Глобальные градиентные методы оптимизации

Судаков Илья

28 января 2023 г.

1 Градиентный спуск

Пусть целевая функция имеет вид: $F(X) : X \to R$

Задача оптимизации задана в следующем виде: $F(X) \to min$

Основная идея метода заключается в том, чтобы идти в направлении наискорейшего спуска, а это направление задаётся антиградиентом $-\nabla F$

и $X^{[i+1]} = X^{[i]} - \lambda^{[i]} \nabla F(X^{[i]})$ где $\lambda^{[j]}$ задает скорость градиентного спуска и может быть выбрана как

- постоянная, тогда метод может не сходиться
- убывать по какому-то закону
- гарантировать наискорейший спуск

1.1 Алгоритм градиентного спуска

- Задают начальное приближение и точность расчёта: X_0, ε
- Рассчитывают $X^{[i+1]} = X^{[i]} - \lambda^{[i]} \nabla F(X^{[i]})$
- Проверяют условие остановки (на усмотрение):
  - $|x^{[j]} - x^{[j+1]}| < \varepsilon$
  - $|F(x^{[j]}) - F(x^{[j+1]})| < \varepsilon$
  - $\nabla F(X^{[i]}) < \varepsilon$
  - иначе переходят к пункту 2

Этот алгоритм будет реализован на python

1.2 Метод наискорейшего спуска

В случае наискорейшего спуска $\lambda^{[j]}$ определяется как:

$$\lambda^{[j]} = argminF(X^{[j+1]}) = argminF(X^{[j]} - \lambda \nabla F(X^{[i]}))$$

чтобы вычислить $\lambda^{[j]}$ будем использовать метод золотого сечения

2 Метод золотого сечения

2.1 Описание

Метод золотого сечения — это эффективная реализации троичного поиска, служащего для нахождения минимума/максимума унимодальной функции на отрезке. Лучше тем, что на каждой итерации вычисляется значение в одной, а не двух точках.

Пусть хотим найти минимум на отрезке $[a, b]$, тогда разобьем его на 3 части 2 точками s_1, s_2, так чтобы при отсекании одного из крайних подотрезков, одна из точек осталась границей подотрезков, а соотношение длин подотрезков оставалось прежним.
$|[a, s_1]| = |[s_2, b]| = l_1$
$|[s_1, s_2]| = l_2$
тогда: $\frac{l_2}{l_1} = \frac{l_1 - l_2}{l_2} \Rightarrow l_1^2 - l_1 \cdot l_2 - l_2^2 = 0 \Rightarrow \frac{l_1}{l_2} = \frac{1+\sqrt{5}}{2} = \phi$

2.2 Алгоритм тернарного поиска

написать ли

2.3 Оценка сложности

Так как на каждой итерации длина рассматриваемого отрезка умножается на $\frac{l_1+l_2}{l_1+l_2+l_2} = \frac{3+\sqrt{5}}{4+2\sqrt{5}} = \phi^{-1}$, то для достижения точности δ потребуется $\frac{\ln(\frac{|[a,b]|}{\delta})}{\ln(\phi)} \approx 2\ln(\frac{|[a,b]|}{\delta})$ итераций.

2.4 Применение

А что если функция не унимодальная, тогда к чему сойдется этот метод? К точке, которая является одним из локальных минимумов функции либо одному из концов, в котором производная неположительная.

3 Глобальная одномерная оптимизация

Алгоритм поиска глобального минимума функции вдоль одного направления будет основываться на интервальном анализе. Перечислим основные определения и теоремы, которые нам пригодятся для посторения алгоритма.

3.1 Определение

Интервалом $[a, b]$ называется следующее множество:

$$[a, b] := \{x \in \mathbb{R} | a \leq x \leq b\}$$

3.2 Арифметические свойства интервалов

- $\mathbf{x} + \mathbf{y} = \underline{\mathbf{x}} + \underline{\mathbf{y}}, \overline{\mathbf{x}} + \overline{\mathbf{y}}$
- $\mathbf{x} - \mathbf{y} = \underline{\mathbf{x}} - \overline{\mathbf{y}}, \overline{\mathbf{x}} - \underline{\mathbf{y}}$
- $\mathbf{x} \cdot \mathbf{y} = min(\underline{\mathbf{x}}\underline{\mathbf{y}}, \underline{\mathbf{x}}\overline{\mathbf{y}}, \overline{\mathbf{x}}\underline{\mathbf{y}}, \overline{\mathbf{x}\mathbf{y}}), max(\underline{\mathbf{x}}\underline{\mathbf{y}}, \underline{\mathbf{x}}\overline{\mathbf{y}}, \overline{\mathbf{x}}\underline{\mathbf{y}}, \overline{\mathbf{x}\mathbf{y}})$
- $\mathbf{x} \div \mathbf{y} = \mathbf{x} \cdot [1/\overline{\mathbf{y}}, 1/\underline{\mathbf{y}}]$, если $0 \notin y$

