Серия «Математика»
2022. Т. 39. С. 17—33

Онлайн-доступ к журналу:
http://mathizv.isu.ru

И З В Е С Т И Я
Иркутского государственного университета

Научная статья

УДК 519.6
MSC 65C05, 65C20
DOI https://doi.org/10.26516/1997-7670.2022.39.17

Модифицированный генетический алгоритм поиска глобального экстремума в сочетании с направленными методами

Д. А. Овсянников[1], Л. В. Владимирова[1], И. Д. Рубцова[1✉], А. В. Рубаник[1], В. А. Пономарев[1]

[1] Санкт-Петербургский государственный университет, Санкт-Петербург, Российская Федерация
✉ i.ribtsova@spbu.ru, rubtsova05@mail.ru

Аннотация. В работе рассмотрен, модифицирован и апробирован стохастический метод поиска глобального экстремума, основанный на моделировании нормального распределения и обеспечивающий адаптацию ковариационной матрицы. Метод является итерационным, на его основе разработан генетический алгоритм. Координаты пробных точек каждого поколения определяются при использовании «лучших» точек предыдущего поколения и значений стандартных нормальных случайных величин. Таким образом на каждом этапе поиска моделируется нормальное распределение, параметры которого – среднее и матрица ковариаций – оцениваются через положения «лучших» точек предыдущего поколения. При этом нет необходимости в вычислении, хранении и преобразовании самой ковариационной матрицы, что является неоспоримым достоинством данного метода.

Практика показала, что эллипсоид рассеяния нормального распределения быстро сжимается с увеличением номера поколения, что может привести к чрезмерному сужению области просмотра и получению локального экстремума вместо глобального. Предложенная модификация метода позволяет избежать этой ситуации. Пробные точки разбиваются на две группы, при моделировании которых используются нормальные случайные величины с различными среднеквадратическими отклонениями, по крайней мере одно из которых больше единицы. Таким образом осуществляется своеобразная мутация популяции пробных точек, позволяющая обеспечить достаточное количество проб как вблизи «наилучшей» точки, так и на отдалении от нее.

Модифицированный генетический алгоритм применен к решению задачи оценки параметров нелинейной параметрической регрессии. Выполнена успешная минимизация многоэкстремальной функции. Стохастический метод используется в сочетании с направленными. Представленные численные результаты подтверждают эффективность введенной модификации генетического алгоритма и позволяют выбрать из двух направленных методов более результативный для рассматриваемой задачи.

Ключевые слова: глобальный экстремум, генетический стохастический алгоритм, адаптация ковариационной матрицы, нелинейная регрессия

Благодарности: Авторы благодарят профессора С. М. Ермакова за ценные советы и постоянное внимание к работе.

Ссылка для цитирования: Овсянников Д. А., Владимирова Л. В., Рубцова И. Д., Рубаник А. В., Пономарев В. А. Модифицированный генетический алгоритм поиска глобального экстремума в сочетании с направленными методами // Известия Иркутского государственного университета. Серия Математика. 2022. Т. 39. С. 17–33. https://doi.org/10.26516/1997-7670. 2022.39.17

Research article

Modified Genetic Algorithm of Global Extremum Search in Combination with Directional Methods

Dmitri A. Ovsyannikov[1], **Liudmila V. Vladimirova**[1], **Irina D. Rubtsova**[1✉], **Alexey V. Rubanik**[1], **Vladimir A. Ponomarev**[1]

[1] Saint Petersburg State University, St. Petersburg, Russian Federation
✉ i.ribtsova@spbu.ru, rubtsova05@mail.ru

Abstract. In the paper, the stochastic method of global extremum search is discussed, modified and tested. The method is based on normal distribution modeling and provides covariance matrix adaptation. The method is iterative; a genetic algorithm has been developed on its basis. The coordinates of the trial points of each generation are determined using the "best" points of the previous generation and the values of standard normal random variables. Thus, at each stage of the search, a normal distribution is simulated, and its parameters (the mean and the covariance matrix) are estimated through the positions of the "best" points of the previous generation. In this case, there is no need to calculate, store and transform the covariance matrix, which is indisputable advantage of this method.
Practice has shown that the dispersion ellipsoid of normal distribution shrinks rapidly with generation number increasing, which can lead to an excessive narrowing the scanning area and obtaining a local extremum instead of a global one. The proposed modification of the method avoids this situation. The trial points are divided into two groups, which are simulated using normal random variables with different standard deviations, at least one of which is greater than 1. Thus, a kind of mutation of the population is carried out, which makes it possible to provide a sufficient number of sample points both near the "best" one and at a distance from it.
The modified genetic algorithm is applied to solving the problem of estimating the parameters of nonlinear parametric regression. A successful minimization of the multiextremal function is performed. The stochastic method is used in combination with directional. The numerical results presented confirm the effectiveness of the introduced modification of the genetic algorithm and make it possible to choose from two directed methods the more efficient one for the problem under consideration.

Keywords: global extremum, genetic stochastic algorithm, covariance matrix adaptation, nonlinear regression

Acknowledgements: The authors are grateful to Professor S.M. Ermakov for valuable advice and permanent care for the work.

For citation: Ovsyannikov D. A., Vladimirova L. V., Rubtsova I. D., Rubanik A. V., Ponomarev V. A. Modified Genetic Algorithm of Global Extremum Search in Combination with Directional Methods. *The Bulletin of Irkutsk State University. Series Mathematics*, 2022, vol. 39, pp. 17–33. (in Russ.) https://doi.org/10.26516/1997-7670.2022.39.17

1. Введение

Многие задачи, возникающие в различных областях математики или в приложениях, требуют отыскания глобального экстремума некоторой целевой функции. Таковы, например, задачи теории оптимального управления, проблемы оптимизации различных динамических процессов и надлежащего выбора параметров технических устройств [1; 2; 6; 8; 10; 11; 12]. Задачам поиска глобального экстремума посвящена обширная литература, например, [2; 3; 4; 5; 7; 19]. В зависимости от свойств функции используются различные подходы и методы решения таких задач. В их числе — генетические алгоритмы и алгоритмы, использующие ковариационную матрицу [1; 2; 3; 4; 9; 13; 14; 15; 18], которые могут быть эффективны в сложных случаях: большая размерность области поиска, овражная функция, много близких локальных экстремумов и т.д.

