

АЛГОРИТМЫ ОПТИМИЗАЦИИ И ФУНКЦИИ ЛЯПУНОВА

(Представил Г. Вайникко)

Нахождение подходящего начального приближения для итерационного процесса оптимизации часто является сложной задачей. Поэтому разработка и изучение итерационных методов, предъявляющих более слабые ограничения на выбор начального приближения, приобретают особое значение.

В данной работе приводится краткий обзор по исследованию дифференциальных методов (наискорейшего спуска, материальной точки и их модификаций) минимизации функционалов [1-4]. Приводятся теоремы об устойчивости — сходимости итерационных методов, основанных на втором методе Ляпунова. Такой подход позволил получить качественную информацию о поведении процесса поиска оптимума и доказать сходимость итерационных методов при любых начальных условиях [2, 5-8].

При этом, применяя к непрерывным аналогам те или иные методы численного интегрирования дифференциальных уравнений, можно прийти к новым итерационным процессам. Одной из особенностей этого подхода является возможность изучить задачи управления алгоритмами, представляющими собой дифференциальные уравнения с разрывной нелинейной правой частью [2, 4, 5, 7, 9].

Применение методов теории устойчивости движения (второго метода Ляпунова) к проблемам нелинейного программирования осуществляется с 1967 г. ([2] с. 235). Оно основано на построении скалярной функции Ляпунова $V(x)$ со следующими свойствами ([5] с. 17):

- 1) $V(X) \in C_1$;
- 2) $V(X) > 0$ для всех $x \in E_n$;
- 3) $V(X) = 0$ тогда и только тогда, когда $x = 0$;
- 4) (а) $\dot{V}(X) = \nabla V \dot{x} \leq 0$ для всех $x \in E_n$;
- (б) $\dot{V}(X) = \nabla V \dot{x} < 0$ для всех $x \in E_n$, кроме $x = 0$.

Свойства 1)–3) не связаны с динамикой изучаемого алгоритма, 4) связано уже с процессом, описываемым алгоритмом. Функция Ляпунова со свойствами 1)–4) (а), если она построена, дает достаточное условие устойчивости — сходимости алгоритма к оптимальной точке — по Ляпунову. Функция Ляпунова со свойствами 1)–4) (б) дает условие асимптотической устойчивости по Ляпунову ([5] с. 20).

Свойство функции Ляпунова

- 5) быть бесконечно большой ([5] с. 19–20), т. е. для любого числа $A > 0$ найдется сфера радиуса R , вне которой будет $V(x) > A$,

дает условие асимптотической устойчивости в целом относительно процесса поиска или при любых начальных условиях поиска.

Ранее ([2] с. 236) доказана лемма 10.1, устанавливающая связь между выпуклыми функциями и бесконечно большими функциями Ляпунова, которая позволила для нескольких алгоритмов поиска минимума функционала построить функции Ляпунова со свойствами 1)–5) ([2] с. 241).

Сходимость к оптимуму целевой функции алгоритмов поиска эквивалентна асимптотической устойчивости в целом относительно положения равновесия соответствующих систем дифференциальных уравнений, описывающих процесс поиска. Кроме того, построенные функции Ляпунова для этих алгоритмов позволили получить модификации алгоритмов градиентного типа и «тяжелого шарика», или материальной точки, путем введения управления процессом поиска оптимума [2, 4]. Предложенные управления превращают алгоритмы в релейные системы со скользящим режимом [7], т. е. с разрывной правой частью. Целесообразность введения скользящего режима объясняется, во-первых, тем, что переходный процесс отличается большим быстрым действием, во-вторых, системы становятся «грубыми» (робастными) по отношению к погрешностям счета и помехоустойчивыми. Перейдем к обзору функций Ляпунова для алгоритмов поиска.

1. Алгоритм наискорейшего спуска ([2] с. 235)

$$\dot{x} = -A \nabla F(x), \quad (1)$$

где $F(x)$ — целевая функция, $x \in E_n$, A — положительно-определенная матрица, $\nabla F(x)$ — градиент функции $F(x)$.

Этот алгоритм и его модификации исследованы в [2] (с. 235—281) для выпуклых функций из класса C_1 и C_2 , в [3] исследованы для кусочно-гладких функций.

Построенная функция Ляпунова имеет вид:

$$V_1(x) = A[F(x) - F(x^*)], \quad (2)$$

где x^* — точка из множества точек минимума функции $F(x)$. Доказана теорема об асимптотической устойчивости в целом относительно положения равновесия системы дифференциальных уравнений (1), т. е. о сходимости метода наискорейшего спуска при любых начальных условиях. Кроме того, на основе построенной функции Ляпунова предложены в [2] (с. 235) управления наискорейшим спуском, в результате чего получены модификации алгоритма вида:

$$\dot{x} = -K \nabla F(x) + b(x)U(x), \quad (3)$$

где K — положительная постоянная, $b(x)$ — вектор-функция, выбираемая в зависимости от управляющей скалярной функции $U(x)$, которая предложена в виде релейной функции типа:

$$a) \quad U(x) = -\text{sign } S(x) = \begin{cases} -1, & S(x) \geq 0, \\ 1, & S(x) < 0, \end{cases}$$

