МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования
**«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМЕНИ Н.Г. ЧЕРНЫШЕВСКОГО»**

Кафедра математического анализа

Градиентные методы на классах выпуклых функций

АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ

Студентки 4 курса 421 группы
направления 02.03.01 Математика и компьютерные науки

механико-математического факультета

Гахраманова Сурхая Галибовича

Научный руководитель
доцент, к.ф.-м.н. М. А. Осипцев
подпись, дата

Заведующий кафедрой
д.ф.-м.н., профессор Д. В. Прохоров
подпись, дата

Саратов 2020

Введение. Данная дипломная работа посвящена изучению выпуклых и слабо выпукных множеств и субградиентным методам решения задач выпуклой минимизации.

Раздел математики, изучающий выпуклые и слабо выпуклые множества и выпуклые экстремальные задачи, называется выпуклым анализом. Понятие выпуклости играет важную роль в различных областях фундаментальной и прикладной математики. Основные факты и понятия выпуклого анализа сформировались еще в 18-м и начале 19-го столетия. В конце 19-го столетия и начале 20-го столетия. Основные понятия выпуклой геометрии, такие как опорная функция, поляра, крайняя точка, сыграли большую роль в создании в начале 20-го века функционального анализа. Примечательной особенностью выпуклых множеств является возможность их двойного описания, прямого (на основе определения, т. е. если две точки принадлежат выпуклому множеству, то и весь отрезок с концами в указанных точках также принадлежит данному множеству) и двойственного (выпуклое множество может быть представлено как пересечение полупространств). В результате для каждого выпуклого множества можно указать двойственное ему множество, называемое полярой. Это свойство позволяет получить двойственное описание и для выпуклых функций. С каждой выпуклой функцией связана двойственная, или сопряженная, получаемая из исходной преобразованием, впервые введенным для выпуклых функций еще в 18-м столетии А.М. Лежандром. Выпуклый анализ находит многочисленные приложения в вариационном исчислении и математической теории управления, в теоретической механике и теории упругости, теории приближений и экономике.

Задачи дипломной работы:

1. Ввести определение выпуклого и слабо выпуклого множества и его свойства.

2. Постановка задачи выпуклого программирования и его решение.

3. Описать субградиентные методы для слабо выпуклых функций.

Градиентные методы решения задач выпуклого программирования

Типичные методы для задач статистики и обработки сигналов следуют двухэтапной стратегии: (1) Найти умеренно точное решение $\hat{x}$ при низкой стоимости сложности выборки (например, с использованием спектральной инициализации), и (2) уточнить $\hat{x}$ итерационным "локальным алгоритмом поиска который быстро сходится при естественных статистических предположениях. Для простых формулировок задач, термин "локальный поиск"почти везде относится к методам градиентного спуска или его близкому варианту. Важными требованиями к этим методам является выпуклость множества на которых рассматривается эксремальная задача и выпуклость функции. Для негладких и невыпуклых задач смысл локального поиска гораздо менее ясен. В этой работе, мы задаем следующий вопрос.

Существует ли общая процедура локального поиска на основе градиента для негладких и невыпуклых задач, которая сходится линейно при стандартных условиях регулярности?

Неудивительно, что наш подход основан на субградиентных методах выпуклой оптимизации. Для того чтобы мотивировать обсуждение, рассмотрим задачу ограниченной оптимизации

$$\min_{x \in X} g(x),$$

где g - выпуклая Липшицева функция на R^n, а X - замкнутое выпуклое множество. Учитывая текущую итерацию x_k, субградиентные методы выполняются следующим образом:

Здесь символ $proj_X(y)$ обозначает ближайшую точку от X до y, а α_k - заданную последовательность шагов. Выбор последовательности α_k определяет поведение схемы и является главной отличительной чертой среди субградиентных методов. В этой работе нас будут интересовать только субградиентные методы, которые линейно сходятся. Как обычно, линейные скорости сходимости итерационных методов требуют выполнения некоторых условий регулярности. Здесь соответствующим условием регулярности является резкость (или, что тоже самое, глобальная ошибка): существует реальная $\mu > 0$, удовлетворяющая

$$g(x) - \min_{x \in X} g \geq \mu \dot{d}ist(x; X^*) \; \forall x \in X,$$

где X^* обозначает множество минимизирующих переменных в (1.1). Предполагая, что резкость имеет место, как субградиентные методы, с разумным выбором α_k, производить итерацию, сходящуюся к X^* при линейной частоте $\sqrt{1 - (\mu/L)^2}$. Результаты этого типа датируются в 60-ых и 70-ых годах.

Различные современные проблемы приводят к постановкам которые действительно актуальны, но являются только слабо выпуклыми и локально Липшицевыми. Напомним, что функция g является ρ-слабо выпуклой, если возмущенная функция $x \longmapsto g(x) + \frac{\rho}{2}|| \cdot ||^2$ является выпуклой для некоторой $\rho > 0$. Заметим, что слабо выпуклые функции не обязательно должны быть гладкими или выпуклыми. Быстрое вычисление, приведенное далее в Лемме 2.1 показывает, что если g является μ-резко и ρ-слабо выпуклым, то существует трубка вокруг множества решений X^*, которая не содержит посторонних стационарных точек:

В данной работе покажем, что стандартные линейно сходящиеся субградиентные методы, первоначально разработанные для выпуклых задач, применяются в этой гораздо большей общности, если они инициализируются в пределах небольшого сжатия трубки $\mathcal{T}$. По существу методы показывают ту же самую линейную частоту сходимости, что и в выпуклом случае, в то время как слабая константа выпуклости только определяет правильность инициализации. Фокусируем свое внимание на трех правилах размера шага: размер шага поляка, геометрически распадающийся шаг и постоянный размер шага. В качестве доказательства концепции иллюстрируем полученные алгоритмы на задачах поиска фазы и оценки ковариации.