а как быть с другими функциями

3.3 Основная теорема интервальной арифметики

Пусть $f(x_1, ..., x_n)$ - рациональная функция вещественных аргументов $x_1, ..., x_n$, и для нее определен результат $\mathbf{F}(\mathbf{X_1}, ..., \mathbf{X_n})$ подстановки вместо аргументов интервалов их изменения $(X_1, ..., X_n) \subset \mathbb{R}^n$ и для $(X_1, ..., X_n)$ операциии выполняются по правилам интервальной арифметики. Тогда

$$\{f\{x_1, ..., x_n\} | x_1 \in \mathbf{X_1}, ..., x_n \in \mathbf{X_n}\} \subset \mathbf{F}(\mathbf{X_1}, ..., \mathbf{X_n})$$

как лучше преобразовать выражение чтобы функция быстрее сходилась

3.4 Утверждение

Пусть $f(x_1, ..., x_n)$ - рациональная функция вещественных аргументов $x_1, ..., x_n$, и $\mathbf{F}(\mathbf{X_1}, ..., \mathbf{X_n})$ соответствующая ей интервальная функция, тогда выполняется монотонность по включению. Пусть $\mathbf{X_1}, ..., \mathbf{X_n}$ и $\mathbf{Y_1}, ..., \mathbf{Y_n}$, такие что $\mathbf{X_1} \subset \mathbf{Y_1}, ..., \mathbf{X_n} \subset \mathbf{Y_n}$, тогда

$$f\{\mathbf{X_1}, ..., \mathbf{X_n}\} \subset f\{\mathbf{Y_1}, ..., \mathbf{Y_n}\}$$

3.5 Определение

Пусть $N > 0$ натуральное число. Тогда если $\langle S_0, ..., S_{n-1}\rangle$ семейство непустых подмножеств множества S, тогда будем называть его разбиением множества S. В частности, обьединение $S_0, ..., S_{n-1}$ равно S , тогда последовательность $\langle S_0, ..., S_{n-1}\rangle$ является покрытием S .

3.6 Теорема

Рассмотрим задачу безусовной глобальной оптимизации. Пусть $\langle B_0, ..., B_{n-1}\rangle$ семейство множеств, содержащее глобальный минимум, такое что упорядочено по возрастанию нижней границы $f(B_i)$ для $i = 0, 1, 2, ..., N-1$. Пусть U наименьшая из верхних значений функции для подмножеств $\langle f(B_0), ..., f(B_{n-1})\rangle$. Тогда интервал $[lb(f(B_0)), U]$ содержит глобальным минимум μ.

используя это напишем алгоритм поиска глобального минимума вдоль направления

Список литературы

Многие методы оптимизации (координатный спуск, градиентный спуск, метод сопряженного гадиента и др.) используют так называемый локальный поиск, т.е. поиск минимума одномерной функции вдоль луча в n-мерном евклидовом пространстве. При этом всегда находится локальный минимум. В проекте предлагается исследовать потенциал применения методов глобальной одномерной оптимизации для организации линейного поиска. В каких случаях удастся найти глобальной минимум многомерной функции, используя глобальный линейный поиск в контексте классических градиентных методов.

1. Программная реализация методов глобального линейного поиска. 2. Автоматическая проекция функции на направление с использованием пакета sympy. 3. Реализация алгоритма оптимизации многомерной функции. 4. Экспериментальное и теоретическое исследование.

"1. Алгоритм поиска глобального минимума одномерной функции. 2. Алгоритм оптимизации многомерной функции, основанный на комбинировании классического градиентного подхода и глобального линейного поиска. 3. Программная реализация алгоритма на языке Python. 4. Результаты вычислительных экспериментов. 5. Возможно написание статьи по результатам проекта."