«НАЦИОНАЛЬНЫЙ ИССЛЕДВАТЕЛЬСКИЙ УНИВЕРСИТЕТ ИТМО»

Факультет информационных технологий и программирования

Кафедра компьютерных технологий

ОТЧЁТ ПО ЛАБОРАТОРНОЙ РАБОТЕ №4

Изучение алгоритмов метода Ньютона и его модификаций, в том числе квазиньютоновских методов.
Вариант 2

Выполнили студенты:

Ефимов Сергей Алексеевич
группа: M3237

Соколов Александр Андреевич
группа: M3234

Проверил:

Свинцов Михаил Викторович

г. Санкт-Петербург

Постановка задачи

Задача лабоpатрной работы – Разработать программы для безусловной минимизации функций многих переменных

Реализовать алгоритмы метода миимизации функции Ньютона:

- Классический
- С одномерным поиском
- С направлением спуска

Также необходимо провести исследование работы методов на различных функциях. В ходе работы необходимо разработать квазиньютоновский метод Бройдена-Флетчера-Шено и метод Пауэлла, проанализировать и сравнить с наилучшим методом Ньютона.

Ход работы

Теория

Рассмотрим задачу итерационной минимизации. Пусть $f(x) \in E^n$ - минимизируемая функция. Также $f(x)$ дважды дифференцируема, тогда, начав с точки x_0 мы можем построить квадратичную аппроксимацию $f(x)$ в окрестности x_0 на основе формулы Тейлора:

$$f(x) = f(x^{k-1}) + \left\langle \nabla f(x^{k-1}), x - x^{k-1} \right\rangle - \frac{1}{2} \left\langle H(x^k)(x - x^{k-1}), x - x^{k-1} \right\rangle + o(|x - x^{k-1}|)^2$$

Пренебрегая остатком в форме Пеано получим:

$$\phi_k(x) = f(x^{k-1}) + \left\langle \nabla f(x^{k-1}), x - x^{k-1} \right\rangle - \frac{1}{2} \left\langle H(x^k)(x - x^{k-1}), x - x^{k-1} \right\rangle$$

Если матрица Гессе является положительно определенной то ϕ_k имеет единственную точку минимума, которая и является следущей точкой итерационной последовательности. Данная точка может быть найдена из следущего условия:

$$\nabla \phi_k(x) = \nabla f(k^{k-1}) + H(x^{k-1})(x - x^{k-1}) = 0$$

Исходя из формулы:

$$\nabla(a^T x) = a$$

$$\nabla(x^T A x) = 2Ax$$

Тогда следущая точка релаксационной последовательности будет вычисляться как:

$$x^k = x^{k-1} - H^{-1}(x^{k-1}) \nabla f(x^{k-1}); k \in N$$

Пусть x'^k - вспомогательная точка релаксационной последовательности, тогда x^k можно найти как:

$$x^k = x^{k-1} + \alpha_k(x'^k - x^{k-1}) = x^{k-1} + \alpha_k p^k$$

$$\alpha_k > 0$$

где: $p^k = x'^k - x^{k-1}$ направление спуска

Классический метод Ньютона

Алгоритм:

Пусть дано x_0 - начальное приближение, ϵ - точность, тогда:

1. Рассчитать $\nabla f(x)$ - Градиент от функции в текущем приближении, H - матрицу Гессе по формуле $\nabla^2 f(x)$
2. Решить СЛАУ $Hp^k = -\nabla f(x)$
3. Вычислить следущий $x^k = x^{k-1} + p^k$
4. Если $\|x^k - x^{k-1}\| < \epsilon$, что эквивалентно $\|p^k\|$, то текущее приближение является искомым решением, иначе повторить алгоритм с 1 пункта

- Минусы:
  - Если начальное приближение выбрать достаточно далеко от минимума, то метод не сходится, такт как не обладает глобальной сходимостью
- Плюсы:
  - Если H удовлетворяет условию Липшица в окрестности решения поставленной задачи, то метод обладает квадратичной сходимостью.

Метод Ньютона с одномерным поиском

Пусть x^{k-1} - одномерный поиск в направлении p^k:

Тогда $\alpha_k = \min_\alpha(f(x^k + \alpha p^k))$ - вычисляется для нового напрвления в вычислении текущего минимума.

Алгоритм очень мохож на предыдущий, но теперь дополнительно вычисляется α_k

Алгоритм:

1. Рассчитать $\nabla f(x)$ - Градиент от функции в текущем приближении, H - матрицу Гессе по формуле $\nabla^2 f(x)$
2. Решить СЛАУ $Hp^k = -\nabla f(x)$
3. $\alpha_k = \min_\alpha(f(x^k + \alpha p^k))$
4. Вычислить следущий $x^k = x^{k-1} + \alpha p^k$
5. Если $\|x^k - x^{k-1}\| < \epsilon$, что эквивалентно $\|p^k\|$, то текущее приближение является искомым решением, иначе повторить алгоритм с 1 пункта

- Минусы:
  - Эффективность алгоритма зависит от того, является ли p_k - направлением спуска
- Плюсы:
  - Алгоритм обладает глобальной сходимостью в отличие от классического метода

Метод Ньютона с направлением спуска

Если p_k - направление спуска:

$$(p^k)^T \nabla f(x^k) < 0$$

Иначе p^k - не напрвление спуска, тогда следует использовать $-\nabla(x^k)$, тогда:

$$H(x^k)p^k = -\nabla f(x^k) \Rightarrow$$

$$p^k = \begin{cases} p^k, & (p^k)^T \nabla f(x^k) < 0 \\ -\nabla f(x^k) & (p^k)^T \nabla f(x^k) > 0 \end{cases}$$

Данный метод позволяет предотвратить неверное направление поиска, которы связан с седловыми точками, а так же точками максимума. Остальные шаги аналогичны предыдущему методу .Метод так же обладает глобальной сходимостью.