В последние годы успешно используется метод ковариационных матриц, предложенный первоначально в 1977 году в работе С. М. Ермакова и Л. В. Митиогловой (Л. В. Владимировой) [4]. Этот метод использовался для нахождения максимума ограниченной положительной кусочно-непрерывной целевой функции $f(X)$ в ограниченной области евклидова пространства E^s. Идея метода заключается в следующем. Рассмотрим последовательность плотностей

$$p(X,n) = \frac{f^n(X)}{\int_D f^n(X)dX}, n = 1, 2, 3, ... \tag{1.1}$$

Предполагается, что интеграл $\int_D f^n(X)dX$ существует и конечен для любого конечного n. Если глобальный максимум функции $f(X)$ единственный, тогда при $n \to \infty$ последовательность (1.1) слабо сходится к δ-функции, сосредоточенной в точке глобального максимума $f(X)$. Поведение плотности вида (1.1) в общем случае может быть сложным, поэтому авторы работы [4] заменили ее на похожую по свойствам плотность и ввели последовательность, более простую для моделирования:

$$\begin{aligned} \bar{p}(X,n) = C_n f(X) N(X, \hat{X}_{max,n}, \hat{B}_n), n = 1, 2, ...; \\ \bar{p}(X,0) = C_0 f(X). \end{aligned} \tag{1.2}$$

Здесь C_n – константа нормировки, функция $N(X, \hat{X}_{max,n}, \hat{B}_n)$ – плотность нормального распределения со средним $\hat{X}_{max,n}$ и матрицей ковариаций $\hat{B}_n$. Формула (1.2) описывает итерационный процесс поиска; значение $\hat{X}_{max,n}$ – оценка точки глобального максимума функции, полученная на предыдущей $(n-1)$-й итерации. На этой же итерации оцениваются элементы матрицы ковариаций $\hat{B}_n$, которая учитывает свойства функции $f(X)$. Понятно, что последовательность средних $\hat{X}_{max,n}$ сходится при $n \to \infty$ к точке глобального максимума $f(X)$.

Недостатком этого метода является необходимость разложения матрицы $\hat{B}_n$ на произведение двух треугольных матриц порядка s^2 для моделирования нормальной плотности $N(X, \hat{X}_{max,n}, \hat{B}_n)$ на каждом этапе поиска. Кроме того, дополнительные трудности могут возникнуть в случае, когда ковариационная матрица близка к вырожденной.

В работе [13] доказывается лемма, на основании которой строится генетический алгоритм, использующий нормальное распределение без вычисления и разложения ковариационной матрицы, которая эволюционирует от поколения к поколению. Этот метод был успешно применен при решении некоторых задач, например, задачи оптимизации динамики пучка в линейном ускорителе, которая была сведена к минимизации критерия качества в области пространства параметров размерности 84 [1; 20]. Однако может возникнуть ситуация, когда с увеличением номера поколения пробных точек эллипсоид рассеяния соответствующего нормального распределения быстро стягивается в точку и поиск практически не происходит. В данной работе предлагается существенная модификация этого алгоритма, позволяющая избежать стягивания в точку эллипсоида рассеяния и получения локального экстремума вместо глобального.

Модифицированный генетический алгоритм применяется при решении задачи оценки параметров нелинейной регрессии, численные результаты подтверждают полезность введенных модификаций.

Уточнить результат глобального поиска можно при использовании направленных методов. При решении задачи оценки параметров функции регрессии применяются и сравниваются два известных метода: градиентный спуск и метод импульса [17].

2. Генетический стохастический алгоритм

Будем рассматривать задачу отыскания точки экстремума положительной ограниченной кусочно-непрерывной функции $f(X)$, заданной в ограниченной области $D \subset E^s$. Предположим, что речь идет о глобальном минимуме, который достигается в единственной точке $X_{min} \in D$:

$$f(X_{min}) = \min_{x \in D} f(X) \tag{2.1}$$

Понятно, что задачу поиска максимума функции легко свести к задаче поиска минимума, заменяя, например, исходную функцию на $1/f(X)$. Общая идея адаптивных методов состоит в организации последовательного выбора точек так, чтобы они сгущались в окрестности точки минимума. Рассмотрим способ построения последовательности поколений пробных точек, предложенный в работе [13]. Алгоритм состоит из ряда этапов.

Нулевой этап. Если мы не имеем специальных априорных сведений о расположении X_{min} и поведении функции $f(X)$, проведем сканирование области D, используя равномерное распределение, и вычислим значения $f(X)$ в точках сканирования. Пусть N_0 — число пробных точек, составляющих нулевое поколение. Выберем среди них m «лучших» точек $X_1, X_2, ..., X_m$, соответствующих наименьшим значениям целевой функции, и найдем точку минимума $X_{min}^{(0)}$ в нулевом поколении («наилучшую» пробную точку):

$$X_{min}^{(0)} = \arg \min_{X_1,...,X_{N_0}} f(X) : f(X_{min}^{(0)}) = f(X_1) < f(X_2) < ... < f(X_m).$$

***k*-й этап.** Осуществляется моделирование N_k пробных точек k-го поколения по формулам:

$$Z_j = \frac{1}{m} \sum_{i=1}^{m} \eta_j^{(i)} (X_i - \bar{X}) + \bar{X}, j = \overline{1, N_k} \tag{2.2}$$

при использовании «лучших» точек $X_1, X_2, ..., X_m$ предыдущего поколения. В формуле (2.2)

$$\bar{X} = \frac{1}{m} \sum_{i=1}^{m} X_i = (\bar{x}^{(1)}, ..., \bar{x}^{(s)}),$$

