывает во сколько раз в среднем возрастает вложенный в проект капитал;
(A_j / V_j) — это доля инвестиций проекта, находящаяся под риском потери;
$\lambda_{риск}$ — множитель Лагранжа, зависящий от риска портфеля.

Список литературы

1. Калмыков С.А., Шокин Ю.И., Юлдашев З.Х. *Методы интервального анализа.* – Новосибирск: Наука,1986.
2. Виленский П.Л., Лившиц В.Н., Смоляк С.А. *Оценка эффективности инвестиционных проектов.* – М.:Дело,2001.
3. Ковалев В.В. *Финансовый анализ.* – М.: ФиС, 1996.
4. Мину М. Математическое программирование. – М.:Наука, 1990.

К проблеме узкой предварительной локализации отрезка интегральной кривой

Меньшиков Г.Г.

Санкт-Петербургский государственный университет
Факультет прикладной математики – процессов управления
miksha@pobox.spbu.ru

Аннотация Предмет доклада назван в заголовке применительно к локализующему поиску решения задачи Коши для ОДУ $y' = f(x,\ y), \quad y(x_0) = y_0$. Компьютерное решение этой задачи предусматривает нахождение локализатора $Y(x)$ искомого решения $y(x)$ – полосы, образованной интервальными значениями и содержащей отрезок интегральной кривой. Имеем в виду тот распространённый образ действий (как для традиционного, так и для интервального исполнения), когда промежуток изменения x узловыми точками x_k разбивается на шаги интегрирования $X_k = [x_k,\ x_{k+1}]$, причём по-очереди вычисляются образы соответствующих отрезков интегральной кривой. Тогда процесс численного интегрирования состоит из последовательности однотипных шагов. Для простоты предположим, что значения x_k формируются точно, что $x_k < x_{k+1}$, а интегрируется одно уравнение первого порядка.
В свою очередь алгоритмику шага можно разделить на *предварительный*, *уточняющий* и, возможно, *контролирующий* этапы. Особую важность представляет предварительный этап, поскольку в его задачи входит выяснение продолжимости решения $y(x)$ на отрезок X_k. Кроме того, важно, что предварительный локализатор (ПЛ) вида прямоугольника в плоскости $(x,\ y)$ с основанием X_k) формируется числами $\underline{Y^*}$ и $\overline{Y^*}$, т.е. минорантой и мажорантой решения на этом отрезке. Он используется на уточняющей стадии (например, на основе формулы Тэйлора) и поэтому влияет на достигаемое уточнение.
В докладе рассматриваются способы получения узкого ПЛ.

1 Введение

На первый взгляд, лучше всего, если ПЛ оказывается равным

$$y(X_k) = \left[\min_{X_k} y(x), \quad \max_{X_k} y(x)\right].$$

Но точное равенство, как правило, невозможно. Всё-таки, оно служит идеалом, к которому можно стремиться при рассмотрении конкретных алгоритмов.

В выпуске 9 конспекта лекций [1] описан $\Delta-$алгоритм получения ПЛ. Суть его такова. Рассматриваем два процесса интегрирования: умозрительный процесс, дающий решение $y(x)$ и реальный локализующий, на компьютере, формирующий локализатор $Y(x)$. Пусть оба процесса пришли к точке x_k. Первый процесс пришёл со значением $y_k = y(x_k)$, второй с значением в виде отрезка $Y_k = Y(x_k)$. Благодаря особенностям алгоритмики локализующего интегрирования, гарантировано включение $y_k \in Y_k$.

Так вот $\Delta-$алгоритм задаёт интервальный итерационный процесс начальным значением $Y^0 = Y_k$ и уравнением

$$Y^{m+1} = Y_k + \Delta x[0,\ 1] \cdot F(X_k,\ Y^m), \tag{1}$$

где обозначено $F(X,Y)$ – интервальное расширение правой части уравнения (1), $\Delta x = x_{k+1} - x_k$. Как доказано в п. 1241 конспекта [1], формируется антивложенная последовательность (АП) $\{Y^m\}$, если по второму аргументу функция $F(X_k,\ Y)$ монотонна по включению.

В п. 1131 доказывается, что в машинном исполнении ограниченная АП стабилизируется, т.е. $Y^{m+1} = Y^m = Y^*$ при некотором m. В п. 1242 доказывается, что в случае стабилизации решение $y(x)$ продолжимо вплоть до точки x_{k+1}, удовлетворяя включению $y(x) \in Y^*$ в пределах отрезка X_k.

После получения ПЛ выполняется уточнение. Вычисляется локализатор решения в точке x_{k+1} в надежде, что он окажется существенно более узким, чем Y^*. Так, локализатор по методу парабол (получаемый разложением Тэйлора с квадратичным остатком) имеет вид:

$$Y(x_{k+1}) = Y_k + \Delta x F_k + \frac{(\Delta x)^2}{2}\Big(F_x(X_k,\ Y^*) + F_y(X_k,\ Y^*)F(X_k, Y^*)\Big),$$

где через $F_x(X_k,\ Y^*)$ и $F_y(X_k,\ Y^*)$ обозначены локализаторы для $f_x(x,\ y)$ и $f_y(x,\ y)$ в прямоугольнике $(X_k,\ Y^*)$, а $F_k = F(x_k,\ Y_k)$.

Перейдём непосредственно к проблеме узкой ПЛ.

Начнём с примера 1244.1 из [1], где взят случай, когда после благополучного окончания рассмотренного этапа ПЛ оказалось:

$$F(X_k,\ Y^*) \not\ni 0. \tag{2}$$

Этим соотношением гарантируется монотонность решений в прямоугольнике $(X_k,\ Y^*)$. Следует заметить, что случай этот — довольно частый в практике численного интегрирования по той причине, что шаги интегрирования имеют тенденцию к измельчению, изменение же знака правой части – явление относительно редкое. Благодаря монотонности, для локализатора решения $y(x)$ на отрезке X_k можно указать иное значение, помимо Y^*. Именно, $y(X_k)$ формируется концевыми значениями $y(x_k)$ и $y(x_{k+1})$. Поэтому можно принять это иное значение локализатора Y^{**} в виде интервальной оболочки отрезков Y_k и Y_{k+1}, где Y_{k+1} подсчитан посредством уточнения первичного локализатора Y^*. Тогда можно принять скорректированный предварительный локализатор равным

$$Y^*_{\text{скорр}} = Y^* \cap Y^{**}.$$

Если Y_{k+1} как отображение от Y^* обладает монотонностью по включению, то более узкий первичный локализатор позволит в результате нового уточнения получить и более узкий Y^*. Это даст возможность получить ещё более узкий $Y(x_{k+1})$. Возникает сужающий итерационный процесс. Этот процесс образует тандем с предшествующим расширяющим. Создавшееся положение подобно тому, что описано в п. 1244.

Однако, всё это было описано в выпуске 9 наших лекций [1].

2 Основные результаты

Что нового мы собираемся предложить?

