

Релаксационные рестарт-методы решения задач безусловной минимизации

А. Г. Бирюков, А. В. Чернов
(Московский физико-технический институт
(Государственный Университет))

Предложены методы решения задачи безусловной минимизации, в которых итерационный процесс перезапускается по определенному правилу через наперед заданное число шагов (рестарт). Приводятся результаты численного эксперимента по сравнению эффективности предложенных методов с методами без рестарта.

Ключевые слова: безусловная минимизация, рестарт, релаксация, модифицированный метод сопряженных градиентов, градиентный метод, метод Флетчера — Ривса.

Введение

Данная статья посвящена анализу численной эффективности новых методов решения задачи (??) безусловной минимизации (БМ). Исследуемые методы представляют собой модификации градиентного метода (ГМ) и метода сопряженных градиентов (МСГ).

$$f(x) \rightarrow \min_{x \in R^n} \quad (1)$$

Итерационный процесс (ИП) решения задачи БМ можно представить в виде

$$x_{k+1} = x_k + \lambda_k p_k, \quad k = 0, 1, 2, \dots, \quad (2)$$

где $p_k \in R^n$ — направление убывания (релаксации) функции (способ выбора p_k определяет название метода); $\lambda_k \in R^1$ — шаг по направлению p_k ; $x_0 \in R^n$ — некоторая начальная точка. В предлагаемых модификациях ГМ и МСГ индекс k представляется в виде

$$k = l(m + 1) + i; \quad i = 1, 2, \dots, m + 1; \quad l = 0, 1, 2, \dots \quad (3)$$

В такой схеме подразумевается, что после каждого блока из m шагов задается по некоторому правилу начальная точка следующего блока из

m шагов. Эта процедура прерывания исходного ИП называется рестартом. Пусть

$$r(l, m) = (l + 1)(m + 1).$$

Другими словами на шаге $k = r(l, m) - 1$ по некоторому правилу определяется вектор p_k , например,

$$\begin{aligned} p_k &= x_{r(l,m)-1} - x_{r(l,m)-(m+1)} \text{ или} \\ p_k &= x_{r(l,m)-1} - x_{r(l,m)-(m+1)-t}, \text{ где } t < m. \end{aligned} \quad (4)$$

Длина шага λ_k из (??) берется так, чтобы $f(x_{k+1}) \leq f(x_k)$ (условие релаксации), что и определяет название методов (их аббревиатура: РРМ или RRM). В нашей работе шаг λ_k определяется из условия решения задачи (??), которая решается приближенно методом золотого сечения [1].

$$\lambda_k = \arg \min_{\lambda > 0} f(x_k + \lambda p_k). \quad (5)$$

Сделаем замечания о скорости сходимости РРМ в случае, когда $f(x)$ — дважды дифференцируемая сильно выпуклая функция.

- 1) Градиентный метод в указанном ИП будет методом наискорейшего спуска и скорость его сходимости будет линейной [2].
- 2) В МСГ при $k \neq r(l, m) - 1$

$$p_k = \begin{cases} -\nabla f(x_k), & k = r(l, m), l = -1, 0, 1, 2, \dots, \\ -\nabla f(x_k) + \beta_k p_{k-1}, & k \neq r(l, m), l = -1, 0, 1, 2, \dots, \end{cases} \quad (6)$$

где β_k — множитель, определяющий модификацию МСГ [3]. Например, $\beta_k = \|\nabla f(x_k)\|^2 / \|\nabla f(x_{k-1})\|^2$ для метода Флетчера — Ривса. При этом длина шага λ_k определяется из (??). Полагая число шагов в блоке МСГ $m < n$ можно утверждать, что метод сходится с линейной скоростью [2].

Очевидно, что погрешности в решении задачи (??) могут влиять на эффективность (время решения задачи БМ) предложенных методов.

