

ОБ ОДНОМ МЕТОДЕ НЕЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ С ПРОИЗВОЛЬНЫМИ ОГРАНИЧЕНИЯМИ

А. Ю. Кучмин^а, канд. техн. наук, ведущий научный сотрудник
^аИнститут проблем машиноведения РАН, Санкт-Петербург, РФ

Введение: при решении современных задач в областях моделирования и идентификации нелинейных и нестационарных систем, а также при анализе и фильтрации данных экспериментов все большее распространение получают методы оптимизации и, в частности, нелинейного программирования. Эти задачи относятся к классу задач многомерной оптимизации при ограничениях. К сожалению, в настоящий момент не существует эффективных численных методов для решения подобных задач в случае, когда число ограничений m многократно превосходит число варьируемых параметров n . **Цель:** разработать новые методы и алгоритмы для решения задач многопараметрической оптимизации при ограничениях. **Результаты:** разработан новый метод нелинейного программирования, который назван «метод многомерных оболочек». Этот метод позволяет решать задачи условной многопараметрической оптимизации для случая $n \ll m$. Он использует как необходимые, так и достаточные условия оптимальности для целевых функций, имеющих вид метрик. Показано, что если единственный минимум (максимум) метрики находится вне области допустимых значений, образованной ограничениями, тогда решение задачи будет находиться на границе области допустимых решений. Тогда задача может иметь несколько решений, которые будут соответствовать одинаковым минимальным (максимальным) значениям целевой функции, и все эти решения могут быть получены с помощью метода многомерных оболочек. Основой данного метода является построение области допустимых решений путем анализа ограничений произвольного вида и их аппроксимации кусочно-полиномиальными ограничениями или рациональными дробями, что позволяет свести исходную задачу к набору взаимосвязанных подзадач с линейными ограничениями, формирующими выпуклые многомерные оболочки, что существенно сужает область поиска. **Практическая значимость:** предложенный метод нелинейного программирования апробирован и найдет применение при решении задач условной оптимизации в областях физической метрологии, оптимальных систем управления и логико-интервальных систем. Данный алгоритм может быть улучшен путем разработки более эффективных алгоритмов для новых типов ограничений.

Ключевые слова — нелинейное программирование, многопараметрическая оптимизация.

Введение

Многие задачи синтеза законов управления, моделирования сложных систем и идентификации объектов могут быть сформулированы как задачи условной оптимизации с ограничениями различных типов. Часто в качестве ограничений выступают уравнения, описывающие исследуемый объект. Они имеют кусочно-полиномиальный, кусочно-экспоненциальный, интервальный вид или являются рациональными дробями.

Такой подход имеет ряд преимуществ при более тонком анализе особенностей различных объектов. Например, в задачах фильтрации экспериментальных данных он позволяет наиболее эффективно разделять физические свойства испытываемого образца и влияние сторонних факторов стенда и внешней среды. А для задач управления и идентификации сочетание оптимизационного подхода и моделей, заданных в виде систем неравенств, может стать аксиоматической альтернативой нечетким системам и нейронным сетям.

Развитие этого направления сдерживает отсутствие эффективных численных методов для решения задач условной оптимизации для большого количества варьируемых переменных n , когда число ограничений m многократно превосходит n .

Цель данной статьи — рассказать о новых идеях в области методов для решения задач условной оптимизации без перехода к задаче безусловной оптимизации, в частности, о разработанном новом методе многомерных оболочек (Multidimensional Patchwork Shell Method).

Постановка задачи условной оптимизации

В математической постановке задача минимизации (максимизации) формулируется в следующем виде:

$$J = \min_x / \max_x / f(\mathbf{x}) \text{ при}$$

$$\mathbf{A}_1 \mathbf{x} = \mathbf{B}_1, \mathbf{A}_2 \mathbf{x} \leq \mathbf{B}_2, \mathbf{C}_1(\mathbf{x}) = \mathbf{0},$$

$$\mathbf{C}_2(\mathbf{x}) \leq \mathbf{0}, \mathbf{L}_1(\mathbf{x}) = \mathbf{0}, \mathbf{L}_2(\mathbf{x}) \leq \mathbf{0}, \quad (1)$$

где J — критерий; \mathbf{x} — вектор варьируемых переменных; f — целевая функция; $\mathbf{A}_1, \mathbf{B}_1, \mathbf{A}_2, \mathbf{B}_2$ — матрицы линейных ограничений, $\mathbf{C}_1, \mathbf{C}_2$ — гладкие матричные функции от соответствующих аргументов; $\mathbf{L}_1, \mathbf{L}_2$ — матричные функции с разрывами первого и второго рода от соответствующих аргументов.

Классический подход в решении задач условной оптимизации — это сведение ее к эквивалентной задаче безусловной оптимизации методом множителей Лагранжа [1] с дальнейшим применением разработанного инструментария

методов, например градиентного [2] или случайного [3] поисков.

Недостатки этого подхода заключаются в следующем.

1. Применение метода множителей Лагранжа корректно, если доказано существование седловой точки для модифицированной целевой функции.

2. Не доказана корректность метода множителей Лагранжа для общего случая нелинейных ограничений типа неравенство.

3. Общее число p варьируемых переменных модифицированной целевой функции многократно увеличивается: $p = n + m_p + 2m_n$, где n — число переменных в задаче условной оптимизации; m_p — количество ограничений типа равенство; m_n — количество ограничений типа неравенство. При этом область допустимых решений изменяется от $-\infty$ до $+\infty$ по всем переменным.

4. Большое количество ограничений разного типа может привести к появлению многих экстремумов, даже для случая монотонной исходной целевой функции f .

5. Ограничения с разрывами не позволяют использовать градиентные методы поиска.

