Лабораторная работа №4
Методы стохастической оптимизации

https://github.com/GooddiLK/AllEDa

Пластинин Алексей M3237
t.me/plstnn

Малков Александр M3237
t.me/AlexM_37

Кинзябулатов Эдуард M3237
t.me/Eduard7000

Цель работы:

Разобрать теоретическое описание и реализовать метод имитации отжига, после чего провести сравнение его эффективности на методах и примерах из предыдущих лабораторных.

Используемые методы:

- Алгоритм имитации отжига
- Генетический алгоритм

Реализация метода отжига

Алгоритм начинает с предоставленной точки с некоторой "температурой" t равной $t_{\max}$. Далее на каждом шаге происходит следующее:

- Алгоритм выбирает случайную точку из куба с длиной ребра t. Если значение функции в этой точки оптимальнее предыдущего принятого значения, то новое значение принимается сразу. В противном случае точка принимается с вероятностью равной $e^{\frac{-\Delta F}{t}}$, где ΔF – разность между значением в новой точке и в текущей.
- t меняется на αt, где α – константа от 0 до 1.

Результатом работы алгоритма считается последняя принятая точка.

Реализация генетического алгоритма

Генетический алгоритм состоит из 4 фаз:
1) Начальная популяция
2) Скрещивания и мутация
3) Селекция
4) Формирование нового поколения, переход к фазе 2.

Если после 4 фазы выполнено некоторое условие, то алгоритм прекращает работу.
Реализованный алгоритм принимает следующие параметры: количество шагов, границы области в которой необхоимо искать минимум, размер поколения, коэффициент мутирования.
Сначала генерируется начальная популяция путем выбора необходимого количества случайных точек из обозначенных границ.
Далее для каждого поколения выбирается $\sqrt{n}$ лучших представителей, где n – размер поколения, после чего создаётся n представителей следующего поколения (скрещивание): выбирается точка в середине отрезка, соединяющего 2 выбранных родителя. Пусть полученная точка – точка A. Для мутирования выбирается случайная точка B в указанных границах. Результатом мутирования считается точка M, такая что $|MB| = \mu|AB|$, где μ – указанный коэффициент мутирования.
Описанный процесс продолжается указанное число шагов, результатом работы алгоритма считается лучшая его точка.

Исследование

- $(x^2 + y - 11)^2 + (x + y^2 - 7)^2$ – Функция Химмельблау
- $20 + (x^2 - 10\cos(2\pi x)) + (y^2 - 10\cos(2\pi y))$ – Функция Растригина

Метод отжига

Стартовая точка для всех функций: $(100, -200)^T$
Параметры: $t_{\max} = 10^8$, $t_{\min} = 10^{-3}$, $\alpha = 0.99$

	Вызовы F	**Расстояние до оптимума**
$x^2 + y^2$	2522	0.14672
$3x^2 - 4xy + 10y^2$	2522	0.04792
Функция Химмельблау	2522	0.00423
Функция Растригина	2522	2.81143

Параметры: $t_{\max} = 10^8$, $t_{\min} = 10^{-3}$, $\alpha = 0.95$

	Вызовы F	**Расстояние до оптимума**
$x^2 + y^2$	495	0.35381
$3x^2 - 4xy + 10y^2$	495	0.40258
Функция Химмельблау	495	0.01828
Функция Растригина	495	1.41419

Параметры: $t_{\max} = 100$, $t_{\min} = 0.1$, $\alpha = 0.97$

	Вызовы F	**Расстояние до оптимума**
$x^2 + y^2$	228	0.47057
$3x^2 - 4xy + 10y^2$	228	0.43186
Функция Химмельблау	228	0.06691
Функция Растригина	228	2.01399

Параметры: $t_{\max} = 1000$, $t_{\min} = 0.1$, $\alpha = 0.95$

	Вызовы F	**Расстояние до оптимума**
$x^2 + y^2$	181	1.62550
$3x^2 - 4xy + 10y^2$	181	0.22678
Функция Химмельблау	181	0.01729
Функция Растригина	181	2.23552

В таблицах представлены лишь некоторые из произведенных запусков, по их результатам можно понять, что метод отжига позволяет значительно (на порядок) менять количество вычислений функции, но тогда значительно повышается вероятность большей ошибки. Так для функции $x^2 + y^2$ для который в первом варианте запуске с большей вероятностью ошибка была в пределах $5 \cdot 10^{-2}$, ошибка, измеряемая единицами, встречается раз в 4-5 запусков.
Результаты работы других методов для сравнения можно посмотреть в отчете по второй лабораторной на GitHub. Основным преимуществом метода отжига является то, что он не вычисляет градиент функции, недостатком же является его случайность, которую можно уменьшить, но ценой увеличения количества вычислений функции.