где $S(x) = (c, \nabla F(x))$, c — произвольный вектор (рекомендации по выбору его даны в [2] с. 281);

$$b) \quad U(x) = -\text{sign } S(x) = \begin{cases} -1, & S(x) > 0, \\ 0, & S(x) = 0, \\ 1, & S(x) < 0. \end{cases}$$

Например, вектор-функция $b(x)$ может иметь вид либо $b(x) = M \nabla S(x)$, M — положительная постоянная, либо $b(x) = M[\nabla F(x+c) - \nabla F(x)]$, $\|c\| \rightarrow \infty$. С такими управлениями система (3) является системой с переменной структурой [7], в которой на поверхности $S(x) = 0$ возникает скользящий режим. Введение скользящего режима улучшает характеристики алгоритма поиска, т. е. делает робастной системой по отношению к выбору шага интегрирования, к изменению начальных условий и возмущений [2, 7].

Один из алгоритмов, основанных на этих идеях ([²] с. 241, 252), имеет вид:

$$\dot{x} = -K \nabla F(x) - M \nabla S(x) \operatorname{sign} S(x), \quad (4)$$

где коэффициент M демпфируется при приближении к множеству точек минимума оптимизируемой функции $F(x)$.

Теорема. Пусть выпуклая функция $F(x) \in C_2$ задана на E_n и имеет минимум в точке $x^* \in D$, где множество

$$D = \{x \in E_n \mid \nabla F(x) + \alpha \nabla S(x) = 0, S(x) = 0\}$$

(α — произвольная постоянная) не содержит точек, отличных от $x = x^*$.

Тогда положение равновесия $x = x^*$ системы (5) асимптотически устойчиво в целом.

Доказательство теоремы проведено в [²] (с. 246) с помощью построения функции Ляпунова

$$V_2(x) = F(x) - F(x^*) + M |S(x)|. \quad (5)$$

2. Алгоритм, предложенный Я. Альбер и С. Альбер ([^{1,2}] с. 269)

$$\dot{x} = -F(x) \frac{\nabla F(x)}{\|\nabla F(x)\|^2}, \quad (6)$$

рассмотрен с произвольной постоянной положительной матрицей A , т. е. в более общем виде:

$$\dot{x} = -AF(x) \frac{\nabla F(x)}{\|\nabla F(x)\|^2}, \quad (7)$$

где $F(x) = |f(x)|^2$, $F(x^*) = 0$.

Теорема. Пусть выпуклая функция $F(x) \in C_1$, A — постоянная положительная матрица.

Тогда положение равновесия системы (7) является асимптотически устойчивым в целом.

Доказательство. Рассмотрим функцию

$$V_3(x) = A |F(x)|. \quad (8)$$

Функция (8) является бесконечно большой в силу леммы 10.1 ([²] с. 236) и удовлетворяет всем условиям 1)–5) функции Ляпунова, так как в силу системы (7) вдоль траекторий ее имеем

$$\begin{aligned} \dot{V}_3(x) &= A \nabla F(x) \operatorname{sign} F(x) \dot{x} = \\ &= -A^2 \nabla F(x) \operatorname{sign} F(x) F(x) \frac{\nabla F(x)}{\|\nabla F(x)\|^2} = -A^2 |F(x)| \leq 0. \end{aligned}$$

Производная $V_3(x)$ обращается в нуль только на множестве решений, т. е. $F(x) = 0$, что и преследуется поиском по алгоритму (7). В силу теоремы 12.2 из [⁵] (с. 46) функция $V_3(x)$ является функцией Ляпунова, удовлетворяющей свойствам 1)–5).

Следовательно, положение равновесия системы (7) асимптотически устойчиво в целом.

3. Рассмотрим двухшаговый алгоритм оптимизации, т. е. алгоритм метода «тяжелого шарика» [⁴]

$$\ddot{x} = -\alpha_1 \dot{x} - \alpha_2 \nabla F(x), \quad (9)$$

где α_1 и α_2 — положительные постоянные.

Дифференциальное уравнение (8) эквивалентно системе дифференциальных уравнений первого порядка

$$\begin{cases} \dot{x} = y, \\ \dot{y} = -\alpha_1 y - \alpha_2 \nabla F(x). \end{cases} \quad (10)$$

Вопрос о сходимости при любых начальных условиях алгоритма поиска минимума функции $F(x)$ сводится к доказательству асимптотической устойчивости в целом относительно соответствующего положения равновесия системы (10).

Теорема. Пусть $F(x)$ — выпуклая функция из класса C_1 , $\alpha_1 > 0$, $\alpha_2 > 0$. Тогда множество

$$M = \{(x^*, 0) \in E_{2n} \mid \nabla F(x^*) = 0\}$$

является асимптотически устойчивым в целом множеством состояний равновесия системы (10).