$\eta_j^{(i)}, i = \overline{1, m}, j = \overline{1, N_k}$ - независимые стандартные нормальные случайные величины, причем

$$M\eta_j^{(i)} = 0, M(\eta_j^{(i)})^2 = 1, M(\eta_j^{(i)}\eta_j^{(l)}) = 0, (i \neq l), i = \overline{1, m}, j = \overline{1, N_k}. \tag{2.3}$$

После этого выбираются m «лучших» точек $Z_1, Z_2, ..., Z_m$ k-го поколения и определяется точка $Z_{min}^{(k)}$ минимума в k-м поколении («наилучшая» пробная точка):

$$Z_{min}^{(k)} = \arg \min_{Z_1,...,Z_{N_k}} f(X), \quad f(Z_{min}^{(k)}) = f(Z_1) < f(Z_2) < ... < f(Z_m).$$

После замены Z на X осуществляется переход к следующей итерации.

Оценка матрицы ковариаций нормального распределения k-го поколения имеет следующий вид:

$$C(Z_j) = \|\frac{1}{m}\sum_{i=1}^{m}(x_i^{(\alpha)} - \bar{x}^{(\alpha)})(x_i^{(\beta)} - \bar{x}^{(\beta)})\|_{\alpha,\beta=1}^{s}, \tag{2.4}$$

где

$$\bar{x}^{(\alpha)} = \frac{1}{m}\sum_{k=1}^{m} x_k^{(\alpha)}, \bar{x}^{(\beta)} = \frac{1}{m}\sum_{k=1}^{m} x_k^{(\beta)}.$$

Действительно, ковариационная матрица случайного вектора Z_j находится по формуле:

$$C(Z_j) = CovZ_j = M((Z_j - MZ_j)^T(Z_j - MZ_j)).$$

Математическое ожидание MZ_j в силу (2.2), (2.3) получим в виде:

$$MZ_j = \frac{1}{m}\sum_{i=1}^{m}(M\eta_j^{(i)})(X_i - \bar{X}) + \bar{X} = \bar{X}.$$

Тогда

$$Z_j - MZ_j = \frac{1}{m}\sum_{i=1}^{m}\eta_j^{(i)}(X_i - \bar{X}). \tag{2.5}$$

С учетом (2.2), (2.3), (2.5) имеем:

$$(Cov(Z_j))^{(\alpha,\beta)} = \frac{1}{m^2}\sum_{i=1}^{m}\sum_{l=1}^{m} M(\eta_j^{(i)}\eta_j^{(l)})(x_i^{(\alpha)} - \bar{x}^{(\alpha)})(x_l^{(\beta)} - \bar{x}^{(\beta)}) =$$
$$= \frac{1}{m}\sum_{i=1}^{m}(x_i^{(\alpha)} - \bar{x}^{(\alpha)})(x_i^{(\beta)} - \bar{x}^{(\beta)}), \alpha, \beta = \overline{1, s}.$$

Полученные результаты позволяют сделать следующие выводы.

1. Параметры нормального распределения k-го поколения оцениваются средним $\bar{X}$ и ковариационной матрицей (2.4), элементы которой зависят только от «лучших» пробных точек предыдущего поколения.

2. Способ моделирования (2.2) случайного вектора Z_j позволяет не вычислять и не хранить матрицу ковариаций, что затруднительно при большой размерности пространства. Не требуется разложения матрицы ковариаций, например, на треугольные (что может быть сложно, если матрица C близка к вырожденной).

3. Способ построения пробных точек по формуле (2.2) обеспечивает возможность эволюции ковариационной матрицы.

Сходимость метода для случая унимодальной функции доказана в работе [13].

3. Модификация метода

Вернемся к формуле (2.2) построения нового поколения точек и рассмотрим способ адаптации ковариационной матрицы.

Замечание 1. Вместо вектора $\bar{X}$ эффективнее использовать вектор $X_{\min}^{(k-1)}$, соответствующий минимальному значению $f(X)$ на $(k-1)$-ом этапе.

Замечание 2. С увеличением номера этапа поиска эллипсоид рассеяния нормально распределенных векторов Z_j может быстро стягиваться в точку, при этом поиск практически не происходит. Поэтому предлагается следующая модификация.

1. На каждом этапе будем домножать стандартные нормально распределенные величины $\eta_j^{(i)}, i = \overline{1,m}, j = \overline{1,N_k}$ в соотношении (2.2) на некоторую константу σ. При этом случайные величины $\sigma\eta_j^{(i)}$, $i = \overline{1,m}, j = \overline{1,N_k}$ имеют уже среднее квадратическое отклонение σ. Практика показывает, что эта константа может зависить от номера этапа: $\sigma = \sigma(k)$. Среднее квадратическое отклонение $\sigma(k)$ можно выбрать так, чтобы эллипсоид рассеяния векторов $Z_j, j = \overline{1,N_k}$ покрывал значительную часть области D. Для каждой целевой функции $f(X)$ подбор величины $\sigma(k)$ индивидуален.

2. При моделировании будем разбивать совокупность векторов Z_j, $j = \overline{1,N_k}$ на две группы, содержащие $N_k^{(1)}$ и $N_k^{(2)}$ векторов соответственно, $N_k^{(1)} + N_k^{(2)} = N_k$. Для первой группы будем применять константу $\sigma^{(1)}(k)$, а для второй – $\sigma^{(2)}(k)$, $\sigma^{(1)}(k) \neq \sigma^{(2)}(k)$. Формулы моделирования пробных точек k-го поколения примут вид:

$$\begin{aligned}
Z_j &= \frac{\sigma^{(1)}(k)}{m}\sum_{i=1}^{m}\eta_j^{(i)}(X_i - X_{min}^{(k-1)}) + X_{min}^{(k-1)}, \quad j = \overline{1,N_k^{(1)}},\\
Z_j &= \frac{\sigma^{(2)}(k)}{m}\sum_{i=1}^{m}\eta_j^{(i)}(X_i - X_{min}^{(k-1)}) + X_{min}^{(k-1)}, \quad j = \overline{1,N_k^{(2)}}.
\end{aligned} \tag{3.1}$$

Затем из $N_k^{(1)}$ полученных проб выбираются m_1 «лучших» значений аргумента, из $N_k^{(2)}$ проб – m_2 «лучших» значений аргумента, причем $m_1 + m_2 = m$. После замены обозначения Z на X имеем набор $X_1, X_2, ..., X_m$ отобранных пробных точек для организации следующего этапа поиска.

