ФОРМАЛИЗАЦИЯ ДИНАМИКИ ОПТИМИЗАЦИОННОЙ ПРОЦЕДУРЫ

УДК 004:383.4:371.69

СОКОЛОВ Андрей Евгеньевич
к.т.н., доцент кафедры информационных технологий
Херсонского национального технического университета.
Научные интересы: компьютеризированные
системы обучения, модели движения информации.

РАСТОРГУЕВ Вячеслав Сергеевич
Аспирант кафедры Экономической кибернетики и управления проектами
Научные интересы: компьютеризированные системы обучения,
модели движения информации.

ВВЕДЕНИЕ

Современные информационные технологии по существу основаны на использовании оптимизационных процедур, обеспечивающих высокую эффективность их реализации [1]. Решаемые реальные задачи могут быть сформулированы как поиск оптимального значения, когда нужно найти значение сложной функции, зависящей от множества входных параметров. В некоторых случаях представляет интерес найти те значения параметров, при которых достигается наилучшее точное значение функции. В других случаях значение точного оптимума находить не требуется или невозможно и решением может считаться любое значение из некоторого множества значений, которые лучше некоторой заданной величины. При этом сложность и связность решаемых задач вызывают появление существенных затрат времени на проведение оптимизации, что проявляется как эффект динамики процедур оптимизации [2]. Следует учитывать, что с развитием и совершенствованием информационных технологий, задачи поиска наилучших решений будут занимать все большую долю среди всех процессов. Стремление повысить производительность и точность процессов обработки информации привело к пересмотру подходов к определению количества информации и перехода к анализу и синтезу информационных систем в информационном пространстве [3].

ПОСТАНОВКА ЗАДАЧИ

Динамические свойства процедуры оптимизации проявляются с уменьшением шага процедуры. Очевидно, что с развитием информационных технологий оптимизационные процедуры будут приобретать свойства самостоятельного динамического объекта. Как следствие, возникает необходимость построения математических моделей динамики оптимизационной процедуры.

СОСТОЯНИЕ ВОПРОСА

На сегодняшний день существует ряд исследований, посвященных вопросу скорости сходимости той или иной процедуры поиска оптимума [4], однако вопрос обоснования методов построения модели оптимизационной процедуры, как динамического объекта, освещен слабо, за исключением нескольких попыток предпринятых в работах [3, 4].

ЦЕЛЬ СТАТЬИ

Целью данной статьи является разработка методов построения математической модели процесса поиска оптимума как модели динамического объекта.

СОДЕРЖАНИЕ ИССЛЕДОВАНИЯ

Для построения модели поведения во времени градиентной процедуры рассмотрим автономный стационарный нелинейный динамический объект первого порядка [5]:

$$\frac{dy}{dx} = f(y). \qquad (1)$$

Выполним линеаризацию правой части уравнения (1). Для этого предположим, что функция в правой стороне уравнения (1), в окрестности точки y^*, удовлетворяет условиям для представления степенным рядом [5]:

$$f(y,x) = f(y^*) + \frac{1}{1!}\frac{\partial f}{\partial y}\Big|_{y^*} \Delta y + ... + R. \qquad (2)$$

После переноса начала координат в точку y^* получаем в (2) первое приближение функции в правой части (1):

$$f(y,x) \approx \frac{1}{1!}\frac{\partial f}{\partial y}\Big|_{y^*} \Delta y. \qquad (3)$$

Следовательно, используя линейное приближение (3), получаем:

$$\frac{dy}{dx} \approx \frac{\partial f}{\partial y}\Big|_{y^*} \Delta y. \qquad (4)$$

Учитывая (4) и возвращаясь к приращениям:

$$\frac{y + \Delta y - y}{\Delta x} \approx \frac{df}{dy}\Big|_{y^*} \Delta y. \qquad (5)$$

После переноса начала координат в точку y^* и перехода в (5) к приращениям:

$$y^* + \Delta y = y^* + \Delta x \frac{df}{dy}\Big|_{y^*} \Delta y. \qquad (6)$$

Обозначив в (6) $y^* = y_n$ и $y^* + \Delta y = y_{n+1}$, получаем модель метода Рунге-Кутты первого порядка [6]:

$$y_{n+1} = y_n + \frac{df}{dy}\Big|_{y_n} y_n \Delta x. \qquad (7)$$

Так как в данном случае линеаризация выполняется в каждой точке траектории, мы имеем случай перехода к новым переменным, и произведение $\Delta x y_n = \alpha$ определяет скорость движения для процедуры.

