Лабораторная работа № 3

ЧИСЛЕННЫЕ МЕТОДЫ ПОИСКА ЭКСТРЕМУМА
ДЛЯ ДИФФЕРЕНЦИРУЕМЫХ ФУНКЦИЙ МНОГИХ ПЕРЕМЕННЫХ

Целью работы являются изучение и моделирование в среде MATLAB различных градиентных методов поиска минимума и сравнение эффективности их применения для конкретных целевых функций нескольких переменных.

1. Краткие теоретические сведения

1.1. Метод градиента с постоянным (заданным) шагом

1. Выбрать начальную точку x^0 и число (величину шага) $\alpha > 0$. Положить $k = 0$.

2. На k-й итерации рассчитать $x^{k+1} = x^k - \alpha \dfrac{\nabla f(x^k)}{\left|\nabla f(x^k)\right|}$, где градиент

$\nabla_k > 0$.

3. Проверить выполнение одного из условий окончания поиска. Возможные варианты останова данного и последующих алгоритмов приведены ниже. Если условие выполняется, то закончить вычисления. В противном случае положить $k=k+1$ и перейти к шагу 2.

1.2. Метод наискорейшего спуска

1. Выбрать начальную точку x^0. Положить $k = 0$.

2. На k-й итерации вычислить $d_k = -\nabla f(x^k)$. Затем найти такое число α_k, чтобы $f(x^k + \alpha_k d_k) = \min\{f(x^k + \alpha d_k)\}$. Положить $x^{k+1} = x^k + \alpha_k d_k$.

3. Проверить выполнение условия окончания поиска. Если условие выполняется, то останов алгоритма. Иначе положить k=k+1 и перейти к шагу 2.

1.3. Метод сопряженного градиента для квадратичных функций

1. Выбрать начальную точку x^0. Определить $g = \nabla q(x^0) = Ax^0 + b$.
Положить $d_0 = - g_0$, $k = 0$.
2. На k-й итерации определить

$$\lambda_k = \frac{g_k^T d_k}{d_k^T A d_k}, \; x^{k+1} = x^k + \alpha_k d_k, \; g_{k+1} = \nabla q(x^{k+1}),$$

$$\beta_k = \frac{g_{k+1}^T A\, d_k}{d_k^T A\, d_k}, \; d_k = - g_k + \beta_k d_k.$$

3. Проверить выполнение условия окончания поиска. Если условие выполняется, то останов алгоритма. Иначе положить k=k+1 и перейти к шагу 2.

1.4. Метод Фленча–Ривза

1. Выбрать начальную точку x^0. Положить $d_0 = - \nabla f(x^0)$, $k = 0$.
2. На k-й итерации выбрать α, минимизирующую функцию $g(\alpha)=$ $=f(x+\alpha d)$, положить $x^{k+1} = x^k + \alpha_k d_k$, где

$$\beta_k = \frac{\left|\nabla f(x^{k+1})\right|^2}{\left|\nabla f(x^k)\right|^2};$$

$$d_k = - \nabla f(x^{k+1}) + \beta_k d_k.$$

3. Проверить выполнение условия окончания поиска. Если условие выполняется, то останов алгоритма. Иначе положить k=k+1 и перейти к шагу 2.

1.5. Метод Ньютона, модифицированный метод Ньютона, метод Марквардта

1. Выбрать начальную точку x^0, $k = 0$.

2. На k-й итерации для метода Ньютона – $x^{k+1} = x^k - [\nabla^2 f(x^k)]^{-1}\nabla f(x^k)$;
для модифицированного метода Ньютона $x^{k+1} = x^k - \lambda_k[\nabla^2 f(x^k)]^{-1}\nabla f(x^k)$;
для метода Марквардта $x^{k+1} = x^k - [\mu I + \nabla^2 f(x^k)]^2\nabla f(x^k)$, где $\mu_0 = 10^4$; $\mu_k = \mu_0/k$, I – единичная матрица размером $n\times n$.

3. Проверить выполнение условия окончания поиска. Если условие выполняется, то останов алгоритма. Иначе положить $k=k+1$ и перейти к шагу 2.

1.6. Алгоритм Давида–Флетчера–Пауэлла (ДФП)

1. Выбрать начальную точку x^0; положить $H_0=I$, где I – единичная матрица; $k=0$.

2. На k-й итерации вычислить $d_k = - H_k\nabla f(x^k)$. Затем найти такое число $\alpha \geq 0$, чтобы

$$f(x^k + \lambda_k d_k) = \min\{f(x^k + \alpha d_k)\}.$$

Положить $x^{k+1} = x^k + \alpha_k d_k$; $\sigma_k = x^{k+1} - x^k$; $\gamma_k = \nabla f(x^{k+1}) - \nabla f(x^k)$;

$$H_k = H_k + \frac{\sigma_k \sigma_k^T}{\sigma_k^T \gamma_k} - \frac{H_k \gamma_k \gamma_k^T H_k}{\gamma_k^T H_k \gamma_k}.$$

3. Проверить выполнение условия окончания поиска. Если условие выполняется, то останов алгоритма. Иначе положить $k=k+1$ и перейти к шагу 2.

