Теорія оптимальних рішень

УДК 519.8

Н.Г. ЖУРБЕНКО

ЧИСЛЕННАЯ ЭФФЕКТИВНОСТЬ ОДНОЙ МОДИФИКАЦИИ r-АЛГОРИТМА*

Рассматривается модификация r-алгоритма – алгоритма минимизации с использованием операции растяжения пространства в направлении разности двух последовательных субградиентов. В отличие от r-алгоритма, значения коэффициентов растяжения в предложенной модификации программно рассчитываются в процессе работы алгоритма. Алгоритм может использоваться с постоянным шагом. Приводятся результаты исследования численной эффективности алгоритма.

Введение. Более 40 лет назад разработан субградиентный алгоритм минимизации с растяжением пространства в направлении разности двух последовательных градиентов – r-алгоритм [1]. Практика использования r-алгоритма показывает, что до настоящего времени – это один из наиболее эффективных алгоритмов негладкой оптимизации. Однако теоретическое исследование эффективности алгоритма далеко не закончено. Основная проблема теоретического обоснования r-алгоритма состоит в согласованном выборе значений коэффициента растяжения пространства и шагового множителя. В работе [2] предложено семейство модификаций r-алгоритма – $r(\sigma)$-алгоритмы. Величины коэффициентов растяжения пространства на итерациях $r(\sigma)$-алгоритмов не постоянны, они вычисляются в процессе его работы. Алгоритмы не требуют использования процедуры одномерного спуска по направлению. В работе [3] приведены результаты исследования численной эффективности конкретных вариантов $r(\sigma)$-алгоритмов. В данной работе предлагается новая модификация r-алгоритма из этого семейства и приводятся результаты численного исследования ее эффективности.

Численная схема r-алгоритма. Рассматривается задача безусловной минимизации субдифференцируемой функции $f(x)$ в R^n.

*Работа выполнена при частичной поддержке Volkswagen Foundation (грант № 90 306).

$\partial f(x)$ обозначим множество субградиентов функции $f(x)$ в точке x. В r-алгоритме используется оператор растяжения пространства [4] $R(\eta) = (\alpha - 1)\eta\eta^T + I$, где $\eta \in R^n$, α – направление и коэффициент растяжения пространства, $|\eta| = 1$, $\alpha \geq 0$. Вычислительная схема r-алгоритма применительно к задаче отыскания безусловного минимума функции $f(x)$ состоит в следующем.

0-ой шаг алгоритма (инициализация). Выбираем начальное приближение x_0 и невырожденное линейное преобразование B_0. Вычисляем: $g(x_0) \in \partial f(x_0)$; $g_0^* = B_0^* g(x_0)$ (субградиент в преобразованном пространстве $Y_0 = B_0^{-1} X \equiv A_0 X$, X – исходное пространство). Пусть на шаге k алгоритма $(k = 0,1,2,\ldots)$ получены определенные значения векторов x_k, g_k^* (субградиент в преобразованном пространстве) и матрицы B_k. ($A_k = B_k^{-1}$ – матрица преобразования пространства).

(k + 1)**-й шаг алгоритма** $(k = 0,1,2,\ldots)$.

Вычисляем:

1) h_{k+1} – шаговый множитель, $h_{k+1} \geq 0$;

2) $x_{k+1} = x_k - h_{k+1} B_k g_k^* / | g_k^* |$;

3) $g(x_{k+1}) \in \partial f(x_k)$, (субградиент в точке x_{k+1});

4) $\tilde{g}_{k+1}^* = B_k^* g(x_{k+1})$ (субградиент в преобразованном пространстве $Y_k = A_k X$);

5) $\eta_{k+1} = (\tilde{g}_{k+1}^* - g_k^*) / | \tilde{g}_{k+1}^* - g_k^* |$ (направление растяжения пространства Y_k);

6) $\alpha_{k+1} \geq 1, \beta_{k+1} = 1/\alpha_{k+1}$ (α_k коэффициент растяжения пространства Y_k);