В данной работе рассматривается адаптация градиентных методов в случае слабовыпуклых функций. В данном разделе мы рассмотрим сущесввующие градиентные методы нахождения экстремальных выпуклых функций на выпуклом множестве, а также адаптацию этих методов в случае если максимизирующая функция будет слабо выпуклой. Текущая работа находится в более широком диапазоне анализа субградиентных и проксимальных методов для слабо выпуклых задач. В частности, доказывается глобальная сублинейная скорость сходимости в терминах естественной меры стационарности (стохастического) субградиентного метода на любой слабо выпуклой функции. Напротив, здесь интересуют субградиентные методы, которые локально линейно сходятся в допущении дополнительной резкости. Аргументы, которые представляем, - это все быстрая модификация доказательств, в уже установленных и доступных выпуклых множествах. Тем не менее, считаем, что сделанные выводы являются интересными и мощными, с помощью этого открываются двери для общих локальных процедур поиска негладких и невыпуклых проблем.

Определения и примеры выпуклых множеств.

Определение 1. Множество $A \subset \mathbb{R}^n$ называется выпуклым, если для любого числа $\lambda \in (0,1)$ и любых точек x_1, $x_2 \in A$ справедливо включение

$$\lambda x_1 + (1-\lambda)x_2 \in A$$

Определение 2. *Выпуклой оболочкой множества A называется пересечение всех выпуклых множеств, содержащих множество A.*

Приведем некоторые примеры выпуклых множеств.

Все пространство $\mathbb{R}^n$ и всякое линейное подпространство пространства являются выпуклыми множествами.

Множество

$$H_p = \{x \in \mathbb{R}^n | \, \langle p, x \rangle = \alpha\}$$

при $p \in \mathbb{R}^n, p \neq 0$ и $\alpha \in R$, называемое гиперплоскостью, а также множества

$$H_p^* = \{x \in \mathbb{R}^n | \, \langle p, x \rangle \ \geq \alpha\}$$

$$H_p^- = \{x \in \mathbb{R}^n | \, \langle p, x \rangle \ \leq \alpha\}$$

называемые полупространствами, очевидно, являются выпуклыми множествами.

Определение 3. Шаром с центром в точке $a \in \mathbb{R}^n$ радиуса $r > 0$ называется множество

$$B(\alpha, r) = \{x \in \mathbb{R}^n : ||x - a|| \leqslant r\}$$

Определение 4. Открытым шаром с центром в точке $a \in \mathbb{R}^n$ радиуса $r > 0$ называется множество

$$B(\alpha, r) = \{x \in \mathbb{R}^n : ||x - a|| < r\}$$

Шары $B(\alpha, r)$ и $B(\alpha, r)^\circ$ являются выпуклыми множествами, так как для любых точек $x_1, x_2 \in B(\alpha, r)$ при любом числе $\lambda \in (0, 1)$ из неравенства треугольника получаем

$$||\lambda x_1 + (1-\lambda)x_2 - a|| \leq \lambda||x_1 - a|| + (1-\lambda)||x_2 - a|| \leq \varepsilon$$

$$\lambda x_1 + (1-\lambda)x_2 \in B(\alpha, r)$$

Важным примером выпуклого множества является аффинное множество. Множество A называется аффинным множеством, если для любых точек $x, y \in A$ и любого числа $\lambda \in R$ справедливо включение

$$\lambda x + (1-\lambda)y \in A$$

, т.е. прямая, проходящая через любые две точки $x, y \in A$, целиком принадлежит множеству A.

Всякое аффинное множество может быть представимо в виде суммы некоторого линейного подпространства L и произвольной точки $a \in A$. В самом деле, возьмем произвольную точку $a \in A$ и определим множество L по формуле $L = A - a$. Ясно, что $0 \in L$. Для любых $b, c \in L$ получаем $b + a \in A$, $c + a \in A$, откуда для любого $\lambda \in R$ имеем

$$\lambda(b+a) + (1-\lambda)(c+a) \in A4$$

т.е. $\lambda b + (1-\lambda)c \in L$, т.е.$L$ есть аффинное множество. Поэтому для любого $\lambda \in R$ получаем $\lambda b = \lambda b + (1-\lambda)0 \in L$, и так как

$$(b+c)/2 = (1/2)b + (1-1/2)c \in L$$

то отсюда получаем, что $c+b \in L$. Итак, показали. что $0 \in L$, для любых $b, c \in L$ и для любого $\lambda \in R$ следует, что L есть линейное подпространство. Также легко проверить, что подпространство $L = A - a$ не зависит от выбора точки $a \in A$, а однозначное определяется аффинным множеством A.

В данном случае подпространство L называется подпространством, параллельным аффинному множеству A. Если подпространство L конечномерно, то размерностью аффинного множества A по определению называют размерность подпространства L.

Выпуклое множество называется строго выпуклым, если его граница не содержит отрезков. Так, например шар $B(\alpha, r) \subset R^n$ является строго выпуклым множеством, а выпуклое множество

$$A = \{x \in R^n | \max_{k \in \overline{1,n}} |x_k| \geq 1\}$$

не является строго выпуклым.

Заключение. В данной работе были рассмотрены основные определения и примеры выпуклых множеств, описан класс задач градиентных методов. Градиентные методы — численные методы решения с помощью градиента задач, сводящихся к нахождению экстремумов функции.