УДК 510.5

Н.В. Панов, С.П. Шарый

Интервальный эволюционный алгоритм поиска глобального оптимума

N.V. Panov, S.P. Shary

Interval Evolutionary Algorithm for Global Optimization

Для поиска глобального экстремума функций непрерывного аргумента на брусе со сторонами, параллельными координатным осям, предлагается интервальная версия эволюционного алгоритма.

Ключевые слова: интервал, интервальное расширение, глобальная оптимизация, стохастические алгоритмы, рандомизированные интервальные алгоритмы, доказательные вычисления, эволюционный алгоритм.

This paper proposes a new interval version of evolutional algorithm to find a global function extreme on a bar with sides parallel to coordinate axis.

Key words: interval, interval expansion, global optimization, interval algorithm, conclusive computations, evolutional algorithm.

1. Постановка задачи. Предмет настоящей работы – поиск глобального оптимума вещественнозначной целевой функции $f:\mathbb{R}^n \to \mathbb{R}$ на некотором прямоугольном брусе $\boldsymbol{X} \subset \mathbb{R}^n$ со сторонами, параллельными координатным осям, который является подмножеством области определения данной функции:

$$\text{найти } \min_{x\in \boldsymbol{X}} f(x). \qquad (1)$$

Чрезвычайно много задач, возникающих в различных сферах человеческой деятельности, может быть сведено к подобной задаче поиска глобального оптимума. Многомерная оптимизация – неотъемлемая часть важнейших этапов моделирования различных (инженерных, экономических и т.д.) систем. В ряде случаев именно сложность возникающей оптимизационной задачи становится тем ограничением, которое не позволяет исследовать общую постановку проблемы.

К настоящему моменту наработан богатый инструментарий поиска глобального оптимума [1–5]. Трудности численного решения оптимизационных задач во многом связаны с видом оптимизируемой целевой функции. Особенно сложно решение задач оптимизации в случаях, когда целевая функция невыпукла, многоэкстремальна или недифференцируема. Далеко не все методы могут справляться с такими задачами, особенно в случае, когда нам не известна какая-либо априорная информация о целевой функции. Методы, о которых пойдет речь в статье, работают и для таких тяжелых постановок задач. Единственное условие, накладываемое в этом случае на целевую функцию, состоит в том, что она должна задаваться явно в виде математического выражения или компьютерной подпрограммы, состоящих из комбинаций символов переменных, арифметических операций и математических функций. Это необходимо для вычисления интервальной оценки области значений функции (подробнее см.: [6]).

В ряде задач, возникающих в практике оптимизации, требуется не просто приближенное численное решение, но еще и гарантия его близости к идеальному математическому оптимуму, а также часто гарантия того, что найденный оптимум действительно является глобальным, а не одним из локальных. Подобные постановки задач обычно характеризуют термином «доказательная глобальная оптимизация»[1], и они являются чрезвычайно трудными. Традиционные подходы к их решению основываются на привлечении той или иной априорной информации о целевой функции (например, выпуклость, унимодальность, удовлетворение условию Липшица). Существенное продвижение в решении задач доказательной глобальной оптимизации связано в последние годы с привлечением методов интервального анализа. Эти методы позволяют успешно решать задачи с осложненными целевыми функциями – нелипшицевыми, недифференцируемыми и т.п. [6].

Традиционно такие алгоритмы основываются на детерминистской вычислительной схеме, что в некоторых случаях приводит к их недостаточной вычислительной эффективности [7–10]. Следуя идеям, впервые сформулированным в [11, 12], авторы надеются улучшить эффективность доказательных интервальных методов оптимизации, введя в них вероятностные и эвристические переходы, но не потеряв при этом гарантированности результатов. Подробнее о недостатках детерминистских методов можно узнать из публикаций [11–13]. Главное направление наших построений – это комбинирование интервальной техники оценивания значений

[1] По терминологии, введенной К.И. Бабенко (Основы численного анализа. М., 1986).