1. Пока не будем опираться на предположение (2). Предположим, что, помимо формы $F(X_k,\ Y^*)$, локализатор правой части дифференциального уравнения ищется по формуле конечных приращений:

$$\begin{aligned} f(x,\ y(x)) = f(x_k,\ y_k) + \Big(f\big(x,\ y(x)\big) - f(x_k,\ y_k)\Big) = \\ = f(x_k,\ y_k) + (x - x_k)\frac{d}{dx}f\big(x,\ y(x)\big)\Big|_{x\,=\,\theta}, \end{aligned}$$

где $x \in X_k,\ \theta \in [x_k,\ x]$. Вычисляя производную, переходя к интервальной арифметике и заменяя производную её локализатором, а число $f(x_k,\ y_k)$ его локализатором $F(x_k,\ Y_k)$, получаем:

$$f(x,\ y(x)) \in F(x_k,\ Y_k) + (X - x_k)\big(F_x(X_k,\ Y^*) + F_y(X_k,\ Y^*)F(X_k, Y^*)\big).$$

Теперь ясно, что правая часть служит локализатором для левой. Здесь $X_k - x_k = \Delta x[0,1]$. Поскольку пересечение локализаторов образует локализатор, то таковым следует признать

$$F(X_k,\ Y^*) \cap \Big(F(x_k,\ Y_k) + \Delta x[0,1]\big(F_x({}_kX,\ Y^*) + F_y(X_k,\ Y^*)F(X_k, Y^*)\big)\Big).$$

После этого для нахождения локализаторов можно ввести интервальный итерационный сужающий процесс:

$$\begin{aligned} F^{m+1}(X_k,\ Y^*) = F^m(X_k,\ Y^*) \cap \\ \cap \Big(F(x_k,\ Y_k) + \Delta x[0,1]\big(F_x(X_k,\ Y^*) + F_y(X_k,\ Y^*)F^m(X_k, Y^*)\big)\Big), \end{aligned} \tag{3}$$

где определим начальное значение как результат первичной предварительной локализации: $F^0(X_k,\ Y^*) = F(X_k,\ Y^*)$.

2. Теперь можно было бы обсудить очевидную возможность совмещения изложенных соображений с привлечением постоянства знака $f(x,\ y)$, т.е. с условием (2).

3. Можно обсудить перспективы применения $\vee$-*подобных* алгоритмов, свободных от требования монотонности по включению.

Эта идея, изложена в п. 1243 книги [1].

Но недостаток места и времени препятствует этим расширениям доклада.

Список литературы

1. Меньшиков Г.Г. Элементы локализующего интегрирования дифференциальных уравнений // *Интервальный анализ и методы вычислений*. Конспект лекций. Выпуск 9. Издание второе. – Санкт-Петербург: ООП НИИ Химии СПбГУ, 2001.

Статистика интервальных данных

А.И. Орлов

МГТУ им. Баумана
Россия, г. Москва
orlov@professor.ru

В статистике интервальных данных элементы выборки — не числа, а интервалы. Это приводит к алгоритмам и выводам, принципиально отличающимся от классических. Доклад посвящен основным идеям и подходам асимптотической статистики интервальных данных. Приведены некоторые результаты, связанные с основополагающими в рассматриваемой области прикладной математической статистики понятиями нотны и рационального объема выборки. Кратко рассмотрен ряд задач оценивания характеристик и параметров распределения, проверки гипотез, регрессионного, кластерного и дискриминантного анализов.

Перспективная и быстро развивающаяся область статистических исследований последних лет — математическая статистика интервальных данных. Речь идет о развитии методов прикладной математической статистики в ситуации, когда статистические данные — не числа, а интервалы, в частности, порожденные наложением ошибок измерения на значения случайных величин. Полученные результаты отражены, в частности, в выступлениях на проведенной в "Заводской лаборатории" дискуссии [1] и в докладах международной конференции ИНТЕРВАЛ-92 [2]. Приведем основные идеи весьма перспективного для вероятностно-статистических методов и моделей принятия решений асимптотического направления в статистике интервальных данных.

В настоящее время признается необходимым изучение устойчивости (робастности) оценок параметров к малым отклонениям исходных данных и предпосылок модели. Однако популярная среди теоретиков модель засорения (Тьюки-Хьюбера) представляется не вполне адекватной. Эта модель нацелена на изучение влияния больших "выбросов". Поскольку любые реальные измерения лежат в некотором фиксированном диапазоне, а именно, заданном в техническом паспорте средства измерения, то зачастую выбросы не могут быть слишком большими. Поэтому представляются полезными иные, более общие схемы устойчивости, в частности, введенные в [3], в которых, например, учитываются отклонения распределений результатов наблюдений от предположений модели.

В одной из таких схем изучается влияние интервальности исходных данных на статистические выводы. Необходимость такого изучения стала очевидной следующим образом. В государственных стандартах СССР по прикладной статистике в обязательном порядке давалось справочное приложение "Примеры применения правил стандарта". При разработке ГОСТ 11.011-83 [4] были переданы для анализа реальные данные о наработке резцов до предельного состояния (в часах). Оказалось, что все эти данные представляли собой либо целые числа, либо полуцелые (т.е. после умножения на 2 становящиеся целыми). Ясно, что исходная длительность наработок искажена. Необходимо учесть в статистических процедурах наличие такого искажения исходных данных. Как это сделать?

Первое, что приходит в голову — модель группировки данных, согласно которой для истинного значения X проводится замена на ближайшее число из множества $\{0{,}5n,\ n=1,2,3,...\}$. Однако эту модель целесообразно подвергнуть сомнению, а также рассмотреть иные модели. Так, возможно, что X надо приводить к ближайшему сверху элементу указанного множества — если проверка качества поставленных на испытание резцов проводилась раз в полчаса. Другой вариант: если расстояния от X до двух ближайших элементов множества $\{0{,}5n,\ n=1,2,3,...\}$ примерно равны, то естественно ввести рандомизацию при выборе заменяющего числа, и т.д.

Целесообразно построить новую математико-статистическую модель, согласно которой **результаты наблюдений — не числа, а интервалы**. Например, если в таблице приведено значение 53,5, то это значит, что реальное значение — какое-то число от 53,0 до 54,0, т.е. какое-то число в интервале [53,5 – 0,5; 53,5 + 0,5], где 0,5 — максимально возможная погрешность. Принимая эту модель, мы попадаем в новую научную область — статистику интервальных данных [5,6]. Статистика интервальных данных идейно связана с интервальной математикой, в которой в роли чисел выступают интервалы (см., например, монографию [7]).

По мнению ряда специалистов, статистика интервальных данных является частью интервальной математики. Впрочем, есть точка зрения, согласно которой такое включение нецелесообразно, поскольку статистика интервальных данных использует несколько иные подходы к алгоритмам анализа реальных данных, чем сложившиеся в интервальной математике (подробнее см. ниже).

В настоящей главе развиваем асимптотические методы статистического анализа интервальных данных при больших объемах выборок и малых погрешностях измерений. В отличие от классической математической статистики, сначала устремляется к бесконечности объем выборки и только потом — уменьшаются до нуля погрешности. В частности, еще в начале 1980-х годов с помощью такой асимптотики были сформулированы правила выбора метода оценивания в ГОСТ 11.011-83 [4].