Данная модификация метода позволяет увеличить объем эллипсоида рассеяния нормального распределения каждого поколения так, чтобы он покрывал достаточную часть области поиска. Тем самым удается избежать быстрого сжатия эллипсоида и фактического прекращения поиска. Использование различных констант $\sigma^{(1)}(k) \neq \sigma^{(2)}(k)$ в формулах (3.1) может рассматриваться как своего рода мутация в популяции пробных точек. Это позволяет иметь подходящее количество проб как вблизи $X_{min}^{(k-1)}$, так и на отдалении от этой точки.

Для уточнения результатов поиска при использовании генетического стохастического алгоритма могут использоваться направленные методы. Приведем два примера таких методов.

4. Используемые направленные методы

Рассмотрим задачу минимизации целевой функции $f(\Theta)$, где Θ – вектор параметров:

$$\min_{\Theta} f(\Theta)$$

Предполагаем, что градиент $\bigtriangledown f(\Theta)$ целевой функции может быть рассчитан или оценен.

Самым простым подходом к минимизации является **градиентный спуск** [16] (GD – Gradient Descent). Пусть t – номер итерации, Θ_t — соответствующий вектор параметров. Шаг метода градиентного спуска:

$$\Theta_{t+1} = \Theta_t + v_t, \text{ где } v_t = -\alpha_t \bigtriangledown f(\Theta_t),$$

α_t — параметр метода (называемый скоростью обучения). Если $f(\Theta)$ является выпуклой и удовлетворяет условию Липшица, то $f(\Theta_t)$ сходится к оптимальному значению f_{min} как $1/t$ [16]. Градиентный спуск популярен, однако альтернативные методы, такие как метод импульса (MOM), могут привести к значительно более быстрой сходимости к точке минимума.

Рассмотрим **импульсный метод или метод моментов** (MOM) [17]. Шаг метода выглядит следующим образом:

$$\Theta_{t+1} = \Theta_t + v_{t+1}, \text{ где } v_{t+1} = \mu_t v_t - \alpha_t \bigtriangledown f(\Theta_t).$$

Здесь α_t — скорость обучения (шаг спуска), $\mu_t \in [0,1]$ — параметр импульса. В отличие от классического градиентного спуска, данный метод пытается сохранить продвижение в том же направлении, предотвращая колебания. Хорошо известно, что GD может сильно замедлиться и иметь зигзагообразный путь оптимизации, если целевая функция имеет гребни. Импульс может облегчить путь к оптимизации, так как компенсирует направления подъема и направления, в которых градиент быстро меняется.

Генетический алгоритм в сочетании с направленными методами GD и MOM использован в задаче выбора параметров функции регрессии. Отметим, что такие задачи возникают при анализе результатов эксперимента в различных областях науки: в математике, физике, химии, биологии и других.

5. Оценка параметров нелинейной параметрической регрессии

Рассмотрим экспоненциальную регрессию

$$\eta(x,\Theta) = \beta_1 \exp(-\lambda_1 x) + \beta_2 \exp(-\lambda_2 x) \tag{5.1}$$

с вектором параметров $\Theta = (\beta_1, \lambda_1, \beta_2, \lambda_2)$ размерности $s = 4$. Пусть вектор $y = (y_1, ..., y_I)$ является результатом наблюдений эксперимента в заданных точках $x_i, i = \overline{1, I}$, где $I > 0$ — заданное число; вектор $x = (x_1, \ldots, x_I)$ — регрессионный план. Будем оценивать вектор Θ параметров функции регрессии (5.1), используя метод наименьших квадратов:

$$\hat{\Theta} = \arg \min_{\Theta \in P \subset E^s} \sum_{i=1}^{I} (y_i - \eta(x_i, \Theta))^2.$$

Здесь P — некоторое множество из пространства E^s. Итак, имеем задачу поиска глобального минимума функции

$$f(\Theta) = \Sigma_{i=1}^{I} (y_i - \eta(x_i, \Theta))^2. \tag{5.2}$$

При заданных плане $x = (x_1, \ldots, x_I)$ и векторе результатов наблюдений $y = (y_1, ..., y_I)$ функция (5.2) зависит только от 4-мерного вектора параметров $\Theta = (\beta_1, \lambda_1, \beta_2, \lambda_2) \in P \subset E^s$. Несмотря на кажущуюся простоту, функция (5.2) является многоэкстремальной, и задача поиска ее глобального минимума достаточно сложна [9; 18].

Представим для этой задачи числовые данные и результаты глобального поиска, полученные при использовании модифицированного генетического алгоритма (MGA) в сочетании с направленными методами (GD и MOM).

6. Численные результаты

Данные эксперимента. Заданный план $x = (x_1, \ldots, x_I)$ и вектор $y = (y_1, \ldots, y_I)$ результатов наблюдений для $I = 14$ представлены в табл. 1. Вектор y был имитирован по формуле:

$$y_i = \eta(x_i, \bar{\Theta}) + \epsilon_i.$$

Здесь $\bar{\Theta}$=(20.0, 0.43, 70.0, 1.27) — заданный вектор параметров; $\epsilon_i, i = \overline{1, I}$ — случайные ошибки измерений — независимые, нормально распределенные и центрированные величины:

$$M\epsilon_i = 0, M\epsilon_i\epsilon_j = 0, i \neq j, D\epsilon_i = M\epsilon_i^2 = \upsilon^2, i, j = \overline{1, I}.$$

Для рассматриваемой задачи $\upsilon = 0.02$.

