

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/283425967>

Review of the particle swarm optimization method (PSO) for a global optimization problem

Article in Science and Education of the Bauman MSTU · September 2015

DOI: 10.7463/00309.0116072

CITATIONS

7

READS

441

2 authors:



[Anatoly Pavlovich Karpenko](#)

Bauman Moscow State Technical University

113 PUBLICATIONS 280 CITATIONS

[SEE PROFILE](#)



[Evgeny Seliverstov](#)

4 PUBLICATIONS 9 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Quality of teaching material's estimation [View project](#)



Students' Meta Competences Evaluation [View project](#)

Обзор методов роя частиц (PSO) для задачи глобальной оптимизации

А.П.Карпенко, Е.Ю.Селиверстов

МГТУ им. Н.Э. Баумана, 105005, Москва, 2-я Бауманская ул., д.5.

Рассматривается метод роя частиц для задачи глобальной многопараметрической оптимизации (PSO), а также известные последовательные и параллельные модификации этого метода.

Ключевые слова: глобальная оптимизация, метод роя частиц, метод PSO.

This paper presents review of particle swarm optimization method (PSO) for global optimization problem and its certain serial and parallel modifications.

Key words: global optimization, particle swarm optimization method, PSO method.

1. Введение

Среди задач непрерывной конечномерной оптимизации самым важным с практической точки зрения и, одновременно, самым сложным является класс задач глобальной условной оптимизации. Методы решения задач этого класса можно разделить на две большие группы:

- методы сведения задачи глобальной условной оптимизации к задаче глобальной безусловной оптимизации с помощью штрафных или барьерных функций;
- методы, специально сконструированные для решения задачи глобальной условной оптимизации.

Рассматриваемый в работе метод роя частиц относится к методам первой группы и, таким образом, предназначен для решения задач глобальной безусловной оптимизации.

Методы решения задачи глобальной безусловной оптимизации делятся на детерминированные методы, стохастические методы и эвристические методы [1].

Эвристические методы являются относительно новыми и быстро развивающимися методами. Среди этих методов выделяются эволюционные и поведенческие (имитационные) методы.

Поведенческие методы решения задачи глобальной безусловной оптимизации основаны на моделировании коллективного поведения самоорганизующихся живых или неживых систем. Взаимодействующие элементы этих систем, в общем случае, называются агентами. Ключевыми идеями поведенческих методов являются децентрализованность, взаимодействие агентов, простота поведения агентов.

Известность получили следующие поведенческие методы решения задачи глобальной безусловной оптимизации: метод поведения пчёл; метод колонии муравьев; рассматриваемый в данной работе метод роя частиц.

В методе оптимизации роем частиц (particle swarm optimization - PSO) агентами являются частицы в пространстве параметров задачи оптимизации. В каждый момент времени (на каждой итерации) частицы имеют в этом пространстве некоторое положение и вектор скорости. Для каждого положения частицы вычисляется соответствующее значение целевой функции, и на этой основе по определенным правилам частица меняет свое положение и скорость в пространстве поиска.

В основу метода PSO положена социально-психологическая поведенческая модель толпы. Существует несколько разновидностей метода. Например, в каноническом методе роя частиц, предложенном в 1995 году в работе Kennedy, Eberhart [2], на каждой итерации при определении следующего положения частицы учитывается информация о наилучшей частице из числа «соседей» данной частицы, а также информация о данной частице на той итерации, когда этой частице соответствовало наилучшее значение целевой функции. Модификация канонической модели FIPS учитывает значения целевой функции, соответствующие всем частицам роя; в некоторых моделях частицы группируются в несколько роев и т.д.

Большинство известных методов роя частиц являются последовательными. Параллельных методов известно немного, и все они появились после 2004 года. В работе рассмотрены, как последовательные, так и параллельные методы PSO.

Основным стимулом для подготовки данной работы явилось почти полное отсутствие русскоязычных исследований и публикаций по методу роя частиц.

2. Канонический метод роя частиц

Рассмотрим задачу глобальной безусловной минимизации целевой функций $\Phi(X)$ в n -мерном арифметическом пространстве R^n :

$$\min_{X \in R^n} \Phi(X) = \Phi(X^*). \quad (1)$$

Множество частиц обозначим $\mathbf{P} = \{P_i, i \in [1: N]\}$, где N - количество частиц в рое (размер популяции). В момент времени $t = 0, 1, 2, \dots$ координаты частицы P_i определяются вектором $X_{i,t} = (x_{i,t,1}, x_{i,t,2}, \dots, x_{i,t,n})$, а ее скорость - вектором $V_{i,t} = (v_{i,t,1}, v_{i,t,2}, \dots, v_{i,t,n})$. Начальные координаты и скорости частицы P_i равны $X_{i,0} = X_i^0$, $V_{i,0} = V_i^0$, соответственно.

Итерации в каноническом методе PSO выполняются по следующей схеме:

$$X_{i,t+1} = X_{i,t} + V_{i,t+1}; \quad (2)$$

$$V_{i,t+1} = \alpha V_{i,t} + U[0, \beta] \otimes (X_{i,t}^b - X_{i,t}) + U[0, \gamma] \otimes (X_{g,t} - X_{i,t}). \quad (3)$$

Здесь $U[a, b]$ представляет собой n -мерный вектор псевдослучайных чисел, равномерно распределенных в интервале $[a, b]$; \otimes - символ покомпонентного умножения векторов; $X_{i,t}^b$ - вектор координат частицы P_i с наилучшим (в смысле (1)) значением целевой функции $\Phi(X)$ за все время поиска; $X_{g,t}$ - вектор координат соседней с данной частицы с наилучшим за время поиска значением целевой функции $\Phi(X)$; α, β, γ - свободные параметры алгоритма.

Пересчет координат частиц по формулам (2), (3) может происходить по синхронной схеме (обновление координат частиц выполняется только после определения текущих скоростей всех N частиц) или по асинхронной схеме (расчет новых координат части производится до завершения указанных вычислений).

В процессе итераций вектор $X_{i,t}^b$ образует, так называемый, собственный путь (private guide) частицы P_i , а вектор $X_{g,t}$ - локальный путь (local guide) этой частицы.

Свободный параметр α определяет «инерционные» свойства частиц (при $\alpha < 1$ движение частиц, очевидно, замедляется). Рекомендуемое значение параметра α равно 0.7298. В процессе оптимизации может быть эффективным постепенное уменьшение коэффициента α от 0.9 до 0.4. При этом большие значения параметра обеспечивают широкий обзор пространства поиска, а малые – точную локализацию минимума целевой функции. Рекомендуемые значения свободных параметров β, γ равны 1.49618.

Важнейшим понятием в методе PSO является понятие соседства частиц, которое определяется соответствующей топологией соседства. В силу важности этого понятия, оно рассмотрено в самостоятельном разделе (см. раздел 5).

Второй компонент в формуле (3) называется «когнитивным» компонентом (по социальной аналогии) и формализует тенденцию частицы вернуться в положение с

минимальным значением целевой функции. Третий компонент в формуле (3) называется «социальным» компонентом. Компонент отражает влияние на данную частицу ее соседей.