Введя в (7) ранее определенное α, получаем выражение (7) в виде:

$$y_{n+1} = y_n + \alpha \frac{df}{dy}\Big|_{y_n}. \qquad (8)$$

Заменив в (8) производную на *gradf*(*y*), что для (1) справедливо, получаем (7) как запись градиентной процедуры [6]:

$$y_{n+1} = y_n + \alpha \cdot gradf\big|_{y_n}. \qquad (9)$$

Следовательно, динамике свободного движения объекта, описываемого линейным дифференциальным уравнением (1), соответствует движение градиентной процедуры оптимизации с корректирующим множителем перед градиентом $\alpha = \Delta x y_n$.

Аналогично, как для одномерного объекта (1), многомерный линейный динамический объект описывается как:

$$\frac{d\mathbf{y}}{dx} = A\mathbf{y}, \qquad (10)$$

Учтем, что матрицу A в (10) можно рассматривать как матрицу линеаризованного объекта

$$A = \begin{pmatrix} \frac{\partial f_1}{\partial y_1} & \cdots & \frac{\partial f_1}{\partial y_n} \\ \cdots & \cdots & \cdots \\ \frac{\partial f_n}{\partial y_1} & \cdots & \frac{\partial f_n}{\partial y_n} \end{pmatrix}_{v^8} = \begin{bmatrix} gradf_1 \\ \vdots \\ gradf_n \end{bmatrix}_{y^*}, \qquad (11)$$

составленную из градиентов функций правой части. Метод Рунге-Кутты первого порядка дает процедуру:

$$\begin{bmatrix} y_{k+1}^1 \\ \vdots \\ y_{k+1}^n \end{bmatrix} = \begin{bmatrix} y_k^1 \\ \vdots \\ y_k^n \end{bmatrix} + \begin{pmatrix} \frac{\partial f_1}{\partial y_1} & \cdots & \frac{\partial f_1}{\partial y_n} \\ \cdots & \cdots & \cdots \\ \frac{\partial f_n}{\partial y_1} & \cdots & \frac{\partial f_n}{\partial y_n} \end{pmatrix}_{v^8} \cdot \begin{bmatrix} y_k^1 \\ \vdots \\ y_k^n \end{bmatrix} \Delta x, \qquad (12)$$

или в векторной записи процедура (12) имеет вид:

$$\mathbf{y}_{k+1} = \mathbf{y}_k + A_{y^{8}}\mathbf{y}_k \Delta x \quad (13)$$

С другой стороны в задаче поиска минимума функции цели, зависящей от вектора

$$\mathbf{y}^* \to \min f(\mathbf{y}) \quad (14)$$

градиентная процедура (9) принимает вид

$$\mathbf{y}_{k+1} = \mathbf{y}_k + \alpha \cdot gradf_{y_k}, \quad (15)$$

или при покомпонентной записи градиентная процедура (15) имеет вид:

$$\begin{bmatrix} y_{k+1}^1 \\ \vdots \\ y_{k+1}^n \end{bmatrix} = \begin{bmatrix} y_k^1 \\ \vdots \\ y_k^n \end{bmatrix} + \begin{pmatrix} \dfrac{\partial f}{\partial y_1} & \cdots & 0 \\ 0 & \cdots & 0 \\ 0 & \cdots & \dfrac{\partial f}{\partial y_n} \end{pmatrix}_{y^{8}} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} \quad (16)$$

