Поиск Глобального Максимума. Случай Многих Равных Эстремумов

Презентация ВКР

Хэ Пин, группа. 20.М03-мм

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

Научный руководитель: д. ф.-м. н., профессор С. М. Ермаков
Рецензент: Директор, ООО «Аптик СПБ» к. ф.-м. н., К. А. Тимофеев

Санкт-Петербург
2022г.

План дальнейшего изложения

Задача поиска глобального максимума

- Классический метод имитации отжига и модификации
- Слабая сходимость алгоритма
- Свойства слабой сходимости

Применение метода: разделение корней нелинейных уравнений

- Связь между глобальными экстремумами и решением уравнений
- Масштабирование уравнения
- Примеры

Классический метод имитации отжига и модификации

Функция метода метрополиса классического алгоритма имитации отжига имеет вид:

Алгоритм Метрополиса классического метода

$$\overline{\mathfrak{M}}(T, X) = C(T) \exp\left(-\frac{f(X)}{T}\right),$$

где C — константа нормировки, зависящая от T.

При $T \to 0$ слабо сходится к δ-функции, сосредоточенной в точке глобального минимума $f(X)$.
Для нахождения глобального максимума использовать другую функцию типа $\overline{\mathfrak{M}}(T, X)$, а именно

Алгоритм Метрополиса модифицированного метода

$$\mathfrak{M}(X, m) = \frac{f^m(X)}{\int\limits_D f^m(X) dX},$$

При $n \to$ функция $\mathfrak{M}(X, m)$ сходится к функции глобального максимума.

Алгоритм модифицированного метода

Положим

$$G_\varepsilon(X) = \begin{cases} 1-\varepsilon, & \text{если } x \geq 1-\varepsilon \\ G(X), & \text{если } f(x) < 1-\varepsilon \end{cases}$$

Пусть $F(X)$ задано в замкнутой области $D \subset R^k$, $0 \leq f(X) \leq M < +\infty$ и M её экстрамальное значение, обозначим его наибольшее Q_1 подобласть D,

$$Q_1 = \{X : F(X) = M - \varepsilon\},$$

и через плотность распределения случайного вектора Ξ_m

$$F_m(X) = \frac{F^m X}{\int\limits_D F^m(X)dx},$$

Лемма

При сделанных предположениях вероятность $P\{\Xi_m \in Q_1\}$ стремится к 1 при $m \to \infty$.

Доказательство

Пусть $Q_2 = D - Q_1$, то $D = Q_2 + Q_1$, $Q_2 = X : f(X) < M - \varepsilon$. Если

$$f_1(X) = \begin{cases} M - \varepsilon, X \in Q_1 \\ 0, X \in Q_2 \end{cases}$$

$$f_2(X) = \begin{cases} f(X), X \in Q_2 \\ 0, X \in Q_1 \end{cases}$$

Алгоритм модифицированного метода

Доказательство.

То $f^m(X) = (f_1(X) + f_2(X))^m = f_1^m(X) + f_2^m(X) = (M-\varepsilon)^m + f_2^m(X)$
$= (M-\varepsilon)^m[1 + (\frac{f_2(x)}{M-\varepsilon})^m]$.
и $\frac{f_2(x)}{M-\varepsilon} < 1$

$$P\{\Xi_m \in Q_1\} = \int\limits_{Q_1} F_m(X)dx = \frac{(M-\varepsilon)^m}{\int\limits_D f_m(X)dX} = \frac{1}{1 + \int\limits_D (\frac{f_2(X)}{M-\varepsilon})^m dX}$$

но $\int\limits_D (\frac{f_2(X)}{M-\varepsilon})^m dX \to 0$, что доказывает лемму. □

Превратить в задачу поиска экстремума

Мы можем преобразовать задачу решения уравнений в задачу нахождения глобального минимума

$$H(X) = \sum_{l=1}^{s} P_l g_l^2(X) = \min,$$

Или нахождение глобального максимума:

$$H(X) = \frac{1}{1 + \sum_{l=1}^{s} P_l g_l^2(X)} = \max,$$

Возможно, потребуется масштабировать уравнение, чтобы иметь надлежащий масштаб для применения нашего алгоритма поиска корней. Вы можете масштабировать или уменьшать масштаб по определенным осям в соответствии с определенным соотношением.

Масштабирование функции

$$H(X) \to H(k_i X),$$

где k_i — коэффициент масштабирования каждой оси.

Он также может быть масштабирован в соответствии с некоторыми функциями, такими как экспоненциальная функция и логарифмическая функция.

Кластеризации DBSCAN

DBSCAN (пространственная кластеризация приложений с шумом на основе плотности) — это алгоритм, который может обнаруживать кластеры произвольной формы. DBSCAN обычно считает кластерами плотные области объектов в пространстве данных, которые разделены низкой плотностью.