Вместо формулы (3) часто используют ее вариант

$$V_{i,t+1} = \alpha(V_{i,t} + \varphi \oplus (Y_{i,t}^g - X_{i,t})). \quad (4)$$

Коэффициент α в этом случае называется ограничивающим фактором (restriction factor),

$$\varphi = U[0, \beta] + U[0, \gamma],$$

$$Y_{i,t}^g = \frac{U[0, \beta] \otimes X_{i,t}^b + U[0, \gamma] \otimes X_{i,t}^g}{\varphi}.$$

Рекомендуемые значения параметров β, γ в этом случае равны 2.05, а их оптимальные значения определяются классом целевой функции $\Phi(X)$ [3].

3. Метод роя частиц с полной информацией (FIPS)

Метод роя частиц с полной информацией FIPS (fully informed particle swarm) предложен в работах [4], [5] и представляет собой обобщение канонического метода PSO. В основе метода FIPS лежит наблюдение, что поведение особи определяется не только наиболее успешным соседом, но и всей совокупностью соседей. Таким образом, основное отличие метода FIPS от канонического метода PSO состоит в том, что в методе FIPS учитывается влияние на данную частицу всех соседних частиц. Ясно, что при этом различные топологии соседства частиц, существенно влияют на процесс оптимизации. В формуле (3) в данном случае исчезает необходимость в слагаемом, содержащем вектор $X_{g,t}$, т.е. «локальный путь» частицы в данном случае не существует.

В итерационной схеме метода FIPS координаты частиц определяются по формуле (2), а их скорости по формуле (4), в которой следующим образом изменяется только схема вычисления члена $Y_{i,t}^g$:

$$Y_{i,t}^g = \frac{\sum_k (\varphi_k \otimes X_{k,t}^b)}{\sum_k \varphi_k}, \quad k \in \hat{N}_i. \quad (5)$$

Здесь \hat{N}_i - множество номеров частиц, являющихся соседями частицы P_i (в смысле используемой топологии соседства);

$$\varphi_k = U\left[0, \beta / N_i\right];$$

N_i - количество элементов во множестве \hat{N}_i .

Метод обычно показывает лучшую сходимость, чем канонический метод PSO. Основной проблемой при использовании метода FIPS является сильное влияние на его

эффективность топологии пространства соседства. Особенностью метода является тенденция его сходимости к центру начального положения частиц X_i^0 , $i \in [1:N]$. В этой связи рекомендуется использовать асимметричное начальное расположение частиц.

3.1. Метод FIPS с весами

Метод FIPS с весами (weighted FIPS) является модификацией канонического метода и заключается в добавлении весов в последнее слагаемое уравнения (4). При соответствующем выборе весов, такая модификация позволяет усилить влияние тех соседних частиц, которые имеют лучшие значения целевой функции. Итерационная формула (4) для метода FIPS с весами имеет вид

$$V_{i,t+1} = \alpha \left(V_{i,t} + \frac{\sum_k w(k,t) U[0, \gamma] \otimes (X_{k,t}^b - X_{i,t})}{\sum_k w(k,t)} \right), \quad k \in \hat{N}_i, \quad (7)$$

где $w(k,t)$ - скалярный весовой множитель, в общем случае меняющийся со временем, который определяет влияние на данную частицу k -ой соседней частицы.

В качестве веса $w(k,t)$ чаще всего используют величину, обратную значению целевой функции $\Phi(X_{k,t})$ (напомним, что рассматривается задача минимизации этой функции):

$$w(k,t) = \frac{1}{\Phi(X_{k,t})}. \quad (8)$$

В роли весов $w(k,t)$ могут использоваться и другие величины, например, расстояние между частицами.

3.2. Ранжированный метод FIPS

Как отмечалось выше, имеет место нестабильность метода FIPS, обусловленная сильным влиянием на рассматриваемую частицу топологии соседства частиц. Кроме того, при использовании в качестве весов величин (8) может проявляться следующий эффект. На поздних итерациях, когда основные улучшения целевой функции произведены и все частицы находятся в окрестности локального минимума, имеют место близкие значения целевой функции, соответствующие всем частицам. Поэтому и значения весов $w(k,t)$ для всех $k \in \hat{N}_i$ оказываются близкими. В результате поиск замедляется или даже останавливается, так что применение в качестве весов значений (8) теряет смысл. Ранжированный метод FIPS (ranked FIPS) [6] позволяет преодолеть данный недостаток.

Идея ранжированного метода FIPS состоит в отказе от использования весов, построенных на основе значений целевой функции. Вместо этого веса $w(k,t)$, $k \in \hat{N}_i$ заменяются рангами r_k , которые определяются по следующему правилу:

$$r_k = 2r_{k+1}; \quad \sum_k r_k = 1; \quad \Phi(X_k) < \Phi(X_{k+1}). \quad (9)$$

Отметим, что в соответствии с этим правилом вес наилучшей частицы в рое равен 0.5.

Таким образом, в ранжированном методе FIPS формула (3) принимает вид

$$V_{i,t+1} = \alpha(V_{i,t} + U[0, \beta] \otimes (X_{i,t}^b - X_{i,t})) + \sum_k r_k U[0, \gamma] \otimes (X_{g,t} - X_{i,t}), \quad k \in \hat{N}_i. \quad (10)$$

Ранжированный метод FIPS обеспечивает также меньшее влияние на данную частицу топологии соседства частиц, так что увеличение количества частиц во множестве соседних частиц не оказывает отрицательного влияния на эффективность поиска.

4. Метод роя частиц, основанный на отношении «значение – расстояние» (FDR PSO)

Метод роя частиц, основанный на отношении «значение – расстояние» FDR PSO (fitness-distance ratio based PSO), предложен в работе [7]. Метод отличается от канонического PSO учетом влияния нескольких окружающих частиц. Заметим, что метод FIPS имеет такое же принципиальное отличие от канонического метода PSO. Однако в методе FDR PSO учет влияния соседних частиц реализуется иначе, чем в методе FIPS.

В методе FDR PSO используется модифицированный канонический итерационный процесс (2), (3). Суть модификации заключается в добавлении в формулу (3) еще одного слагаемого, равного

$$k_3 U[0, \delta] \otimes (X_{q,t} - X_{i,t}), \quad (11)$$

где δ - свободный параметр метода.

Таким образом, в методе FDR-PSO скорости частиц определяются по формуле

$$V_{i,t+1} = \alpha V_{i,t} + k_1 U[0, \beta] \otimes (X_{i,t}^b - X_{i,t}) + k_2 U[0, \gamma] \otimes (X_{g,t} - X_{i,t}) + k_3 U[0, \delta] \otimes (X_{q,t} - X_{i,t}), \quad (12)$$

где k_1, k_2, k_3 - свободные параметры, $X_{q,t}$ - вектор координат некоторой частицы $P_q \in \mathbf{P}$, определяемой по следующим правилам:

- частица P_q является соседней с частицей P_i ;
- частице P_q соответствует лучшее значение целевой функции $\Phi(X)$ по сравнению со всеми соседними частицами.

Введем в рассмотрение функцию $FDR(X_j, X_i)$ положений частиц P_j, P_i из популяции \mathbf{P} :

$$FDR(X_j, X_i) = \frac{\Phi(X_j) - \Phi(X_i)}{\|X_j - X_i\|}, \quad (13)$$

где $\|\cdot\|$ - некоторая векторная норма.

