

ИССЛЕДОВАНИЕ ОДНОГО ВАРИАНТА СУБГРАДИЕНТНОГО МЕТОДА

Н. С. Самойленко, В. Н. Крутиков, В. В. Мешечкин

RESEARCH OF ONE VARIANT OF SUBGRADIENT METHOD

N. S. Samoylenko, V. N. Krutikov, V. V. Meshechkin

В работе исследуется субградиентный метод с выбором шага на основе известного минимального значения функции. Показано, что он является аналогом метода минимальных ошибок для решения систем линейных уравнений. Для последовательности минимальных значений функции на текущем множестве итераций метода получена оценка скорости сходимости в виде неравенства.

The subgradient step selection method based on the known minimal value of function is studied in the paper. The authors show that it is an analogue of the method of minimal errors for solving linear equation systems. For a sequence of the minimum function values on the current set of method iterations, the estimate for the rate of convergence in the form of inequality is received.

Ключевые слова: субградиент, выпуклая функция, линейная алгебра, минимум функции, скорость сходимости.

Keywords: subgradient, convex function, linear algebra, minimum of function, rate of convergence.

Введение

Во многих случаях функции, которые нужно минимизировать, являются недифференцируемыми. Подобные функции возникают в задачах декомпозиции, при использовании штрафных функций, в задачах наилучшего приближения и многих других задачах минимизации. Субградиентные методы являются расширением градиентных методов на недифференцируемые функции [1 – 3].

Градиентные методы линейной алгебры [4; 5], с точки зрения дальнейшего анализа их аналогий в методах оптимизации, можно подразделить на два вида. В первом из них минимизируется непосредственно квадратичная функция, т. е. соблюдается релаксационность метода относительно минимизируемой функции. К методам этого типа относятся метод скорейшего спуска и метод сопряженных градиентов. Второй вид методов обеспечивает релаксационность убывания расстояния до экстремума. К таким методам линейной алгебры можно отнести метод минимальных ошибок и метод минимальных итераций [3; 4].

Градиентные аналоги релаксационных по функции методов линейной алгебры получили широкое развитие в оптимизации [2; 3] и, в частности, в недифференцируемой оптимизации [2 – 3; 6 – 15]. Субградиентные методы, основанные на минимизации расстояния до экстремума [1 – 3], в оптимизации появились независимо от методов линейной алгебры, и аналогии с линейной алгеброй в этих методах не устанавливались. Установление аналогий и заимствование полезно сказывается на развитии теории минимизации. В свое время это позволило получить метод сопряженных градиентов [2] и развить квазиньютоновские методы [2]. Можно ожидать, что установление аналогий между методами линейной алгебры и методами оптимизации, основанными на минимизации расстояния до экстремума, также окажется полезным.

Методы оптимизации традиционно разрабатывались для дифференцируемых функций и основывались на возможности локальной аппроксимации функции линейной либо квадратичной моделью [2]. Для решения задач негладкой оптимизации первоначально был предложен субградиентный метод с фиксированным шагом для оптимизации недифференцируемых функций [3]. Позднее были разработаны субградиентные методы с изменяющимся шагом на итерации (см. [1 – 2] и цитируемую там литературу). Изучаемый в работе метод принадлежит классу методов негладкой оптимизации, основное свойство которых заключается в сокращении расстояния до

экстремума минимизируемой функции $f(x)$, $x \in R^n$. Итерация подобного субградиентного метода имеет вид: $x_{k+1} = x_k - \gamma_k g(x_k)$.