целевой функции с так называемыми эволюционными вычислениями, эксплуатирующими бионические идеи.

Бионические алгоритмы – это эвристические алгоритмы поиска, так или иначе использующие для решения задач оптимизации случайный подбор, комбинирование и вариации входных параметров на основе механизмов, напоминающих биологическую эволюцию. Идеи применить биологические механизмы или, скорее, принципы их функционирования для создания новых алгоритмов возникли с небольшими вариациями практически одновременно у нескольких авторов. В 1966 г. Л. Фогель, А. Оуэнс, М. Уолш предложили схему эволюции логических автоматов, решающих задачи прогноза [14]. В 1975 г. вышла основополагающая книга Дж. Холланда «Адаптация в естественных и искусственных системах» [15], в которой был предложен законченный вариант генетического алгоритма. Примерно в это же время немецкие ученые И. Рехенберг и Х.-П. Швефель начали разработку так называемой эволюционной стратегии вычислений [16, 17]. Эти работы заложили основы прикладного эволюционного моделирования или эволюционных алгоритмов. В нашей стране исследования по прикладному эволюционному моделированию, идейно близкие к упомянутым работам Л. Фогеля с сотрудниками, были разносторонне развиты, в частности, И.Л. Букатовой [18].

Уже первые исследования показали, что идея использовать принципы биологической эволюции оказа-лась плодотворной. В настоящее время множество разновидностей бионических алгоритмов (главным образом, «генетические» и «поведенческие») используется в самых разнообразных ситуациях, в том числе и для решения задач глобальной оптимизации [3, 19, 20]. Все они, как правило, используют точечное оценивание целевой функции. Впервые идея встраивания интервальной техники в генетические алгоритмы была высказана, по-видимому, в [21], но, к сожалению, в этой работе все ограничилось краткой словесной формулировкой общей идеи, а схема (и тем более детали реализации) получающегося алгоритма при этом представлена не была. Еще одна интересная попытка комбинирования генетических и интервальных алгоритмов – это работа [22], в которой генетический алгоритм используется, главным образом, для уточнения верхней границы глобального оптимума. Настоящая статья является продолжением работ авторов [8–13] и посвящена описанию нового интервального бионического алгоритма, использующего эволюционную организацию вычислений вкупе с интервальными методами оценивания значений функций. В нем стратегия генетического алгоритма применяется для глубокой модификации схемы дробления области определения, лежащей в основе всех традиционных интервальных алгоритмов глобальной оптимизации.

2. Общее описание алгоритма и основ интервального анализа. В эволюционных алгоритмах каким-либо образом (чаще всего случайно) задается начальная «*популяция*» – некое множество объектов. Они оцениваются с помощью «*функции приспособленности*», в результате чего каждому объекту присваивается определенное значение (*приспособленность*), которое определяет вероятность выживания организма, представленного данным объектом. После этого с использованием полученных значений приспособленности выбираются объекты (производится *селекция*), допущенные к «*размножению*». Также к этим объектам могут применяться «*генетические операторы*» (в большинстве случаев это так называемое *скрещивание – crossover* и *мутация – mutation*). Особи следующего поколения также оцениваются, затем снова производится *селекция*, применяются генетические операторы и т.д. Так моделируется «*эволюционный процесс*», продолжающийся в течение нескольких жизненных циклов (*поколений*), до тех пор пока не будет выполнен критерий остановки алгоритма. На листинге 1 приведены основные этапы эволюционного алгоритма.

Листинг 1. Эволюционный алгоритм.

1. Создание начальной популяции.
2. Вычисление функций приспособленности для особей популяции (оценивание).