Тогда частица, наилучшим образом удовлетворяющая сформулированным правилам, доставляет минимум функции FDR :

$$FDR(X_q, X_i) = \max_j FDR(X_j, X_i), j \in \hat{N}_i.$$

Рекомендованные в работе [7] значения параметров k_1, k_2, k_3 равны 1, 1, 2, соответственно. Экспериментально показано, что метод малочувствителен к значениям этих параметров - даже их обнуление не приводит к существенному ухудшению скорости сходимости метода. Основной вклад в скорость сходимости метода обеспечивает компонент (11).

Отметим, что в том виде, в котором метод FDR PSO рассматривается в работе [7], случайные коэффициенты $U[a, b]$ не используются, что эквивалентно принятию $U[a, b] = const = 1$. Однако в той же работе указывается, что в качестве этих коэффициентов имеет смысл использовать случайные величины, равномерно распределенные в одинаковом диапазоне $[0, 2.05]$.

5. Топологии соседства частиц

Эффективность метода PSO в значительной мере зависит от топологии соседства частиц (population topology, neighbourhood topology, swarm topology, sociometry). Топология соседства определяется неориентированным графом, вершины которого соответствуют частицам роя, а ребра связывают непосредственных соседей.

В вычислительной практике чаще всего используются следующие топологии соседства частиц [8]:

- клика (gbest-топология - глобально оптимальная топология);
- кольцо (lbest-топология - локально оптимальная топология);
- двумерный тор (топология фон Неймана) - двумерный тор;
- кластерная топология.

В топологии «клика» (полносвязном графе) соседями каждой из частиц P_i , $i \in [1, N]$ являются остальные $(N - 1)$ частицы. Диаметр полносвязного графа равен единице [9].

В топологии «кольцо» соседями каждой из частиц P_i являются 2 частицы. Например, в графе, приведенном на Рис. 1, соседями частицы P_3 являются частицы P_2 , P_4 . Диаметр соответствующего графа равен $\sim N/2$.

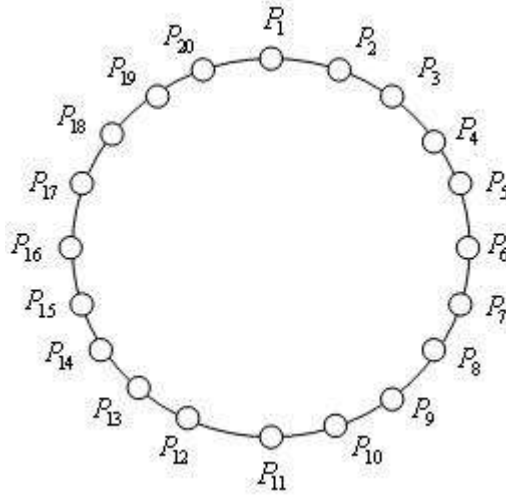


Рис. 1. Топологии соседства частиц «кольцо»: $N = 20$

В топологии «двумерный тор» соседями каждой из частиц P_i , $i \in [1, N]$ являются 4 частицы. Например, в графе, приведенном на Рис. 2, соседями частицы P_3 являются частицы P_2, P_4, P_8, P_{18} . Если граф типа «двумерный тор» построен на основе $m \times m$ решетки, то диаметр этого графа равен $2\lceil m/2 \rceil$, где $\lceil * \rceil$ - символ ближайшего целого меньшего; $N = m^2$.

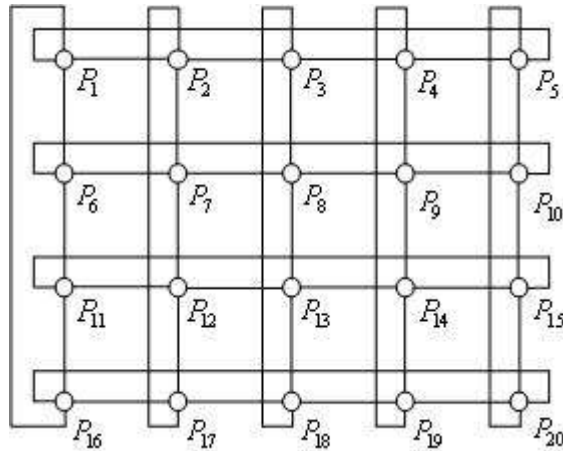


Рис. 2. Топологии соседства частиц «двумерный тор»: $N = 20$

В кластерной топологии граф имеет в качестве узлов клики из C_{clique} узлов. Клик числом N_{clique} объединяются в полносвязный граф, ребра которого связаны с теми узлами в кликах, степени которых равны $(C_{clique} - 1)$. Очевидно, что диаметр такого графа равен 3. Соседями каждой из частиц P_i , $i \in [1, N]$, являются C_{clique} частиц. К примеру, в графе, приведенном на Рис. 3, $N_{clique} = 4$, $C_{clique} = 5$, диаметр равен 3, соседями частицы P_3 являются 5 следующих частиц: $P_1, P_2, P_4, P_5, P_{15}$.

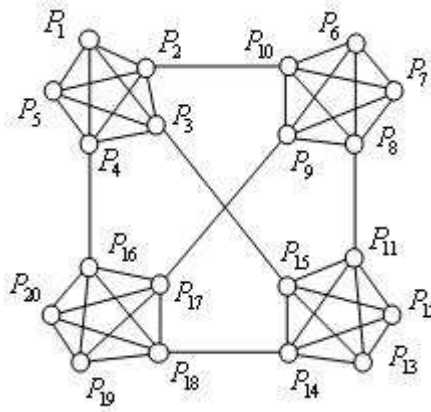


Рис. 2. Кластерная топологии соседства частиц: $N = 20$

Диаметр графа, соответствующего используемой роем топологии соседства частиц, определяет скорость распространения информации в рое. Поэтому в рое с топологией соседства «клика» лучшее значение целевой функции, достигнутое той или иной частицей, сразу становится известным всем остальным частицам роя. Из числа рассмотренных топологий самую низкую скорость распространения информации обеспечивает топология «кольцо». Промежуточные значения скорости распространения информации достигаются в роях, обладающих топологией соседства частиц «двумерный тор» и кластерной топологией.

Известно значительное количество исследований эффективности и сходимости метода PSO и его модификаций при различных топологиях соседства частиц [4], [5], [8], [10].

Результаты исследований показывают, что топология типа «клика» применима для поиска минимума функции $\Phi(X)$ с малыми градиентами. Топология обеспечивает высокую эффективность, но может привести к преждевременной сходимости метода к некоторому локальному минимуму функции $\Phi(X)$.

Напротив, топология типа «кольцо» эффективна при оптимизации сложных многоэкстремальных целевых функций. Это обстоятельство обусловлено тем, что частицы роя, принадлежащие «далеким» частям кольца, слабо связаны друг с другом и могут эффективно исследовать различные области пространства поиска. Топология обеспечивает слабое «притяжение» частиц локальными минимумами целевой функции и позволяет избежать преждевременной сходимости.

Топология «двумерный тор» и кластерная топология обладают промежуточными свойствами. Они обеспечивают высокую вероятность локализации глобального минимума целевой функции и достаточно высокую эффективность поиска.