Здесь x_k – текущее приближение на k -й итерации, γ_k – шаговый множитель, $g(x_k)$ – произвольный субградиент субградиентного множества $\partial f(x_k)$ в точке x_k , т. е. $g(x_k) \in \partial f(x_k)$. В работах Б. Т. Поляка [1 – 2] предложен и исследован вариант субградиентного метода минимизации для выпуклой, не обязательно дифференцируемой функции, в котором длина шага вычисляется (а не находится подбором) на основании известного минимального значения функции f^* :

$$x_{k+1} = x_k - \frac{\gamma[f(x_k) - f^*]}{(g(x_k), g(x_k))} g(x_k). \quad (1)$$

Здесь γ – шаговый множитель, определяемый свойствами минимизируемой функции $f(x)$, $g(x_k)$

– субградиент функции в точке x_k , f^* – минимальное значение функции $f(x)$, которое считается известным. В [2] также обосновывается применимость метода (1) для минимизации однородных и близких к ним функций.

В настоящей работе анализируются свойства субградиентного метода (1). В частности, устанавливается взаимосвязь этого метода с известным в линейной алгебре методом минимальных ошибок. Получена оценка скорости сходимости в виде неравенства для последовательности минимальных значений функции на текущем множестве итераций метода. Некоторые результаты исследования авторов в этом направлении приведены в работах [16 – 17].

Аналогия между методом минимальных ошибок и субградиентным методом

Установим аналогию между процессом (1) и методом минимальных ошибок линейной алгебры [3 – 4]. Для этой цели нам необходимо сформулировать целевую функцию минимизации так, чтобы метод минимальных ошибок в терминах характеристик этой функции был идентичным процессу (1).

Рассмотрим метод наискорейшего спуска для решения системы уравнений:

$$Ax = b, \quad (2)$$

где A – произвольная несимметричная невырожденная матрица. В методе наискорейшего спуска в качестве направления поиска выбирается вектор, направление которого противоположно направлению вектора градиента функции.

Использование метода наискорейшего спуска, в силу его применимости только для систем уравнений с симметричной и положительно определенной матрицей, становится возможным после преобразования системы посредством второй трансформации Гаусса [4]:

$$AA^T y = b, \quad (3)$$

где использована взаимосвязь $x = A^T y$.

Метод наискорейшего спуска [4, 5] для системы уравнений (3) относительно y имеет вид:

$$y_{k+1} = y_k - \frac{(AA^T y_k - b, AA^T y_k - b)}{(AA^T y_k - b, AA^T (AA^T y_k - b))} (AA^T y_k - b).$$

Домножим обе части последнего выражения на A^T . После возврата к старым переменным x получим:

$$x_{k+1} = x_k - \frac{(Ax_k - b, Ax_k - b)}{(Ax_k - b, AA^T (Ax_k - b))} A^T (Ax_k - b). \quad (4)$$

Этот итерационный метод называется методом минимальных ошибок [3 – 4].

Приведем (4) к виду (1). Обозначим ошибку $r = Ax - b$. Введем квадратичную функцию

$$f(x) = \frac{(A_x - b, A_x - b)}{2} + f^* = \frac{(r, r)}{2} + f^*, \quad (5)$$

точка минимума x^* которой совпадает с решением системы (2). Ее градиент $g(x) = A^T (Ax - b) = A^T r$.

С учетом приведенных обозначений преобразуем (4):

$$x_{k+1} = x_k - \frac{2[f(x_k) - f^*]}{(g(x_k), g(x_k))} g(x_k). \quad (6)$$

Метод минимальных ошибок (4) представлен в терминах характеристик минимизируемой функции (5). Для квадратичных функций в (1) величина шагового множителя $\gamma = 2$. Выбор величины шагового множителя γ в алгоритме (1) определяется свойствами функции [2]. Когда известно минимальное значение функции f^* , метод (1) при $\gamma = 1$ применяется для решения выпуклых задач минимизации общего вида, не обязательно квадратичных.

Таким образом, мы показали, что субградиентный метод (1) можно рассматривать как аналог в оптимизации известного метода минимальных ошибок.

Оценка скорости сходимости

Рассмотрим итерацию процесса (1) для выпуклых функций:

$$x_{k+1} = x_k - \frac{[f(x_k) - f^*]}{(g(x_k), g(x_k))} g(x_k). \quad (7)$$

Относительно сходимости процесса (7) известны следующие результаты.