Кластеризации DBSCAN

$$\frac{N_{ab}}{(N_a + N_b)/2} \leq \alpha_{\text{merge}}$$

Решение простой нелинейной системы по сравнению с классическим моделируемым отжигом

Возьмем в качестве примера простое нелинейное уравнение.

$$\begin{cases} f_1(x,y) = y - \sin(x^2) \\ f_2(x,y) = y - 0.5x - 2. \end{cases}$$

Используя классической имитации отжига, решаемая задача может быть преобразована в

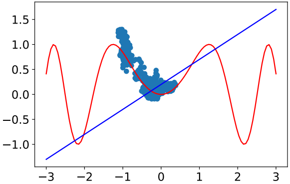
$$H(x,y) = (y - \sin(x^2))^2 + (y - 0.5x - 2)^2 = \min,$$

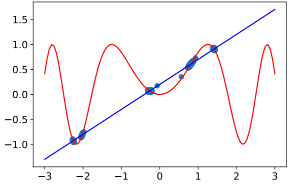
Используя модифицированной имитации отжига, решаемая задача может быть преобразована в

$$H(x,y) = \frac{1}{1 + (y - \sin(x^2))^2 + (y - 0.5x - 2)^2} = \max,$$

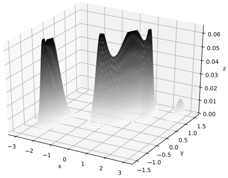
Решение простой нелинейной системы по сравнению с классическим моделируемым отжигом

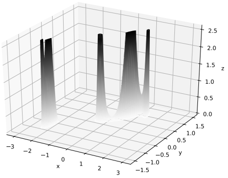
Сравнение двух методов: на первом изображении используется классический метод имитации отжига, а на втором изображении используется модифицированный алгоритм имитации отжига.





Разделение корней. Первая картинка — случай m=10, вторая картинка — случай m=100





Мы используем гиперболическую функцию в качестве примера модели масштабирования. Когда корни будут очень близко, использование имитации отжига напрямуючасто не позволяет разделить корни

$$\begin{cases} f_1(x, y) = x^2 + 5y^2 - 2.1x - y - 4.9 \\ f_2(x, y) = 2x^2 - y^2 - 4.2x + 0.2y + 3.4, \end{cases}$$

Используя классической имитации отжига, решаемая задача может быть преобразована в

$$m = 100, \varepsilon = 0.00000001,$$

Мы можем решить систему $f(x/10, y)$, чтобы получить более широкую модель.

Корени	Масштабированный x	Настоящий x	y
Корень №1	10.004614	1.000461	1.191932
Корень №2	10.005746	1.000575	-1.000014
Корень №3	10.986438	1.098644	1.199998
Корень №4	10.991736	1.099174	-1.000017

Таблица: Результат кластеризации DBSCAN гиперболической системы

Будем решать многомерную нелинейную систему и произведем анализ точности.

$$f_i(X) = \sum_{j=1}^{n} \vec{a}_{i,j} x_i^2 + \vec{b}_i, i = 1, 2, \ldots n,$$

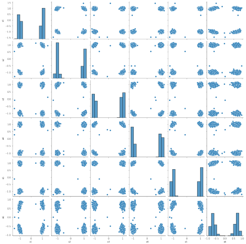
$$\vec{a}_{i,j} = \begin{pmatrix} 0.1 & 0.2 & 0.3 & 0.4 & 0.5 & 0.6 \\ 0.2 & 0.3 & 0.4 & 0.5 & 0.6 & 0.1 \\ 0.3 & 0.4 & 0.5 & 0.6 & 0.1 & 0.2 \\ 0.4 & 0.5 & 0.6 & 0.1 & 0.2 & 0.3 \\ 0.5 & 0.6 & 0.1 & 0.2 & 0.3 & 0.4 \\ 0.6 & 0.1 & 0.2 & 0.3 & 0.4 & 0.5 \end{pmatrix},$$

$$\vec{b}_i = \begin{pmatrix} 1.8 & 2.05 & 2.0 & 1.95 & 1.9 & 1.85. \end{pmatrix}$$

Решение системы должно быть $\pm 1, \pm 1, \pm 1, \pm 1, \pm 1, \pm\sqrt{0.5}$, $2^6 = 64$ всего 64 корня.

Решение простой нелинейной системы по сравнению с классическим моделируемым отжигом

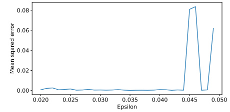
Точечная диаграмма результата имитации отжига



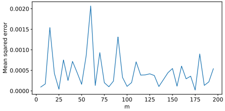
Поскольку каждое решение симметрично, рассмотрим корень с положительными координатами и вычислим среднеквадратичную ошибку всех координат

$$\mathrm{MSE} = \frac{1}{6}\sum_{i=1}^{n}(x_i - \hat{x}_i)^2$$

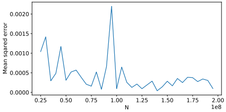
Среднеквадратические ошибки при различных ε



Среднеквадратические ошибки при различных m



Среднеквадратические ошибки при различных N



Заключение

В работе указана теоретическую возможность поиска глобального максимума в случае многих равных экстремумов и решению нелинейных уравнений в качестве практической проверки. Результаты показывают высокую точность. Чем больше m и N и чем меньше ε, тем стабильнее результат решения.

Список литературы

- С.М. Ермаков Д.В. Куликов С. Н. Л. К анализу метода имитации отжига в много экстремальном случае // Вестник Санкт-Петербургского университета. Математика. Механика. Астрономия. — 2017. — Vol. 4(62), no. 2. — P. 220–226.
- Henderson D., Jacobson S. H., Johnson A. W. The Theory and Practice of Simulated Annealing // Handbook of Metaheuristics / ed. by Glover F., Kochenberger G. A. — Boston, MA : Springer US, 2003. — P. 287–319. — ISBN: 978-0-306-48056-0.
- Equation of State Calculations by Fast Computing Machines / Metropolis N., Rosenbluth A. W., Rosenbluth M. N., Teller A. H., and Teller E. // The Journal of Chemical Physics. — 1953. — June. — Vol. 21, no. 6. — P. 1087–1092.
- El-Sonbaty Y., Ismail M., Farouk M. An efficient density based clustering algorithm for large databases // 16th IEEE International Conference on Tools with Artificial Intelligence. — IEEE Comput. Soc.