Эффективность метода PSO в значительной степени зависит от размера популяции N . Эта зависимость может иметь различный характер для различных топологий соседства частиц.

Для топологии «клика» и большого размера популяции канонический метод PSO может привести к ранней сходимости. В тех же условиях в методе FIPS большая популяция может охватывать несколько локальных минимумов целевой функции, и значительное количество частиц могут колебаться между этими минимумами. Такие же колебания частиц возможны и в каноническом методе PSO.

Топология «кольцо» более устойчива к большому размеру популяции, так как среди рассмотренных топологий является наименее связной и обеспечивает самую низкую скорость передачи информации.

С точки зрения эффективности метода PSO при больших популяциях в качестве универсальной топологии рекомендуется использовать топологию «двумерный тор», которая обеспечивает достаточно высокую эффективность любого из рассмотренных методов.

6. Методы, использующие динамические топологии соседства частиц

Рассмотренные методы PSO используют статические топологии соседства частиц, т.е. соседство частиц для заданного роя задается до начала итераций и во время итерационного процесса не меняется. Наряду с этим известно значительное количество методов PSO, использующих топологии соседства частиц, которые меняются в процессе итераций - динамические топологии соседства.

6.1. Метод роя частиц с организацией в клубы

Метод роя частиц с организацией в клубы CB PSO (clubs-based PSO) предложен в работе [11]. Суть метода состоит в следующем. Частицы роя образуют динамические группы, называемые клубами. Каждая частица может входить в несколько клубов. В процессе итераций частица может мигрировать между клубами. Соседними частицами считаются частицы из одного клуба. Скорости частиц определяются по формулам (2), (3) канонического метода PSO.

Количество клубов M_i , в которые входит частица P_i , называется уровнем вхождения частицы в клубы (membership level). В начальный момент времени каждая из частиц случайным образом распределяется по N_{cl} клубам, где N_{cl} - так называемый стандартный уровень вхождения в клубы. При росте величины N_{cl} скорость сходимости метода CB PSO возрастает. Обычно ограничиваются выбором уровня вхождения в пределах 4 - 15.

В методе CB PSO на каждой итерации оптимизационного процесса находят частицы P_{best} и P_{worse} с наилучшим и наихудшим значениями целевой функции, соответственно. Частица P_{best} исключается из произвольного клуба, к которому она принадлежит, а частица P_{worse} - включается в произвольный клуб. Каждые rr итераций для частиц, не являющихся наилучшими или наихудшими в клубе, производится изменение их уровней вхождения в клубы в направлении стандартного уровня вхождения. Такой прием улучшает стабильность метода и равномерность распределение частиц по клубам.

На параметр M_i накладывается ограничение $M_i \in [M^{\min}, M^{\max}]$, так что каждая из частиц P_i , $i \in [1 : N]$ во всех случаях входит не менее чем в M^{\min} и не более чем в M^{\max} клубов. Величина M^{\max} не должна быть меньше N , чтобы частица могла состоять в разных клубах. В работе [11] для роя из 20 частиц используются значения $M^{\min} = 5$, $M^{\max} = 33$.

Исключение частицы P_{best} из клуба имеет целью предотвратить преждевременную сходимость метода, так как эта частица оказывает сильное влияние на соседние частицы в клубе. Включение частицы P_{worse} в некоторый другой клуб позволяет увеличить количество соседних с нею частиц. Так как частицы в новых клубах в общем случае не хуже частиц, находящихся в исходном клубе, то это дает возможность частице P_{worse} улучшить значение целевой функции.

Как отмечалось выше, в методах PSO, основанных на статической топологии соседства частиц, ранняя сходимость является типичной проблемой. Причина этого эффекта заключается в том, что частица, нашедшая какой-либо локальный минимум целевой функции, оказывает сильное и постоянное влияние на соседние частицы, стремящееся локализовать эти частицы в окрестности того же локального минимума. В результате уменьшается вероятность того, что эти частицы могут исследовать другие возможные локальные минимумы. В методе CB PSO эта проблема решается путем исключения из числа соседей наиболее успешных частиц. При этом на ранних стадиях поиска частицы имеют довольно мало соседей, и нахождение отдельных локальных минимумов слабо влияет на остальные частицы.

В работе [11] показано, что метод CB PSO по сравнению с каноническим методом PSO дает лучшие результаты с точки зрения сходимости и выхода из локальных минимумов целевой функции, как для топологии соседства «клика», так и для топологии «кольцо».

6.2. Метод роя частиц со стереотипированием

Метод роя частиц со стереотипированием (PSO with stereotyping) предложен в работе [12]. Основная идея метода состоит в статическом разделении частиц по фиксированному количеству кластеров. Критерием разделения является геометрическая близость частиц в пространстве параметров R^n . В качестве центра кластера используется геометрический центр координат частиц, образующих этот кластер. В формулах вычисления скорости частицы P_i , $i \in [1:N]$ ее собственные координаты заменяются координатами центра кластера, которому принадлежит эта частица. Например, в формуле вычисления скоростей частиц для канонического метода PSO (3) указанным образом заменяется компонент $X_{i,t}$.

Метод может быть скомбинирован, как с каноническим методом PSO, так и с его модификацией FIPS.

6.3. Метод роя частиц с дополнением графа соседства частиц

Метод PSO с дополнением графа соседства частиц DS PSO (dynamic sociometry PSO) предложен и исследован в работах [13], [14], [15]. Метод реализует технику так называемого «перехода от обзора к поиску» (switch from exploration to exploitation).

Как отмечалось выше, разреженные графы соседства частиц обеспечивают высокую эффективность при обзоре, а плотные графы – при поиске. Идея метода DS PSO состоит в следующем. На начальных итерациях оптимизации используется сильно разреженный граф соседства частиц. Через фиксированное количество итераций k в этот граф регулярным или случайным образом добавляется новое ребро. Например, в топологии соседства частиц «кольцо» после каждых k итераций в текущий граф соседства частиц добавляется ребро (P_j, P_{j+i}) . В результате, в конце концов, при достаточном количестве итераций получается граф соседства типа «клика».

Допустим, что в рое N частиц, предполагается решение задачи за I итераций и получить полносвязную топологию требуется после выполнения $0.8I$ итераций. Тогда добавлять ребро в текущий граф соседства частиц следует $(N - 2)$ раз, т. е. каждые

$$k = \frac{0.8I}{N - 2} \quad (15)$$

итераций.

На ранних стадиях поиска метод DS PSO реализует стратегию глобального исследования пространства параметров (обзор), а на поздних стадиях – производит уточнение полученных результатов (поиск).

7. Многороевые методы

Все рассмотренные методы PSO, как со статической, так и с динамической топологией соседства частиц, использовали один рой. Известны также методы, использующие несколько роев и миграцию частиц между ними.

7.1. Многороевый метод с миграцией частиц

Многороевый метод с миграцией частиц MPSO (PSO with migration) предложен в работе [6]. Идея метода состоит в формировании нескольких роев частиц (subswarms) и организации миграции частиц между роями. При этом каждый из роев использует, вообще говоря, свои значения свободных параметров и свою топологию соседства частиц. При переходе частицы из одного роя в другой, топология соседства взаимодействующих роев не меняется. Метод является адаптивным и может применяться при неполных знаниях о минимизируемой функции $\Phi(X)$. Напомним, что для метода FIPS, например, топология соседства частиц должна выбираться в зависимости от класса функции соответственно эмпирически полученным рекомендациям.