Разработана [8] общая схема исследования, включающая расчет нотны (максимально возможного отклонения статистики, вызванного интервальностью исходных данных) и рационального объема выборки (превышение которого не дает существенного повышения точности оценивания). Она применена к оцениванию математического ожидания и дисперсии [1], медианы и коэффициента вариации [9], параметров гамма-распределения [4, 10] и характеристик аддитивных статистик [8], при проверке гипотез о параметрах нормального распределения, в т.ч. с помощью критерия Стьюдента, а также гипотезы однородности с помощью критерия Смирнова [9]. Изучено асимптотическое поведение оценок метода моментов и оценок максимального правдоподобия (а также более общих — оценок минимального контраста), проведено асимптотическое сравнение этих методов в случае интервальных данных, найдены общие условия, при которых, в отличие от классической математической статистики, метод моментов дает более точные оценки, чем метод максимального правдоподобия [11].

Разработаны подходы к рассмотрению интервальных данных в основных постановках регрессионного, дискриминантного и кластерного анализов [12]. В частности, изучено влияние погрешностей измерений и наблюдений на свойства алгоритмов регрессионного анализа, разработаны способы расчета нотн и рациональных объемов выборок, введены и исследованы новые понятия многомерных и асимптотических нотн, доказаны соответствующие предельные теоремы [12,13]. Начата разработка интервального дискриминантного анализа, в частности, рассмотрено влияние интервальности данных на показатель качества классификации [12,14]. Основные идеи и результаты рассматриваемого направления в статистике интервальных данных приведены в публикациях обзорного характера [5,6].

Как показала, в частности, международная конференция ИНТЕРВАЛ-92, в области асимптотической математической статистики интервальных данных мы имеем мировой приоритет. По нашему мнению, со временем во все виды статистического программного обеспечения должны быть включены алгоритмы интервальной статистики, "параллельные" обычно используемым алгоритмам прикладной математической статистики. Это позволит в явном виде учесть наличие погрешностей у результатов наблюдений, сблизить позиции метрологов и статистиков.

Многие из утверждений статистики интервальных данных весьма отличаются от аналогов из классической математической статистики. В частности, не существует состоятельных оценок; средний квадрат ошибки оценки, как правило, асимптотически равен сумме дисперсии оценки, рассчитанной согласно классической теории, и некоторого положительного числа (равного квадрату т.н. нотны — максимально возможного отклонения значения статистики из-за погрешностей исходных данных) — в результате метод моментов оказывается иногда точнее метода максимального правдоподобия [11]; нецелесообразно увеличивать объем выборки сверх некоторого предела (называемого рациональным объемом выборки) — вопреки классической теории, согласно которой чем больше объем выборки, тем точнее выводы.

В стандарт [4] был включен раздел 5, посвященный выбору метода оценивания при неизвестных параметрах формы и масштаба и известном параметре сдвига и основанный на концепциях статистики интервальных данных. Теоретическое обоснование этого раздела стандарта опубликовано лишь через 5 лет в статье [10].

Следует отметить, что хотя в 1982 г. при разработке стандарта [4] были сформулированы основные идеи статистики интервальных данных, однако из-за недостатка времени они не были полностью реализованы в ГОСТ 11.011-83, и этот стандарт написан в основном в классической манере. Развитие идей статистики интервальных данных продолжается уже в течение 20 лет, и еще много чего надо сделать! Большое значение статистики интервальных данных для современной прикладной статистики обосновано в [15,16].

Ведущая научная школа в области статистики интервальных данных — это школа проф. А.П. Вощинина, активно работающая с конца 70-х годов. Полученные результаты отражены в ряде монографий (см., в частности, [17,18,19]), статей [1, 20, 21], докладов, в частности, в трудах [2] Международной конференции ИНТЕРВАЛ-92, диссертаций [22,23]. В частности, изучены проблемы регрессионного анализа, планирования эксперимента, сравнения альтернатив и принятия решений в условиях интервальной неопределенности. Рассматриваемое ниже направление отличается нацеленностью на асимптотические результаты, полученные при больших объемах выборок и малых погрешностях измерений, поэтому оно и названо **асимптотической статистикой интервальных данных.**

Сформулируем основные идеи асимптотической математической статистики интервальных данных.

Пусть существо реального явления описывается выборкой $x_1, x_2, ..., x_n$. В вероятностной теории математической статистики, из которой мы исходим (см. терминологическую статью [24]), выборка — это набор независимых в совокупности одинаково распределенных случайных величин. Однако беспристрастный и тщательный анализ подавляющего большинства реальных задач показывает, что статистику известна отнюдь не выборка $x_1, x_2, ..., x_n$, а величины

$$y_j = x_j + \varepsilon_j \ , \ j = 1, 2, \dots , n \ ,$$

где $\varepsilon_1, \varepsilon_2, ..., \varepsilon_n$ – некоторые погрешности измерений, наблюдений, анализов, опытов, исследований (например, инструментальные ошибки).

Одна из причин появления погрешностей — запись результатов наблюдений с конечным числом значащих цифр. Дело в том, что для случайных величин с непрерывными функциями распределения событие, состоящее в попадании хотя бы одного элемента выборки в множество рациональных чисел, согласно правилам теории вероятностей имеет вероятность 0, а такими событиями в теории вероятностей принято пренебрегать. Поэтому при рассуждениях о выборках из нормального, логарифмически нормального, экспоненциального, равномерного, гамма — распределений, распределения Вейбулла-Гнеденко и др. приходится принимать, что эти распределения имеют элементы исходной выборки x_1 , x_2 , ..., x_n, в то время как статистической обработке доступны лишь искаженные значения $y_j = x_j + \varepsilon_j$.

Введем обозначения

$$x = (x_1 , x_2 , ..., x_n),\ y = (y_1 , y_2 , ..., y_n),\ \varepsilon = \varepsilon_1 + \varepsilon_2 + ... + \varepsilon_n .$$

Пусть статистические выводы основываются на статистике $f : R^n \to R^1$, используемой для оценивания параметров и характеристик распределения, проверки гипотез и решения иных статистических задач. Принципиально важная для статистики интервальных данных идея такова: СТАТИСТИК ЗНАЕТ ТОЛЬКО *f(y),* НО НЕ *f(x).*

Очевидно, в статистических выводах необходимо отразить различие между *f(y)* и *f(x).* Одним из двух основных понятий статистики интервальных данных является понятие нотны.

Определение. Величину максимально возможного (по абсолютной величине) отклонения, вызванного погрешностями наблюдений ε, известного статистику значения *f(y)* от истинного значения *f(x)*, т.е.

$$N_f(x) = \sup | f(y) — f(x) | ,$$

где супремум берется по множеству возможных значений вектора погрешностей ε (см. ниже), будем называть НОТНОЙ .