(*Начало цикла*)
(a) Выбор индивидов из текущей популяции (селекция).
(b) Формирование нового поколения.
(c) Применение генетических операторов (например, мутации).
(d) Вычисление функций полезности для всех особей.
(e) Если достигнуто условие останова, то решение найдено (конец работы), иначе тело цикла повторяется.
(*Конец цикла*)

Этап мутаций не обязателен. С одной стороны, он позволяет повысить рандомизацию метода, увеличить вариабельность популяции и избежать застоя вблизи локальных оптимумов. Кроме того, в классических генетических алгоритмах операции мутации и скрещивания могут порождать новые решения, которые никогда не встречались в предыдущих поколениях (подробнее о генетических операторах см.: [19, 20]). Интервальный алгоритм, описываемый в настоящей статье, их не использует. Тем не менее в отличие от классических вариантов он позволяет найти именно глобальный оптимум. Прежде чем подробнее рассмотреть вычислительную схему алгоритма, напомним ряд необходимых понятий из интервального анализа.

Интервал $\boldsymbol{x}$ – это замкнутый отрезок вещественной оси $\mathbb{R}$, так что

$$\boldsymbol{x} = [\underline{\boldsymbol{x}}, \overline{\boldsymbol{x}}] = \{x \in \mathbb{R} \,|\, \underline{\boldsymbol{x}} \le x \le \overline{\boldsymbol{x}}\},$$

а $\underline{\boldsymbol{x}}$ и $\overline{\boldsymbol{x}}$ называются, соответственно, левым (или нижним) и правым (или верхним) концами интервала. Многомерные интервалы (также называемые брусами) определяются как векторы (вектор-столбцы или вектор-строки) с интервальными компонентами. Интервальные величины мы будем обозначать жирным шрифтом.

Задача об определении области значений функции на том или ином подмножестве области ее определения эквивалентна двум задачам оптимизации: для непрерывной функции

$$\operatorname{ran}_{\boldsymbol{x}} f = [\min_{x\in\boldsymbol{x}} f(x), \max_{x\in\boldsymbol{x}} f(x)],$$

где через $\operatorname{ran}_{\boldsymbol{x}} f := \{f(x) \mid x \in \boldsymbol{x}\}$ обозначена область значений функции f на брусе $\boldsymbol{x}$. В интервальном анализе задача определения области значений сводится к задаче о вычислении так называемого интервального расширения функции. Обозначим через $\mathbb{I}D$ множество всех интервальных векторов-брусов, содержащихся во множестве $D \subset \mathbb{R}^n$. Напомним

Определение 1.1. *Пусть D – непустое подмножество пространства* $\mathbb{R}^n$. *Интервальная функция* $\boldsymbol{f}: \mathbb{I}D \to \mathbb{I}\mathbb{R}^m$ *называется интервальным продолжением вещественной функции* $f: D \to \mathbb{R}^m$, *если* $\boldsymbol{f}(x) = f(x)$ *для всех* $x \in D$.

Определение 1.2 [5–7, 23–25]. *Пусть D – непустое подмножество пространства* $\mathbb{R}^n$. *Интервальная функция* $\boldsymbol{f}: \mathbb{I}D \to \mathbb{I}\mathbb{R}^m$ *называется интервальным расширением вещественной функции* $f: D \to \mathbb{R}^m$, *если*

1) $\boldsymbol{f}(\boldsymbol{x})$ – интервальное продолжение $f(x)$ на D,

2) $\boldsymbol{f}(\boldsymbol{x})$ монотонна по включению, т.е. $\boldsymbol{x} \subseteq \boldsymbol{y} \Rightarrow \boldsymbol{f}(\boldsymbol{x}) \subseteq \boldsymbol{f}(\boldsymbol{y})$ на $\mathbb{I}D$.

Таким образом, если $\boldsymbol{f}(\boldsymbol{x})$ – интервальное расширение функции $f(x)$, то для области значений f на брусе $\boldsymbol{x} \subset \mathbb{R}^n$ мы получаем следующую внешнюю (с помощью объемлющего множества) оценку:

$$\{f(x) \mid x \in \boldsymbol{x}\} \subseteq \boldsymbol{f}(\boldsymbol{x}).$$

В частности, если $\boldsymbol{f}$ – интервальное расширение целевой функции f из задачи (1), то для ее решения справедлива оценка

$$\underline{\boldsymbol{f}(\boldsymbol{x})} \le \min_{x\in\boldsymbol{x}} f(x).$$