В простейшем случае миграция частиц между роями осуществляется через фиксированное количество итераций k . Для каждого из роев S_i на основе текущих значений целевой функции $\Phi(X)$ вычисляется значение функции оптимальности этого роя $\Psi(S_i)$, равное значению функции $\Phi(X)$ для лучшего собственного пути в рое S_i . Таким образом, функция оптимальности роя отражает наилучшие достижения частиц в рое.

Каждые k итераций осуществляется переход одной частицы из роя с минимальным значением функции Ψ в рой с максимальным значением этой функции. Здесь k - свободный параметр метода. В работе [6] использовалось значение $k = 60$.

7.2. Модифицированный многороевый метод с миграцией частиц

Схема модифицированного метода MPSO имеет следующий вид [6].

1). Для каждой пары роев S_i, S_j через каждые k итераций вычисляется отношение значений их модифицированных функций оптимальности $\Psi^*(S_i), \Psi^*(S_j)$

$$R(S_i, S_j) = R_{i,j} = \frac{\Psi^*(S_i)}{\Psi^*(S_j)}. \quad (16)$$

2). Находится множество роев $\{S_k\}$, для которых справедливо неравенство $R_{i,j} > R_{\lim}$.

3). Каждому из роев $\{S_k\}$ назначается штрафной балл B_k (penalty).

4). Из числа роев $\{S_k\}$ определяется новое множество роев $\{S^*\}$ таких, что для каждого из них имеет место неравенство $B_k > B_{\lim}$.

5). Рои множества $\{S^*\}$ сортируются по возрастанию их штрафных баллов.

6). Из каждого из роев отсортированного множества $\{S^*\}$ (кроме последнего роя) мигрирует произвольно выбранная частица в следующий по порядку рой.

Назначение функции $\Psi^*(S_i)$ состоит в классификации множества роев на «успешные» и «неуспешные». В качестве этой функции может использоваться относительное улучшение функции оптимальности за некоторое количество итераций d :

$$\Psi^*(S_i) = \frac{\Psi_t(S_i)}{\Psi_{t-d}(S_i)}. \quad (17)$$

В модифицированном методе MPSO параметры R_{\lim} , B_{\lim} , d являются свободными. Для тестирования метода авторы применяли $R_{\lim} = 2$, $B_{\lim} = 60$, $d = 200$. Возможна оптимизация этих параметров для конкретных классов целевых функций.

8. Параллельные методы роя частиц

Идеи, лежащие в основе метода PSO, дают основания надеяться на существование эффективных параллельных аналогов метода. Действительно, известно значительное количество хорошо масштабируемых параллельных методов PSO. Сразу заметим, что в своей массе параллельные методы алгоритмически отличаются от последовательных аналогов.

Для генетических методов, гораздо более исследованных по сравнению с методами PSO, выделяется три следующих класса параллельных методов [16]:

- методы, основанные на глобальной модели параллелизма;
- миграционные методы, основанные на островной модели параллелизма;
- методы, основанные на диффузной модели параллелизма.

По аналогии с этой классификацией, для PSO выделяется три рассматриваемых ниже группы параллельных методов [17].

8.1. Методы, основанные на глобальной модели параллелизма

Параллельные методы PSO, относящиеся к данной группе методов, представляют собой параллельные аналоги рассмотренных выше последовательных методов. Эти методы используют параллелизм по данным и ориентированы на организацию параллельных вычислений по типу «master – slave». Мастер-процесс выполняет работу по поиску глобального минимума целевой функции $\Phi(X)$, а подчиненные процессы производят вычисления значений целевой функции, а также скоростей и координат частиц.

Рабочие процессы после каждой итерации отсылают мастер-процессу значения полученных координат частиц и их скоростей, а также соответствующие значения целевой функции. Перед каждой итерацией рабочие процессы получают от мастер-процесса координаты «лучших» частиц. Мастер-процесс на основе данных, полученных от рабочих процессов, вычисляет координаты «лучших» частиц и посылает их рабочим процессам.

Параллельные вычисления в методах данного класса могут быть, как синхронными, так и асинхронными.

Достоинство методов, основанных на глобальной модели параллелизма, состоит в использовании информации о глобально «лучших» частицах. Результат оптимизации при использовании данных методов получается точно таким же, как при аналогичных последовательных вычислениях. Недостатком методов являются большие накладные расходы на коммуникации (при большом диаметре коммуникационной сети).

8.2. Методы, основанные на островной модели параллелизма

Методы данного класса используют параллелизм по данным и ориентированы на реализацию на кластерных вычислительных системах. Весь рой из N частиц делится на m островов (по количеству процессоров в системе) и частицы, принадлежащие каждому из островов, обрабатываются на своем процессоре. После каждой k независимых итераций, острова обмениваются между собой лучшими частицами (в соответствие с используемой топологией соседства частиц).

При большом k метод имеет малые коммуникационные расходы, но недостаточно эффективно использует информацию о лучших частицах из других островов. Существуют различные стратегии замещения частиц в островах [18].

8.3. Методы, основанные на диффузной модели параллелизма

Методы этого класса можно считать частным случаем методов, основанных на островной модели параллелизма (когда количество частиц N равно количеству островов m и, тем самым, количеству используемых процессоров вычислительной системы).

Основная идея диффузных методов состоит в том, что «лучшие» частицы определяются параллельно и только среди соседних частиц (в смысле используемой топологии соседства частиц). Для этого на каждой итерации процессор, на котором обрабатывается частица P_i , получает от всех тех процессоров, на которых обрабатываются частицы, соседние с данной, их координаты, скорости и соответствующие значения целевой функции $\Phi(X)$. Такие коммуникации выполняются параллельно всеми N процессорами.

Коммуникационные расходы в диффузной модели напрямую определяются используемой топологией соседства частиц и являются высокими в случае, если соответствующий граф имеет высокую связность. Методы данного класса обеспечивают высокую производительность только при невысокой связности указанного графа и/или высокой вычислительной сложности целевой функции.

В целом, метод ориентирован на массивно-параллельные вычислительные системы с сильно связными (например, гиперкуб или двумерный тор) и быстрыми коммуникационными сетями.

8.4. Параллельный синхронный метод

Параллельный синхронный метод предложен в работе [19] (в этой работе метод называется parallel PSO method). Метод реализует глобальную модель параллелизма, при которой параллельно вычисляются лишь значения целевой функции.

На каждой итерации мастер-процессор рассылает по рабочим процессорам текущие координаты частиц и ожидает возвращения соответствующих значений целевой функции $\Phi(X)$. Текущая итерация завершается только после того, как мастер-процессором будут получены значения целевой функции от всех рабочих процессоров (т.е. используется глобальная барьерная синхронизация). Обновление состояния роя (скоростей и координат частиц) происходит только после получения мастер-процессором указанной информации обо всех частицах роя. В методе используется статическая балансировка загрузки процессоров.