Если функция *f* имеет частные производные второго порядка, а ограничения на погрешности имеют вид

$$| \varepsilon_i | \leq \Delta,\ i = 1,2,..., n, \qquad (1)$$

причем Δ мало, то приращение функции *f* с точностью до бесконечно малых более высокого порядка описывается главным линейным членом, т.е.

$$f(y) - f(x) = \sum_{1 \leq i \leq n} \frac{\partial f(x)}{\partial x_i} \varepsilon_i + O(\Delta^2).$$

Чтобы получить асимптотическое (при $\Delta \to 0$) выражение для нотны, достаточно найти максимум и минимум линейной функции (главного линейного члена) на кубе, заданном неравенствами (1). Легко видеть, что максимум достигается, если положить

$$\varepsilon_i = \begin{cases} \Delta, \dfrac{\partial f(x)}{\partial x_i} \geq 0, \\ -\Delta, \dfrac{\partial f(x)}{\partial x_i} < 0, \end{cases}$$

а минимум, отличающийся от максимума только знаком, достигается при $\varepsilon_i^{'} = -\varepsilon_i$. Следовательно, *нотна* с точностью до бесконечно малых более высокого порядка имеет вид

$$N_f(x) = (\sum_{1 \leq i \leq n} | \frac{\partial f(x)}{\partial x_i} |)\Delta.$$

Это выражение назовем *асимптотической нотной*.

Условие (1) означает, что исходные данные представляются статистику в виде интервалов $[y_i - \Delta; y_i + \Delta], i = 1,2,..., n$ (отсюда и название этого научного направления). Ограничения на погрешности могут задаваться разными способами — кроме абсолютных ошибок используются относительные или иные показатели различия между *x* и *y*.

Если задана не предельная абсолютная погрешность Δ, а предельная относительная погрешность δ, т.е. ограничения на погрешности вошедших в выборку результатов измерений имеют вид

$$| \varepsilon_i | \leq \delta | x_i |, i = 1,2,..., n,$$

то аналогичным образом получаем, что нотна с точностью до бесконечно малых более высокого порядка, т.е. асимптотическая нотна, имеет вид

$$N_f(x) = (\sum_{1 \le i \le n} | x_i \frac{\partial f(x)}{\partial x_i} |)\delta.$$

При практическом использовании рассматриваемой концепции необходимо провести тотальную замену символов x на символы y. В каждом конкретном случае удается показать, что в силу малости погрешностей разность $N_f(y) - N_f(x)$ является бесконечно малой более высокого порядка сравнительно с $N_f(x)$ или $N_f(y)$.

Основные результаты в вероятностной модели. В классической вероятностной модели элементы исходной выборки $x_1, x, ..., x_n$ рассматриваются как независимые одинаково распределенные случайные величины. Как правило, существует некоторая константа $C > 0$ такая, что в смысле сходимости по вероятности

$$\lim_{n\to\infty} N_f(x) = C\Delta. \qquad (2)$$

Соотношение (2) доказывается отдельно для каждой конкретной задачи.

При использовании классических эконометрических методов в большинстве случаев используемая статистика *f(x)* является асимптотически нормальной. Это означает, что существуют константы a и σ^2 такие, что

$$\lim_{n\to\infty} P\left(\sqrt{n}\,\frac{f(x)-a}{\sigma} < x\right) = \Phi(x),$$

где $\Phi(x) -$ функция стандартного нормального распределения с математическим ожиданием 0 и дисперсией 1. При этом обычно оказывается, что

$$\lim_{n\to\infty} \sqrt{n}(Mf(x) - a) = 0$$

и

$$\lim_{n\to\infty} nDf(x) = \sigma^2,$$

а потому в классической эконометрике средний квадрат ошибки статистической оценки равен

$$M(f(x)-a)^2 = (Mf(x)-a)^2 + Df(x) = \frac{\sigma^2}{n}$$

с точностью до членов более высокого порядка.

В статистике интервальных данных ситуация совсем иная — обычно можно доказать, что средний квадрат ошибки равен

$$\max_{\{\varepsilon\}} M(f(y)-a)^2 = \frac{\sigma^2}{n} + N_f^2(y) + o(\Delta^2 + \frac{1}{n}). \qquad (3)$$

Из соотношения (3) можно сделать ряд важных следствий. Прежде всего отметим, что правая часть этого равенства, в отличие от правой части соответствующего классического равенства, не стремится к 0 при безграничном возрастании объема выборки. Она остается больше некоторого положительного числа, а именно, квадрата нотны. Следовательно, статистика *f(x)* не является состоятельной оценкой параметра a. Более того, состоятельных оценок вообще не существует.

Пусть доверительным интервалом для параметра a, соответствующим заданной доверительной вероятности γ, в классической математической статистике является интервал $(c_n(\gamma); d_n(\gamma))$. В статистике интервальных данных аналогичный доверительный интервал является более широким. Он имеет вид $(c_n(\gamma) - N_f(y); d_n(\gamma) + N_f(y))$. Таким образом, его длина увеличивается на две нотны. Следовательно, при увеличении объема выборки длина доверительного интервала не может стать меньше, чем $2C\Delta$ (см. формулу (2)).

В статистике интервальных данных методы оценивания параметров имеют другие свойства по сравнению с классической математической статистикой. Так, при больших объемах выборок метод моментов может быть заметно лучше, чем метод максимального правдоподобия (т.е. иметь меньший средний квадрат ошибки — см. формулу (3)), в то время как в классической математической статистике второй из названных методов всегда не хуже первого.

Рациональный объем выборки. Анализ формулы (3) показывает, что в отличие от классической математической статистики нецелесообразно безгранично увеличивать объем выборки, поскольку средний квадрат ошибки остается всегда большим квадрата нотны. Поэтому представляется полезным ввести понятие "рационального объема выборки" n_{rat}, при достижении которого продолжать наблюдения нецелесообразно.

Как установить "рациональный объем выборки"? Можно воспользоваться идеей "принципа уравнивания погрешностей", выдвинутой в монографии [3]. Речь идет о том, что вклад погрешностей различной природы в общую погрешность должен быть примерно одинаков. Этот принцип дает возможность выбирать необходимую точность оценивания тех или иных характеристик в тех случаях, когда это зависит от исследователя. В статистике интервальных данных в соответствии с "принципом уравнивания погрешностей" предлагается определять рациональный объем выборки n_{rat} из условия равенства двух величин — метрологической составляющей, связанной с нотной, и статистической составляющей — в среднем квадрате ошибки (3), т.е. из условия

$$\frac{\sigma^2}{n_{rat}} = N_f^2(y), \quad n_{rat} = \frac{\sigma^2}{N_f^2(y)}.$$

Для практического использования выражения для рационального объема выборки неизвестные теоретические характеристики необходимо заменить их оценками. Это делается в каждой конкретной задаче по-своему.

Исследовательскую программу в области статистики интервальных данных можно "в двух словах" сформулировать так: для любого алгоритма анализа данных (алгоритма прикладной статистики) необходимо вычислить нотну и рациональный объем выборки. Или иные величины из того же понятийного ряда, возникающие в многомерном случае, при наличии нескольких выборок и при иных обобщениях описываемой здесь простейшей схемы. Затем проследить влияние погрешностей исходных данных на точность оценивания, доверительные интервалы, значения статистик критериев при проверке гипотез, уровни значимости и другие характеристики статистических выводов. Очевидно, классическая математическая статистика является частью статистики интервальных данных, выделяемой условием $\Delta = 0$.