Данные недостатки могут быть продемонстрированы на простом примере: необходимо найти минимум целевой функции

$$J = \min_x \{f\}, f = ax^2 + bx + c \text{ при } ax^2 + bx + c \geq d, \quad (2)$$

где J — функционал; f — целевая функция; x — варьируемая переменная; a, b, c, d — постоянные коэффициенты.

Решим эту задачу, анализируя ограничения и целевую функцию. Найдем интервалы аргумента, которые соответствуют ограничению в (2)

$$ax^2 + bx + c - d = 0;$$

$$x_{1,2} = \frac{-b \mp \sqrt{b^2 - 4a(c-d)}}{2a};$$

$$-\infty \leq x \leq +\infty, \text{ если } b^2 - 4a(c-d) < 0 \text{ и } a > 0;$$

$$x \in \emptyset, \text{ если } b^2 - 4a(c-d) < 0 \text{ и } a < 0;$$

$$x \leq x_1, x \geq x_2, \text{ если } b^2 - 4a(c-d) \geq 0 \text{ и } a > 0;$$

$$x_1 \leq x \leq x_2, \text{ если } b^2 - 4a(c-d) \geq 0 \text{ и } a < 0, \quad (3)$$

где $x_{1,2}$ — корни уравнения.

Исследуем целевую функцию на монотонность для всех интервалов аргумента, полученных в (3).

1. Для случая $b^2 - 4a(c-d) < 0$ и $a > 0$ целевая функция f монотонно убывает на интервале $x < -b/2a$, на интервале $x \geq -b/2a$ функция монотонно возрастает. Отсюда следует, что f имеет глобальный минимум в точке $x = -b/2a$.

2. Для случая $b^2 - 4a(c-d) < 0$ и $a < 0$ аргумент принадлежит пустому множеству, и задача не имеет решения.

3. Для случая $a > 0$ и $b^2 - 4a(c-d) \geq 0$ целевая функция f монотонно убывает на интервале $x \leq x_1$, на интервале $x \geq x_2$ функция монотонно возрастает. Отсюда следует, что f имеет два минимума с одинаковым значением f в точках x_1 и x_2 .

4. Для случая $a < 0$ и $b^2 - 4a(c-d) \geq 0$ целевая функция f монотонно возрастает на интервале $x_1 \leq x < -b/2a$, на интервале $-b/2a \leq x \leq x_2$ функция монотонно убывает. Отсюда следует, что f имеет два минимума с одинаковым значением в точках x_1 и x_2 .

Количество варьируемых переменных осталось неизменным, а область поиска была существенно сокращена для случаев 2, 3 и 4. В случае 1 было доказано, что ограничения можно не учитывать.

Для решения задачи (2) воспользуемся переходом от условной оптимизации к безусловной с применением метода множителей Лагранжа:

$$\Phi = \min_{x, \lambda, \xi} \{g\},$$

$$g(x, \lambda, \xi) = f + \lambda(-ax^2 - bx - c + d + \xi^2), \quad (4)$$

где Φ — критерий безусловной оптимизации; λ — множитель Лагранжа; ξ — свободный параметр для учета неравенства.

Если доказано существование седловой точки для g , то справедлива следующая система уравнений:

$$\begin{cases} \frac{\partial g}{\partial x} = 0 \\ \frac{\partial g}{\partial \lambda} = 0 \\ \frac{\partial g}{\partial \xi} = 0 \end{cases} \Rightarrow \begin{cases} 2ax + b - \lambda 2ax - \lambda b = 0 \\ -ax^2 - bx - c + d + \xi^2 = 0 \\ 2\lambda \xi = 0 \end{cases}$$

$$\Rightarrow \begin{cases} \lambda = 1 \\ x_{1,2} = \frac{-b \mp \sqrt{b^2 - 4a(c-d)}}{2a} \\ \xi = 0 \end{cases} \quad (5)$$

Полученное решение (5) не описывает случай 1, когда $b^2 - 4a(c-d) < 0$ и $a > 0$. Численное решение данной задачи градиентными методами даст только одно значение аргумента, а в случае $b^2 - 4a(c-d) < 0$ и $a < 0$ правильный ответ может быть получен только аналитически.

Количество варьируемых переменных увеличилось в 3 раза, область поиска равна R^3 .

Для использования метода множителей Лагранжа необходимо доказать существование седловой точки модифицированной целевой функции, что в случае большого количества переменных является сложной задачей, сравнимой по трудоемкости с нахождением глобального оптимума.

Данный пример показывает преимущества методов, основанных на анализе ограничений, по сравнению с методами, использующими переход от условной оптимизации к безусловной, а также демонстрирует нетождественность такого перехода.

Уменьшение числа переменных путем анализа линейных ограничений типа равенство

Ограничения типа равенство можно использовать для редукции числа варьируемых переменных и исключения ограничений типа равенство, которые дублируют друг друга.

Рассмотрим ограничения

$$A_1 x = B_1, \tag{6}$$

где матрица A_1 имеет размерности $m_p \times n$; вектор x имеет размерность $n \times 1$, а B_1 имеет размерности $m_p \times 1$. Здесь возможны три случая.

1. Если число линейно независимых уравнений больше n , то ограничения противоречивы и задача не имеет решения.

2. Если число линейно независимых уравнений равно n и найденное значение $x_0 = B_1 A_1^{-1}$ не противоречит другим ограничениям, то x_0 — это единственное решение задачи. Если x_0 противоречит другим ограничениям, то задача не имеет решения.

3. Если число линейно независимых уравнений n_1 меньше n , тогда необходимо определить базисные ограничения, а все линейно зависимые ограничения должны быть исключены. Полученная система уравнений позволяет определить линейную замену варьируемых переменных и сократить их число до $n - n_1$.