Условиями высокой производительности параллельного синхронного метода являются гомогенность вычислительной системы, одинаковое время вычисления значений целевой функции в любой точке пространства параметров, а также равномерное распределение вычислений по процессорам системы. Обеспечить выполнение всех этих условий обычно не удается. Так, часто приходится использовать гетерогенные вычислительные системы, в реальных оптимизационных задачах вычислительная

сложность целевой функции $\Phi(X)$ зависит от значений компонентов вектора варьируемых параметров X .

8.5. Параллельный асинхронный метод

Параллельный асинхронный метод PAPSO (parallel asynchronous PSO) предложен в работах [20], [21]. Метод также реализует глобальную модель параллелизма.

В отличие от синхронного метода, в этом случае мастер-процессор получает данные от рабочих процессоров не после глобальной синхронизации, а в любой момент времени по мере готовности этих данных. На основе полученной информации мастер-процессор обновляет скорости и координаты соответствующих частиц роя и немедленно возвращает их свободным рабочим процессорам для продолжения итераций.

Метод реализует динамическую балансировку загрузки процессоров. Номера необработанных частиц хранятся в очереди FIFO. Мастер-процесс обновляет координаты и скорость первой частицы из этой очереди и посылает первому свободному процессору.

8.6. Сравнение эффективности синхронного и асинхронного методов

Применение асинхронного метода оправдано, когда вычисление целевой функции $\Phi(X)$ при разных значениях вектора X происходит за различные промежутки времени, т.е. когда вычислительная сложность этой функции переменна.

В работе [20] исследовано влияние дисперсии вычислительной сложности целевой функции и гетерогенности вычислительной системы на ускорение и эффективность синхронного и асинхронного методов. Вычислительная сложность функции $\Phi(X)$ полагалась случайной величиной с заданным законом распределения и моделировалась с помощью соответствующих временных задержек.

Исследование показало, что при увеличении дисперсии вычислительной сложности функции $\Phi(X)$ от нуля (что эквивалентно постоянной вычислительной сложности) эффективность синхронного метода линейно падает. Асинхронный метод на малом количестве процессоров (менее 15) показывает в этих условиях низкую эффективность, но при дальнейшем увеличении количества процессоров эффективность метода асимптотически возрастает до 93 процентов. Такая зависимость эффективности от количества процессоров не является типичной для параллельных вычислений и объясняется малой долей коммуникаций в процессе вычислений, а также распределением их по времени (в отличие от единовременных коммуникаций в синхронном методе). Вариации вычислительной сложности целевой функции практически не оказывают влияния на эффективность асинхронного метода. Указанные достоинства асинхронного метода обуславливают его высокую эффективность при использовании сильно гетерогенных кластерных вычислительных систем.

Таким образом, метод PAPSO неэффективен при малом количестве процессоров, но обеспечивает высокую эффективность при большом количестве процессоров (вплоть до N). Для обеспечения эффективной работы асинхронного метода при бо́льшем N количестве процессоров, необходимо распараллеливать вычисление значений целевой функции $\Phi(X)$.

Важно отметить, что при параллельной реализации одного и того же метода PSO в синхронном и асинхронном вариантах, при прочих равных условиях, результаты оптимизации могут быть различными.

9. Тестовые функции

Для тестирования метода PSO чаще всего применяют перечисленные ниже функции.

1). Функция Экли (Ackley)

$$\Phi(X) = \sum_{i=1}^n \left(e^{-0.2 \sqrt{x_i^2 + x_{i+1}^2}} + 3(\cos 2x_i + \sin 2x_{i+1}) \right) - \text{(см. Рис. 4)}.$$

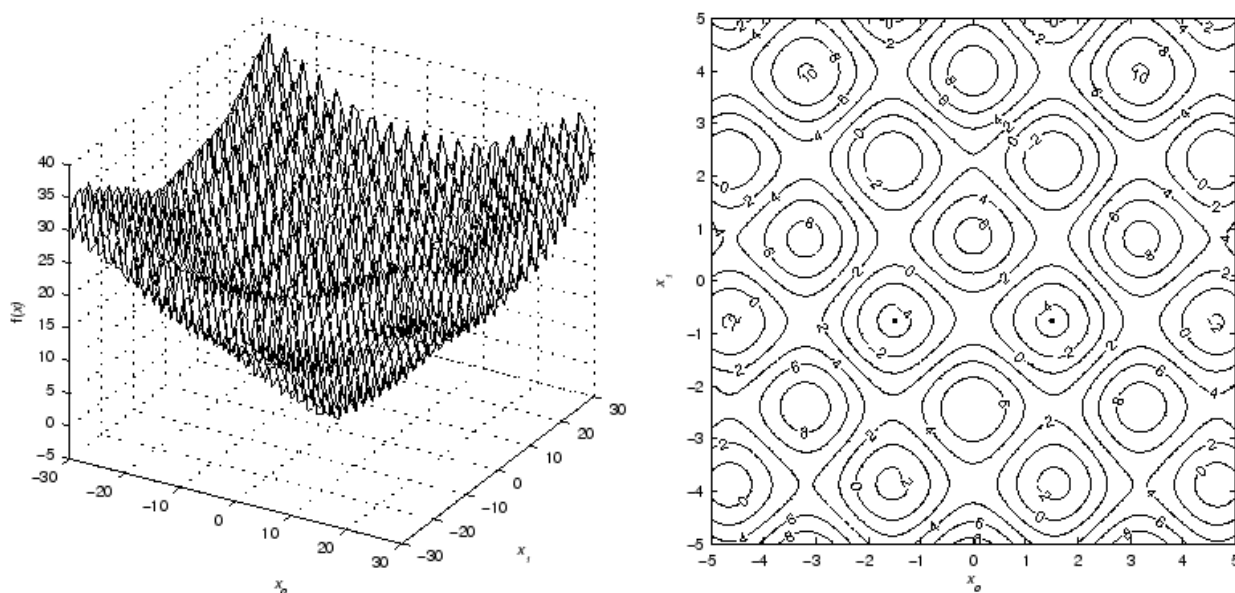


Рис. 4. Поверхность и линии уровня функции Экли: $n = 2$

2). Функция Химмельблау (Himmelblau)

$$\Phi(X) = (x_1^2 + x_2 - 11)^2 + (x_1 + x_2^2 - 7)^2 - \text{см. Рис. 5}.$$

3). Функция Растригина (Rastrigin)