Теорема 1 [2]. Пусть $f(x)$ – выпуклая функция на R^n , множество точек минимума X^* которой непусто. Тогда в методе (7) $x_k \rightarrow x^* \in X^*$. Оценка скорости сходимости следующая: для произвольной функции $f \lim_{k \rightarrow \infty} \sqrt{k} (f(x_k) - f^*) = 0$.

Для последовательностей процесса (1) справедливы следующие ограничения [2]:

$$\|x_k - x^*\| \leq M, \quad \|g(x_k)\| \leq C, \quad k = 0, 1, 2, \dots, \quad (8)$$

где M и C – некоторые фиксированные константы [2].

Обозначим достигнутое «рекордное» значение функции на множестве итераций (7):

$$\varphi_k = \min_{i=0, \dots, k} f(x_i).$$

В следующей теореме дана оценка скорости сходимости последовательности «рекордных» значений функции φ_k .

Теорема 2. Пусть $f(x)$ – выпуклая функция на R^n , множество точек минимума X^* которой непусто. Тогда для последовательности «рекордных» значений функции φ_k , генерируемых процессом (7),

справедлива оценка $|\varphi_k - f^*| \leq \frac{MC}{\sqrt{k}}$, где M и C – фиксированные константы, определенные в (8).

Доказательство. Пусть $x^* \in X^*$ – произвольная точка минимума функции $f(x)$. Тогда для процесса (7), аналогично [2], получим:

$$\begin{aligned} \|x^{k+1} - x^*\|^2 &\leq \|x^k - x^*\|^2 - \\ &\quad - \frac{(f(x_k) - f^*)^2}{\|g(x_k)\|^2} \leq \\ &\leq \|x^{k-1} - x^*\|^2 - \frac{(f(x_k) - f^*)^2}{\|g(x_k)\|^2} - \\ &\quad - \frac{(f(x_{k-1}) - f^*)^2}{\|g(x_{k-1})\|^2} \leq \dots \\ &\dots \leq \|x_0 - x^*\|^2 - \sum_{i=0}^k \frac{(f(x_i) - f^*)^2}{\|g(x_i)\|^2}. \end{aligned} \quad (9)$$

Из (9), с учетом (8), получим цепочку неравенств:

Литература

1. Поляк Б. Т. Минимизация негладких функционалов // Журн. вычислит. математики и матем. физики. 1969. Т. 9. № 3. С. 507 – 521.
2. Поляк Б. Т. Введение в оптимизацию. М.: Наука, 1983. 384 с.
3. Шор Н. З. Методы минимизации недифференцируемых функций и их приложения. Киев: Наукова думка, 1979. 199 с.
4. Фадеев Д. К., Фадеева В. Н. Вычислительные методы линейной алгебры. М.: Физматгиз, 1960. 656 с.
5. Воеводин В. В., Кузнецов Ю. А. Матрицы и вычисления. М.: Наука, 1984. 329 с.
6. Демьянов В. Ф., Васильев Л. В. Недифференцируемая оптимизация. М.: Наука, 1981. 384 с.
7. Немировский А. С., Юдин Д. Б. Сложность задач и эффективность методов оптимизации. М.: Наука, 1979. 340 с.
8. Крутиков В. Н. Обучающиеся методы безусловной оптимизации и их применение. Томск: Изд-во Том. гос. педаг. ун-та, 2008. 264 с.
9. Крутиков В. Н. Релаксационные методы безусловной оптимизации, основанные на принципах обучения. Кемерово: КемГУ, 2004. 171 с.
10. Крутиков В. Н., Петрова Т. В. Релаксационный метод минимизации с растяжением пространства в направлении субградиента // Экономика и мат. методы. 2003. Т. 39. Вып. 1. С. 33 – 49.
11. Крутиков В. Н., Горская Т. А. Семейство релаксационных субградиентных методов с двухграновой коррекцией матриц метрики // Экономика и мат. методы. 2009. Т. 45. Вып. 4. С. 37 – 80.
12. Крутиков В. Н., Вершинин Я. Н. Алгоритмы обучения на основе ортогонализации последовательных векторов // Вестник КемГУ. 2012. Вып. № 2(50). С. 37 – 42.
13. Крутиков В. Н., Вершинин Я. Н. Многошаговый субградиентный метод для решения негладких задач минимизации высокой размерности // Вестник Томск. гос. ун-та. Матем. и мех. 2014. № 3. С. 5 – 19.