Генетический алгоритм

Исследуемая область: прямоугольник $[(-100, -100)^T; (50, 50)^T]$
Параметры: количество шагов – 10, размер поколения – 100, коэффициент мутирования – $5 \cdot 10^{-3}$

	Вызовы F	**Расстояние до оптимума**
$x^2 + y^2$	1100	0.07640
$3x^2 - 4xy + 10y^2$	1100	0.02843
Функция Химмельблау	1100	0.02305
Функция Растригина	1099	1.01044

Для указанных параметров метод для функций 1-3 почти всегда ошибается в пределах 10^{-1}. Для функции 4 ошибка лежит в диапазоне от 1 до 2. Если при тех же параметрах изменить коэффициент мутирования на $5 \cdot 10^{-4}$, то ошибка для функции 4 не изменится, а её порядок для функций 1-3 возрастёт до -1. Если же изменить количество шагов на 10^3, то ошибка для функций практически не поменяется (количество вызовов функции увеличится в 100 раз!), однако же если применить оба описанных изменения параметров, т.е. вызывать функцию с параметрами: количество шагов – 1000, размер поколения – 100, коэффициент мутирования – $5 \cdot 10^{-4}$, то порядок ошибок для функций 1-3 уменьшится на 1, а для функции 4 погрешность может увеличится на 1. Увеличение коэффициент мутирования с одной стороны на старте позволяет получить больше представителей приближенных к оптимуму, с другой стороны в конце мешает получить более точный результат.

Если при неизменных остальных параметрах увеличить в 10 раз размер поколения, то порядок ошибки для функций 1-3 уменьшится на 1, а для функции 4 – на 2 (однако всё ещё достаточно часто ошибка измеряется в единицах, причём уменьшение в 10 раз коэффициента мутирования значительно снижает вероятность происхождения этого). Если всё при тех же параметрах ещё раз увеличить в 10 раз размер поколения, то никакого видимого эффекта это не окажет. В случае одновременного уменьшения коэфициента мутирования в 10 раз, порядок ошибки для всех функций уменьшится ещё на 1, т.е. станет равным -4.

Попробуем приблизить количество вычислений функции к тому же порядку, что получался при использовании методов из предыдущих лабораторных: с 2 шагами, 100 представителями в одном поколении и коэффициентом мутирования 0.005 оно будет равно 300. Порядок погрешности у метода будет -1. Преимущества и недостатки этого метода в целом такие же как у метода отжига: мы не вычисляем градиент, но для того, чтобы повысить вероятность более оптимального ответа должны во много раз увеличивать вычисление функции.

Применение методов стохастической оптимизации

Рассмотрим задачу оптимизации инвестиционного портфеля с фиксированными транзакционными издержками.
Цель: Инвестор с начальным вектором долей капитала $w^0 = [w_1^0, w_2^0, ..., w_n^0]^T$ $(\Sigma_{i=1}^n w_i^0 \leq 1)$ хочет перераспределить капитал между активами и получить следующий вектор долей капитала $w = [w_1, w_2, ..., w_n]^T$ $(\Sigma_{i=1}^n w_i = 1)$, минимизируя риск (дисперсию доходности) при заданной целевой ожидаемой доходности (математическом ожидании) и учитывая фиксированные издержки f_i за смену w_i^0 на w_i. Пересмотр портфеля происходит единоразово.
Задача сводится к комбинаторной стохастической оптимизации.
Можно накладывать множество дополнительных ограничений, например, ограничение на сумму, которую можно потратить на издержки.
Пример задачи:

Актив	Доходность	Риск (σ)
Акции A	10%	15%
Облигации Б	5%	10%
Золото	3%	8%

Если инвестор стремится к доходности 7%, оптимальное распределение выглядит так: Акции A: 40%, Облигации Б: 50%, Золото: 10%. *Источник*