ИП для РРМ (??), (??), (??), (??) для ГМ идейно близок схеме итерационного партан-метода (ПМ) [4], в котором $m = n$, и модифицированного ПМ, который рассматривается для $n \equiv 2$ [4]. Некоторую аналогию методам РРМ для ГМ и МСГ можно найти в так называемом «овражном методе» [2], но в нем способ выбора шага λ_k отличается от (??) и число шагов m не фиксируется.

Достаточно гибкая схема РРМ (??), (??), (??), (??) позволяет надеяться, что на её основе можно строить эффективные по времени алгоритмы решения задачи БМ.

Численные эксперименты

Эксперименты проводились на ЭВМ с процессором Intel Core i5, 2.5 ГГц и оперативной памятью 6 Гб в среде Matlab 2012® (8.0) под управлением операционной системы Microsoft Windows 7.

Предложенные методы исследовались на примере двух задач:

- 1) Задача энтропийно-линейного программирования (ЭЛП) (??).

$$\begin{aligned} f(x) &= \sum_{i=1}^n x_i \ln(x_i/\xi_i) \rightarrow \min_{x \in G}; \\ G &= \{x \in S_n(1) : Ax = b\}, \quad b \in R^q, \quad A \in R^{q \times n}; \\ S_n(1) &= \{x \in R_{++}^n : \sum_{i=1}^n x_i = 1\}. \end{aligned} \quad (7)$$

Исследуемые методы применялись для решения выпуклой задачи (??) двойственной к ЭЛП (??):

$$\begin{aligned} \phi(y) &= \langle y, b \rangle + \ln \left[\sum_{i=1}^n \xi_i \exp(-[C^T y]_i) \right] \rightarrow \min_{y \in R^m}; \\ x_i(y) &= \frac{\xi_i \exp(-[C^T y]_i)}{\sum_{j=1}^n \xi_j \exp(-[C^T y]_j)}, \quad i = \overline{1, n}; \quad \nabla \phi(y) = b - Ax(y). \end{aligned} \quad (8)$$

Условия окончания имеют вид:

$$\begin{aligned} |f(x_k) - f(x^*)| &\leq |\phi(y_k) + f(x(y_k))| \leq \varepsilon_f; \\ \|\nabla \phi(y_k)\| &= \|b - Ax(y_k)\| \leq \varepsilon_g; \end{aligned}$$

Здесь $n = 3^4 \div 9^4$, а $q = 3^2 \cdot 2 \div 9^2 \cdot 2$.

- 2) Рассматривалась модификация (??) невыпуклой функции Розенброка [2], где $\alpha > 0$ (в экспериментах $\alpha = 10; 20$).

$$f(x) = (x_1 - 1)^2 + \alpha \sum_{i=2}^n (x_i - x_{i-1}^2)^2 \rightarrow \min_{x \in R^n}. \quad (9)$$

Решение такой задачи известно: $x_i^* = 1, i = \overline{1, n}; f(x^*) = 0$. Условия окончания:

$$|f(x_k) - f(x^*)| \leq \varepsilon_1, \quad \|\nabla f(x_k)\| \leq \varepsilon_2, \quad \|x_k - x_{k-1}\| \leq \varepsilon_3.$$

Результаты численных экспериментов для решения задачи ЭЛП (??) и задачи БМ (??) приведены на рисунках ??, ??, ?. Метод без рестартов представлен на графиках значением шага рестарта 0.

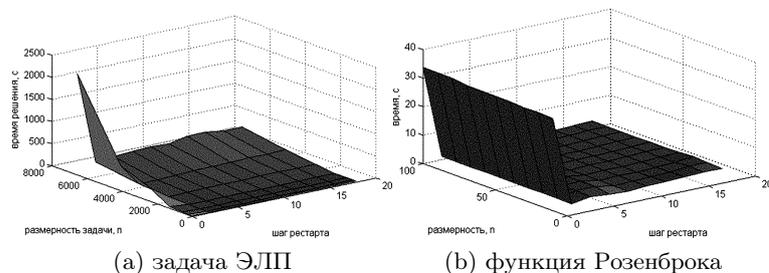


Рис. 1. Градиентный метод с рестартами.

