Методы стохастической оптимизации

Арсланов Николай Адельевич, гр. 622

Санкт-Петербургский государственный университет
Математико-механический факультет
Кафедра статистического моделирования

Научный руководитель — д.ф.-м.н. **С. М. Ермаков**
Рецензент — к.ф.-м.н. **К. А. Тимофеев**

Санкт-Петербург
2022г.

Введение

Стохастические методы оптимизации — это класс методов, в которых используется случайность в процессе решения оптимизационных задач. Данная особенность позволяет эффективно работать со сложными целевыми функциями (с большим количеством переменных, экстремумов, разрывов и пр.).

Эволюционные алгоритмы — это разновидность стохастических методов оптимизации, основная идея которых заключается в том, что решения могут быть улучшены путем случайного преобразования параметров, выделении новых решений и выбора наилучших из них.

В работе рассматриваются следующие эволюционные алгоритмы:

1. алгоритм CMA-ES;
2. модифицированный эволюционный алгоритм;
3. алгоритм (μ, λ)-ES.

Постановка цели и задач

Цель работы — выполнение следующих экспериментов:

1. Определение зависимости между средним количеством итераций для схождения алгоритмов и количеством используемых наилучших решений для генерации новых решений;
2. Применение эволюционных алгоритмов для решения систем линейных уравнений с плохой обусловленностью и систем нелинейных уравнений с высокой чувствительностью к решениям.

Основные задачи:

1. Привести общие сведения об эволюционных алгоритмах;
2. Описать методы оптимизации (μ, λ)-ES, CMA-ES и ее модификации;
3. Определить зависимость числа итераций от количества наилучших решений;
4. Решить системы линейных / нелинейных уравнений;
5. Проанализировать полученные результаты.

Функция сферы:

$$f\left(x_1,\ldots,x_n\right)=\sum_{i=1}^{n}x_i^2$$

- глобальный минимум $f(0,\ldots,0)=0$
- область генерации начального решения $[-100,100]$

Функция Растригина:

$$f\left(x_1,\ldots,x_n\right)=10n+\sum_{i=1}^{n}\left[x_i^2-10\cos\left(2\pi x_i\right)\right]$$

- глобальный минимум $f(0,\ldots,0)=0$
- область генерации начального решения $[-5.12,5.12]$

Функция Розенброка:

$$f\left(x_1,\ldots,x_n\right)=\sum_{i=1}^{n-1}\left[100\left(x_{i+1}-x_i^2\right)^2+\left(1-x_i\right)^2\right]$$

- глобальный минимум $f(1,\ldots,1)=0$
- область генерации начального решения $[-5,5]$

Параметры моделей

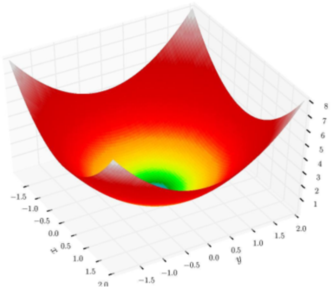
Основные параметры моделей:

1. Количество генерируемых решений — 5000
2. Максимальное число итераций — 5000
3. Количество наилучших решений — меняем от 25 до 5000 с шагом 100
4. Количество моделей для усреднения — 5

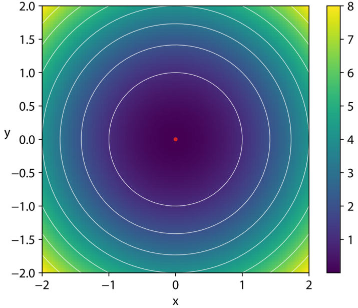
Критерии остановки алгоритмов:

1. Достижение максимального числа итераций
2. Найденное значение целевой функции $f(X) < 1e - 6$
3. Относительное изменение наилучших значений функции $< 1e - 10$