Доказательство. Функция, предложенная в [4]

$$V_4(x, y) = \frac{1}{2} \|y\|^2 + \alpha_2 [F(x) - F(x^*)], \quad (11)$$

является бесконечно большой в силу условий теоремы и построения $V_4(x, y)$.

Полная производная по времени в силу системы (10) имеет вид:

$$\dot{V}_4 = (x, y) = (y, -\alpha_1 y - \alpha_2 \nabla F(x)) + (c, \nabla F(x), y) = -\alpha_1 \|y\|^2 \leq 0$$

является знакоотрицательной функцией и обращается в нуль лишь на множестве $\{y=0\}$, которое не содержит целых полутраекторий системы (10), за исключением точек множества M . В силу теоремы 12.3 из [5] с. 47 и [8] можно утверждать, что множество M является асимптотически устойчивым в целом множеством состояний равновесия системы (10) и функция $V_4(x, y)$ является для нее функцией Ляпунова. Таким образом, доказана сходимость алгоритма (9) при любых начальных условиях.

4. Алгоритм, являющийся модификацией алгоритма (10), с управлением вида

$$U_1 = -Kc \operatorname{sign} S(x), \quad U_2 = 0, \quad (12)$$

где $S(x) = (c, \nabla F(x))$, K — положительная постоянная, исследован в [4], где приведены и другие виды управления, выбираемые из условия, обеспечивающего наибольшую скорость убывания функции Ляпунова $V_4(x, y)$. Существование функции Ляпунова обеспечивает асимптотическую устойчивость в целом относительно положения равновесия системы (10) с вводимым в нее управлением $U(x)$.

Так, при управлении (12) функция Ляпунова (11) имеет знакоотрицательную полную производную по времени

$$\dot{V}_4(x, y) = -\alpha_1 \|y\|^2 - K\alpha_2 |S(x)| \leq 0,$$

в которой равенство нулю достигается на множестве $M = \{(x, y) \in E_{2n} \mid \alpha_1 \|y\|^2 + K\alpha_2 |S(x)| = 0\}$, не содержащем целых полутраекторий системы (10) с управлением (12).

Итак, прямой метод Ляпунова — один из основных методов исследования устойчивости систем автоматического регулирования, позволяет исследовать условия сходимости при любых начальных условиях алгоритмов поиска в нелинейном программировании и выбирать управления, улучшающие алгоритмы поиска.

ЛИТЕРАТУРА

1. Альбер С. И., Альбер Я. И. // Ж. выч. математики и матем. физики, 1967, 7, № 1, 14—32.
2. Геращенко Е. И., Геращенко С. М. Метод разделения движений и оптимизация нелинейных систем. М., Наука, 1975.
3. Геращенко С. М. // Уч. зап. СГПИ. Алгебра и анализ, 1974, № 229, 166—170.
4. Геращенко С. М. // Доклады и сообщения V симпозиума «Computers in Chemical Engineering». Высокие Татры, 1977.
5. Барбашии Е. А. Введение в теорию устойчивости. М., Наука, 1967.
6. Ваарманн О. М. // Доклады и сообщения II симпозиума по методам решения нелинейных уравнений и задач оптимизации, I. Таллин, Валгус, 1981, 18—26.
7. Уткин В. И. Скользящие режимы и их применения в системах с переменной структурой. М., Наука, 1974.
8. La-Salle, J. P. // J. Diff. Equat., 1968, 4, № 1.
9. Геращенко С. М., Тяхт Т. К. // Тр. ТПИ, 1987, вып. 645, 153—159.

Институт кибернетики
Академии наук Эстонской ССР

Поступила в редакцию
1/VII 1988

Таллинский политехнический институт

O. VAARMANN, Svetlana GERASTSENKO-TONISMÄGI

OPTIMEERIMISLOGARITMID JA LJAPUNOVI FUNKTSIOONID

On antud lühiülevaade autorite töödest, mis on ilmunud alates 1967. aastast Ljapunovi funktsioonide kohta ning esitatud iteratsioonimeetodite koonduvuse ja stabiilsuse teoreemid. Viimased baseeruvad Ljapunovi teisel meetodil ja neid kasutatakse siin mitme muutujaga funktsiooni minimeerimisprotsessi kirjeldavate harilike diferentsiaalvõrrandite asümptootilise stabiilsuse näitamiseks. Funktsioonide klassi $F(x) \in C_1$ tarvis on konstrueeritud neli Ljapunovi funktsiooni.

O. VAARMANN and Svetlana GERASHCHENKO-TONISMÄGI

ALGORITHMS FOR OPTIMIZATION AND LYAPUNOV FUNCTIONS

A brief review is given of the results obtained by the authors since 1967 concerning Lyapunov functions. The latter are constructed to verify convergence stability and controllability of algorithms for optimization. Their convergence is equivalent to the global asymptotic stability of the corresponding equilibrium state for ordinary differential equations describing the process of minimizing a function in several variables. The second Lyapunov method is used to demonstrate the global asymptotic stability of some algorithms for optimization. Four Lyapunov functions are constructed for a class of functions $F(x) \in C_1$.