$$\begin{aligned} \frac{1}{C^2} \sum_{i=0}^k (f(x_i) - f^*)^2 &\leq \sum_{i=0}^k \frac{(f(x_i) - f^*)^2}{\|g(x_i)\|^2} \leq \\ &\leq \|x_0 - x^*\|^2 \leq M^2. \end{aligned} \quad (10)$$

Неравенство (10) приведем к следующему виду:

$$\frac{1}{k+1} \sum_{i=0}^k (f(x_i) - f^*)^2 \leq \frac{M^2 C^2}{k+1}. \quad (11)$$

В левой части (11) находится среднее арифметическое слагаемых. Следовательно, хотя бы одно из этих слагаемых с индексом i_k должно быть не больше среднего, что приводит к неравенству:

$$(f(x_{i_k}) - f^*)^2 \leq \frac{M^2 C^2}{k+1}.$$

Отсюда, с учетом определения φ_k следует оценка теоремы

$$(\varphi_k - f^*)^2 \leq (f(x_{i_k}) - f^*)^2 \leq \frac{M^2 C^2}{k}. \quad (12)$$

Теорема доказана.

Заключение

В работе проведено исследование субградиентного метода с выбором шага на основе известного минимального значения функции. Показано, что изучаемый алгоритм является аналогом метода минимальных ошибок линейной алгебры. Для последовательности минимальных значений функции на текущем множестве итераций метода получена оценка скорости сходимости.

14. Крутиков В. Н., Вершинин Я. Н. Субградиентный метод минимизации с коррекцией векторов спуска на основе пар обучающих соотношений // Вестник КемГУ. 2014. Вып. № 1(57). Т. 1. С. 46 – 54.
15. Крутиков В. Н. Методы оптимизации: учебное пособие. Кемерово: КемГУ, 2011. 93 с.
16. Самойленко Н. С., Крутиков В. Н., Мешечкин В. В. Об аналогии между методом минимальных ошибок и субградиентным методом Поляка // Научное творчество молодежи. Математика. Информатика: материалы XVIII Всероссийской научно-практической конференции (24 – 25 апреля 2014 г.). Ч. 1. Томск: Издательство ТГУ, 2014. С. 67 – 68.
17. Самойленко Н. С., Крутиков В. Н., Мешечкин В. В. Об оценке сходимости субградиентного метода // Образование, наука, инновации: вклад молодых исследователей // Образование, наука, инновации – вклад молодых исследователей: материалы IX (XLI) Международной научной конференции студентов, аспирантов и молодых ученых. Кемерово: 2014. Вып. 15. С. 1518 – 1519.

Информация об авторах:

Самойленко Наталья Сергеевна – аспирант кафедры математической кибернетики КемГУ, nostienataly@mail.ru.

Natalya S. Samoylenko – post-graduate student at the Department of Mathematical Cybernetics, Kemerovo State University.

(Научный руководитель – **В. Н. Крутиков**).

Крутиков Владимир Николаевич – доктор технических наук, профессор кафедры математической кибернетики КемГУ, krutikovvn@rambler.ru.

Vladimir N. Krutikov – Doctor of Technical Science, Professor at the Department of Mathematical Cybernetics, Kemerovo State University.

Мешечкин Владимир Викторович – кандидат физико-математических наук, доцент кафедры математической кибернетики КемГУ, vvm@kemsu.ru.

Vladimir V. Meshechkin – Candidate of Physics and Mathematics, Assistant Professor at the Department of Mathematical Cybernetics, Kemerovo State University.

Статья поступила в редколлегию 30.04.2015 г.