Методы адаптации в случайном поиске

Грицай Дмитрий Анатольевич, гр. 522

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

Научный руководитель: д. ф.-м. н., профессор Ю. А. Сушков
Рецензент: Г. С. Тамазян

Санкт-Петербург
2015г.

Цели работы

Цели работы:

- Нахождение минимума функции методом случайного поиска.
- Исследование различных законов сужения перспективной области.
- Классификация тестовых функций.
- Оптимизация параметров случайного поиска.
- Сравнение эффективности поиска для различных законов сужения перспективной области.

Постановка задачи

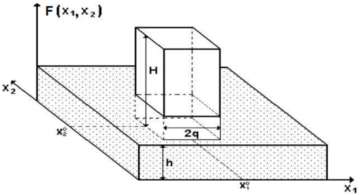
Постановка задачи:

- $x^* = \operatorname{argmin} F(x)$;
- x^j — текущий минимум за j шагов;
- n_{step} — число шагов работы алгоритма;
- $\mathsf{P}(||x^* - x^{n_{\text{step}}}|| < \varepsilon) \to \max$.

Метод случайного поиска

Пусть $X = [0,1]^n$.
Обозначения:

- $I^j \subset [0,1]$ — перспективная область для x^* на j-ом шаге;
- $s_j = (2q_j)^n$ — объем I^j;
- $(x^j_{0,i}, \ldots, x^j_{0,n})$ — центр I^j;
- $p_j = \mathsf{P}(x^j \in I^j)$;
- H_j — плотность внутри I^j;
- h_j — плотность вне I^j;
- $p_{\min}$, $s_{\min} = (2q_{\min})^n$ — гиперпараметры.



Распределение вектора x в случае $n = 2$.

Алгоритм:

1. $x_i^0 = 0.5, i = 1 \ldots n;$
2. пока $j \leq n_{\text{step}}$:
   1. вычисляем q_j;
   2. вычисляем p_j;
   3. $x_{0,i}^j = \max\{q_j, \min\{x_i^{j-1}, 1 - q_j\}\};$
   4. моделируем x^j;
   5. $x_j := \begin{cases} x_j, & \text{если } F(x_j) < F(x_{j-1}; \\ x_{j-1}, & \text{если } F(x_j) \geqslant F(x_{j-1}). \end{cases}$

Метод случайного поиска

$$p_j = \begin{cases} \frac{s_j(p_{\min}-1)}{s_{\min}} + 1, & \text{если } 0 \le s_j \le s_{\min}, \\ \frac{s_j(1-p_{\min})}{1-s_{\min}} + \frac{p_{\min}-s_{\min}}{1-s_{\min}}, & \text{если } s_{\min} \le s_j \le 1; \end{cases}$$

$$H_j = \begin{cases} \frac{p_{\min}-1}{s_{\min}} + \frac{1}{s_j}, & \text{если } 0 \le s_j \le s_{\min}, \\ \frac{1-p_{\min}}{1-s_{\min}} + \frac{p_{\min}-s_{\min}}{s_j(1-s_{\min})}, & \text{если } s_{\min} \le s_j \le 1; \end{cases}$$

$$h_j = \begin{cases} \frac{(1-p_{\min})s_j}{(1-s_j)s_{\min}}, & \text{если } 0 \le s_j \le s_{\min}, \\ \frac{1-p_{\min}}{1-s_{\min}}, & \text{если } s_{\min} \le s_j \le 1. \end{cases}$$

Пока $s_j \ge s_{\min}$, h_j не меняется; $p_{\min} = \min p_j$.

Шкалы

Шкала — это функция, отображающая множество качественных оценок превосходства Λ в множество положительных вещественных чисел.

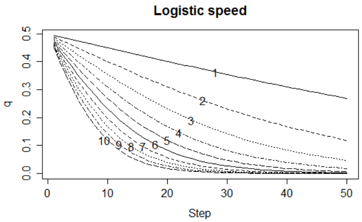
$$\begin{aligned}
\text{Шкала Саати: } & \varphi_S(\lambda) = (1 + x_S|\lambda|)^{\operatorname{sign}(\lambda)}, \lambda \in \Lambda; \\
\text{Шкала Брука: } & \varphi_B(\lambda) = c_B + \lambda x_B, \lambda \in \Lambda; \\
\text{Логистическая шкала: } & \varphi_{log}(\lambda) = 2/(1 + \exp(-\mu\lambda)), \lambda \in \Lambda; \\
\text{Шкала Лутсма: } & \varphi_L(\lambda) = c^{\lambda}, \lambda \in \Lambda; \\
\text{Шкала Ма-Зенга: } & \varphi_{MZ}(\lambda) = (K/(K - |\lambda|))^{\operatorname{sign}(\lambda)}, \lambda \in \Lambda, |\lambda| < K.
\end{aligned}$$

Модификация шкал для случайного поиска

В методе случайного поиска $q_0 = 0.5$, q монотонно убывает. Изменим шкалы так, чтобы они удовлетворяли этим свойствам, и возьмем их в качестве законов сужения перспективной области.