Для случая 3 замена переменных может быть выполнена по рекуррентной формуле

$$z_{n-n_1+k} = b_k - \sum_{i=1}^{n-n_1} a_{k,i} z_i - \sum_{j=1}^{k-1} a_{k,n-n_1+j} z_{n-n_1+j}, \tag{7}$$

где $a_{k,i}$ и b_k — коэффициенты матриц A_{1z} и B_{1z} соответственно. Для переменных z_i с порядковым номером, большим чем $n - n_1$, имеем

$$\begin{aligned} z_{n-n_1+1} &= b_1 - \sum_{i=1}^{n-n_1} a_{1,i} z_i; \\ z_{n-n_1+2} &= b_2 - \sum_{i=1}^{n-n_1} a_{2,i} z_i - a_{2,n-n_1+1} z_{n-n_1+1} = \\ &= b_2 - a_{2,n-n_1+1} b_1 - \sum_{i=1}^{n-n_1} [a_{2,i} + a_{2,n-n_1+1} a_{1,i}] z_i; \\ z_{n-n_1+3} &= b_3 - \sum_{i=1}^{n-n_1} a_{3,i} z_i - a_{3,n-n_1+1} z_{n-n_1+1} - \\ &\quad - a_{3,n-n_1+2} z_{n-n_1+2} = b_3 - a_{3,n-n_1+2} b_2 - \end{aligned}$$

$$\begin{aligned} &- [a_{3,n-n_1+1} + a_{3,n-n_1+2} a_{2,n-n_1+1}] b_1 - \\ &- \sum_{i=1}^{n-n_1} \left[\begin{aligned} &a_{3,i} + a_{3,n-n_1+2} a_{2,i} + \\ &+ [a_{3,n-n_1+1} + a_{3,n-n_1+2} a_{2,n-n_1+1}] a_{1,i} \end{aligned} \right] z_i \text{ и т. д.,} \end{aligned}$$

если в (6) матрица A_1 методом Гаусса и перестановками столбцов приведена к нижней треугольной A_{1z} вида

$$A_{1z} z = B_{1z}, \quad x = Tz, \tag{8}$$

$$A_{1z} = \begin{bmatrix} a_{1,1} & \dots & a_{1,n-n_1} & 1 & 0 & \dots & 0 \\ a_{2,1} & \dots & a_{2,n-n_1} & a_{2,n-n_1+1} & 1 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n_1,1} & \dots & a_{n_1,n-n_1} & a_{n_1,n-n_1+1} & a_{n_1,n-n_1+2} & \dots & 1 \end{bmatrix},$$

где z — вектор x с измененным порядком следования компонентов; B_{1z} — преобразованный вектор B_1 ; T — матрица перестановок. Учесть ограничения (6) можно линейной заменой

$$x = T \Lambda_1 z_{n-n_1} + T \beta_1, \tag{9}$$

где Λ_1 — матрица размерности $n \times (n - n_1)$, коэффициенты которой находятся согласно (7); z_{n-n_1} — новый вектор варьируемых переменных; β_1 — вектор размерности $n \times 1$, коэффициенты которого находятся согласно (7).

Построение области кандидатов в решения методом анализа линейных ограничений типа неравенство

Другой широко распространенный вид ограничений — это линейные неравенства, которые задаются выражением вида

$$A_2 x \leq B_2, \tag{10}$$

где матрица A_2 имеет размерности $m_n \times n$; вектор x имеет размерности $n \times 1$, а B_2 имеет размерности $m_n \times 1$.

Первый шаг: определить взаимоисключающие ограничения, что может быть выполнено по алгоритму 1.

1. Начало алгоритма, ввод исходных данных: матрица A_2 и вектор B_2 .
2. Цикл: изменять индекс i от 1 до $m_n - 1$ с шагом 1, пропустить номер из запретного списка: выполнять от п. 3 до п. 12.
3. Переменной флаг1 присвоить 0.
4. Цикл: изменять индекс j от $i + 1$ до m_n с шагом 1, пропустить номер из запретного списка: выполнять от п. 5 до п. 12.
5. Если коэффициенты выбранных неравенств $a_{i,1} \neq 0$ и $a_{j,1} \neq 0$ или выбранные неравенства имеют нулевые коэффициенты от 1 до $s - 1$, то рассчитать коэффициент $k = a_{i,1}/a_{j,1}$ или $k = a_{i,s}/a_{j,s}$, иначе перейти в п. 4.

6. Если для всех p от 1 до n хотя бы одно из $a_{i,p} - a_{j,p}k \neq 0$, то перейти в п. 4.
7. Если $b_i - b_jk = 0$, то исключить ограничение j , заменить ограничение i на $a_{i,1}x_1 + a_{i,2}x_2 + \dots + a_{i,n}x_n = b_i$, добавить в список исключенных ограничений индекс j , добавить в запретный список индекс j , переменной флаг1 присвоить 1, перейти в п. 4.
8. Если переменная флаг1 равна 1 и $b_i - b_jk \neq 0$, то противоречие в ограничениях, задача не имеет решения, добавить в список несовместных ограничений индексы i и j , перейти в п. 13.
9. Если $k > 0$ и $b_i - b_jk > 0$, то исключить ограничение j , добавить в список исключенных ограничений индекс j , добавить в запретный список индекс j , перейти в п. 4, не изменяя значения j .
10. Если $k > 0$ и $b_i - b_jk < 0$, то заменить ограничение i на j , добавить в список исключенных ограничений индекс i , добавить в запретный список индекс j , перейти в п. 4.
11. Если $k < 0$ и $b_i - b_jk > 0$, то параллельные ограничения, добавить в список параллельных ограничений индексы i и j , перейти в п. 4.
12. Если $k < 0$ и $b_i - b_jk < 0$, то противоречие в ограничениях, задача не имеет решения, добавить в список несовместных ограничений индексы i и j , перейти в п. 13.
13. Окончание алгоритма, выходные данные: список исключенных ограничений, список параллельных ограничений, список несовместных ограничений, задача имеет / не имеет решения, формируется новая система ограничений

$$\mathbf{A}_2^* \mathbf{x} \leq \mathbf{B}_2^*. \quad (11)$$

Второй шаг: определить границу области, формируемой ограничениями, или убедиться, что такой области не существует. Известные алгоритмы решения подобной задачи [4, 5] адаптируем для нашего случая.