Развитие эффективных методов построения интервальных расширений функций – это важнейшая задача интервального анализа и его приложений. Поиски ее различных решений продолжаются и в настоящее время. Приведем в рамках нашего беглого обзора некоторые общезначимые результаты на эту тему. Первый из них часто называют «основной теоремой интервальной арифметики»:

Теорема [1, 4, 5, 7, 23, 24]. *Если для рациональной функции* $f(x)$ *на интервале* $\boldsymbol{x} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots \boldsymbol{x}_n)$ *определен результат* $\boldsymbol{f}_\lozenge(\boldsymbol{x})$ *подстановки вместо ее аргументов* $(x_1, x_2, \ldots x_n) = x$ *интервалов их изменения и выполнения всех действий над ними по правилам интервальной арифметики, то*

$$\{f(x) \mid x \in \boldsymbol{x}\} \subseteq \boldsymbol{f}_\lozenge(\boldsymbol{x}),$$

т.е. $\boldsymbol{f}_\lozenge(\boldsymbol{x})$ *содержит множество значений функции* $f(x)$ *на* $\boldsymbol{x}$.

Нетрудно понять, что по отношению к рациональной функции $f(x)$ интервальная функция $\boldsymbol{f}_\lozenge(\boldsymbol{x})$, о которой идет речь в теореме, является интервальным расширением. Оно называется *естественным интервальным расширением* и вычисляется совершенно элементарно.

Важной особенностью интервальных оценок области значений функции является свойство их асимптотической точности (доказательство заинтересованный читатель может найти, к примеру, в [4–6]): если $\boldsymbol{f}(\boldsymbol{x})$ – какое-либо интервальное расширение функции $f(x)$ на брусе $\boldsymbol{x}$, то, как правило,

$$\operatorname{dist}(\boldsymbol{f}(\boldsymbol{x}), \operatorname{ran}_{\boldsymbol{x}} f) \to 0 \qquad \text{при } \operatorname{wid} \boldsymbol{x} \to 0, \qquad (2)$$

где dist – расстояние между интервалами, определяемое как

$$\operatorname{dist}(\boldsymbol{u}, \boldsymbol{v}) = \max\{|\underline{\boldsymbol{u}} - \underline{\boldsymbol{v}}|, |\overline{\boldsymbol{u}} - \overline{\boldsymbol{v}}|\}.$$

Это свойство означает, что при уменьшении размеров области определения точность интервального расширения функции увеличивается. При этом не следует ожидать, что уменьшение размеров конкретного бруса области определения в какое-то число раз приведет к пропорциональному улучшению реальной точности конкретной интервальной оценки. Приведенное соотношение является, во-первых, всего лишь оценкой сверху и, во-вторых, носит асимптотической характер. Тем не менее отмеченный факт может быть положен в основу процедуры уточнения интервальной оценки области значений функции.

В самом деле, если разбить исходный брус $\boldsymbol{x}$ на два подбруса $\boldsymbol{x}'$ и $\boldsymbol{x}''$, дающие в объединении весь $\boldsymbol{x}$, т.е. такие, что $\boldsymbol{x}' \cup \boldsymbol{x}'' = \boldsymbol{x}$, то

$$\{f(x) \mid x \in \boldsymbol{x}\} = \{f(x) \mid x \in \boldsymbol{x}'\} \cup \{f(x) \mid x \in \boldsymbol{x}''\}.$$

Соответственно, можно вычислить интервальные расширения по каждому подбрусу, и в качестве новой оценки минимума целевой функции на $\boldsymbol{x}$ взять

$$\min\{\underline{\boldsymbol{f}(\boldsymbol{x}')}, \underline{\boldsymbol{f}(\boldsymbol{x}'')}\}.$$