$$\Phi(X) = \sum_{i=1}^n (x_i^2 - 10 \cos 2\pi x_i + 10) - \text{см. Рис. 6}.$$

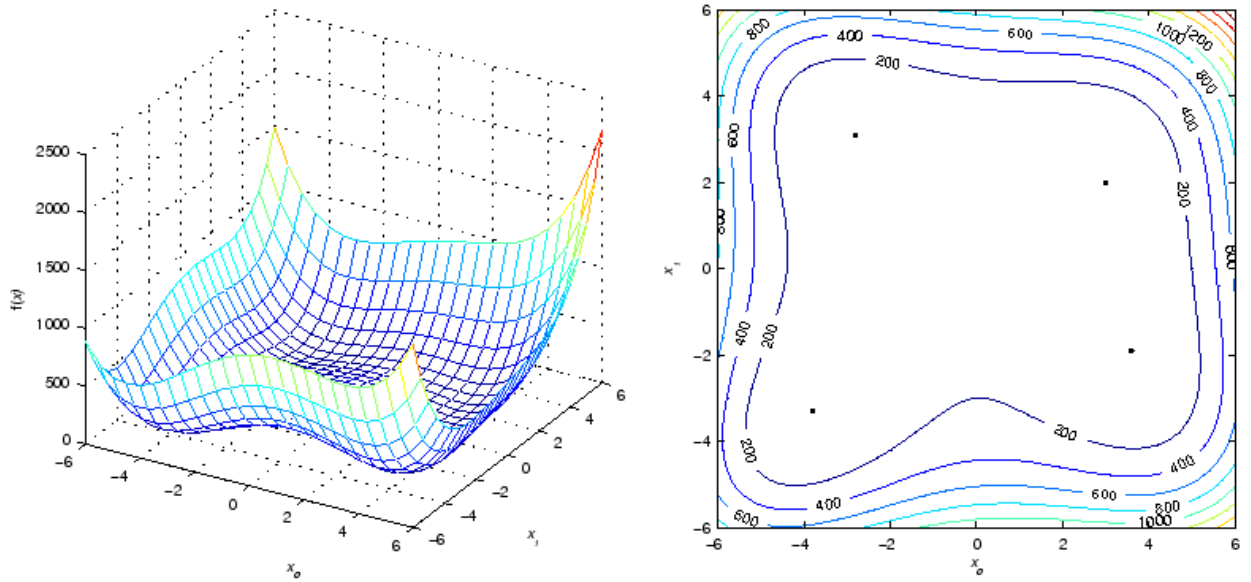


Рис. 5. Поверхность и линии уровня функции Химмельблау

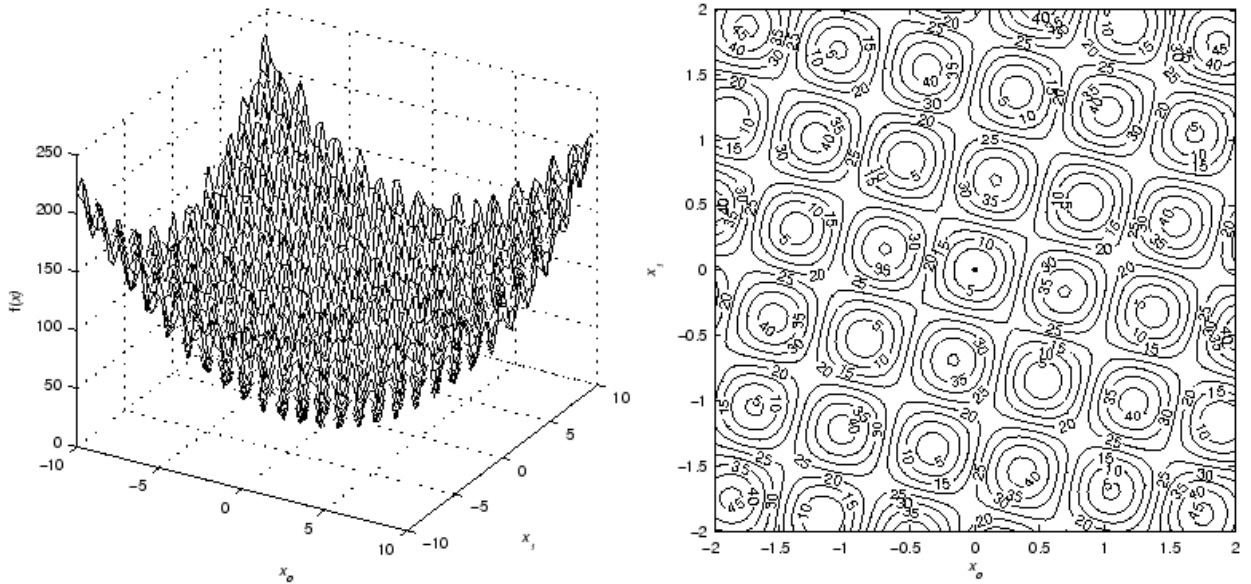


Рис. 6. Поверхность и линии уровня функции Растригина: $n = 2$

4). Функция Розенброка (Rosenbrock)

$$\Phi(X) = \sum_{i=1}^{n-1} \left(100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2 \right) - \text{см. Рис. 7.}$$

5). Функция Шекеля (Shekel)

$$\Phi(X) = \sum_{i=1}^n \frac{1}{c_i + \sum_{j=1}^n (x_j - a_{j,i})^2} - \text{см. Рис. 8.}$$

Для тестирования метода PSO широко применяются также пакеты тестовых функций CES [22].