Функция сферы



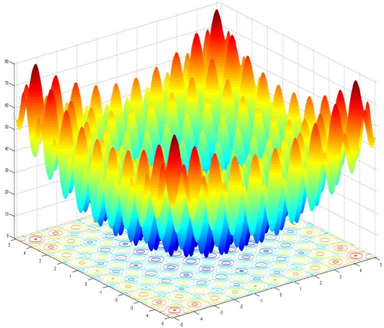
(a) График функции $f(X1, X2)$



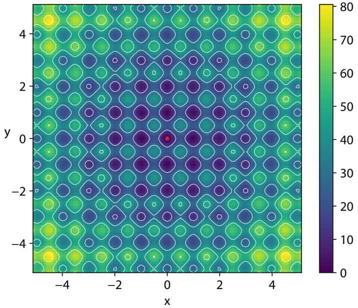
(b) Тепловая карта функции

Рис. 1: Графики функции сферы

Функция Растригина



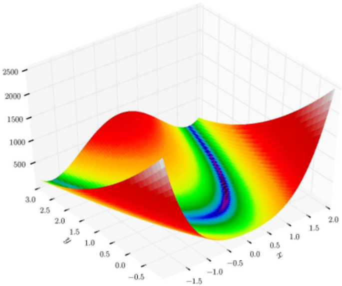
(a) График функции $f(X1, X2)$



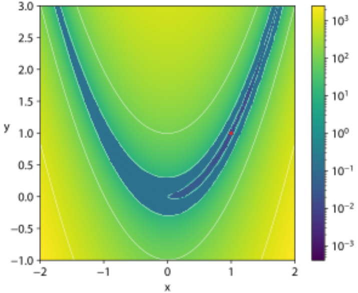
(b) Тепловая карта функции

Рис. 2: Графики функции Растригина

Функция Розенброка



(a) График функции $f(X1, X2)$



(b) Тепловая карта функции

Рис. 3: Графики функции Розенброка

Результаты (на примере функции Розенброка)

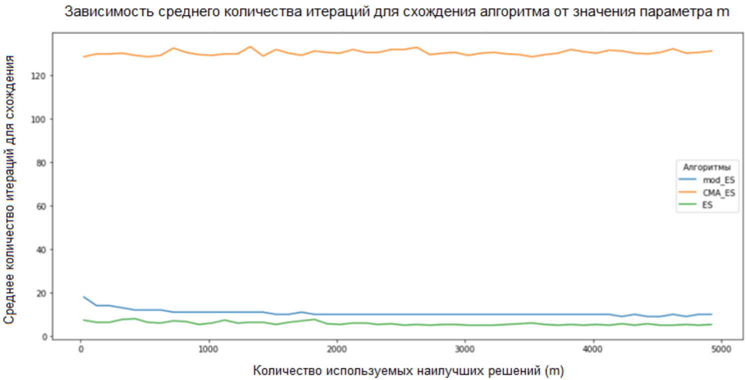


Рис. 4: Среднее количество итераций от значения параметра m

Результаты (на примере функции Розенброка)

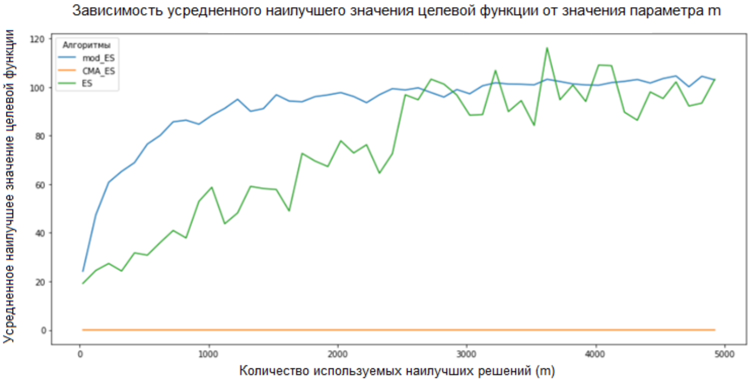


Рис. 5: Наилучшие значения целевой функции от значения параметра m

Решение системы линейных уравнений

Предположим, что задана система линейных уравнений:

