Лабораторная работа 2: Метод Ньютона

Численное нахождение градиента и гессиана

Численное дифференцирование — это метод приближённого вычисления производных функции, когда аналитическое вычисление затруднено или невозможно. Основная идея заключается в использовании разностных схем для аппроксимации производных.

Первая производная (градиент)

Для функции дифференцируемой $f(x)$ одной переменной первая производная может быть вычислена с помощью **центральной разностной схемы**:

$$f'(x) \approx \frac{f(x+h) - f(x-h)}{2h},$$

где h — малый шаг. Данное приближение разумно обосновывается, если рассмотреть формулу Тейлора для $f(x-h)$ и $f(x+h)$ при разложении в точке x.

Данный результат с легкостью обобщается на случай n переменных и поиск градиента, достаточно лишь научиться искать частную производную. Для удобства записи, примем обозначение e_i за i-ый координатный орт. Тогда

$$\frac{\partial f}{\partial x_i} \approx \frac{f(\mathbf{x} + h\mathbf{e}_i) - f(\mathbf{x} - h\mathbf{e}_i)}{2h}$$

Осталось вспомнить, что $\nabla f(\mathbf{x})$ - вектор частных производных.

Вторая производная (гессиан)

Для функции $f(x)$ одного переменного вторая производная вычисляется по формуле:

$$f''(x) \approx \frac{f(x+h) - 2f(x) + f(x-h)}{h^2}.$$

Аналогично первой производной, приближение подтверждается разложением $f(x+h)$ и $f(x-h)$ в точке x.

Для формулы в частных производных ничего неожиданного не происходит:

$$\frac{\partial^2 f}{\partial x_i \partial x_j} \approx \frac{f(\mathbf{x} + h\mathbf{e}_i + h\mathbf{e}_j) - f(\mathbf{x} + h\mathbf{e}_i - h\mathbf{e}_j) - f(\mathbf{x} - h\mathbf{e}_i + h\mathbf{e}_j) + f(\mathbf{x} - h\mathbf{e}_i - h\mathbf{e}_j)}{4h^2}.$$

Из вторых производных можно составить гессиан $\nabla^2 f(\mathbf{x})$. Ради некоторой вычислительной оптимизации стоит вспомнить, что для дважды непрерывно дифференцируемых функций, гессиан симметричен.

Практическое задание

1. Реализуйте методы для численной оценки градиента и гессиана.

2. Оцените зависимость точности градиента и гессиана от выбора шага h. Постройте график нормы разности полученных значений от найденных аналитически (шкала логарифмическая по обоим осям). Оценку производить на функциях:

   (a) Комбинация из пары тригонометрических функций;

   (b) Полином нескольких переменных степени 3;

   (c) Композиция, содержащая exp, ln, arctan и возведение в степень.

   Выбор точек, в которых производится оценка, остаётся также на ваш выбор.

3. Предложите некоторый способ создания функции $f(x)\colon \mathbb{R}^n \to \mathbb{R}$ с произвольным n так, чтобы вычислительная сложность росла пропорционально n (функция не меняла характер в зависимости от n). Произведите оценку времени, необходимого на вычисление градиента и гессиана в зависимости от размерности n, постройте график. Сравните результат с теоретическими ожиданиями.

Метод Ньютона

1. Реализуйте метод Ньютона используя удобный вам язык программирования. Учтите возможность выбора коэффициента α как $\alpha = 1$, так и с подбором согласно стратегии Армихо-Вульфа.

2. Протестируйте метод на квадратичной функциях вида $x^T A x + b^T x$ из предыдущей лабораторной. Опишите поведение метода и сходимость с теоретическими ожиданиями.

3. Протестируйте метод на функции $f(x, y) = \dfrac{-1}{1 + (x-a)^2 + (y-b)^2}$, коэффициенты a и b возьмите как среднее гармоническое и среднее геометрическое из номеров ИСУ членов вашей команды, деленное на 100000. Для неё:

   (a) Предложите некоторый набор начальных точек (не меньше 3х), когда метод с $\alpha = 1$ даёт сходимость. Изобразите график ошибки - нормы разности между настойщим оптимумом и координатами текущего шага - в зависимости от числа итераций. Оцените скорость сходимости и сравните с теоретической

   (b) Эмпирически оцените область, в которой метод сходится и границу этой области (учтите симметрию функции). Обоснуйте сходимость в этой области аналитически.

   (c) Примените стратегию с подбором шага α. Как изменилась область? Продемонстрируйте графики сходимости, оцените скорость сходимости и сравните с теоретической.

4. Исследуйте функцию $f(x, y) = -9x - 10y + 10(-\ln(100 - x - y) - \ln(x) - \ln(y) - \ln(50 - x + y))$.

   (a) Опишите область определения данной функции (4 неравенства);

   (b) Постройте график данной функции (учтите, возможно, стоит несколько отдалиться по со значениями функции);

   (c) Попробуйте метод Ньютона с постоянным шагом для следующих начальных точек:

   $$(8, 90), (1, 40), (15, 68.69), (10, 20)$$

   Какое поведение наблюдается? Приложите иллюстрации этого поведения.

   (d) Повторите эксперимент из предыдущего пункта, поменяв стратегию выбора шага. Какое поведение наблюдается теперь? Приложите иллюстрации.

Отчет

Отчет должен состоять из:

1. Кода для каждого из разделов;
2. Предложенных к построению графиков;
3. Аналитических выкладок;
4. Выводов к каждому разделу.

Может пригодится

- Вычисление нормы средставами numpy
- Библиотека для оценки времени работы небольших кусочков кода