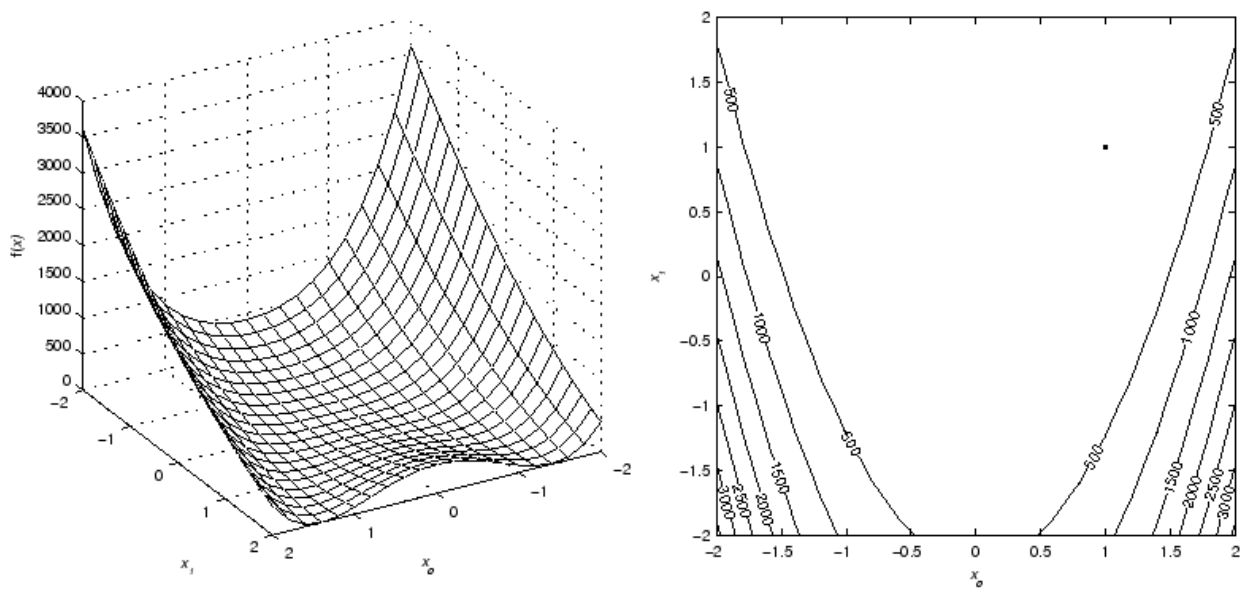


Рис. 7. Поверхность и линии уровня функции Розенброка: $n = 2$

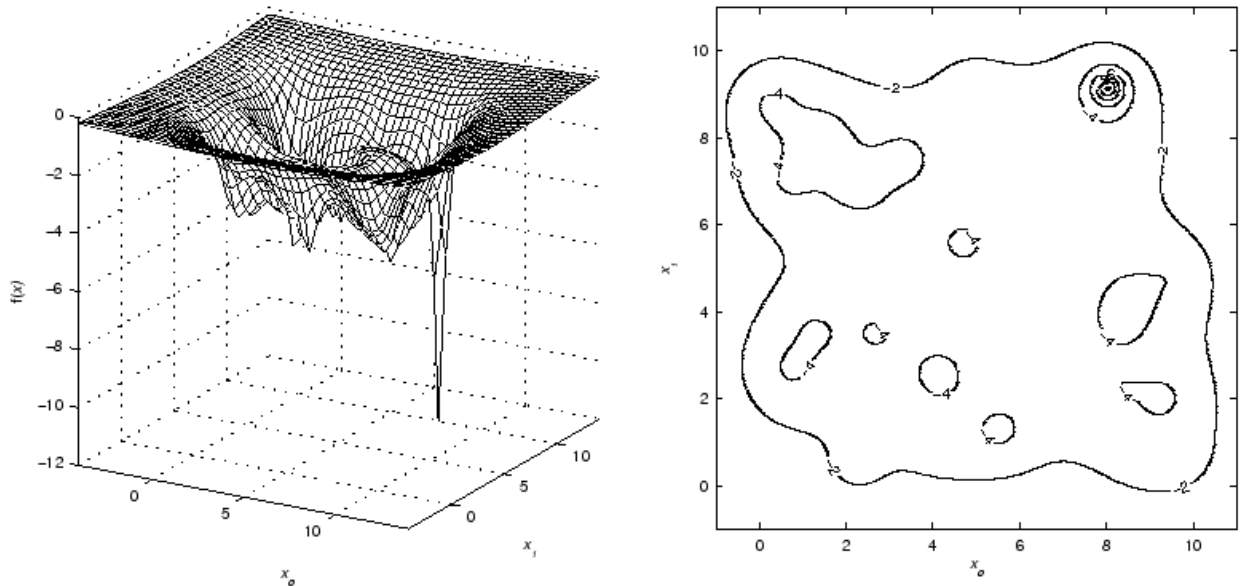


Рис. 8. Поверхность и линии уровня функции Шекеля: $n = 2$

Литература

1. T. Weise. Global Optimization Algorithms - Theory and Application: Ph.D Thesis. - University of Kassel, 2008.
2. J Kennedy, R Eberhart. Particle swarm optimization. // Proceedings of IEEE International conference on Neural Networks. – 1995, pp. 1942 - 1948.
3. X. Li, A. P. Engelbrecht. Particle swarm optimization: an introduction and its recent developments. // Proceedings of the 2007 GECCO conference companion on Genetic and evolutionary computation. – London, ACM, 2007, pp. 3391 - 3414.

4. R. Mendes, J. Kennedy, J. Neves. The fully informed particle swarm: Simpler, maybe better. // IEEE Transactions on Evolutionary Computation. - 2004, v. 8, pp. 204 - 210.
5. R. Mendes, J. Kennedy, J. Neves. Watch thy neighbor or how the swarm can learn from its environment. // Proceedings of the Swarm Intelligence Symposium 2003. – IEEE, 2003, pp. 88 - 94.
6. J. Jordan, S. Helwig, R. Wanka. Social interaction in particle swarm optimization, the ranked FIPS, and adaptive multi-swarms. // Proceedings of the 10th annual conference on Genetic and evolutionary computation. - Atlanta, USA, ACM, 2008, pp. 49 - 56.
7. K. Veeramachaneni, Th. Peram, Ch. Mohan, L. A. Osadciw. Optimization Using Particle Swarm with Near Neighbor Interactions. // Lecture Notes Computer Science. - Springer Verlag, 2003.
8. J. Kennedy, R. Mendes. Population structure and particle swarm performance. // Proceedings of the 2002 Evolutionary Computation Congress. - Washington, IEEE Computer Society, pp. 1671 – 1676.
9. В. В. Воеводин, Вл. В. Воеводин. Параллельные вычисления. - БХВ-Петербург, 2002.
10. J. Kennedy, R. Mendes. Neighborhood topologies in fully informed and best-of-neighborhood particle swarms. // Systems, Man, and Cybernetics. - 2006, v. 36, pp. 515—519.
11. W. Elshamy, H. M. Emara, A. Bahgat. Clubs-based Particle Swarm Optimization. // Swarm Intelligence Symposium. - 2007, pp. 289 – 296.
12. J. Kennedy. Stereotyping: improving particle swarm performance with cluster analysis. // Proceedings of the 2000 Congress on Evolutionary Computation. - 2002, v.2, pp. 1507—1512.
13. P. N. Suganthan. Particle swarm optimizer with neighborhood operator. // Proceedings of the 1999 Congress on Evolutionary Computation. - 1999, v.3, pp. 1958-1962.
14. M. Richards, D. Ventura. Dynamic Sociometry in Particle Swarm Optimization. // Proceedings of the Joint Conference on Information Sciences. - 2003, pp. 1557—1560.
15. A. S. Mohais, R. Mendes, Ch. Ward, Ch. Posthoff. Neighborhood Re-structuring in Particle Swarm Optimization. // Proceedings of the Australian Conference on Artificial Intelligence. - 2005, pp. 776-785.
16. M. Dorigo, V. Maniezzo. Parallel Genetic Algorithms: Introduction and Overview of Current Research. // Parallel Genetic Algorithms: Theory and Applications. / Ed. by J. Stenders. - Amsterdam, IOS Press, 1992.

17. M. Belal, T. El-Ghazawi. Parallel Models for Particle Swarm Optimizers. // International Journal of Intelligent Computing and Information Sciences. - 2004, v. 1, pp. 100-111.
18. Shu-Chuan Chu, Jeng-Shyang Pan. Intelligent Parallel Particle Swarm Optimization Algorithms. // Parallel Evolutionary Computations. - Springer Berlin, Heidelberg, 2006, v. 22, pp. 159-175.
19. J. F. Schutte, J. A. Reinbolt, B. J. Fregly, R. T. Haftka, A. D. George. Parallel global optimization with particle swarm algorithm. // Journal of Numerical Methods in Engineering. – 2004, v. 61, pp. 2296—2315.
20. B. Koh, A. D. George, R. T. Haftka, B. J. Fregly. Parallel asynchronous particle swarm optimization. // International Journal for Numerical Methods in Engineering. - 2006, v. 67(4), pp. 578—595.
21. G. Venter, J. Sobieszczanski-sobieski. A parallel particle swarm optimization algorithm accelerated by asynchronous evaluations. // Journal of Aerospace Computing, Information, and Communication. - 2006, v. 3, pp. 123 – 137.
22. K. Tang, X. Yao, P. N. Suganthan, C. MacNish, Y. P. Chen, C. M. Chen, Z. Yang. Benchmark Functions for the CEC'2008 Special Session and Competition on Large Scale Global Optimization. - Nature Inspired Computation and Applications Laboratory, USTC, China, 2007.