7) $B_{k+1} = B_k R_{\beta_{k+1}}(\eta_k)$, (обратный оператор преобразования пространства $Y_{k+1} = A_{k+1} X = B_{k+1}^{-1} X$);

8) $g_{k+1}^* = R_{\beta_{k+1}}(\eta_{k+1}) \tilde{g}_{k+1}$ $(g_{k+1}^* = B_{k+1}^* g(x_k)$, (субградиент в преобразованном пространстве $Y_{k+1} = A_{k+1} X = B_{k+1}^{-1} X$).

Переходим к $(k+2)$-му шагу алгоритма или прекращаем работу при выполнении критерия останова.

В r-алгоритме значения коэффициентов растяжения пространства (параметр r-алгоритма) выбираются одинаковыми на всех итерациях: $\alpha_k = \alpha > 1$. На практике рекомендуется это значение выбирать порядка 2.0. Величина шагового множителя определяется процедурой минимизации («спуска») по направлению $-B_k g_k^* / | g_k^* |$. Обычно применяемые процедуры «спуска» являются достаточно грубой реализацией алгоритма локализации минимума по направлению. Наиболее часто используется процедура адаптивной регулировки шаговых множителей [1, 4]. Эта процедура определяется параметрами $0 < q_1 \leq 1$, $q_2 \geq 1$ и целым числом $L \geq 2$. Эти величины являются параметрами (константами) алгоритма.

Вычислительная схема $r\,(\sigma)$**-алгоритмов** соответствует приведенной схеме r-алгоритма. Отличие состоит лишь в следующем. Вместо оператора растяжения $R_{\alpha}(\eta)$ используем оператор

$$\tilde{R}(\tilde{\eta}) = (\alpha - 1)\tilde{\eta}\tilde{\eta}^T + I, \tag{1}$$

где $\tilde{\eta} \in R^n$, σ – нормирующий множитель, $\sigma \in R^1$, $\sigma > 0$. В отличие от оператора $R_{\alpha}(\eta)$, вектор $\tilde{\eta}$ не нормирован, т. е. выполнение условия $|\tilde{\eta}| = 1$ не требуется. Различные варианты алгоритма будут определяться выбором нормирующего множителя σ.

Отметим простейшие свойства оператора $\tilde{R}_{\sigma}(\tilde{\eta})$: $\tilde{R}^* = \tilde{R}$; $\tilde{R}^{-1}(\tilde{\eta}) = -\left(\sigma/(1+\sigma\tilde{\eta}^2)\right)\tilde{\eta}\tilde{\eta}^T + I$; $\tilde{R}_{\sigma}(0) = I$. Пусть $|\tilde{\eta}| \neq 0$ и $\eta = \tilde{\eta}/|\tilde{\eta}|$. Тогда $\tilde{R}(\tilde{\eta}) = R_{\alpha}(\eta) + I$, где

$$\alpha = 1 + \sigma|\tilde{\eta}|^2. \tag{2}$$

Таким образом, если $\tilde{\eta} \neq 0$, то оператор $\tilde{R}_{\sigma}(\tilde{\eta})$ является оператором растяжения по направлению $\tilde{\eta}/|\tilde{\eta}|$. Значение коэффициента растяжения определяется формулой (2). Значение нормирующего множителя σ будет определяться на основании субградиентов (в преобразованном пространстве) $\tilde{g}^*_{k+1}, g^*_k$: $\sigma_{k+1} = \sigma(\tilde{g}^*_{k+1}, g^*_k)$. Естественным требованием на функцию $\sigma(g_1, g_2)$ будет выполнение условия (условия «однородности») $\sigma(\mu g_1, \mu g_2) = \sigma(g_1, g_2)/\mu^2$, где $\mu \in R^1$, $\mu > 0$. Это условие обеспечивает независимость работы алгоритма от множителя на целевую функцию.