Введем следующие определения и допущение.

1. Каждое ограничение типа неравенство из (11) имеет *граничную плоскость*, которая определяется уравнением $a_{i,1}x_1 + a_{i,2}x_2 + \dots + a_{i,n}x_n - b_i = 0$.

2. *Узлом (вершиной)* будем называть точку пересечения n граничных плоскостей, и эта точ-

ка должна удовлетворять всем ограничениям из (11); узел будем обозначать как $V_{i,j,\dots,s}$, где i, j, \dots, s — код узла, состоящий из n индексов и формируемый номерами ограничений в (11); $V_{i,j,\dots,s}$ — вектор координат узла с размерностью $n \times 1$; индексы упорядочиваются таким образом, чтобы $i < j < \dots < s$.

3. *Ребром* будем называть отрезок, соединяющий два узла с числом общих индексов, равным двум; ребро будем обозначать как $W_{i,j,\dots,s}^{i,j,\dots,s}$, где верхний и нижний индексы — коды узлов, которые соединяет ребро.

4. *Гранью* будем называть область граничной плоскости, которой принадлежат все узлы, для которых один из индексов кодов i является общим, и эта область сформирована ребрами, соединяющими эти узлы; грань будем обозначать как P_i , где i — это номер соответствующего ограничения из (11).

5. Все варьируемые переменные x изменяются в соответствующих интервалах $x_{\min,i} \leq x_i \leq x_{\max,i}$, что обуславливает реализуемость решения задачи; тогда следует, что система не имеет решения, когда нет ни одного узла; в случае одного узла система имеет единственное решение, если оно не противоречит ограничениям в (1); если есть два узла, то область кандидатов в решение — это ребро; если три узла, то область кандидатов в решение — это грань и т. д.

Центральным вопросом определения области кандидатов в решение является отыскание узлов для системы (11), что может быть выполнено по алгоритму 2.

1. Начало алгоритма, ввод исходных данных: матрица \mathbf{A}_2^* и вектор \mathbf{B}_2^* .
2. Цикл: пока не выбраны все возможные последовательности из n неповторяющихся индексов $i < j < \dots < s$ такие, что не могут включаться индексы двух взаимно параллельных граничных плоскостей одновременно, выполнять от п. 2 до п. 6.
3. Из выбранных различных n граничных плоскостей с индексами $i < j < \dots < s$ сформировать систему уравнений $\mathbf{A}_{i,j,\dots,s} \mathbf{x} = \mathbf{B}_{i,j,\dots,s}$.
4. Если $\det(\mathbf{A}_{i,j,\dots,s}) = 0$, то перейти в п. 1.
5. Вычислить $V_{i,j,\dots,s} = (\mathbf{A}_{i,j,\dots,s})^{-1} \mathbf{B}_{i,j,\dots,s}$.
6. Если ограничения (11) от $V_{i,j,\dots,s}$ выполняются, то добавить $V_{i,j,\dots,s}$ в список узлов, перейти в п. 2.
7. Если список узлов пустой, то задача не имеет решения.
8. Если число узлов равно 1, то задача имеет единственное решение.
9. Окончание алгоритма, выходные данные: список узлов, задача имеет / не имеет решения.

Если найдены узлы для (11), то исходя из определений 3 и 4 простыми средствами могут быть построены все ребра и грани, образующие область кандидатов в решения.

Анализ полиномиальных ограничений

Другой широко распространенный вид ограничений — это полиномиальные уравнения. Разработано много методов и алгоритмов для их решения [6–8]. Следует отметить, что полиномиальные аппроксимации могут быть использованы для описания с некоторой точностью более сложных нелинейных ограничений в выбранной области аргументов. Это позволяет сводить для некоторых областей значений аргумента задачу условной оптимизации с произвольными нелинейными ограничениями к задачам условной оптимизации с полиномиальными ограничениями при заданной точности решения.

Полиномиальные ограничения встречаются двух видов — полиномиальные равенства и неравенства, которые имеют различные методы в решении. Универсальным подходом в анализе ограничений и первого, и второго вида может считаться разложение на множители порядка не старше двух полинома ограничения φ :

$$\varphi(x_1, x_2, \dots, x_n) = \prod_{i_1} \left(\sum_{j_1} a_{1,i_1,j_1} x_{j_1} + b_{1,i_1} \right) \times \prod_{i_2} \left(\sum_{j_2} a_{2,i_2,j_2} x_{j_2}^2 + \sum_{j_3, j_4} a_{3,i_2,j_3,j_4} x_{j_2} x_{j_4} + \sum_{j_5} a_{4,i_2,j_5} x_{j_5} + b_{2,i_2} \right), \quad (12)$$

где $a_{1,i_1,j_1}, b_{1,i_1}, a_{2,i_2,j_2}, a_{3,i_2,j_3,j_4}, a_{4,i_2,j_5}, b_{2,i_2}$ — коэффициенты разложения на множители.