$$\begin{cases} 1.008x_1 + 0.955x_2 + 0.984x_3 + 1.060x_4 + 0.900x_5 + 0.924x_6 = 20.201 \\ 1.034x_1 + 1.065x_2 + 0.927x_3 + 1.010x_4 + 1.078x_5 + 0.941x_6 = 21.048 \\ 0.937x_1 + 0.921x_2 + 0.943x_3 + 1.095x_4 + 1.062x_5 + 0.934x_6 = 20.912 \\ 1.063x_1 + 0.954x_2 + 0.986x_3 + 1.088x_4 + 1.063x_5 + 0.967x_6 = 21.404 \\ 0.935x_1 + 0.974x_2 + 0.901x_3 + 0.950x_4 + 1.059x_5 + 0.903x_6 = 20.103 \\ 1.019x_1 + 1.020x_2 + 0.921x_3 + 0.976x_4 + 0.907x_5 + 1.078x_6 = 20.735 \end{cases}$$

- Она имеет единственное решение $X^* = (1, 2, 3, 4, 5, 6)^{\mathrm{T}}$
- Определитель матрицы коэффициентов $\mathbb{A} = 3.8e - 05$
- Число обусловленности $cond(\mathbb{A}) = 184.8224$

Полученные решения:

- CMA-ES:
  $X^* = (1.005, 1.997, 3.006, 3.994, 5.001, 5.999)^{\mathrm{T}}$, $f(X^*) = 1.835e - 07$
- $(\mu,\ \lambda)$-ES:
  $X^* = (1.32, 2.485, 1.989, 4.925, 4.613, 5.565)^{\mathrm{T}}$, $f(X^*) = 0.006$

Решение системы нелинейных уравнений

Предположим, что задана система нелинейных уравнений:

$$\begin{cases} 0.998x_1 + 0.945x_2 + 0.951x_3 + 0.979x_4 + 0.975x_5^3 + 1.099x_6^5 = 8.845 \\ 0.981x_1 + 1.054x_2 + 1.052x_3 + 0.962x_4 + 0.969x_5^3 + 0.970x_6^5 = 9.077 \\ 0.929x_1 + 1.094x_2 + 1.081x_3 + 1.012x_4 + 0.962x_5^3 + 1.077x_6^5 = 9.263 \\ 1.034x_1 + 0.978x_2 + 1.001x_3 + 1.004x_4 + 1.085x_5^3 + 1.014x_6^5 = 9.133 \\ 1.033x_1 + 0.910x_2 + 0.965x_3 + 0.911x_4 + 0.936x_5^3 + 1.085x_6^5 = 8.751 \\ 1.087x_1 + 1.042x_2 + 1.046x_3 + 0.992x_4 + 1.086x_5^3 + 0.981x_6^5 = 9.413 \end{cases}$$

- Она имеет единственное решение $X^* = (2, 2, 2, 1, 1, 1)^{\mathrm{T}}$
- Определитель матрицы коэффициентов $\mathbb{A} = 3.2518e - 06$
- Число обусловленности $cond(\mathbb{A}) = 743.516$
- Определитель матрицы Якоби $\mathbb{J} = 4.8777e - 05$
- Число обусловленности $cond(\mathbb{J}) = 1696.8693$

Полученные решения:

- CMA-ES:
  $X^* = (2.001, 1.997, 2.001, 1.004, 0.999, 0.999)^{\mathrm{T}},\ f(X^*) = 5.66e - 09$
- метод Ньютона:
  $X^* = (2, 2, 2, 1, 3.346, 4)^{\mathrm{T}},\ f(X^*) = 7.26e + 06$

Заключение

В результате получилось, что:

- Алгоритм CMA-ES не зависит от количества используемых наилучших решений;
- Модифицированный эволюционный алгоритм и (μ, λ)-ES с увеличением количества наилучших решений начинают чуть быстрее сходиться, однако они теряют свою точность;
- Алгоритм CMA-ES достаточно быстро и точно решает системы линейных уравнений с плохой обусловленностью и системы нелинейных уравнений с высокой чувствительностью.