Она будет, вообще говоря, более точна, чем исходная оценка $\boldsymbol{f}(\boldsymbol{x})$, так как у брусов $\boldsymbol{x}'$ и $\boldsymbol{x}''$ размеры меньше, чем у исходного $\boldsymbol{x}$. Брусы-потомки $\boldsymbol{x}'$ и $\boldsymbol{x}''$ можно, в свою очередь, опять разбить на более мелкие части, найти для них интервальные расширения и далее уточнить оценку для минимума, потом снова повторить процедуру и т.д. По такой схеме фактически действуют все интервальные методы глобальной оптимизации, основанные на адаптивном дроблении исходной области определе-

3. Интервальный эволюционный алгоритм. В интервальном эволюционном алгоритме исходная область поиска по-прежнему разбивается на непересекающиеся подобласти – брусы. Каждый такой брус в описываемой здесь разновидности эволюционного алгоритма выступает в качестве «особи». При этом, как и в классическом случае, при запуске работы алгоритма каким-либо способом требуется задать начальную популяцию. Сделать это можно совершенно произвольным образом, так как выбор начальной популяции практически никак не сказывается на дальнейшей работе алгоритма. При желании этот шаг вообще можно пропустить, приняв брус исходной области поиска за единственную существующую особь.

В качестве меры приспособленности особи можно выбрать нижнюю границу (для определенности рассматриваем поиск минимума) интервальной оценки целевой функции на соответствующем брусе. Чем лучше мера приспособленности (чем меньше нижняя граница в привычных терминах), тем больше шансов у данной особи размножиться (данному брусу быть раздробленным) и тем многочисленнее будет потомство (брус может быть раздроблен на большее количество подбрусов).

Для сохранения доказательности вычислений (которая является характерной чертой интервальных методов) мы не исключаем из рассмотрения никакие подобласти исходной области поиска до тех пор, пока не будет доказано, что они гарантированно не содержат оптимум. Подробнее о различных подходах к выявлению бесперспективности брусов (их называют еще «критериями отбраковки») можно прочитать, например, в [13]. Таким образом, в алгоритме особи либо рождаются нежизнеспособными (не выдерживают критериев отбраковки, применяемых сразу после дробления), либо погибают в результате «*эпидемий*» – срабатывания уточненных критериев отбраковки при дальнейшем исполнении алгоритма.

Листинг 2. Схема интервального генетического алгоритма.

Задается начальная популяция и передается в основной цикл

Основной цикл

{

1. **Вычислить значения функции приспособ-лен-ности новорожденных особей** (этот шаг включает вычисление интервального расширения це-левой функции по новым подбрусам как необходимое для определения приспособленности).
2. **N из наиболее приспособленных брусов с вероятностью P_n порождают от L_n до U_n потомков.**
3. **M из неприспособленных брусов с вероятностью P_m порождают от L_m до U_m потомков.**
4. **Потомки проверяются на жизнеспособность** (применяются интервальные критерии отбраковки).
5. **Если критерий отбраковки был уточнен, возможно, случается эпидемия** (улучшенные критерии применяются ко всем особям).

}

В этой схеме имеется несколько параметров, существенным образом влияющих на работу алгоритма:

– вероятности P_n и P_m;

– максимальные количества потомков L_n и L_m;

– минимальное количество потомков U_n и U_m;

– величина N, определяющая количество объектов, начиная с самого приспособленного, которые могут оставить потомство;

– брусы дробятся на равные части или разбиение происходит в случайной пропорции, «разновесные» дети;

– количество «особей» в начальной популяции.

Заметим, что если скорректировать условия работы алгоритма таким образом, чтобы потомки были равноправны и их было обязательно двое, но оставить потомство мог лишь «вожак» (т.е. самый приспособленный), то в более привычных терминах это будет звучать как «на каждой итерации дробится брус с наименьшей нижней оценкой». В этом случае реализуется детерминистская вычислительная схема классического интервального адаптивного дробления, в основе которой лежит «метод ветвей и границ».