Рассмотренные выше соображения оказались полезными для решения чисто экономических проблем, в частности, в инвестиционном менеджменте [25]. Основные идеи статистики интервальных данных включены в учебник по эконометрике [26]. Подробное изложение статистики интервальных данных содержится в учебнике по теории принятия решений [27].

Список литературы

1. Дискуссия по анализу интервальных данных // Заводская лаборатория. 1990. Т.56. No.7, с.75-95.
2. Сборник трудов Международной конференции по интервальным и стохастическим методам в науке и технике (ИНТЕРВАЛ-92). Тт. 1,2. — М.: МЭИ, 1992, 216 с., 152 с.
3. Орлов А.И. Устойчивость в социально-экономических моделях. — М.: Наука, 1979. 296 с.
4. ГОСТ 11.011-83. Прикладная статистика. Правила определения оценок и доверительных границ для параметров гамма-распределения. — М.: Изд-во стандартов, 1984, 53 с.
5. Orlov A.I. // Interval Computations, 1992, No.1(3), p.44-52.
6. Орлов А.И. // Наука и технология в России. 1994. No.4(6). С.8-9.
7. Шокин Ю.И. Интервальный анализ. Новосибирск: Наука, 1981, 112 с.
8. Орлов А.И. — В сб.: Статистические методы оценивания и проверки гипотез. Межвузовский сборник научных трудов. Пермь: Изд-во Пермского государственного университета, 1990, с..89-99.
9. Орлов А.И. — В сб.: Статистические методы оценивания и проверки гипотез. Межвузовский сборник научных трудов. Пермь: Изд-во Пермского государственного университета, 1991, с.77-86.
10. Орлов А.И. — В сб.: Статистические методы оценивания и проверки гипотез. Межвузовский сборник научных трудов. Пермь: Изд-во Пермского государственного университета, 1988, с.45-55.
11. Орлов А.И. — В сб.: Статистические методы оценивания и проверки гипотез. Межвузовский сборник научных трудов. — Пермь: Изд-во Пермского государственного университета, 1995, с.114-124.
12. Орлов А.И. — В сб.: Статистические методы оценивания и проверки гипотез. Межвузовский сборник научных трудов. Пермь: Пермский государственный университет, 1993, с.149-158.
13. Биттар А.Б. Метод наименьших квадратов для интервальных данных. Дипломная работа. — М.: МЭИ, 1994. 38 с.
14. Пузикова Д.А. // Наука и технология в России. 1995. No.2(8). С.12-13.
15. Орлов А.И. // Надежность и контроль качества, 1991, № 8, с.3-8.
16. Орлов А.И. // Заводская лаборатория. 1998. Т.64. № 3. С.52-60.
17. Вощинин А.П. Метод оптимизации объектов по интервальным моделям целевой функции. — М.: МЭИ, 1987. 109 с.
18. Вощинин А.П., Сотиров Г.Р. Оптимизация в условиях неопределенности. — М.: МЭИ — София: Техника, 1989. 224 с.
19. Вощинин А.П., Акматбеков Р.А. Оптимизация по регрессионным моделям и планирование эксперимента. — Бишкек: Илим, 1991. 164 с.
20. Вощинин А.П. // Заводская лаборатория. 2000. Т.66, № 3. С.51-65.
21. Вощинин А.П. // Заводская лаборатория. 2002. Т.68, № 1. С.118-126.
22. Дывак Н.П. Разработка методов оптимального планирования эксперимента и анализа интервальных данных. Автореф. дисс. канд. технич. наук. — М.: МЭИ, 1992. 20 с.

23. Симов С.Ж. Разработка и исследование интервальных моделей при анализе данных и проектировании экспертных систем. Автореф. дисс. канд. технич. наук. — М.: МЭИ, 1992. 20 с.
24. Орлов А.И. // Заводская лаборатория. 1999. Т.65. № 7. С.46-54.
25. Алешин Д.Н. Экономическое обоснование эффективности инвестиционных проектов на предприятиях на основе применения эконометрического метода интервальной оценки. Автореферат дисс. … канд. экон. наук. – М.: МГТУ им. Н.Э.Баумана, 2001. — 16 с.
26. Орлов А.И. Эконометрика. Учебник для вузов. – 2-е изд., перераб. и доп.– М.: Издательство «Экзамен», 2003.
27. Орлов А.И. Теория принятия решений. Учебник для вузов. — М.: Издательство «Экзамен», 2003.

Интервальная природа числа в физике как следствие минимизации погрешности измерения

Л.С. Терехов

Омский филиал Института математики СО РАН им. С.Л. Соболева,
644099, г. Омск, ул. Певцова, 13, ОФ ИМ СО РАН
terekhov@iitam.omsk.net.ru

В экспериментах по радиоволновому зондированию ионосферного слоя плазмы, проводимых одновременно с Земли и со спутника, наблюдается явление, не находящее приемлемого объяснения [1]. Поиск причины расхождения эксперимента и теории приводит к необходимости ввести дополнение к известной, ставшей уже канонической, форме соотношения неопределённостей [2]. В настоящем сообщении предложено обобщение соотношения неопределённостей в случае, не нагруженном ионосферной спецификой, а также - следствия такого обобщения. Минимизация погрешности измерений, допускаемая формой обобщённого соотношения неопределённостей, оказывается достижимой при условии локального квантования его сомножителей. В отличие от следствий известных форм соотношения неопределённостей основной полученный результат состоит в том, что: 1) минимум функции, представляющей локальную погрешность измерений, достигается лишь в том случае, если её аргумент есть интервал определённой ненулевой длины; 2) число, представляющее погрешность измерения, не может быть точным даже потенциально.

Введение

Известны эквивалентность квантовомеханических и соответствующих форм соотношения неопределённостей (СН) классической физики, которые использовались ещё задолго до создания квантовой механики. Часто используемая форма СН классической физики в переменных время t - частота f имеет вид

$$\Delta f * \Delta t \approx 1\,. \tag{1}$$

В волновых задачах, в которых удаётся располагать характеристиками сигнала, более точные оценки можно получить на основе так называемого радиолокационного СН:

$$\Delta f * \Delta t \approx 1/\alpha\sqrt{\mu}\,, \tag{2}$$

где α - определяется формой сигнала (для простых сигналов $\alpha \sim 1$); μ -отношение сигнал/шум сигнала по мощности. Наличие величины μ указывает на случайную природу сомножителей Δf и Δt в (2). Оценка случайной погрешности измерения частоты, обозначаемой далее $\Delta fr(\mu,\Delta t^{1})$ на практике используется, согласно (2), в виде:

$$\Delta f_r(\mu, \Delta t^{-1}) \approx 1/\alpha * \Delta t\sqrt{\mu} \tag{3}$$

В начале развития теории СН сомножителям Δf и Δt придавался смысл как неопределённостям измеряемых величин, позже - как среднеквадратическим отклонениям (как, например, в форме (2)). Однако СН может быть также получено на основе прямого и обратного преобразований Фурье, применимых и к детерминированным функциям. При этом сомножители Δf и Δt находятся как области интегрирования из условия локализации в них преобладающей доли энергии сигнала. В итоге СН можно рассматривать как отвечающее сомножителям и случайной, и детерминированной природы.