Таблица 1

Регрессионный план и вектор результатов наблюдений

i	1	2	3	4	5	6	7
x_i	0.07	0.11	0.14	0.18	0.24	0.27	0.38
y_i	88.44	79.95	77.43	74.20	6.63	67.48	60.22
i	8	9	10	11	12	13	14
x_i	0.43	0.44	0.50	0.65	0.96	1.28	1.65
y_i	57.17	56.56	53.24	45.80	33.92	25.31	18.48

Значения параметров функции регрессии (5.1) принадлежат компакту P:

$$\Theta \in P = \{5.0 \leq \beta_i \leq 100; 0.075 \leq \lambda_i \leq 1.925, i = \overline{1, 2}\}.$$

Параметры генетического алгоритма.

Число особей в каждом поколении: $N_k = 1000$; k — номер поколения.

Число «лучших» точек в каждом поколении: $m = 20$.

Количество пробных точек в группах: $N_k^{(1)} = N_k/4$, $N_k^{(2)} = 3N_k/4$.

Постоянные множители в формулах (3.1): $\sigma^{(1)}(k) = 1, \sigma^{(2)}(k) = 2k\sqrt{k}$.

Число «лучших» точек в каждой группе: $m_1 = m/2, m_2 = m/2$.

Критерий останова поиска: $\epsilon = 10^{-5}$; если $f_{min}(k-1) - f_{min}(k) \leq \epsilon$, поиск прекращается. Здесь $f_{min}(k)$ — минимальное значение функции (5.2), полученное на k-й итерации.

Результаты оптимизации. Представим графики, иллюстрирующие работу генетического алгоритма. При использовании алгоритма без модификаций область расположения пробных точек быстро сжимается с ростом номера поколения (см. рис. 1). На плоскости (λ_1, λ_2) представлены пять поколений пробных точек ($k = 0, 1, 2, 3, 4$). Уже при $k = 3$ область просмотра практически стягивается в точку.

На рис. 2 представлен результат применения модифицированного генетического алгоритма (MGA). На плоскости (λ_1, λ_2) также показаны пять поколений пробных точек. Область, занимаемая популяцией, в данном случае не сжимается с ростом номера поколения. Модификация дает возможность получать на каждом этапе поиска достаточно обширную область просмотра.

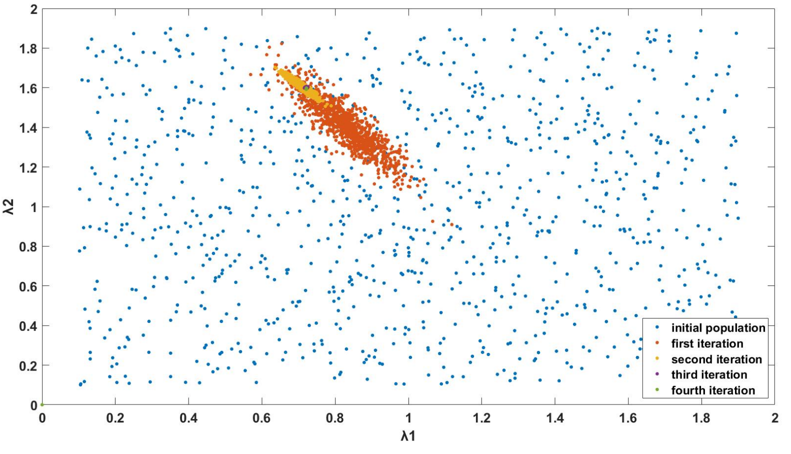

Рис. 1. Генетический алгоритм без модификаций: пять поколений пробных точек (k=0,1,2,3,4) представлены соответственно голубым, оранжевым, желтым, фиолетовым и зеленым цветами.

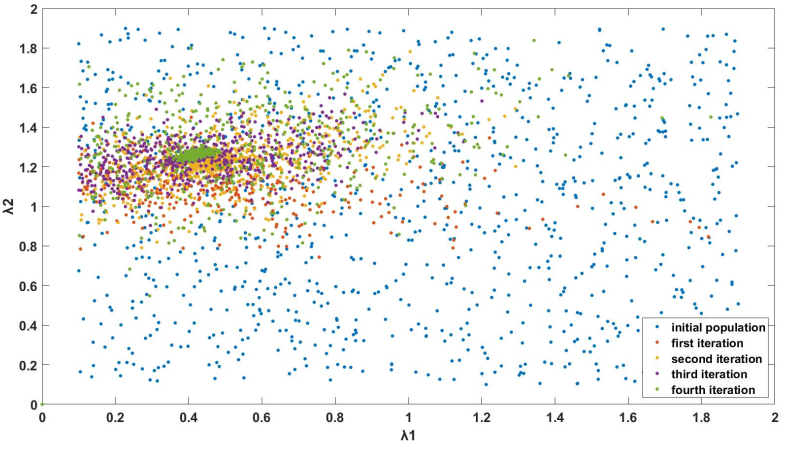

Рис. 2. Алгоритм MGA: пять поколений пробных точек (k=0,1,2,3,4) представлены соответственно голубым, оранжевым, желтым, фиолетовым и зеленым цветами.