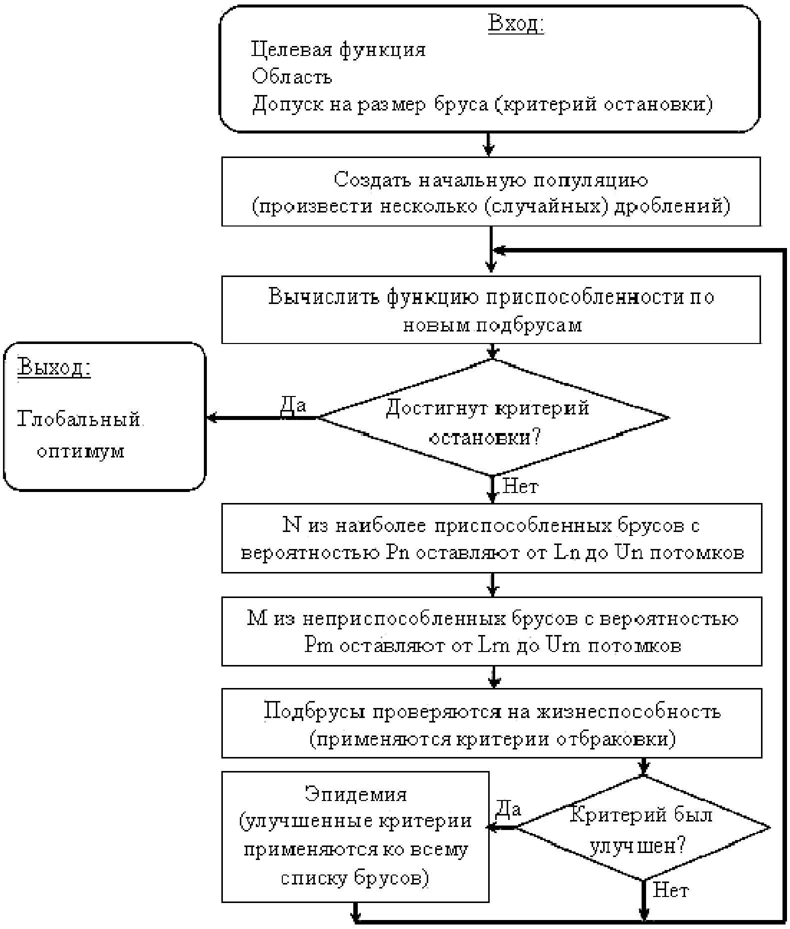
В качестве меры приспособленности мы использовали в алгоритме нижнюю границу интервальной оценки функции на подбрусе (оценку оптимума на подбрусе снизу). С тем же успехом можно опираться и на ширину интервальной оценки или же на размер бруса. Несмотря на то, что все три метода очень похожи и критерии, которыми они руководствуются, направлены на достижение одной и той же цели, работают они по-разному. Таким образом, следующим логичным шагом является объединение этих методов в один, использующий преимущества каждого.

Блок-схема объединенного алгоритма показана на рисунке.

Функция приспособленности бруса (при поиске минимума) приобретает следующий вид:

$$Fit(\boldsymbol{b}) = \sum\nolimits_{i-1}^{n} (\alpha_i \cdot \mathrm{wid}\, \boldsymbol{b}_i) + \beta \cdot \boldsymbol{f}(\underline{\boldsymbol{b}}) + \gamma \cdot \mathrm{wid}\, \boldsymbol{f}(\boldsymbol{b}) \cdot \quad (3)$$

Здесь $\boldsymbol{f}(\boldsymbol{b})$ – интервальная оценка области значений целевой функции f на брусе $\boldsymbol{b}$; $\boldsymbol{f}(\underline{\boldsymbol{b}})$ – нижняя граница интервальной оценки (оценка минимума снизу); $\mathrm{wid}\, \boldsymbol{f}(\boldsymbol{b})$ – ширина (точность) интервальной оценки области значений; wid $\boldsymbol{b}_i$ – ширина i-й компоненты бруса области определения; $\alpha_1, ..., \alpha_n, \beta, \gamma$ – соответствующие весовые коэффициенты.