Например, логистический закон сужения перспективной области:

$$\psi(\lambda, \mu) = 1 - 1/(1 + \exp(-\mu\lambda)),$$
$$\lambda = k/n_{\text{step}}.$$

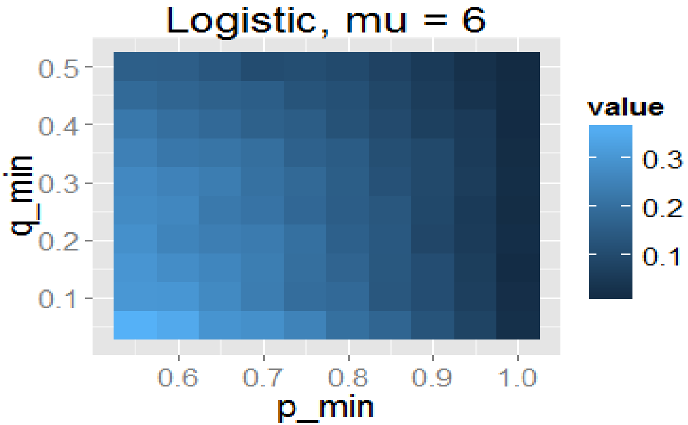


Оптимизация параметров $q_{\min}$ и $p_{\min}$

Оптимизация параметров:

- Оценивалась вероятность попадания в ε-окрестность точки глобального минимума функции за n_{step} шагов.
- Оптимизация велась на сетке с шагом 0.01 по $q_{\min}$ и $p_{\min}$.
- Выделились 4 класса тестовых функций.
- Внутри одинаковых классов тестовых функций оптимальные параметры совпали.
- Вид закона сужения перспективной области не важен.

Оптимизация параметров $q_{\min}$ и $p_{\min}$



Логистический закон с $\mu = 6$ для f_4.

Классы тестовых функций

Классы рассматриваемых тестовых функций:

- Одноэкстремальные.
- Многоэкстремальные:
  - Первый класс:
    - Много локальных минимумов, в которых функция принимает близкие к минимальному значения.
  - Второй класс:
    - Мало локальных минимумов, в которых функция принимает близкие к минимальному значения;
    - Ярко выраженный тренд.
- Овражные.

Овражные функции

Функция $F(x)$ называется овражной на множестве $X \in \mathbf{R}^n$, если для собственных значений $\lambda_i(x)$ матрицы Гессе

$$F^{''}(x) = \left\| \frac{d^2 F(x)}{dx_i dx_j} \right\|, \qquad i, j = 1 \dots n,$$

упорядоченных в любой точке $x \in X$, справедливо неравенство

$$0 < |\min_i \lambda_i(x)| \ll \lambda_1, \qquad x \in X.$$

Поверхность овражной функции $F(x)$ напоминает по форме овраг. Плохо поддаются градиентной оптимизации.

Оптимальные $q_{\min}$ и $p_{\min}$

Классы рассматриваемых тестовых функций:

- Одноэкстремальные: $q_{\min} = 0.45$ и $p_{\min} = 0.95$.
- Многоэкстремальные:
  - Первый класс: $q_{\min} = 0.05$ и $p_{\min} = 0.95$;
  - Второй класс: $q_{\min} = 0.05$ и $p_{\min} = 0.55$.
- Овражные: $q_{\min} = 0.05$ и $p_{\min} = 0.95$.

Оптимизация параметров законов сужения перспективной области

Оптимизация параметров:

- Зафиксируем наилучшие гиперпараметры $q_{\min}$ и $p_{\min}$.
- n_{search} раз запустим процедуру случайного поиска.
- Оценим вероятность попадания в ε-окрестность глобального минимума за n_{step} шагов.
- Внутри одинаковых классов тестовых функций оптимальные параметры близки.

Наилучшие параметры законов сужения перспективной области

Одноэкстремальные:

- Саати ($x_S = 10$);
- Лутсма ($c = 2^{10/4}$);
- Логистический ($\mu = 6$);
- Показательный;
- Ма-Зенга;
- Брука.

Овражные:

- Логистический ($\mu = 1$);
- Саати ($x_S = 1$);
- Брука;
- Лутсма ($c = 2^{1/4}$);
- Показательный;
- Ма-Зенга.

Наилучшие параметры законов сужения перспективной области

Многоэкстремальные I:

- Лутсма ($c = 2^{10/4}$);
- Логистический ($\mu = 4$);
- Саати ($x_S = 8$);
- Показательный;
- Брука;
- Ма-Зенга.

Многоэкстремальные II:

- Логистический ($\mu = 1$);
- Саати ($x_S = 1$);
- Брука;
- Лутсма ($c = 2^{1/4}$);
- Показательный;
- Ма-Зенга.

Оптимизация параметров законов сужения перспективной области с помощью случайного поиска

Методом случайного поиска ищем минимум функции:

$$f(\mu) = -\,\mathsf{P}(A),$$

где A — событие, соответствующее попаданию в ε-окрестность глобального минимума за n_{step} шагов при фиксированных параметрах $q_{\min}$ и $p_{\min}$, μ — параметр скорости сужения перспективной области.

Значения функции f для разных аргументов можно сравнивать как с пересчётом вероятностей, так и без него.

Оптимизация параметров законов сужения перспективной области с помощью случайного поиска

- $f(\mu)$ —многоэкстремальная функция второго класса.
- Гиперпараметрами внешнего поиска (поиска верхнего уровня) будут $q_{\text{step}} = 0.05$ и $p_{\text{step}} = 0.55$.
- Во внешнем случайном поиске используется логистический закон сужения перспективной области с $\mu = 4$.
- Результаты в обоих способах очень близки к предыдущим результатам.
- При пересчете вероятностей дисперсия меньше.

Итоги

В ходе работы были получены следующие основные результаты:

- На основе статистического исследования найдены оптимальные значения параметров алгоритма случайного поиска для исследуемых классов тестовых функций.
- Рассмотрены наиболее эффективные с точки зрения выбранных критериев законы сужения перспективной области случайного поиска и найдены их оптимальные параметры.
- Был введен случайный поиск верхнего уровня, подтвердивший выбор наилучших параметров.
- Проведено сравнение результатов для разных законов сужения перспективной области для классов рассмотренных тестовых функций.