Следовательно, для моды $M_0(X)=k_0$ выполняется неравенство $np-q\le k_0\le np+p$ или, когда k_0 принадлежит отрезку $[np-q;np+p]$, который имеет длину, равную $(np+p)-(np-q)=p+q=1$.

Таким образом, если $(np+p)\notin N$, то распределение имеет моду $M_0(X)=[np+p]$, где $[np+p]$ – целая часть числа $(np+p)$, и распределение является **унимодальным**.

Если же $(np+p)\in N$, то распределение имеет две моды $M_0(X)=np+p-1=$ $=np-q$ и $M_0(X)=np+p$, и распределение является **бимодальным**.

Заметим, что распределение является бимодальным, если $np+p=m\in N$ или когда $p=\frac{m}{n+1}$, где $m\in N$ и $m<n+1$.

Список цитированных источников

1. Вентцель, Е.С. Теория вероятностей / Е.С. Вентцель. – М.: Высшая школа, 1999. – 576 с.
2. Корн, Г. Справочник по математике для научных работников и инженеров / Г. Корн, Т. Корн – М.: Наука, 1977. – 831 с.

УДК 519.853.3
Мокин А.А., Бруцкий В.Р.
Научный руководитель: к.ф.-м.н., доцент Ракецкий В.М.

ЧИСЛЕННОЕ СРАВНЕНИЕ ПРЯМОГО ОПОРНОГО МЕТОДА С МЕТОДАМИ СОПРЯЖЕННЫХ НАПРАВЛЕНИЙ

1. Введение. Целью настоящей работы является экспериментальное исследование прямого опорного метода для минимизации выпуклых функций [1] и его сравнение с другими известными методами. В качестве объекта для проведения численного эксперимента рассматривается задача безусловной оптимизации

$$f(x)\to min, x\in R^n, \qquad (1)$$

где $f(x)$- строго выпуклая функция, т.е. в задаче (1) всегда существует единственное решение x^* – точка минимума функции $f(x)$, которая удовлетворяет необходимым и достаточным условиям оптимальности

$$\Delta_j(x)=\frac{d}{dx_j}f(x)=0,\quad j=\overline{1,n}. \qquad (2)$$

Для выяснения эффективности прямого опорного метода и получения данных для анализа и сравнения в эксперименте использовались:

1) метод сопряженных градиентов Флетчера-Ривса [2]. Выбор этого известного метода в качестве одного из эталонов для сравнения обусловлен его алгоритмической простотой и низкой трудоемкостью итерации, что обеспечивает высокую эффективность метода на задачах невысокой размерности. Его недостатком является чувствительность к вычислительным погрешностям, что снижает его эффективность особенно для задач с большим числом переменных;

2) метод Давидона-Флетчера-Пауэлла [2]. Этот метод относится к группе квазиньютоновских методов, которые для организации итерационных вычислений используют специальную вспомогательную матрицу, в которой накапливается информация о структуре минимизируемой функции. Итерация квазиньютоновских методов отличается большей трудоёмкостью в сравнении с методом сопряженных градиентов. Однако, благодаря ис-

пользованию информации о структуре функции и устойчивости к вычислительным погрешностям, квазиньютоновские методы показали высокую эффективностью для задач вида (1). Метод Давидона-Флетчера-Пауэлла отличается наиболее высокими характеристиками.

2. Алгоритмы методов. Приведем краткое описание алгоритмов методов, которые были программно реализованы и приняли участие при проведении численного эксперимента.

2.1. Прямой опорный метод. Пусть $\{x, J_{оц}\}$ – опорный план [1] задачи (1), известный к началу итерации, ε – заданная точность выполнения условий оптимальности (2). Предположим, что опора целевой функции $J_{оц}$ состоит из индексов $\{i_1, i_2, \ldots, i_m\}$, множество неопорных индексов $J_н$ состоит из индексов $\{j_1, j_2, \ldots, j_{n-m}\}$, j_k – неопорный индекс, для которого предстоит проверить условия оптимальности на текущей итерации, $J_н^0$ – множество индексов из $J_н$, для которых на предыдущих итерациях был установлен факт выполнения условий оптимальности с заданной точностью, т.е.

$$|\Delta_j(x)| \le \varepsilon, \quad j \in J_н^0,$$

и опорный план $\{x, J_{оц}\}$ не менялся.

2.1.1. Проверим, выполняются ли условия оптимальности (2) на индексе j_k с заданной точностью:

$$|\Delta_{j_k}(x)| \le \varepsilon. \qquad (3)$$

Если условие (3) не выполняется, то перейдем к п. 2.1.3. В противном случае положим $J_н^0 = J_н^0 \bigcup \{j_k\}$. Возможны 2 случая:

а) $J_н^0 = J_н$, т.е. все неопорные индексы удовлетворяют с заданной точностью условиям оптимальности. В этом случае перейдем к проверке согласованности опорного плана.