Область применимости СН в его канонических формах (1) и (2), в явлениях классической физики во внимание не принимаемая, ограничивается взаимной независимостью входящих в него величин. Так, оценки погрешности $\Delta\tau$ времени распространения сигнала в среде могут оказываться неверными, если среда (например, ионосферный слой плазмы) обладает заметной частотной дисперсией: $\tau = \tau(f)$, $d\tau/df \neq 0$. В настоящей работе, не осложнённой ионосферной спецификой [2], рассматривается обобщение СН на примере зависимости частоты от времени: $f = f(t)$, $df/dt \neq 0$. Иллюстрация необходимости учёта зависимости $f(t)$ при построении обобщения СН может быть следующая. При измерении частоты f, зависящей от времени t, возникает проблема: какой длительности должно быть время анализа Δt (иначе - время измерения, накопления), при котором погрешность измерения частоты Δf оставалась бы приемлемой, например, $\Delta f \ll f$. В случае быстрых изменений частоты, таких, что за время Δt успевает смениться несколько характерных времён $\{\tau i\}$, и $\tau i \ll \Delta t$, разрешение по времени теряется вследствие того, что неверными оказываются сами оценки погрешностей, полученные из (2). Искомая зависимость изменения частоты $f(t)$, «свёрнутая» по интервалу усреднения Δt к единственному значению – отсчёту, оказывается потерянной. Измерения в подобных экспериментах теряют смысл. Если же выбрать $\tau i \gg \Delta t$, то, хотя оценки согласно (1) или (2) остаются верными, но теряется разрешение по частоте. Практически, до начала измерений, устанавливают величину Δt как некоторое компромиссное значение для всего ожидаемого ряда $\{\tau i\}$, что, однако, при большом разбросе значений τi этого ряда, ведёт всё-таки к потере τi, заметно отличающихся от Δt. И названная выше проблема, существующая из-за невыявленной взаимной зависимости величин, входящих в СН (например, частоты, зависящей от времени), остаётся актуальной.

Цель настоящего сообщения - иллюстрация порождения чисел интервальной природы в физическом процессе измерения.

Постановка задачи

Приведённый выше краткий анализ трудностей измерения частоты, зависящей от времени, свидетельствует о существовании общей проблемы - неполноты СН в его классической форме. Распространение оценок погрешности, получаемых согласно (2) и (3), на нестационарные явления с зависимостью $f = f(t)$ ведёт к дополнительной и скрытой погрешности измерения частоты. Основное допущение в постановке задачи по обобщению СН состоит в следующем. Погрешность измерения частоты $\Delta f(\mu, \Delta t, f'(t))$, называемая далее полной, полагается равной линейной сумме случайного $\Delta f_r(\mu, \Delta t^{-1})$ и систематического $\Delta f_s(\mu, \Delta t, f'(t))$ компонентов, входящих в полную погрешность с равными весами:

$$\Delta f(\mu, \Delta t, f'(t)) = \Delta f_r(\mu, \Delta t^{-1}) + \Delta f_s(\mu, \Delta t, f'(t)). \qquad (4)$$

Выбор показателя степени суммируемых независимых погрешностей измерения при представлении полной погрешности в (4) не имеет принципиального значения. Поэтому здесь принята наиболее простая линейная форма.

В теории сеток погрешностью, аналогичной рассматриваемой случайной, является погрешность округления, а аналогом систематической погрешности является погрешность аппроксимации [3,4].

Первое слагаемое правой части (4) - это погрешность измерения частоты, даваемая СН в его известных, ставших уже каноническими, формах (2) и (3), то есть - без учёта зависимости $f(t)$. Второе слагаемое (4) построено при ограничении вида функции $f = f(t)$ детерминированной зависимостью и представлено как приращение Δfs в тейлоровском разложении при приращении времени Δt:

$$\Delta f_s(\mu, \Delta t, f'(t), f''(t)) = f'(t)\Delta t + \frac{f''(t)}{2!}\Delta t^2 + \ldots \qquad (5)$$

Ограничиваясь первым слагаемым приращения в (5), получаем выражение для полной погрешности в явном и расширенном сравнительно с (3) виде:

$$\Delta f(\mu, \Delta t, f'(t)) = 1/(\alpha * \Delta t\sqrt{\mu}) + f'(t)\Delta t. \qquad (6)$$

Решение задачи

В отличие от выражения для Δfr из (3), в котором учтена зависимость величины Δfr от времени анализа Δt и отношения μ, полная погрешность $\Delta f(\mu, \Delta t, f'(t))$ измерения частоты зависит в (6) также от производной $f(t)'$ в каждый момент времени t. Минимизируя выражение для полной погрешности $\Delta f(\mu, \Delta t, f'(t))$ в (6) как функцию времени анализа Δt, находим для каждого момента времени, исключая экстремумы, оптимальное время анализа Δt^*:

$$\Delta t^*(\mu, f'(t)) = \frac{1}{\sqrt{\alpha\sqrt{\mu}\,|f'(t)|}} \ . \tag{7}$$

Соответствующая оптимальному времени (7) минимальная полная погрешность измерения частоты, зависящей от времени, имеет вид:

$$\Delta f_{\min}(\Delta t^*)) = \sqrt{\frac{4}{\alpha\sqrt{\mu}}|f'(t)|} \ . \tag{8}$$

В численных методах известен результат по оптимизации шага и соответствующей минимизации погрешности [3-5].

Минимизация полной погрешности $\Delta f(\mu, \Delta t, f'(t))$ приводит к «свёртыванию» расширенной формы СН (6). То есть, если в качестве *Δf* и *Δt* выступают соответственно *Δfmin* и *Δt**, то их произведение приводит к сохранению канонической формы СН:

$$\Delta t^*(\mu, f'(t)) * \Delta f_{\min}(\Delta t^*) = \frac{1}{(\alpha/2)\sqrt{\mu}} . \tag{9}$$

Возможность определения оптимального интервала при каждом отсчёте означает последовательную, по мере производства отсчётов, расстановку узлов, образующих одномерную сетку. Сетка в общем случае неравноотстоящих узлов строится следующим образом. Поскольку алгоритм нахождения оптимальной длины текущего интервала известен, а минимальная погрешность одинакова в пределах интервала, то целесообразно, с целью уменьшения погрешности в междоузлиях до величины, не большей погрешности в соседних узлах, каждый последующий шаг по времени приравнивать текущему значению оптимальной длины интервала $\Delta t^*(t_i)$. Упорядочиванию интервалов и узлов сетки отвечает при переходе от текущего момента времени t_i к следующему моменту t_{i+1} переменный шаг $\Delta t^*(t_i)$:

$$t_{i+1} = t_i + \Delta t^*(t_i) \, . \tag{10}$$

Обсуждение

Согласно принципу физики, который называют операционным, в экспериментальных науках считается целесообразным вводить в научный язык только те величины, которые могут быть определены операционно, кратким описанием эксперимента, необходимого для измерения этих величин. То есть, в состав физического понятия «определение величины» процесс измерения должен входить неотъемлемым компонентом.