Результаты работы модифицированного генетического алгоритма и их уточнение направленными методами показаны на рис. 3. На горизонтальной оси цифрами отмечены номера (k) этапов минимизации: итераций метода MGA и шагов направленных методов. По вертикальной оси откладывается минимальное значение целевой функции, полученное на том или ином этапе. Таким образом, график представляет зависимость $f_{min}(k)$. Синими квадратами отмечены значения $f_{min}(5)$, $f_{min}(6)$,

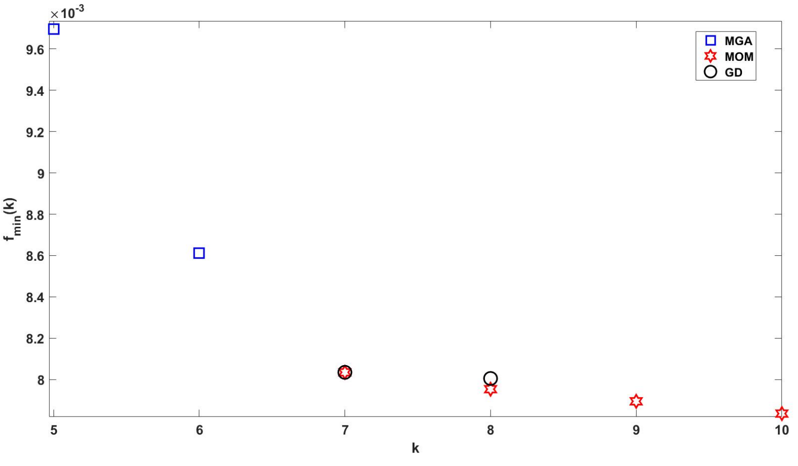

Рис. 3. График зависимости $f_{min}(k)$ при использовании алгоритма MGA (синие квадраты) в сочетании с методами GD (черные круги) и MOM (красные звездочки)

полученные на пятой и шестой итерациях алгоритма MGA. Вектор параметров, соответствующий f_{min} (6), является начальным вектором для методов GD и MOM. Как видно из графика, использование комбинации модифицированного генетического алгоритма и направленных методов оказалось успешным. При этом метод импульса (MOM), использующий на каждом шаге информацию двух предыдущих шагов, лучше уточняет решение и является более походящим для данной задачи.

При расчетах использовались следующие параметры направленных методов. Параметр импульса для метода MOM: $\mu_t = 0.4$. Шаг спуска: $\alpha_t = 4$ при $t = 0$ для обоих методов. Далее осуществлялся автоматический выбор шага для получения релаксационной последовательности пока $\alpha_t > 0.0001$.

В таблице 2 представлены полученные при оптимизации разными способами значения параметров функции регрессии и соответствующие значения f_{min} целевой функции. В строке MGA даны результаты использования модифицированного генетического алгоритма без уточнения, в строке MGA+MOM — с уточнением методом импульса, в строке MGA+GD — с уточнением методом градиентного спуска.

Данные таблицы и графиков подтверждают эффективность работы алгоритма MGA. Направленные методы незначительно уточняют решение.

Таблица 2

Итоговое сравнение результатов, полученных методом MGA без уточнения, с уточнением методом MOM и с уточнением методом GD

	Параметры функции регрессии				f_{min}
	β_1	λ_1	β_2	λ_2	
MGA	78.5602	1.1983	11.3690	0.2407	0.00861
MGA+MOM	78.5603	1.1979	11.3691	0.2412	0.00784
MGA+GD	78.5602	1.1980	11.3690	0.2407	0.00801

7. Заключение

В работе представлен генетический стохастический алгоритм поиска глобального экстремума и предложена его модификация. Алгоритм использует нормальное распределение с адаптацией ковариационной матрицы. Модификация сводится к разбиению популяции на две группы, при моделировании которых используются нормальные случайные величины с различными среднеквадратическими отклонениями, хотя бы одно из которых больше 1. Модификация позволяет осуществить мутацию популяции и обеспечить на каждой итерации подходящее количество проб как вблизи центра распределения, так и на периферии. Это дает возможность избежать слишком быстрого сжатия эллипсоида рассеяния и получения локального экстремума вместо глобального.

Модифицированный генетический алгоритм успешно применен к решению многоэкстремальной задачи оценки параметров функции регрессии. Стохастический метод дополнен направленными методами, которые незначительно уточняют решение. Представленные результаты подтверждают эффективность предложенной модификации и позволяют выбрать градиентный метод импульса как более подходящий для локального уточнения экстремума рассматриваемой функции.

Список источников

1. Владимирова Л. В., Жданова А. Ю., Рубцова И. Д. Использование генетического алгоритма глобального поиска в задаче оптимизации динамики пучка // VI Международная конференция «Лазерные, плазменные исследования и технологии - ЛаПлаз-2020» : сб. науч. тр. Ч. 1. М. : НИЯУ МИФИ, 2020. С. 91–92.
2. Владимирова Л. В., Овсянников Д. А., Рубцова И. Д. Методы Монте-Карло в прикладных задачах. СПб. : Изд-во ВВМ, 2015. 167 с.
3. Ермаков С. М. Метод Монте-Карло и смежные вопросы. М. : Наука, 1975. 472 с.
4. Ермаков С. М., Митиоглова Л. В. Об одном методе поиска глобального экстремума функции, основанном на оценивании ковариационной матрицы // Автоматика и вычислительная техника. 1977. № 5. С. 38–41.