Квазиньютоновские методы

Квазиньютоновские методы - методы оптимизации, основанные на накоплении информации о кривизне целевой функции по наблюдениям за изменением градиента, чем принципиально отличаются от ньютоновских методов. Класс квазиньютоновских методов исключает явное формирование матрицы Гессе, заменяя её некоторым приближением.

Квазиньютоновские методы:

- Объединяют в себе достоиства от наискорейшего спуска и метода Ньютона
- Не требуют обращение к матрице H
- Сохраняют высокую сходимость итерационной последовательности

Общий вид релаксационной последовательности

$$x^k = x^{k-1} + \alpha_k p^k$$

где p^k - направление спуска

$$p^k = G_k w^k, k \in N$$

$$w^k = -\nabla f(x^{k-1})$$

G_k - Положительно определенная матрица (n x n) специального вида

Вичисление матрицы G_k происходит следущим образом, Она должна сходится к обратной матрица Гессе при достаточно больших k, то есть:

$$G_{k \longrightarrow \infty} \longrightarrow H^{-1}(x^*)$$

где x^* - точка минимума

За счет этого метод может гарантировать высокую сходимость, присущую метода Ньютона

α_k выбирается одним из типовых способов:

1. Константное значение $\alpha_k = 1$
2. Дробление шага
3. Частоиспользуемый вариант выбора α_k использование исчерпывающего спуска напралении p^k

Идеей квазиньютоновских методов является удачный выбор апроксимации, который может занчительно сократить объем вычислений по сравнению с обращением H, тем самым упростить процедуру построения p^k

Методы обладают глобальной сходимостью.

Метод Бройдена-Флетчера-Шено

Метод Бройдена-Флетчера-Шено - один из наиболее широко применяемых квазиньютоновских методов

Свойства присущие данному методу:

- G_k - сохраняет положительную определенность
- Если G_k - симметричная, то G_{k+1} - тоже симметричная
- При минимизации квадаратичной функции с положительно опреленной матрицей A метод сводится к методу сопряженных направлений, точное решение не более чем за n итераций
- Матрицы G_k связанны равенством

  $$G_k \cdot A \cdot p_i = p_i$$

  Следовательно G_k - обратная матрица к Гессиану
- Если целевая функция не квадратичная, то метод не позволяет найти решение за конечое кол-во итераций. Для уменьшения ошибки принято первые n итераций $G_k = I$(единичная матрица)
- Если целевая функция квадаратичная то

  $$H^{-1} = \sum_{i=1} n \frac{\triangle x_i (\triangle x_i)^T}{i(\triangle x_i)^T \triangle w_i}$$

На первой итерации $G1 = I$ $w_1 = -\nabla f(x_0)$
$p_1 = w_1$
$\alpha_1 = \min_\alpha f(x_0 \alpha p_1)$
$x_1 = x_0 + \alpha_1 p_1$
$\triangle x_1 = x_1 - x_0$
Если k > 0, то $w_k = -\nabla f(x_{k-1})$
$\triangle w_k = w_k - w_{k-1}$
$p_k = G_k \cdot w_k$
$\alpha_k = \min_\alpha f(x_{k-1} \alpha_k p_k)$
$x_k = x_{k-1} + \alpha_k p_k$
$\triangle x_k = x_k - x_{k-1}$
Условие остановы: $\|\triangle x_k\| < \epsilon$

$$G_{k+1} = G_k - \frac{\triangle x_i (\triangle x_i)^T}{\langle \triangle w^k, \triangle x^k \rangle} - \frac{G_k \triangle w^k (\triangle w^k)^T G_k^T}{\rho_k} + \rho_k r^k (r^k)^T$$

$$r^k = \frac{G_k \triangle w^k}{\rho_k} - \frac{\triangle x^k}{\langle \triangle x^k, \triangle w^k \rangle}$$

$$\rho_k = \langle G_k \triangle w^k, \triangle w^k \rangle$$

Метод Пауэлла

Очень сильно похож на предыдущий метод, за исключением определения G_k:

$$G_{k+1} = G_k - \frac{\triangle x'_k (\triangle x'_k)^T}{\langle \triangle w_k, \triangle x'_k \rangle}$$

Демострация методов на различных функциях

Проведем исследование на двух функциях:

1. $\sin(x)$
2. $10x^2 + 2xy + 12y^2$

- – Классический метод Ньютона: $\sin(x)$

  Начальное приближение: $x_0 = 1$

Кол-во итераций	x_k	α_k	p_k	$f(x_k)$
0	1.6721	1	0.5623	0.9812215
1	1.5707	1	-0.6782	1
2	1.5707	1	0.0015	1
3	1.5707	1	0	1
4	1.5707	1	0	1

  Очевидно, что 1 не мининиум функции $\sin(x)$. Это наглядно показывает что классический метод Ньютона не имеет глобальной сходимости.