Легко видеть, что алгоритм $r(\sigma_0)$: $\sigma_0(g_1, g_2) = 1/|g_2 - g_1|^2$ фактически является r-алгоритмом с коэффициентом растяжения равным 2. Отметим, что именно это значение рекомендуется на практике использования r-алгоритма. Выбирая различные нормирующие множители σ, мы будем получать различные алгоритмы рассматриваемого класса.

В данной работе рассматривается $r(\sigma_1)$-алгоритм с нормирующим множителем $\sigma_1(g_1, g_2) = 1/(|g_2||g_1|)$. Основная особенность этого нормирующего множителя состоит в следующем. Согласно (2) значение коэффициента растяжения определяется формулой

$$\alpha = 1 + |g_2|/|g_1| + |g_1|/|g_2| - 2\cos(\varphi), \tag{3}$$

где φ – угол между субградиентами g_2, g_1. Отсюда следует, что значение коэффициента растяжения не ограничено сверху, оно может быть сколь угодно большим. Коэффициент может принимать большие значения в случае когда норма одного из субградиентов существенно больше нормы другого субградиента. Пусть, например, $|g_2| \gg |g_1|$. Тогда направление растяжения пространства будет близким к направлению субградиента g_2. Поэтому операция растяжения пространства будет существенно уменьшать значение $|g_2|/|g_1|$. То есть операция преобразования пространства будет направлена не только на уменьшение угла между субградиентами, но и на выравнивание их норм. Следует отметить, что это свойство может быть полезным в ситуации, когда по одним направлениям в точке минимума производная по направлению равна нулю, а по другим направлениям терпит разрыв (образно выражаясь, по одним направлениям функция гладкая, а по другим существенно не дифференцируемая).

Численная эффективности $r(\sigma_1)$-алгоритма. Приведем результаты численных исследований эффективности $r(\sigma_1)$-алгоритма в сравнении с r-алгоритмом. Результаты численных исследований эффективности других вариантов алгоритм $r(\sigma)$ приведены в [1, 2].

В качестве тестовых задач рассматривались задачи минимизации двух следующих функций: $f_1(x)=\sum_{i=1}^{n}\rho_n^{i-1}x_i^2$, $f_2(x)=\sum_{i=1}^{n}\rho_n^{i-1}|x_i|$, где параметр ρ_n выбирался в зависимости от размерности задачи n по формуле $\rho_n=10^{6/(n-1)}$. Таким образом, степень вытянутости линий уровня (степень «овражности») функций определяется значением параметра $\rho_n^{n-1}=10^6$, она одинакова для всех функций независимо от числа переменных. Начальная точка $x_i=1.0,\ i=1,2,...n.$ Критерий останова: $f_k \le 10^{-6}$, где f_k – значение функции на итерации останова.

Результаты решения тестовых задач минимизации функций $f_1(x)$, $f_2(x)$ приведены в табл. 1 и 2 соответственно, где приняты следующие обозначения: $r(\sigma_1)$ – алгоритм с регулировкой шаговых множителей r-алгоритма; $r^*(\sigma_1)$ – алгоритм с постоянным шагом; n – размерность пространства переменных; k – номер итерации, на которой алгоритм прекратил работу; k_g – количество вычислений субградиента; $\alpha_{\max}$ – максимальное значение коэффициента растяжения; α_{avrg} – среднее значение коэффициента растяжения; r – результаты работы r-алгоритма с параметрами $q_1=q_2=1$. В алгоритмах отмеченных символом “ **” используется постоянная величина шага (в преобразованном пространстве). В алгоритмах, не отмеченных символом “ **”, используется адаптивная регулировка шаговых множителей r-алгоритма.