1.7. Алгоритм Бройдена–Флетчера–Гольдфарба–Шанно (БФГШ)

1. Выбрать начальную точку x^0. Положить $H = I$, где I- единичная матрица; $k = 0$.
2. На k-й итерации вычислить $d_k = -H_k \nabla f(x^k)$. Затем найти такое $\alpha \geq 0$, чтобы $f(x^k + \alpha_k d_k) = \min\{f(x^k + \alpha d_k)\}$.

Положить $x^{k+1} = x^k + \alpha_k d_k$; $\sigma_k = x^{k+1} - x^k$; $\gamma_k = \nabla f(x^{k+1}) - \nabla f(x^k)$,

$$H_{k+1} = H_k + \left[1 + \frac{\gamma_k^T H_k \gamma_k}{\sigma_k^T \gamma_k}\right] \frac{\sigma_k \sigma_k^T}{\sigma_k^T \gamma_k} - \frac{\sigma_k \gamma_k^T H_k + H_k \gamma_k \sigma_k^T}{\sigma_k^T \gamma_k}.$$

3. Проверить выполнение условия окончания поиска. Если условие выполняется, то останов алгоритма. Иначе положить k=k+1 и перейти к шагу 2.

2. Наиболее употребительные критерии для определения окончания процесса вычислений

1. $\max\left|\frac{\partial f}{\partial x_1}\right| < \varepsilon$ ($\varepsilon > 0$ задано).
2. $|\nabla f|^2 = \sum_{i=1}^{n}\left(\frac{\partial f}{\partial x_i}\right) < \varepsilon$ ($\varepsilon > 0$ задано).
3. $|f(x^{k+1}) - f(x^k)| < \eta$ ($\eta > 0$ задано).

Выбранный критерий останова алгоритма должен быть проверен на нескольких последовательных итерациях; в лабораторной работе выполнить проверку как минимум на трех итерациях.

3. Порядок выполнения лабораторной работы

3.1. Изучить предлагаемые градиентные методы конечномерной безусловной оптимизации, используя дополнительную литературу и конспект лекций, если необходимо.

2.2. В соответствии с вариантом задания, определенным преподавателем, в среде MATLAB составить программы, реализующие вышеописанные методы поиска (градиента с постоянным шагом, наискорейшего спуска, сопряженных направлений, сопряженного градиента для квадратичных функций, Флетчера–Ривза, Ньютона, Давида–Флетчера–Пауэлла, Бройдена–Флетчера–Гольдфарба–Шанно), и найти точку минимума функции $f(x)$ с заданной точностью ε, изменяемой в ходе исследования. Варианты задания взять в соответствующей таблице для лабораторной работы № 2 (с.7 настоящих методических указаний).

2.3. Оформить отчет о выполнении задания с приведением условия задачи, алгоритмов и программ указанных методов поиска, графиков зависимостей количества итераций от точности решения, таблицы результатов сравнения рассмотренных методов, заключения по результатам сравнения методов.

4. Контрольные вопросы

1. В чем основная особенность методов градиента с заданным шагом?
2. Приведите основной недостаток метода наискорейшего спуска на примере некоторых типов функций.
3. Как расположены последовательные направления перемещения метода наискорейшего спуска?
4. Почему метод минимизации должен обеспечить быструю сходимость на квадратичных функциях?

5. Какие направления называются сопряженными?

6. В чем заключается суть методов сопряженных направлений?

7. За какое количество шагов обеспечивается сходимость метода сопряженного градиента для квадратичных функций?

8. В чем отличие метода Флетчера–Ривза от метода сопряженного градиента?

9. Какую информацию необходимо хранить в памяти ЭВМ в методе Флетчера–Ривза?

10. За какое количество итераций сохранится метод Ньютона при применении к строго выпуклой квадратичной функции?

11. Обеспечивается ли метод Ньютона для произвольных функций?

12. Какому требованию должен удовлетворять гессиан в методе Ньютона?

13. Назовите способы получения положительно определенной матрицы в методе Ньютона.

14. Какой общий подход положен в основу квазиньютоновских методов?

15. Какому требованию отвечает аппроксимация обращения гессиана функций в квазиньютоновских методах?

16. Какие формулы коррекции использованы в методах ДФП и БФГШ?

17. Как ведет себя алгоритм ДФП к неточностям в подзадачах одномерной минимизации?

18. В чем отличие алгоритма Бройдена-Флетчера-Гольдфарба-Шанно от алгоритма Давида-Флетчера-Пауэлла?

19. Чему пропорциональна загрузка памяти в квазиньютоновских методах?

20. Оцените объем промежуточных вычислений в квазиньютоновских методах.