СН в его классических формах устанавливает существование неопределённости значений двух одновременно измеряемых величин, что согласуется с опытом. Но при этом из СН также следует, что любая из величин пары, может быть определена с неограниченной точностью. Можно видеть, что достижение бесконечно высокой точности измерения одного из сомножителей СН есть следствие допускаемой континуальности обоих сомножителей СН. Указанное следствие в немалой степени оказалось теоретической базой понятий «истинное» и «точное» значения. Погрешность измерения, определяемая как отличие измеренного значения от точного, при указанном выше допущении, противоречия не содержит. Хотя пользоваться определением, основа которого – «точное» значение – оказывается величиной неизмеряемой, оказывается затруднительно. Требования практики измерений привели в метрологии в настоящее время к определению погрешности измерений, в котором ненаблюдаемому «точному» значению места не оставлено. Погрешность выражается посредством величин измеряемых, через среднеквадратические отклонения.

Наличие в (9) жестко детерминированных на каждой частоте сомножителей свидетельствует о существовании ограничения потенциальной точности измерения даже одного из сомножителей СН. Этот результат отличается от классического результата, следующего из (2), имеющего континуальную форму и не накладывающего предела на потенциальную точность измерений и согласуется с операционным принципом физики.

Делается предположение, что интервальное число более соответствует описанию явлений природы, чем число континуальное.

Список литературы

1. Пиггот В.Р., Равер К. Руководство URSI по интерпретации и обработке ионограмм. – М.: Наука, 1977.
2. Терехов Л.С. О полной погрешности радиоволновых измерений параметров неоднородного слоя плазмы // Геомагнетизм и аэрономия. – 1998. – Т. 38, N 6. – С.142–148.
3. Мак-Кракен Д., Дорн У. Численные методы и программирование на ФОРТРАНе. – М.: "Мир", 1977.
4. Мудров А.Е. Численные методы для ПЭВМ на языках БЕЙСИК, ФОРТРАН и ПАСКАЛЬ. – Томск: МП "РАСКО", 1992.
5. Бахвалов Н.С., Жидков Н.П., Кобельков Г.М. Численные методы. – М.: Лаборатория Базовых Знаний, 2001. – С. 83–84.

Abstracts of the papers of the volume

Akhmerov R.R. **Interval-affine arithmetic with controlled accuracy**

The work is devoted to developing interval-affine arithmetic with controlled accuracy, which is a new technique for computing outer estimates of the ranges of functions. We consider some problems inherent to the classical approach and possible ways to cope with them. Advantages of the new technique are presented along with the difficulties arising in the course of its application. Finally, we show how one can control the accuracy of the estimates obtained by the interval-affine arithmetic. The results of extensive computational tests are presented and discussed.

Bozorov M.B. **An interval algorithm for determining soil static pressure on multilined stiff round pipes in a high embankment**

We propose an interval technique for computing soil static pressure upon round single and multilined pipes in a high embankment. Kaucher complete interval arithmetic is used for the computation. We describe a software package designed for the calculation of the immersed multilined pipes by interval finite element method. The numerical results obtained for two- and multilined pipes are compared with known theoretical and experimental results.

Bozorov M.B., Berdiev B.Kh. **Towards the efficiency comparison of interval algorithms solving nonlinear equations systems**

We discuss general concepts that underlie comparing efficiency of interval methods for the solution of nonlinear equations systems. Some parameters determining the characteristic features of every equations system are are drawn up, several specific interval methods are considered briefly as well as their comparison criteria. Finally, we formulate recommendations on how to choose methods for solution of the main problem types.

Brevnov E.V. **Interval approach to the solution of the interval constraint satisfaction problem**

The work considers solving global optimization problems with the help of constraint propagation methods and interval techniques. Separate attention is paid to integer-valued optimization problems, and we involve interval branch-and-bound strategy for their solution. The results of numerical experiments are presented showing the efficiency of the approach we have developed.

Ershov A.G. **Guaranteed suboptimal solutions for linear optimization problems**

The paper inquires into the possibility of efficient calculation, on digital computers, of suboptimal solutions to linear optimization problems that satisfy feasibility conditions with guarantee. We introduce a class of problems, named {\em conditionally correct linerar optimization problems}, for which computing such solutions is correct in a sense. Presented is a new algorithm for calculating a suboptimal guaranteed feasible solution of the linear optimization problem. It is based on an interior point method and uses ``small perturbation'' technique. We describe how to construct, with our algorithm, a subminimal outer and submaximal inner interval estimates of the set of all the suboptimal solutions to the linear optimization problems.

Ivlev R.S. **Exponential stability of a class of nonlinear interval dynamical systems**

The work considers a class of nonlinear interval dynamic systems, its mathematical model being defined in the state space, with the nonlinearity of the so-called ``sector type''. Relying on the direct Lyapunov method, we derive sufficient conditions that the equlibrium state at zero is exponentially stable.

Janybekov B.S., Shary S.P. **Optimal outer estimation of the solution sets to interval linear systems**

The paper is devoted to computing optimal (exact) interval outer estimates of AE-solution sets to interval linear algebraic systems $\mathbf{A}x = \mathbf{b}$. We generalize, to a wider context, recently proposed Jansson's method that enables to find optimal eclosures of the solution sets in p calls of a linear programming algorithm, where p is the number of nonempty intersections of the solution set with separate orthants.

Kashevarova T.P., Semenov A.L. **On the solutions to systems of nonlinear equations**

The work is devoted to the solution of nonlinear systems of equations with the use of interval techniques. The numerical algorithms we propose are efficient for solving equations systems of moderate dimensions and can be applied to the so-called "creeping solutions" as well. We apply an approach based on combination of interval techniques and constraint propagation methods for the problem statements involving interval input data.

***Kinsht N.V., Petrunko N.N.* On possibilities to describe a connection diagram of an interval electrical circuit**

We consider possibilities of heuristic construction of the matrices of the main cut-sets and circuits associated with the trees of the electrical circuit graph of a special form, that are constructed taking into account interval characteristics of the regimes and parameters of separate branches.

***Kleimenov A.E.* Constructing cooperative solvers on the basis of constraint propagation methods**

The paper describes applying cooperative approach to the solution of mathematical modelling problems with complex structure. We propose both a technology for construction of the cooperative solvers and software environment facilitating this process. Architecture of the environment and its current implementation are presented along with specific examples of how it can be efficiently applied to the solution of some classes of problem.

***Kozina G.L., Kudermetov R.* Linear criterion for graph optimization problems with interval parameters**

We propose a new criterion for comparing solutions in discrete optimization problems. In particular, minimal spanning tree problem for graphs with interval edge weights is thoroughly considered.

***Kuntashev P.A.* On selecting investment project**

The problem of selecting project portfolio structure providing for the maximum effect (NPV) of the portfolio has been studied. The project risk connected with a money flow uncertainty is treated with the use of interval analysis methods.