Видно, что ограничение с полиномом (12) сводится к системе линейных и квадратичных ограничений, что приводит к разделению исходной задачи на несколько подзадач, которые могут решаться независимо.

В общем случае разложение полиномиальных ограничений — это сложная проблема, исследования в этом направлении сейчас идут очень интенсивно.

Рассмотрим случай систем ограничений, приводимых к треугольному виду:

$$\begin{cases} v_1^{m_1} + \varphi_1(v_1, v_2, \dots, v_k) = 0 \\ v_2^{m_2} + \varphi_2(v_2, v_3, \dots, v_k) = 0 \\ \dots \\ v_k^{m_k} + \varphi_k(v_k) = 0 \end{cases}, \quad \mathbf{v} = \mathbf{T}_v \mathbf{x} + \mathbf{B}_v, \mathbf{v} = [v_1, v_2, \dots, v_k]^T, k \leq n, \quad (13)$$

где m_1, m_2, \dots, m_k — степени соответствующих переменных; k — количество уравнений; φ_i — полиномы от переменных v_i, \dots, v_k , причем показатели при переменной v_i в многочлене φ_i меньше, чем m_i ; $\mathbf{T}_v, \mathbf{B}_v$ — матрицы параметров.

В последнее уравнение в (13) входит только v_k , и существуют эффективные алгоритмы нахождения корней полинома одной переменной [8]. Рассматриваются только вещественные корни полинома уравнения k , которые могут быть подставлены в уравнение $k - 1$, которое после замены станет зависеть от переменной v_{k-1} , и его корни могут быть найдены аналогично уравнению k и т. д. Если все корни хотя бы одного из уравнений в (13) комплексные, то задача (1) не имеет решения. Множество вещественных корней системы (13) формирует наборы линейных ограничений вида

$$\mathbf{T}_v \mathbf{x} + \mathbf{B}_v = \Lambda_{v,i}, \quad (14)$$

где $\Lambda_{v,i}$ — i -й набор вещественных корней системы (13), и разделяет исходную задачу на несколько параллельно решаемых подзадач. Ограничения (14) анализируются аналогично (6).

Для ограничений типа неравенство рассмотрим случай

$$a_i v_i^{m_i} + \varphi_i(v_i) \leq 0, i = 1, 2, \dots, k, \quad \mathbf{v} = \mathbf{T}_v \mathbf{x} + \mathbf{B}_v, \mathbf{v} = [v_1, v_2, \dots, v_k]^T, \quad (15)$$

где количество неравенств может быть больше количества переменных, а каждый полином φ_i зависит только от одной переменной v_i . Найдя корни полинома, представим (15) как произведение слагаемых только с вещественными корнями λ_j :

$$a_i \prod_j (v_i - \lambda_j) \leq 0. \quad (16)$$

Из (16) находятся интервалы для v_i , которые преобразуются в набор систем линейных неравенств от \mathbf{x} , и необходимо рассмотреть все возможные комбинации этих систем для всех неравенств из (15) и решить их с использованием алгоритмов из предыдущего раздела.

Анализ ограничений с разрывами

Часто встречающийся на практике тип ограничений — это ограничения с разрывами. К ним относятся ограничения в виде кусочно-линейных функций, дробей, полиномов (например, гиперболы) и т. д. Наиболее удобным способом представления таких ограничений является описание в виде систем равенств или неравенств, заданных на интервалах, границами которых являются особые точки, например переход с одной ветви

функции на другую. Применим данный подход для дроби

$$\frac{a}{\prod_{i,j} (v_i - \lambda_j)} \leq 0, \quad (17)$$

где вещественные корни для каждой переменной v_i упорядочены согласно правилу $\lambda_1 < \lambda_2 < \dots < \lambda_p$, тогда (17) можно представить как системы неравенств, соответствующие интервалам λ . Для случая (17) при $a < 0$ необходимо, чтобы выполнялись условия $v_i - \lambda_j > 0$ для всех компонент или количество условий $v_i - \lambda_j < 0$ было четным для выбранных интервалов v_i . Анализ систем неравенств выполняется аналогично (6)–(9) и алгоритмам 1 и 2 при условии, что x , удовлетворяющие уравнениям $v_i - \lambda_j = 0$, не могут быть решениями. Если узел, ребро или грань удовлетворяет этим условиям, то в качестве кандидата в решения необходимо выбрать точку, линию или плоскость, лежащую в области кандидатов в решение и отстоящую на точность расчета ε .

Разбиение на интервалы приводит к возникновению нескольких задач, которые могут решаться независимо, и все их возможные комбинации должны быть рассмотрены. Наиболее простым способом описания подобного разбиения является представление в виде логической структуры, например конечного автомата [9].

Анализ ограничений с разрывами должен выполняться после всех других типов ограничений, так как в этом случае повышается вероятность исключения подобных ограничений или областей аргументов, которые соответствуют разрывам.

Анализ произвольных ограничений

Каждый тип ограничений требует разработки соответствующего алгоритма анализа. Наиболее простым способом решения этой задачи является аппроксимация произвольного нелинейного ограничения кусочно-полиномиальной моделью при заданной точности расчета ε . В этом случае можно воспользоваться подходом, описанным в двух предыдущих разделах.

При синтезе оптимальных систем управления в качестве ограничений используются уравнения объекта управления, компонентами которых являются экспоненциальные функции. Подобный пример подробно описан в работе [10], а ограничения имеют вид

$$x_i e^{-b_i x_1} - a_{1,i} e^{-x_1 x_i} - a_{2,i} x_i + a_{3,i} = 0; \quad (18)$$

$$a_j e^{-b_j x_j} - x_j e^{-x_i x_j} = 0; \quad (19)$$

$$\frac{e^{-b_i x_i} - e^{-x_i x_j}}{\frac{a_{1,j}}{x_i} - a_{2,j}} \leq c_i, \quad (20)$$

где $b_i, a_{1,i}, a_{2,i}, a_{3,i}, a_j, b_j, a_{1,j}, a_{2,j}, c_i$ — постоянные коэффициенты, $i \neq j$.