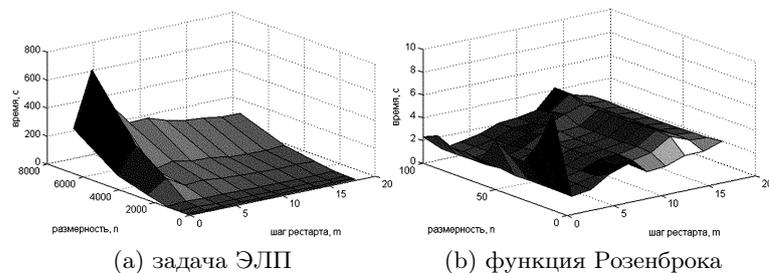


Рис. 2. Метод Флетчера—Ривса с рестартами.

Из рисунка ?? видно, что градиентный метод без рестартов существенно уступает методу с рестартами причем эффективность последнего возрастает с размерностью. Причем эффективность градиентного метода с рестартом существенно выше исходного метода уже при шаге рестарта 2, так и при больших значениях шага рестарта.

Рисунок ?? отражает специфику применения шага рестарта к методу Флетчера—Ривса. Видно, что методы с малым значением шага рестарта (2, 3, 4) не эффективнее оригинального метода, а скорее наоборот. Это объясняется тем, что слишком частые рестарты не дают возможность методу в полной мере воспользоваться результатами предыдущих шагов, что сказывается на его эффективности. Однако у методов с большим значением шага рестарта наблюдается увеличение эффективности до двух раз.

На рисунке ??, отражающем результаты численных экспериментов по введению рестартов в модифицированный МСГ [3], наблюдаются схожие с методом Флетчера—Ривса результаты, в частности методы с достаточно большим значением номера шага рестарта оказывались сравнимы с исходным методом без рестартов, а во многих случаях были эф-

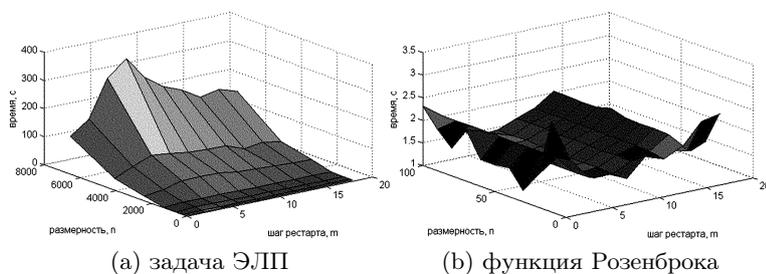


Рис. 3. Модифицированный МСГ с рестартами.

фективнее. Одновременно методы с небольшим значением номера шага рестарта были менее эффективны чем исходный метод.

Выводы

Результаты численных экспериментов показали, что применение методики рестартов к градиентному методу как в случае выпуклой задачи так и невыпуклой приводит к многократному увеличению эффективности метода, причем последний оказывается сравним с МСГ. Можно отметить, что полученные модификации метода сопряженных градиентов при относительно небольших значениях шага рестарта сравнимы с аналогичного методом без рестартов, а при увеличении номера шага рестарта оказывались лучше.

Таким образом можно предположить, что предложенные модификации методов не менее эффективны чем их исходные варианты при решении задач БМ большой размерности.

Список литературы

- [1] Сухарев А. Г., Тимохов А. В., Федоров В. В. Курс методов оптимизации. — М.: Физматлит, 2008.
- [2] Поляк Б. Т. Введение в оптимизацию. — М.: Ленанд, 2014.
- [3] Бирюков А. Г., Чернов А. В. О модификациях метода сопряженных градиентов // Модели и методы обработки информации. — М.: МФТИ, 2016.
- [4] Химмельблау Д. Прикладное нелинейное программирование. — М.: Мир, 1975.