***Menshikov G.G.* To the problem of narrow initial enclosing of trajectory segment**

The paper deals with validated solution of initial value problem for differential equations of the form $y'=f(x, y)$, $y(x_0)=y_0$. Computer solution of the problem implies finding a preliminary enclosure $Y(x)$ for the function $y(x)$, i.e., a strip formed by interval values and containing a segment of the trajectory. The latter is necessary in the step-wise solution methods, whose algorithms can be splitted into *preliminary*, *refining* and, possibly, *testing* stages. The preliminary stage is of special importance, since it has to clarify whether the solution $y(x)$ can be extended to the interval X_k or not. Our report advances several ways to compute narrow preliminary enclosures for the solution of differential equations.

***Nazin S.A., Polyak B.T.* Interval technique in the parameter estimation problem**

The problem of paramteric estimation for a linear model under interval uncertainty is considered. We design an algorithm of optimal estimation taking the solution process for interval linear algebraic system as a basis.

***Noskov S.I.* A point approximation of the solution set to interval system of linear algebraic equation**

For interval linear algebraic systems with nonsquare matrices, the solution set may sometimes turn out empty or badly shaped. We propose a procedure of chosing a single ``solution point'' that represents, in some sense, the whole solution set or substitutes it in case of it is empty.

***Orlov A.I.* Statistics of interval data**

In the statistics of interval data, sampling elements are intervals rahter than numbers. That leads to techniques and conclusions which are principally different from the classical ones. Our paper gives a survey of the main ideas and approaches of the asymptotical statistics of interval data. We present some results related to the concepts of notna and rational sampling size, which are fundamental for the field of applied statistics under consideration. Also, some problems of estimation of distribution parameters, of hypothesis testing, of regression analysis, cluster analysis and discriminant analysis are briefly considered.

***Pertsev N.V.* Construction of the attracting sets for stable solutions of differential equations with the use of monotone technique and M-matrices**

In the paper, construction of attracting sets for stable equilibrium states for a class time-delay differential equations systems is considered. We advance a technique for constructing such regions that is based on two-sided approximation method and properties of nonsingular M-matrices. For a given equilibrium state x^*, the attracting region is built as a multi-dimensional parallelepiped with the center x^*. The bounds of the parallelepiped $x^* - z^1$? x ? $x^* + z^2$ are determined by vectors z^1, z^2, that can be computed as starting points of a special iterative process.

***Petrov E.S.* A symbolic-interval heuristic for minimization under boundary constraints**

Global optimization subject to bound constraints helps answering many practical questions in chemistry, molecular biology, economics. Most of algorithms for the solution of global optimization problems are a combination of interval methods and exhuastive search. The efficiency of such algorithms is characterized by their ability to detect and eliminate sub-optimal feasible regions. This ability is increased by availability of a good upper bound on the global minimum. In this paper, we present a symbolic-interval algorithm for calculation of upper bounds in bound-constrained global minimization problems and report the results of some experiments.

Proskurin A.V., Sagalakov A.M. **On spectral Orr-Sommerfeld problem**

The work deals with automation of complex spectra computations in the problem of parallel flows stability. We use argument principle. Apart from high reliability, the technique provides us with additional information about the problem features.

Pushkov S.G., Krivoshapko S.Yu. **On the state space realization problem for interval dynamic systems**

We consider state space realization problem for interval dicrete-time dynamic systems, analyzing its possible formulations and ways of solution. A sufficient criterion of algebraic realizability of impulse sequence of interval matrices has been stated and proved. Also, we propose a technique for finding algebraic realizations of either totally nonegative or totally nonpositive systems. A numerical example illustrates our results.

Rogalev A.N. **Problems of practical (interval) stability with a given set of marginary deviations**

Our report describes how guaranteed estimation of the solution sets to differential equations can be applied to practical stability problems. The numerical values of the guaranteed bounds of the solution sets resulting from our technique takes into account the influence of permanent perturbations represented by multivalued functions. Knowing such bounds enables one to formulate and substantiate mathematically rigorous results about practical stability of specific dynamic systems.

Semenov A.L. **Constraint propagation methods: basic concepts**

The paper considers the basic notions and algorithms of constraint programming - one of the fields of artificial intelligence. Conceptions of consistencies for finite and continuous domains are given and some algorithms of achieving consistencies are presented. Brief descriptions of some systems based on the approach are given as well.

Shary S.P. **Solving interval linear systems with tied data**

The work elaborates *parameter partition methods* (PPS-methods) for optimal outer estimation of the solution sets to interval systems of equations. We consider interval linear algebraic systems whose elements are not allowed vary independently within the prescribed intervals, being subject to some ties (constraints). In particular, symmetric interval linear systems are treated more thoroughly, while the solution of some other problems — with persymmetric, skew-symmetric, Hankel and Töplitz matrices — is outlied briefly.

Sokolova S.P. **Immunocomputing for complex interval systems**

The work is devoted to extending the so-called immunocomputing approach for a class of complex systems with parameter uncertainty of the interval type. Using the notions and methods of interval analysis, we develop singular-value decomposition for interval matrices, learning with an expert and self-learning, classification and representation of the results in the image space. The paper includes examples of interval immune systems for plague monitoring and information technologies security.

Terekhov L.S. **Interval nature of number in physics as a consequence of minimizing measurement errors**

Experiments of radiowave exploration of ionospheric plasma layers conducted simultaneously from both the Earth and satellites demonstrate a phenomenon that cannot be explained in a satisfactory way. Searching for the reasons why the experiments and theory contradict to each other leads to the necessity of introducing an addendum to the well-known form of uncertainty (indeterminacy) principle. In the present note, we propose a generalization of the uncertainty relationship for the case not connected with the specific ionospheric character as well as consequences of such a generalization. Minimization of the measurement errors allowed by the new form of uncertainty principle turns out attainable provided that its multipliers are locally quantized.

Utyubaev G.Sh. **On interval methods for differential problems with data uncertainty in the form of convex sets**

The work consists of five mutually connected parts devoted to 1) properties of non-standard interval arithmetic, 2) general properties of quasilinear spaces, 3) properties of interval functions, 4) convergence of interval methods, 5) applications. We introduce the concepts of complete quasilinear space, midpoint and width of an element from the space, quasilinear subspace, inner (outer) approximation, stability, inner (outer) convergence of single-step interval methods solving initial value problems with interval data. An estimate of the solution produced by interval approximating method is obtained. Also, we prove a theorem about inner and outer convergence for dynamic systems with interval input data. An approximate method in the space of n-dimensional ellipsoids is constructed for linear systems of ordinary differential equations, having the second convergence order, as well as a similar modified parallelotopes method.

Пятая международная конференция
ПЕРСПЕКТИВЫ СИСТЕМ ИНФОРМАТИКИ

8–9 июля 2003 года, Новосибирск, Академгородок, Россия

Рабочее совещание

Интервальная математика
и методы распространения ограничений

Доклады и тезисы

Рукопись поступила в редакцию 26.06.2003
Обложка М. Миронова
Ответственные за выпуск А. Л. Семенов, С. П. Шарый

Подписано в печать 01.07.2003
Тираж 40 экз

НФ ООО ИПО "Эмари" РИЦ, 630090, г. Новосибирск, пр.Акад. Лаврентьева, 6