5. Жиглявский А. А. Математическая теория глобального случайного поиска. Л. : Изд-во ЛГУ, 1985. 296 с.
6. Котина Е. Д., Овсянников Д. А. Математическая модель совместной оптимизации программного и возмущенных движений в дискретных системах// Вестник Санкт-Петербургского университета. Прикладная математика. Информатика. Процессы управления. 2021. Т. 17. Вып. 2. С. 213-224. https://doi.org/10.21638/11701/spbu10.2021.210
7. Нестеров Ю. Е. Методы выпуклой оптимизации. М. : Изд-во МЦНМО, 2010. 281 с.
8. Оптимизация динамики пучков траекторий с использованием гладких и негладких функционалов. Часть 1 / Д. А. Овсянников, М. А. Мизинцева, М. Ю. Балабанов, А. П. Дуркин, Н. С. Едаменко, Е. Д. Котина, А. Д. Овсянников // Вестник Санкт-Петербургского университета. Прикладная математика. Информатика. Процессы управления. 2020. Т. 16, вып. 1. С. 73–84. https://doi.org/10.21638/11702/spbu10.2020.107
9. Рубаник А. В. Решение экспоненциальной регрессионной задачи с использованием генетического алгоритма // Процессы управления и устойчивость. 2020. Т. 7, №1. С. 64-68.
10. Срочко В. А., Аксенюшкина Е. В. Параметризация некоторых задач управления линейными системами // Известия Иркутского государственного университета. Серия Математика. 2019. Т. 30. С. 83–98.
11. Срочко В. А., Аксенюшкина Е. В., Антоник В. Г. Решение линейно-квадратичной задачи оптимального управления на основе конечномерных моделей // Известия Иркутского государственного университета. Серия Математика. 2021. Т. 37. С. 3–16. https://doi.org/10.26516/1997-7670.2021.37.3
12. Метод кодифференциального спуска в задаче нахождения глобального минимума кусочно-аффинного целевого функционала в линейных системах управления / А. В. Фоминых, В. В. Карелин, Л. Н. Полякова, С. К. Мышков, В. П. Трегубов // Вестник Санкт-Петербургского университета. Прикладная математика. Информатика. Процессы управления. 2021. Т. 17. Вып. 1. С. 47–58. https://doi.org/10.21638/11701/spbu10.2021.105
13. Ermakov S. M., Semenchikov D. N. Genetic global optimization algorithms // Communications in Statistics Part B: Simulation and Computation. 2019. https://doi.org/10.1080/03610918.2019.1672739
14. Hansen N., Ostermeier A. Adapting arbitrary normal mutation distributions in evolution strategies: The covariance matrix adaptation // Proceedings of 1996 IEEE Conference on Evolutionary Computation (ICEC'96). Berlin, Germany, 1996. P. 312–317.
15. Igel C., Hansen N., Roth S. Covariance matrix adaptation for multi-objective optimization. // Evolutionary Computation. 2007. Vol. 15, N 1. P. 1–28.
16. Nocedal J., Wright S. J. Numerical optimization. Springer Series in Operations Research and Financial Engineering. Springer, Berlin, 2006. 634 p.
17. Qian N. On the momentum term in gradient descent learning algorithms // Neural networks. 1999. Vol. 12, N 1. P. 145-151. https://doi.org/10.1016/s0893-6080(98)00116-6
18. Vladimirova L., Fatyanova I. Construction of regression experiment optimal plan using parallel computing // International Conference "Stability and Control Processes" in Memory of V. I. Zubov, SCP 2015 : Proceedings. 2015. P. 361–363. 7342140
19. Vladimirova L. V., Ovsyannikov D. A. Random search for global extremum of a function using Markov chains simulation // Journal of Physics: Conference Series. 2019. Vol. 1238, N 1. 012073.

20. Genetic Stochastic Algorithm Application in Beam Dynamics Optimization Problem / L. V. Vladimirova, A. Y. Zhdanova, I. D. Rubtsova, N. S. Edamenko // Stability and Control Processes : Proceedings of the 4th International Conference Dedicated to the Memory of Professor Vladimir Zubov. Springer International Publishing (Lecture Notes in Control and Information Sciences - Proceedings), 2020.

References

1. Vladimirova L.V., Zhdanova A.Y., Rubtsova I.D. Application of the Genetic Global Search Algorithm in Beam Dynamics Optimization Problem. *VI International Conference on Laser&Plasma researches and technologies – LaPlas-2020: Proceedings, part 1.* Moscow, National Research Nuclear University MEPhI, 2020, pp.91–92. (in Russian)
2. Vladimirova L.V., Ovsyannikov D.A., Rubtsova I.D. *Metody Monte-Karlo v prikladnykh zadachakh* [Monte-Carlo Methods in Applied Problems]. St. Petersburg, VVM Publ., 2015, 167 p. (in Russian)
3. Ermakov S.M. *Metod Monte-Karlo i smezhnyye voprosy* [Monte-Carlo Method and Related Issues]. Moscow, Nauka Publ., 1975, 472 p. (in Russian)
4. Ermakov S.M., Mitioglova L.V. Ob odnom metode poiska global'nogo ekstremuma funktsii, osnovannom na otsenivanii kovariatsionnoy matritsy. [On Extreme Search Method Based on the Estimation of the Covariance Matrix]. *Avtomatika i vychislitel'naya tekhnika* [Computer Engineering], 1977, no. 5, pp. 38-41. (in Russian)
5. Zhiglyavsky A.A. *Matematicheskaya teoriya global'nogo sluchaynogo poiska* [Mathematical Theory of Global Random Search]. Leningrad, Leningrad St. Univ. Publ., 1985, 296 p. (in Russian).
6. Kotina E.D., Ovsyannikov D.A. Mathematical model of joint optimization of programmed and perturbed motions in discrete systems. *Vestnik of Saint Petersburg University. Applied Mathematics. Computer Science. Control Processes*, 2021, vol. 17, iss. 2, pp. 213-224. https://doi.org/10.21638/11701/spbu10.2021.210
7. Nesterov Yu.Ye. *Metody vypukloy optimizatsii.* [Convex Optimization Methods]. Moscow, MCCME Publ., 2010, 281 p. (in Russian)
8. Ovsyannikov D.A., Mizintseva M.A., Balabanov M.Yu., Durkin A.P., Edamenko N.S., Kotina E.D., Ovsyannikov A.D. Optimization of dynamics of trajectory bundles using smooth and nonsmooth functionals. Part 1. *Vestnik of Saint Petersburg University. Applied Mathematics. Computer Science. Control Processes*, 2020, vol. 16, iss. 1, pp. 73–84. https://doi.org/10.21638/11702/spbu10.2020.107
9. Rubanik A.V. Resheniye eksponentsial'noy regressionnoy zadachi s ispol'zovaniyem geneticheskogo algoritma [Solving an Exponential Regression Problem Using a Genetic Algorithm]. *Control Processes and Stability*, 2020, vol. 7, no. 1, pp. 64-68. (In Russian)
10. Srochko V.A., Aksenyushkina E.V. Parameterization of Some Control Problems by Linear Systems. *The Bulletin of Irkutsk State University.* Series Mathematics, 2019, vol. 30, pp. 83-98. https://doi.org/10.26516/1997-7670.2019.30.83 (In Russian)
11. Srochko V.A., Aksenyushkina E.V., Antonik V.G. Resolution of a Linear-quadratic Optimal Control Problem Based on Finite-dimensional Models. *The Bulletin of Irkutsk State University. Series Mathematics*, 2021, vol. 37, pp. 3-16. https://doi.org/10.26516/1997-7670.2021.37.3 (In Russian)
12. Fominyh A.V., Karelin V.V., Polyakova L.N., Myshkov S.K., Tregubov V.P. The codifferential descent method in the problem of finding the global minimum of