Нетрудно заметить в (18), что переменные x_i могут быть выражены через x_1 , для этого введем замену $y_i = x_1 x_i$, тогда (18) примет вид

$$e^{-y_i} = -\mu_i y_i + \sigma_i, \quad \mu_i(x_1) = \frac{a_{2,i} - e^{-b_i x_1}}{a_{1,i} x_1},$$

$$\sigma_i = \frac{a_{3,i}}{a_{1,i}}. \quad (21)$$

Уравнение (21) имеет два одинаковых решения в случае

$$\frac{\partial e^{-y_i}}{\partial y_i} = -e^{-y_i} = \frac{\partial(-\mu_i y_i + \sigma_i)}{\partial y_i} = -\mu_i, \quad (22)$$

что позволяет уточнить нижнюю границу для σ_i :

$$y_i = -\ln \mu_i, \quad \mu_i = \mu_i \ln \mu_i + \sigma_i, \quad \sigma_i = \mu_i - \mu_i \ln \mu_i. \quad (23)$$

Учтем (23) в уравнении (21):

$$e^{-y_i} = -\mu_i y_i + \sigma_i, \quad \mu_i > 0, \quad \mu_i - \mu_i \ln \mu_i - \sigma_i < 0. \quad (24)$$

Перегруппировав (24): $-\frac{1}{\mu_i} = \left(y_i - \frac{\sigma_i}{\mu_i}\right) e^{y_i}$, и введя замену $z_i = y_i - \frac{\sigma_i}{\mu_i}$, получим

$$z_i e^{z_i} = -\frac{1}{\mu_i} e^{\frac{\sigma_i}{\mu_i}}. \quad (25)$$

Решением (25) является функция Ламберта ψ^{-1} :

$$z_i = \psi^{-1} \left(-\frac{1}{\mu_i} e^{\frac{\sigma_i}{\mu_i}} \right),$$

$$y_i = \frac{\sigma_i}{\mu_i} + z_i = \frac{\sigma_i}{\mu_i} + \psi^{-1} \left(-\frac{1}{\mu_i} e^{\frac{\sigma_i}{\mu_i}} \right). \quad (26)$$

Тогда переменные x_i находятся из соотношения

$$x_i = \frac{\sigma_i}{\mu_i(x_1) x_1} + \frac{1}{x_1} \psi^{-1} \left(-\frac{1}{\mu_i(x_1)} e^{\frac{\sigma_i}{\mu_i(x_1)}} \right) = g_i(x_1),$$

$$\mu_i(x_1) > 0, \quad \mu_i(x_1) - \mu_i(x_1) \ln \mu_i(x_1) - \sigma_i < 0. \quad (27)$$

Нетрудно заметить, что переменные x_j могут быть выражены через $g(x_1)$ с использованием ограничений (19):

$$\frac{a_j}{g_i(x_1)} = e^{(b_j - g_i(x_1)) x_j} \Rightarrow$$

$$\Rightarrow x_j = \frac{1}{b_j - g_i(x_1)} \ln \left(\frac{a_j}{g_i(x_1)} \right) = h_i(x_1), \quad (28)$$

$$\frac{a_j}{g_i(x_1)} > 0.$$

Тогда с учетом (28) ограничения (20) будут зависеть только от x_1 и примут вид

$$\frac{e^{-b_i g_i(x_1)} - e^{-g_i(x_1) h_i(x_1)}}{\frac{a_{1,j}}{g_i(x_1)} - a_{2,j}} \leq c_i. \quad (29)$$

Если учесть выражения (21)–(29), то ограничения (18)–(20) сводятся к эквивалентной системе ограничений от переменной x_1 :

$$\begin{aligned} \mu_i(x_1) > 0, \mu_i(x_1) - \mu_i(x_1) \ln \mu_i(x_1) - \sigma_i < 0, \\ x_1 \neq 0, g_i(x_1) \neq 0, b_j - g_i(x_1) \neq 0, \\ \frac{e^{-b_i g_i(x_1)} - e^{-g_i(x_1) h_i(x_1)}}{\frac{a_{1,j}}{g_i(x_1)} - a_{2,j}} \leq c_i, \frac{a_j}{g_i(x_1)} > 0. \end{aligned} \quad (30)$$

Нелинейные функции в (30) могут быть аппроксимированы с помощью кусочно-полиномиальной модели от одной переменной x_1 и проанализированы, как это рассмотрено выше.

Анализ целевой функции

После построения границы области допустимых решений и определения ее узлов путем анализа ограничений необходимо определить оптимальные значения целевой функции, которые принадлежат области допустимых решений. На основании примера, описанного в первом разделе, сформулируем общий вывод: если решается задача условной минимизации (максимизации), целевая функция гладкая и ее экстремумы могут быть определены аналитически, тогда возможны три варианта.

1. Если глобальный минимум (максимум) находится в области допустимых значений, образованной ограничениями, тогда аргумент, соответствующий этому минимуму (максимуму), является решением поставленной задачи.

2. Если область допустимых решений — это пустое множество, тогда поставленная задача не имеет решения.

3. Если глобальный минимум (максимум) находится вне области допустимых значений, образованной ограничениями, тогда решение задачи будет находиться на границе области допустимых решений или это будут аргументы локальных минимумов (максимумов), которые находятся внутри области допустимых решений. В данном случае задача может иметь несколько решений, которые будут соответствовать одинаковым минимальным (максимальным) значениям целевой функции.