Таким образом, процедура Рунге–Кутты первого порядка для системы (12) отличается от градиентной процедуры для задачи (16), во-первых тем, что матрица A для градиентной процедуры диагональная, и, во-вторых, тем, что множитель **α** в градиентной процедуре выбирается из соображений сходимости, а в (12) множитель определен. Исходя из этого, можно утверждать, что моделью динамики градиентной процедуры является модель свободного движения несвязанного линейного объекта, если выполняется условие (17):

$$\begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} y_k^1 \Delta x \\ \vdots \\ y_k^n \Delta x \end{bmatrix}. \quad (17)$$

В общем случае, с учетом (17), модель динамики градиентной процедуры (16) описывается процедурой Рунге-Кутты первого порядка, причем коррекцию процедуры оптимизации описываем как управление **u** с матрицей *B:*

$$\mathbf{y}_{k+1} = \mathbf{y}_k + A_{y^{8}}\mathbf{y}_k \Delta x + B\mathbf{u}_k \Delta x \quad (18)$$

Для единичной матрицы управления процедура (18) принимает вид:

$$\begin{bmatrix} y_{k+1}^1 \\ \vdots \\ y_{k+1}^n \end{bmatrix} = \begin{bmatrix} y_k^1 \\ \vdots \\ y_k^n \end{bmatrix} + \begin{pmatrix} \dfrac{\partial f}{\partial y_1} & \cdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \cdots & \dfrac{\partial f}{\partial y_n} \end{pmatrix}_{y^{8}} \cdot \begin{bmatrix} y_k^1 \\ \vdots \\ y_k^n \end{bmatrix} \Delta x + \begin{pmatrix} 1 & \cdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \cdots & 1 \end{pmatrix}_{y^{8}} \cdot \begin{bmatrix} u_k^1 \\ \vdots \\ u_k^n \end{bmatrix} \Delta x. \quad (19)$$

Следовательно, динамика простой градиентной процедуры с множителем α, удовлетворяющим условию (17), описывается системой со структурой (рис.1).

При этом последовательность переходов градиентной процедуры, определяющая движение системы, связана с переходами в пространстве переменных и изменением состояния – значения функции цели. Собственно изменение значения функции цели и определяет переходный процесс в модели (рис. 2).

Как видно из рис. 2 переходный процесс в динамической системе совпадает с процессом изменения значения функции цели в градиентной процедуре.

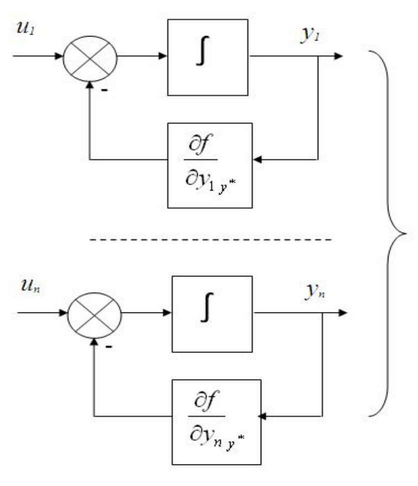
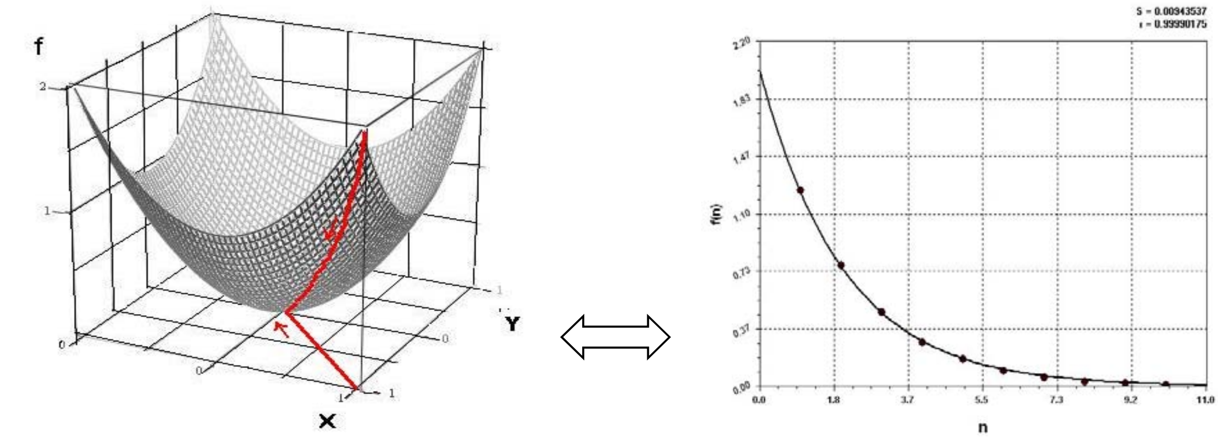


Рис. 1. Модель динамики градиентной процедуры.