б) $J_н^0 \subset J_н$, $J_н^0 \ne J_н$, т.е. возможно, что среди неопорных индексов есть такие, для которых условия оптимальности не выполняются с заданной точностью. В этом случае положим $k = (k+1) \bmod (n-m)$ и снова проверим условие (3) для нового индекса (повторим пункт 2.1.1).

2.1.2. Проверим условия согласования опорного плана

$$|\Delta_j(x)| \le \varepsilon, \quad j \in J_{оц} \qquad (4)$$

Если условия согласования (4) выполняются, то решение задачи (1) окончено, план **x** принимаем в качестве оптимального плана. В противном случае полагаем $J_{оц} = \varnothing$ и переходим к п.2.1.1. В качестве индекса j_k используем первый из индексов опорного плана, для которого не выполнены условия согласования (4).

2.1.3. Построим направление для улучшения опорного плана $\{x, J_{оц}\}$

$$l_{j_k} = -sign\Delta_{j_k}, \quad l_j = 0, \quad j \in J_н \setminus \{j_k\}, \quad l_{оц} = G_{оц}\, p_{оц}\, sign\Delta_{j_k}, \qquad (5)$$

где $p_{оц} = \{p_j, j \in J_{оц}\}$, $p_j = \dfrac{\partial^2 f(x)}{\partial x_j \partial x_{j_k}}$.

2.1.4. Вычислим шаг θ вдоль направления l:

$$\theta = \begin{cases} |\Delta_{j_k}| / \alpha, \text{если } \alpha > 0, \\ \infty, \text{если } \alpha \le 0, \end{cases} \quad \alpha = l'_{оц} p_{оц} sign\,\Delta_{j_k} + d_{j_k j_k}, \; d_{j_k j_k} = \frac{\partial^2 f(x)}{\partial x_{j_k} \partial x_{j_k}}$$

Если $\theta = \infty$, то перейдем к п.2.1.6.

2.1.5. Проверим условие

$$f(x + \theta l) < f(x) - \eta. \qquad (6)$$

Возможны 2 варианта:

а) условие (6) выполняется. В этом случае построим новый опорный план $\{\tilde{x},\tilde{J}_{оц}\}$, где $\tilde{x}=x+\theta l$, $\tilde{J}_{оц}=J_{оц}\bigcup\{j_k\}$, и пересчитаем матрицу $G_{оц}$:

$$\tilde{G}_{оц}=\begin{pmatrix} G_{оц}+\dfrac{l_{оц}l'_{оц}}{\alpha} & -\dfrac{l_{оц}}{\alpha}sign\Delta_{j_0} \\ -\dfrac{l'_{оц}}{\alpha}sign\Delta_{j_0} & \dfrac{1}{\alpha} \end{pmatrix}$$

Если $|\tilde{J}_{оц}|=n$, то $\tilde{J}_н=\varnothing$. Поэтому следующую итерацию начнем сразу с п. 2.1.2 – проверки условий согласования опорного плана $\{\tilde{x},\tilde{J}_{оц}\}$. Если же $|\tilde{J}_{оц}|<n$, то перейдем к п. 2.1.1. При этом в проверку условий оптимальности начнем с индекса j_{k+1} (следующего по порядку после j_k в множестве $J_н$), приняв, что $\tilde{J}^0_н=\varnothing$ (если j_k - последний из индексов множества $J_н$, т.е. $k=n-m$, то проверку условий оптимальности начнем с индекса j_1 – первого индекса в множестве $\tilde{J}_н$);

б) условие (6) не выполняется. Продолжим итерацию, для чего перейдем к п. 2.1.6.

2.1.6. Найдем шаг θ вдоль направления l с помощью метода одномерной оптимизации (например, метода «золотого» сечения) и проверим условие

$$f(x+\theta l)<f(x)-\eta/8 \qquad (7)$$

Возможны 2 исхода:

а) условие (7) выполняется. В этом случае уточним значение параметра $\eta=(f(x)-f(x+\theta l))/2$ и перейдем к следующей итерации по правилам пункта 2.1.5 а), приняв в качестве значения α величину $\alpha=|\Delta_{j_0}|/\theta$.

б) условие (7) не выполняется. В этом случае прервем текущую итерацию и проверим условия согласования опорного плана $\{x, J_{оц}\}$, т.е. перейдем к п. 2.1.2.

2.2. Метод сопряженных градиентов.

1. Пусть x^0 – начальная точка, тогда в качестве направления выберем антиградиент $S^0=-\Delta(x^0)$ и найдем минимум вдоль направления S^0.