Обычно целевая функция имеет простой вид и ее оптимумы известны, поэтому проверка вариантов 1 и 2 не вызывает трудностей. Для про-

извольной целевой функции проверка варианта 3 очень трудоемка даже на небольшой области допустимых решений и может быть проведена стандартными методами поиска, например методом сеток, методом управляемого случайного поиска.

Рассмотрим случай, когда целевая функция является метрикой; имеет не более одного экстремума; все точки, в которых эта функция имеет одинаковое значение, ограничивают фигуру таким образом, что отрезок, соединяющий любые две точки ее поверхности, принадлежит этой фигуре; выполняются следующие соотношения.

Если определена функция ρ такая, что $\rho = \rho(\mathbf{x})$, и существуют такие точки $\mathbf{x}(i)$, в которых $\rho(\mathbf{x}(i))$ имеют одинаковые значения, то должны выполняться следующие условия.

Если ищется минимум функции f и эта функция имеет максимум в точке, где $\rho = \rho_0$, то

$$\begin{aligned} f(\rho_0) = \max, f(\rho_0) > f(\rho_1) > \dots > f(\rho_n), \\ \rho_0 < \rho_1 < \dots < \rho_n; \end{aligned} \quad (31a)$$

если ищется максимум функции f и эта функция имеет минимум в точке, где $\rho = \rho_0$, то

$$\begin{aligned} f(\rho_0) = \min, f(\rho_0) < f(\rho_1) < \dots < f(\rho_n), \\ \rho_0 > \rho_1 > \dots > \rho_n. \end{aligned} \quad (31b)$$

Тогда решением будут узлы, в которых целевая функция имеет максимальное/минимальное значение для задачи максимизации/минимизации.

Если ищется минимум функции f и эта функция имеет минимум в точке, где $\rho = \rho_0$, то

$$\begin{aligned} f(\rho_0) = \min, f(\rho_0) < f(\rho_1) < \dots < f(\rho_n), \\ \rho_0 < \rho_1 < \dots < \rho_n; \end{aligned} \quad (32a)$$

если ищется максимум функции f и эта функция имеет максимум в точке, где $\rho = \rho_0$, то

$$\begin{aligned} f(\rho_0) = \max, f(\rho_0) > f(\rho_1) > \dots > f(\rho_n), \\ \rho_0 > \rho_1 > \dots > \rho_n. \end{aligned} \quad (32b)$$

Тогда решением будут узлы или точки граней, если эти грани являются касательными к поверхности ρ в этих точках, где целевая функция имеет максимальное/минимальное значение для задачи максимизации/минимизации.

Алгоритм оптимизации

Укрупненно алгоритм оптимизации может быть описан следующим образом.

1. Начало алгоритма, ввод данных об ограничениях и целевой функции.

2. Если выделены линейные равенства, то выполнить анализ линейных неравенств согласно (6)–(9).
3. Если не найдены противоречия в ограничениях в п. 2 и выделены линейные неравенства, то выполнить анализ линейных неравенств согласно алгоритмам 1 и 2.
4. Если не найдены противоречия в ограничениях в п. 3 и выделены полиномиальные ограничения, то выполнить анализ полиномиальных ограничений согласно (12)–(16).
5. Если не найдены противоречия в ограничениях в п. 4 и выделены ограничения с разрывами, то выполнить анализ полиномиальных ограничений согласно разделу «Анализ ограничений с разрывами».
6. Если не найдены противоречия в ограничениях в п. 5 и выделены произвольные ограничения, то выполнить анализ ограничений согласно разделу «Анализ произвольных ограничений».
7. Если найдены противоречия в ограничениях, то задача не имеет решения, перейти в п. 12.
8. Найти все узлы области допустимых решений.
9. Вычислить значения целевой функции в узлах – найти ее максимальные/минимальные значения.
10. Если целевая функция удовлетворяет условиям (31a) и (31b), то исходя из задачи максимизации/минимизации выбрать узлы с максимальными/минимальными значениями, полученными в п. 9, перейти в п. 12.
11. Если целевая функция удовлетворяет условиям (32a) и (32b), то исходя из задачи максимизации/минимизации выбрать узлы с максимальными/минимальными значениями, полученными в п. 9, проверить, являются ли выбранные узлы точками касания поверхности p граней области допустимых значений, вычислить точки касаний граней, образующих узел с максимальным/минимальным значением, выбрать точки с максимальным/минимальным значением целевой функции.
12. Окончание алгоритма, выходные данные: список решений, задача имеет / не имеет решения.

Заключение

В статье рассмотрены вопросы решения задач условной оптимизации большой размерности, которые часто встречаются в областях науки и техники, таких как теория управления, распознавание образов, моделирование и т. д.

Получены следующие результаты.

1. Показано, что переход от условной оптимизации к безусловной в случае ограничений типа неравенство не тождественен.

2. Сформулированы необходимые и достаточные условия оптимальности для задач условной оптимизации с целевыми функциями в виде метрики, имеющими единственный оптимум.

3. Рассмотрен вопрос построения областей допустимых решений для ограничений произвольного вида путем их аппроксимации кусочно-полиномиальными функциями или рациональными дробями, что позволяет свести исходную задачу к набору взаимосвязанных подзадач с линейными ограничениями, формирующими выпуклые многомерные оболочки, что существенно сужает область поиска.

4. Разработан новый метод нелинейного программирования, позволяющий решать задачи высокой размерности с большим количеством ограничений различных типов. Предложенный метод при нахождении оптимальных значений не использует операцию дифференцирования, что позволяет применять его при анализе стохастических и интервальных данных.