а б

Рис. 2. Процесс минимизации функции а) и переходный процесс динамической модели б).

АНАЛИЗ РЕЗУЛЬТАТОВ

Анализ сделанных предположений выполним, используя стандартную квадратичную функцию цели.

Так для приведенных на рис. 2 процессов целевая функция *f*(*x*,*y*) выбрана сильно выпуклой:

$$f(x,y)=x^2+y^2, \tag{20}$$

и использована градиентная процедура:

$$\begin{bmatrix} y_{k+1}^1 \\ \vdots \\ y_{k+1}^n \end{bmatrix} = \begin{bmatrix} y_k^1 \\ \vdots \\ y_k^n \end{bmatrix} + \begin{pmatrix} \dfrac{\partial f}{\partial y_1} & \cdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \cdots & \dfrac{\partial f}{\partial y_n} \end{pmatrix}_{y^8} \cdot \Delta \mathbf{x} \tag{21}$$

Для выбранной функции (20) цели и постоянного шага $\Delta\mathbf{x}$ получаем:

$$\left.\begin{aligned} x_{k+1} &= x_k - 2x_k \cdot \Delta x \\ y_{k+1} &= y_k - 2y_k \cdot \Delta y \end{aligned}\right\}. \tag{22}$$

Таким образом, модель динамики градиентной процедуры в виде дискретной модели Рунге – Кута первого порядка для автономного движения объекта:

$$\left.\begin{aligned} \frac{dx}{dt} &= -2x \\ \frac{dy}{dt} &= -2y \end{aligned}\right\}. \tag{23}$$

Решение несвязной системы (23) дает комбинацию экспонент:

$$\left.\begin{aligned} x(t) &= 1 \cdot e^{-2t} \\ y(t) &= 1 \cdot e^{-2t} \end{aligned}\right\}. \tag{24}$$

После подстановки координат в (20) получаем траекторию целевой функции:

$$f(x,y)=x(t)^2+y(t)^2=2\cdot e^{-4t}. \tag{25}$$

Таким образом, учитывая достаточные условия оптимума [2], для окрестности точки оптимума динамика градиентной процедуры точно описывается свободным движением линейной динамической системы.

Для множителя α, содержащего координату, получаем процедуру с нелинейностью и, как следствие, переходный процесс теряет экспоненциальность. Так для процедуры:

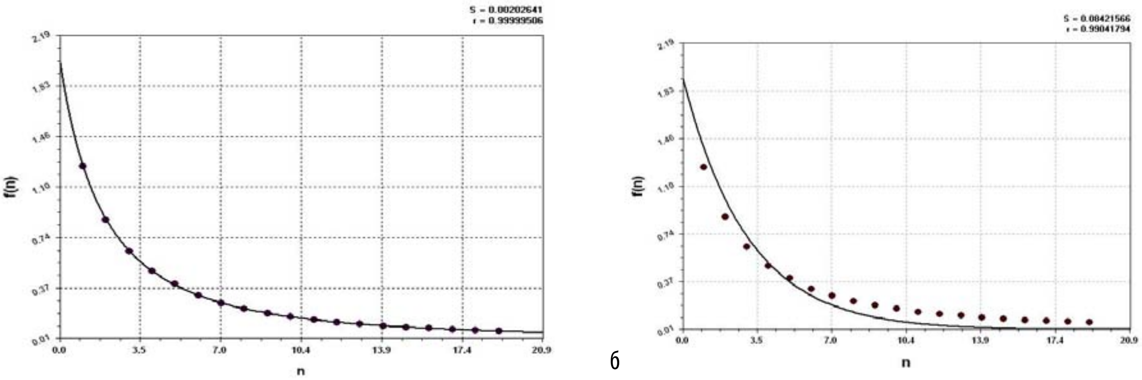
$$\left.\begin{aligned} x_{k+1} &= x_k - \frac{\partial f}{\partial x}_{x_k} x_k \cdot \Delta x \\ y_{k+1} &= y_k - \frac{\partial f}{\partial y}_{y_k} y_k \cdot \Delta y \end{aligned}\right\}. \tag{26}$$