Блок-схема интервального генетического алгоритма

Необходимо заметить, что поведение алгоритма очень сильно зависит от величин коэффициентов и авторам представляется перспективной реализация подобного алгоритма с использованием интервальных параметров, равно как и разработка мета-алгоритма, адаптивно оптимизирующего параметры эволюционного алгоритма в процессе решения конкретных задач.

Библиографический список

1. Сергеев Я.Д., Квасов Д.Е. Диагональные методы глобальной оптимизации. – М., 2008.

2. Diwekar U. Introduction to Applied Optimization. – Springer, 2008.

3. Zhigljavsky A., Zilinskas A. Stochastic Global Optimization. – Springer, 2008.

4. Hansen E., Walster G. Global Optimization Using Interval Analysis. – New York, 2004.

5. Moore R.E., Kearfott R., Cloud M.J. Introduction to Interval Analysis. – SIAM, 2009.

6. Шарый С.П. Конечномерный интервальный анализ. – Новосибирск, 2011. – Электронная книга, доступная на http://www.nsc.ru/interval/index.php?j=Library/InteBooks/index.

7. Калмыков С.А., Шокин Ю.И., Юлдашев З.Х. Методы интервального анализа. – Новосибирск, 1986.

8. Панов Н.В. Гибкость и гарантированность. Интервальные стохастические методы поиска оптимума // IX Всероссийская конференция молодых ученых по математическому моделированию и информационным технологиям. – Кемерово, 2008.

9. Панов Н.В. Объединение стохастических и интервальных подходов для решения задач глобальной оптимизации функций // Вычислительные технологии. – 2009. – Т. 14, №5.

10. Панов Н.В., Шарый С.П. Развитие стохастических подходов в интервальной глобальной оптимизации // VIII Всероссийская конференция молодых ученых по математическому моделированию и информационным технологиям. – Новосибирск, 2007.

11. Шарый С.П. Стохастические подходы в интервальной глобальной оптимизации // Труды XIII Байкальской международной школы-семинара «Методы оптимизации и их приложения». Т. 4: Интервальный анализ. – Иркутск, 2005.

12. Шарый С.П. Рандомизированные алгоритмы в интервальной глобальной оптимизации // Сибирский

журнал вычислительной математики. – 2008. – Т. 11, №4.

13. Панов Н.В., Шарый С.П. Стохастические подходы в интервальных методах глобальной оптимизации // Всероссийское (с международным участием) совещание по интервальному анализу и его приложениям ИНТЕРВАЛ-06. – СПб., 2006.

14. Фогель Л., Оуэнс А., Уолш М. Искусственный интеллект и эволюционное моделирование. – М., 1969.

15. Holland J. Adaptation in Natural and Artificial Systems. – Michigan, 1975.

16. Rechenberg I. Evolutionstrategie: Optimierung technischer Systeme nach Prinzipien der biologischen Evolution. – Stuttgart, 1973. (1993 – 2nd edn.).

17. Schwefel H.-P. Numerische Optimierung von Computer-Modellen mittels der Evolutionsstrategie. – Basel, 1977.

18. Букатова И.Л. Эволюционное моделирование и его приложения. – М., 1979.

19. Гладков Л.А., Курейчик В.В., Курейчик В.М. и др. Биоинспирированные методы в оптимизации. – М., 2009.

20. Fogel D.B. Evolutionary Computation: Towards a New Philosophy of Machine Intelligence. – New York, 2006.

21. Alander J. On interval factorial genetic algorithm in global optimization // APIC’95. – Texas, 1995.

22. Zhang X., Liu S. A new interval-genetic algorithm // Third International Conference on Natural Computation ICNC. – Haikou, 2007.

23. Kearfott R.B. Rigorous global search: Continuous problems. – Dordrecht, 1996.

24. Neumaier A. Interval methods for systems of equations. – Cambridge, 1990.

25. Ratschek H., Rokne J. New computer methods for global optimization. – Chichester; New York, 1988.