2. Допустим, что на шаге k мы находимся в точке x^k и $-\Delta(x^k)$ – направление антиградиента. Направление на данном шаге будем выбирать согласно формуле: $S^k=-\Delta(x^k)+\omega_k S^{k-1}$, где $\omega_k=\dfrac{\|\Delta(x^k)\|^2}{\|\Delta(x^{k-1})\|^2}$ (стандартный алгоритм – Флетчера-Ривса).

3. Новую точку x^{k+1} вычислим по формуле $x^{k+1}=x^k+\theta_k S^k$, где $\theta_k=\arg\min\limits_{\theta} f(x^k+\theta S^k)$.

Если $\|\Delta(x^{k+1})\|<\varepsilon$, то останавливаем алгоритм.

Иначе, если $(k+1)<n$, то переходим к пункту 2. Иначе выполняем пункт 2 при $\omega_k=0$.

2.3. Метод Давидона-Флетчера-Пауэлла.

1. Задаем начальную точку x^0 и погрешность ε.

2. Вычисляем $H^0=\left[\dfrac{\partial^2 f(x^0)}{\partial x^2}\right]^{-1}$.

3. Допустим, что на шаге k мы находимся в точке x^k и в этой точке построена матрица H^k. Тогда новые координаты точки находятся следующим образом: $x^{k+1}=x^k+\theta_k S^k$, где $S^k=-H^k\Delta(x^k)$, $\theta_k=\arg\min\limits_{\theta} f(x^k+\theta S^k)$.

4. Пересчет H^{k+1} производится по формуле:

$$H^{k+1}=H^k-\frac{H^k y^k \left(y^k\right)^T H^k}{\left(y^k\right)^T H^k y^k}+\theta_k \frac{S^k \left(S^k\right)^T}{\left(S^k\right)^T y^k},$$

где $y^k=\Delta(x^{k+1})-\Delta(x^k)$.

5. Если $\left\|\Delta(x^{k+1})\right\|<\varepsilon$, то останавливаем алгоритм.

Если **$(k+1)<n$**, то переходим к пункту 3.

Иначе полагаем $x^0=x^n$ и выполняем пункт 2 .

При программной реализации описанных выше алгоритмов для реализации одномерного поиска использовался метод «золотого» сечения [2].

3. Эксперимент. В эксперименте целевая функция (1) имела вид

$$f(x)=\frac{1}{2}x^T Dx+c^T x+\sum_{i=1}^{m} e^{\sum_{j=1}^{n} a_{ij}x_j}, \qquad (8)$$

где D – симметричная положительно определенная (n*n)- матрица, c – n-вектор, $a_{ij}, i=\overline{1,m}, j=\overline{1,n}$ – известные величины. Для автоматизации процесса тестирования был разработан генератор задач вида (8). В ходе эксперимента генерировались задачи одного размера сериями по 10 штук. Каждая задача из серии по очереди решалась тремя вышеописанными алгоритмами. Сравнение алгоритмов проводилось по двум параметрам: времени решения и количеству итераций. В качестве результатов эксперимента принимались средние результаты по серии, из которых отбрасывалось одно наибольшее и одно наименьшее значение (т.е. сумма оставшихся делилась на 8). В табл. 1 приведены результаты эксперимента по решению задач с размерами от 3 до 20 переменных.

Таблица 1 – Результаты эксперимента (ПМ – прямой метод, СГ – метод сопряженных градиентов, ДФП – метод Давидона-Флетчера-Пауэлла)

Количество переменных	Количество итераций			Время решения (мс)		
	ПМ	СГ	ДФП	ПМ	СГ	ДФП
3	10.0	7.8	5.5	2.4	3.4	12.6
4	12.0	11	5.9	4.8	9.4	25.0
5	14.1	17.3	5.5	3.8	15.1	31.6
10	34.8	131.9	10.8	48.5	436.3	238.4
15	70.4	-	15.8	343.8	-	784.6
20	96.0	-	22.5	1042.6	-	4515.3

Результаты эксперимента подтверждают известные выводы о численной неустойчивости метода сопряженных градиентов и высокой эффективности метода Давидона-Флетчера-Пауэла. Метод сопряженных градиентов плохо решал или не решал вовсе с заданной точностью задачи с количеством переменных больше 10.

Проведенный численный эксперимент свидетельствует о высокой эффективности прямого опорного метода: хотя прямой опорный метод уступает методу Давидона-Флетчера-Пауэла по количеству итераций, однако существенно превосходит его по скорости решения.

Список цитироанных источников

1. Ракецкий, В.М. К минимизации выпуклых функций с простыми ограничениями // Вестник Брестского государственного технического университетата. – 2011. – № 5(71): Физика, математика, информатика. – С. 108–110

2. Банди, Б. Методы оптимизации. Вводный курс. – М.: Радио и связь, 1988. – 128 с.