Получаем переходный процесс с аппроксимацией переходного процесса дробно рациональной функцией а) и экспонентой б):

$$a) \quad x=i, \quad f(x)=\frac{2-0.02x}{1+0.5x+0.05x^2};$$

$$b) \quad x=i \quad f(x)=2e^{-0.3x}.$$

Графики переходного процесса показаны на рис. 3.



а б

Рис. 3. Аппроксимации переходного процесса, а) степенная функция, б) экспонента.

Как видно из приведенных графиков для рассмотренной процедуры переходный процесс описывается дробно рациональной функцией.

Существенным преимуществом рассматриваемого подхода является возможность определения необходимого количества шагов для достижения заданной точности определения минимального значения функции цели. Действительно, из (25) следует, что при начальном значении функции цели $f=f_0$ и заданном значении $f=f^*$ при заданном λ получаем:

$$]i_m[= -\frac{1}{\lambda}\ln\frac{f^*}{f_0}.$$

Полученное выражение позволяет оценить время поиска оптимального решения.

ВЫВОДЫ

Таким образом, можно сделать следующие выводы:

1. Градиентная процедура поиска экстремума функции цели может быть описана автономным движением динамической системы.

2. Для строго выпуклой функции цели, согласно достаточным условиям оптимальности, процедура оптимизации описывается динамикой автономного движения стационарного линейного несвязанного динамического объекта.

3. Выбор множителя для градиента существенно влияет на динамику процесса, и для строго выпуклой функции цели множитель α равен вектору приращений Δx.

4. Использование динамической модели позволяет определить необходимое число шагов для достижения заданной точности, что упрощает организацию вычислительного процесса.

REFERENCES

1. Esipov B.A. Metodi optimizacii i issledovanie operacii Konspekt lekcii: uceb. posobie/ B.A.Esipov.- Samara:Izd-vo Samar. gos. aerokosm. un-ta, 2009. – 180s.
2. Chernoruckii I.G. Metodi optimizacii v teorii upravlenia: Uchebnoe posobie/ I.G. Chernoruckii – SPb.: Piter, 2014.-256s.
3. A. Sokolov. Modeling the Information Accumulation Process in Artificial and Natural Systems / A. Sokolov, O. Sokolova// Project management: State and Perspectives:X International Scientific-Practical Conference, 16-19 of September, 2014, materials. – Mykolaiv, 2014. – p. 276-279.
4. Panteleev A.V. Metodi optimizacii v primerah i zadachah: Uchebnoe posobie / A.V. Panteleev, T.A. Letova. – M: Vissh. Shk., 2012. -544s.
5. Kim D.P. Teoria avtomaticheskogo upravlenia. T.1. Lineinie sistemi/ D.P. Kim. – M.: FIZMATLIT, 2013. – 288s.
6. Vasilev F.P. Chislenie metodi reshenia ekstremalnih zadach: Ucheb. posobie dlya vuzov. – 2-e izd., pererab. I dop. / F.P. Vasilev.-M.:Nauka, 2008. -552s.
7. Hansen E., Global Optimization Using Interval Analysis, New York: Dekker, 2004, -133s.
8. Goldberg, D. E. Genetic Algorithms in Search, Optimization and Machine Learning. Addison–Wesley Pub. Company, 1989. -121s.
9. P.-J. Laurent. Approximation et optimization. Paris, Hermann, 1972. 496 p.
10. Brosowski B. Nichtlineare Approximation in normierten Vektorraumen, Tagung, Oberwolfach, Juli 1968. S. 24-28

Рецензент: *д.т.н., проф. Рудакова Г. В.*
Херсонський національний технічний університет