Литература

1. **Tabak D., Kuo B. C.** Optimal Control by Mathematical Programming. — Englewood Cliffs, New Jersey: Prentice-Hall, 1971. — 280 p.
2. **Zangwill W. I.** Nonlinear Programming. A Unified Approach. — Englewood Cliffs, New Jersey: Prentice-Hall, 1969. — 312 p.
3. **Растрюгин Л. А.** Случайный поиск в задачах оптимизации многопараметрических систем. — Рига: Зинатне, 1965. — 212 с.
4. **Motzkin T., Raiffa H., Thompson G., Thrall R. M.** The Double Description Method // Contributions to the Theory of Games. — Princeton: Princeton University Press, 1953. — P. 51–73.
5. **Шевченко В. Н., Груздев Д. В.** Модификация алгоритма Фурье — Моцкина для построения триангуляции // Дискретный анализ и исследование операций. 2003. Т. 10. № 1. С. 53–64.
6. **Computer Algebra: Symbolic and Algebraic Computation** / Ed. by B. Buchberger, G. Collins, R. Loos. — N. Y.: Springer-Verlag Wien, 1983. — 291 p.
7. **Davenport J., Siret Y., Tournier E.** Computer Algebra: Systems and Algorithms for Algebraic Computation. — London, Academic Press, 1993. — 313 p.
8. **Cox D., Little J., O’Shea D.** Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra. — Springer, 2006. — 551 p.
9. **Дубаренко В. В.** Принципы логического управления динамическими объектами // Информационно-управляющие системы. 2003. № 5(6). С. 2–11.
10. **Артеменко Ю. Н., Агапов В. А., Дубаренко В. В., Кучмин А. Ю.** Групповое управление актуаторами контррефлектора радиотелескопа // Информационно-управляющие системы. 2012. № 4(59). С. 2–9.

UDC 681.5

doi:10.15217/issn1684-8853.2016.2.2

A Nonlinear Programming Method with Arbitrary Restrictions

Kuchmin A. Yu.^a, PhD, Tech., Senior Researcher, radiotelescope@yandex.ru

^aInstitute of Problems of Mechanical Engineering of RAS, 61, Bol'shoi Pr. V. O., 199178, Saint-Petersburg, Russian Federation

Introduction: Modeling and identification of nonlinear and non-stationary systems, as well as the analysis and filtration of experimental data often require the use of optimization methods, in particular, nonlinear programming. These problems belong to the class of problems of multidimensional optimization with restrictions. Unfortunately, at the moment there are no efficient numerical methods to solve such problems for the case when the number of restrictions m exceeds manifold the number of the varied parameters n . **Purpose:** The goal is to develop new methods and algorithms for multiple-parameter optimization with restrictions. **Results:** A new nonlinear programming method is developed which is called "Multidimensional Patchwork Shell Method" (MPSM). It allows you to solve problems of conditional multiple-parameter optimization for the case of $n \ll m$. The method uses both necessary and sufficient optimality conditions for metric-like criterion functions. It has been shown that if the only minimum (maximum) of a metric is out of the area of admissible values formed by the restrictions, the solution of the problem will be found on the border of the area of admissible decisions. In this case, the problem can have several solutions which will correspond to identical minimum (maximum) values of the criterion function, and all these solutions can be obtained by MPSM. The basis of MPSM is creating an area of admissible decisions by analyzing restrictions of any type and approximating them with piecewise and polynomial restrictions or with rational fractions. This reduces the initial problem to a set of interconnected subtasks with the linear restrictions forming multidimensional shells, significantly narrowing the search area. **Practical relevance:** The proposed method of nonlinear programming has been tested. It will find application in conditional optimization in such fields as physical metrology, optimal control systems and logical-interval systems. This algorithm can be improved by developing more efficient algorithms for new types of restrictions.

Keywords — Nonlinear Programming, Multiple-Parameter Optimization.

References

1. Tabak D., Kuo B. C. *Optimal Control by Mathematical Programming*. Englewood Cliffs, New Jersey, Prentice-Hall, 1971. 280 p.
2. Zangwill W. I. *Nonlinear Programming. A Unified Approach*. Englewood Cliffs, New Jersey, Prentice-Hall, 1969. 312 p.
3. Rastrigin L. A. *Sluchainyi poisk v zadachakh optimizatsii mnogoparametricheskikh sistem* [Random Search in Problems of Optimization of Multiple Parameter Systems]. Riga, Zinatne Publ., 1965. 212 p. (In Russian).
4. Motzkin T., Raiffa H., Thompson G., Thrall R. M. The Double Description Method. In: *Contributions to the Theory of Games*. Princeton, Princeton University Press, 1953. Pp. 51–73.
5. Shevchenko V. N., Gruzdev D. V. Modification of Algorithm of Fourier — Motzkin for Creation of a Triangulation. *Diskretnyi analiz i issledovanie operatsii*, 2003, vol. 10, no. 1, pp. 53–64 (In Russian).
6. *Computer Algebra: Symbolic and Algebraic Computation*. Ed. by B. Buchberger, G. Collins, R. Loos. New York, Springer-Verlag Wien, 1983. 291 p.
7. Davenport J., Siret Y., Tournier E. *Computer Algebra: Systems and Algorithms for Algebraic Computation*. London, Academic Press, 1993. 313 p.
8. Cox D., Little J., O'Shea D. *Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra*. Springer, 2006. 551 p.
9. Dubarenko V. V. Principles of Logic Automatic Control by Dynamic Objects. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2003, no. 5(6), pp. 2–11 (In Russian).
10. Artemenko Yu. N., Agapov V. A., Dubarenko V. V., Kuchmin A. Yu. Co-operative Control of Subdish Actuators of Radio-Telescope. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2012, no. 4(59), pp. 2–9 (In Russian).