a piecewise affine objective functional in linear control systems. *Vestnik of Saint Petersburg University. Applied Mathematics. Computer Science. Control Processes*, 2021, vol. 17, iss. 1, pp. 47–58. https://doi.org/10.21638/11701/spbu10.2021.105 (In Russian)

13. Ermakov S.M., Semenchikov D.N. Genetic global optimization algorithms. *Communications in Statistics Part B: Simulation and Computation*, 2019. https://doi.org/10.1080/03610918.2019.1672739
14. Hansen N., Ostermeier A. Adapting arbitrary normal mutation distributions in evolution strategies: The covariance matrix adaptation. *Proceedings of 1996 IEEE Conference on Evolutionary Computation (ICEC'96)*, Berlin, Germany, 1996, pp. 312–317.
15. Igel C., Hansen N., Roth S. Covariance matrix adaptation for multi-objective optimization. *Evolutionary Computation*, 2007, vol. 15, no. 1, pp. 1–28.
16. Nocedal J., Wright S.J. *Numerical Optimization.* Springer Series in Operations Research and Financial Engineering. Springer, Berlin, 2006, 634 p.
17. Qian N. On the momentum term in gradient descent learning algorithms. *Neural networks*, 1999, vol. 12, no. 1, pp. 145-151. https://doi.org/10.1016/s0893-6080(98)00116-6
18. Vladimirova L., Fatyanova I. Construction of regression experiment optimal plan using parallel computing. *2015 International Conference "Stability and Control Processes"in Memory of V.I. Zubov, SCP 2015 - Proceedings*, 2015, pp. 361–363, 7342140
19. Vladimirova L.V., Ovsyannikov D.A. Random search for global extremum of a function using Markov chains simulation. *Journal of Fhysics: Conference Series*, 2019, vol. 1238, no. 1, 012073. https://iopscience.iop.org/article/10.1088/1742-6596/1238/1/012073
20. Vladimirova L.V., Zhdanova A.Y., Rubtsova I.D., Edamenko N.S. Genetic Stochastic Algorithm Application in Beam Dynamics Optimization Problem. *Stability and Control Processes - Proceedings of the 4th International Conference Dedicated to the Memory of Professor Vladimir Zubov, Springer International Publishing (Lecture Notes in Control and Information Sciences - Proceedings)*, 2020.

Об авторах

Овсянников Дмитрий Александрович, д-р физ.-мат. наук, проф., Санкт-Петербургский государственный университет, Российская Федерация, 199034, г. Санкт-Петербург, d.a.ovsyannikov@spbu.ru, https://orcid.org/0000-0002-0829-2023.

About the authors

Dmitri A. Ovsyannikov, Dr. Sci. (Phys.-Math.), Prof., Saint Petersburg State University, St. Petersburg, 199034, Russian Federation, d.a.ovsyannikov@spbu.ru, https://orcid.org/0000-0002-0829-2023.

Владимирова Людмила Васильевна, канд. физ.-мат. наук, доц., Санкт-Петербургский государственный университет, Российская Федерация, 199034, г. Санкт-Петербург l.vladimirova@spbu.ru, https://orcid.org/0000-0002-7612-8719.

Liudmila V. Vladimirova, Cand. Sci. (Phys.Math.), Assoc. Prof., Saint Petersburg State University, St. Petersburg, 199034, Russian Federation, l.vladimirova@spbu.ru, https://orcid.org/0000-0002-7612-8719.

Рубцова Ирина Деонисовна, канд. физ.-мат. наук, доц., Санкт-Петербургский государственный университет, Российская Федерация, 199034, г. Санкт-Петербург, i.ribtsova@spbu.ru, https://orcid.org/0000-0002-8531-8522.

Irina D. Rubtsova, Cand. Sci. (Phys.Math.), Assoc. Prof., Saint Petersburg State University, St. Petersburg, 199034, Russian Federation, i.ribtsova@spbu.ru, https://orcid.org/0000-0002-8531-8522.

Рубаник Алексей Витальевич, Санкт-Петербургский государственный университет, Российская Федерация, 199034, г. Санкт-Петербург, st040092@student.spbu.ru, https://orcid.org/0000-0003-2826-8933.

Alexey V. Rubanik, Saint Petersburg State University, St. Petersburg, 199034, Russian Federation, st040092@student.spbu.ru, https://orcid.org/0000-0003-2826-8933.

Пономарев Владимир Андреевич, ассистент, Санкт-Петербургский государственный университет, Российская Федерация, 199034, г. Санкт-Петербург, v.a.ponomarev@spbu.ru, https://orcid.org/0000-0002-1958-4822.

Vladimir A. Ponomarev, Assistant, Saint Petersburg State University, St. Petersburg, 199034, Russian Federation, v.a.ponomarev@spbu.ru, https://orcid.org/0000-0002-1958-4822.

Поступила в редакцию / Received 10.01.2022
Поступила после рецензирования / Revised 30.01.2022
Принята к публикации / Accepted 21.02.2022