ТАБЛИЦА 1

Параметры / Алгоритм	n	k	k_g	$\alpha_{\max}$	α_{avrg}
$r(\sigma_1)$	10	73	106	9.95	4.42
$r^*(\sigma_1)$	10	78	79	9.39	3.94
r	10	135	168	2.00	2.00
$r(\sigma_1)$	100	603	659	14.43	4.92
$r^*(\sigma_1)$	100	629	630	12.36	4.83
r	100	1197	1382	2.00	2.00
$r(\sigma_1)$	300	1775	1951	15.03	5.08
$r^*(\sigma_1)$	300	1841	1842	15.03	4.98
r	300	3416	3898	2.00	2.00
$r(\sigma_1)$	500	2972	3306	22.99	5.15
$r^*(\sigma_1)$	500	3041	3042	27.43	5.00
r	500	6275	6946	2.00	2.00

ТАБЛИЦА 2

Параметры / Алгоритм	n	k	k_g	$\alpha_{\max}$	α_{avrg}
$r(\sigma_1)$	10	158	305	7.86	4.15
$r^*(\sigma_1)$	10	181	182	7.52	3.64
r	10	304	350	2.00	2.00
$r(\sigma_1)$	100	1536	1554	5.38	4.49
$r^*(\sigma_1)$	100	1540	1541	5.76	4.49
r	100	3256	3264	2.00	2.00
$r(\sigma_1)$	300	4669	4678	5.29	4.57
$r^*(\sigma_1)$	300	4672	4673	5.22	4.57
r	300	10015	10023	2.00	2.00
$r(\sigma_1)$	500	7847	7873	5.19	4.58
$r^*(\sigma_1)$	500	7850	7851	5.19	4.58
r	500	16850	16864	2.00	2.00

Выводы. $r(\sigma_1)$-алгоритм – это модификация r-алгоритма. Вычислительная схема $r(\sigma_1)$-алгоритма с постоянным шагом существенно проще схемы r-алгоритма. Величины коэффициентов растяжения пространства на итерациях $r(\sigma_1)$-алгоритма не постоянны, они вычисляются в процессе его работы. Алгоритм может использоваться с постоянным шаговым множителем. Численные эксперименты показали достаточно высокую эффективность $r(\sigma_1)$-алгоритма.

М.Г. Журбенко

ЧИСЕЛЬНА ЕФЕКТИВНІСТЬ ОДНІЄЇ МОДИФІКАЦІЇ *r*-АЛГОРИТМУ

Розглядається модифікація *r*-алгоритму – алгоритму мінімізації з використанням операції розтягування простору в напрямку різниці двох послідовних субградієнтів. На відміну від *r*-алгоритму, значення коефіцієнтів розтягування у запропонованій модифікації програмно розраховуються в процесі роботи алгоритму. Алгоритм може використовуватися з постійним кроком. Наводяться результати дослідження чисельної ефективності алгоритму.

N.G. Zhurbenko

NUMERICAL EFFICIENCY OF ONE MODIFICATION OF R-ALGORITHM

The modification of r-algorithm, the minimization algorithm using space dilation operation along the direction of the difference of two successive subgradients, is considered. In contrast to *r*-algorithm, in the proposed modification the values of dilation coefficients are calculated during the execution of the algorithm. The algorithm can be used with a constant step size. The results of the study of the numerical efficiency of the algorithm are presented.

1. *Шор Н.З., Журбенко Н.Г.* Метод минимизации, использующий операцию растяжения пространства в направлении разности двух последовательных градиентов. *Кибернетика*. 1971. № 3. С. 51 – 59.
2. *Журбенко Н.Г.* Об одной модификации *r*-алгоритма. Материалы III Международной конференции. *Математическое моделирование, оптимизация и информационные технологии. Кишинеу*: Эврика, 2012. С. 355 – 361.
3. *Журбенко Н.Г., Чумаков Б.М.* Программное управление коэффициентами растяжения *r*-алгоритма. *Теорія оптимальних рішень*. К.: Ін-т кібернетики імені В.М. Глушкова НАН України, 2012. С. 113 – 118.
4. *Шор Н.З.* Методы минимизации недифференцируемых функций и их применение. Киев: Наук. думка, 1979. 200 с.

Получено 10.03.2017