

Использование генетических алгоритмов для нахождения оптимума функции

Проект по предмету: методы оптимизации

Горская Елена, 695 группа

20 декабря 2018

Постановка задачи

Рассмотрим следующую постановку задачи: у нас имеется функция (в данной работе рассматриваются функции двух переменных), для которой необходимо найти минимум. Одним из способов нахождения оптимума являются генетические алгоритмы. В данной работе рассмотрена реализация некоторых вариаций генетических алгоритмов и продемонстрирована их работа на нескольких примерах. Также произведен анализ поведения и эффективности генетических алгоритмов, рассмотрены достоинства и недостатки каждой из реализаций и класса генетических алгоритмов в целом.

Основные понятия

Для начала попытаемся понять, что из себя представляют генетические алгоритмы. Они являются одной из разновидностей алгоритмов поиска (вообще говоря, это эвристический алгоритм поиска) и используются в том числе и для решения задач оптимизации. При решении данных задач используются методы, схожие с теми, которые встречаются в естественном отборе:

- отбор: мы выбираем из популяции долю тех, кто "пойдёт дальше".
- наследование: передача некоторых свойств, признаков, информации следующему поколению.
- скрещивание: создание нового поколения.
- мутация: из всей популяции выбирается некоторое количество особей (особь = пробное решение), характеристики которых изменяются в соответствии с определенными операциями - происходит их мутация.

Идея данных алгоритмов такая же, как и в естественном отборе в ходе эволюции: у нас есть какое-то количество особей, находящихся во враждебной среде, которые могут скрещиваться и передавать потомкам часть своих генов, а уже наиболее приспособленные потомки будут выживать. В задаче оптимизации у нас задана функция приспособленности (fitness function). Это вещественная либо целочисленная функция одной или нескольких переменных, которую нам и нужно оптимизировать в ходе генетического алгоритма. Она направляет эволюцию в сторону правильного (оптимального) решения. Функция приспособленности является одним из частных случаев целевой функции.

Итак, попробуем формализовать сказанное выше и описать, из чего состоит простейший генетический алгоритм. Пусть у нас есть функция приспособленности f , заданная на множестве P , которую нужно минимизировать. Алгоритм будет устроен так:

1. Мы выбираем функцию скрещивания C и функцию мутации M , где функция C принимает на вход два аргумента (в более общем случае, функция скрещивания может принимать 3 и более аргументов), а выдает один, в то время как функция M принимает на вход один аргумент и выдает также один аргумент:

$$C : P \times P \rightarrow P, M : P \rightarrow P$$

2. Выбираем начальную популяцию $A \subseteq P$, такую что $|A| = m$

Обычно начальная популяция генерируется случайно. Важно, чтобы у нас было относительное разнообразие особей, чтобы популяция не пришла просто в ближайший экстремум.

3. Строим множество потомков: $D = \{C(x, y) \mid x, y \in A\}$

4. Строим множество мутантов: $X = \{M(x) \mid x \in A\}$

5. Рассматриваем множество $A \cup D \cup X$ и сортируем его по возрастанию значений функции f , а затем отбираем первые (лучшие) m элементов, которые составят новую популяцию A . После этого вновь переходим к третьему шагу.

Какие мы можем рассматривать критерии остановки:

- Значение функции f достаточно мало: $\min f(x) < \epsilon$, где $x \in A_{current}$. В данном случае мы рассматриваем текущее множество A (текущую популяцию).
- Полученная на очередном шаге популяция оказалась не намного лучше той, которая была на предыдущем шаге:
$$\min_{x \in A_{current}} f(x) \geq 0.99 \min_{x \in A_{previous}} f(x).$$
- Мы сделали уже довольно большое количество шагов (исчерпали число поколений, отпущенных на эволюцию).

Какие функции скрещивания мы можем использовать:

- Так как довольно часто мы говорим про R^n , то можно брать $C(x, y) = \frac{x+y}{2}$
- Если мы говорим, например, про R^2 , то можно брать $C((x_1, y_1), (x_2, y_2)) = (x_1, y_2)$
- Очень часто зависит от задачи: "хорошие гены" мы определяем в зависимости от условия

Какие функции мутации мы можем использовать:

- Например, если мы ищем минимум в R^n , то в качестве функции мутации можно выбирать замену произвольной случайно выбранной координаты (x_1, \dots, x_n) на другую произвольную координату.
- Можно расположить координаты в противоположном порядке: $M(x_1, \dots, x_n) = (x_n, \dots, x_1)$

Когда мы можем применять мутацию:

- До скрещивания - то есть только к начальной популяции.
- После скрещивания - то есть еще и к потомкам.

Более того, иногда может мутировать вся популяция, а иногда лишь определенная (и заранее заданная) ее часть. В целом, оператор мутации по сравнению с оператором скрещивания играет второстепенную роль. В классическом генетическом алгоритме скрещивание используется практически всегда, в то время как мутации могут и не использоваться (более того, используются мутации не слишком часто). Мы рассмотрим как алгоритм с использованием функции мутации, так и без использования данной функции.

Классический генетический алгоритм без использования мутаций.

Рассматриваем функцию 2-х переменных, например $func_1(x, y) = x^2 + (y - 1)^2 + 1$. В первой части нашего исследования мы реализуем классический генетический алгоритм без использования мутаций.

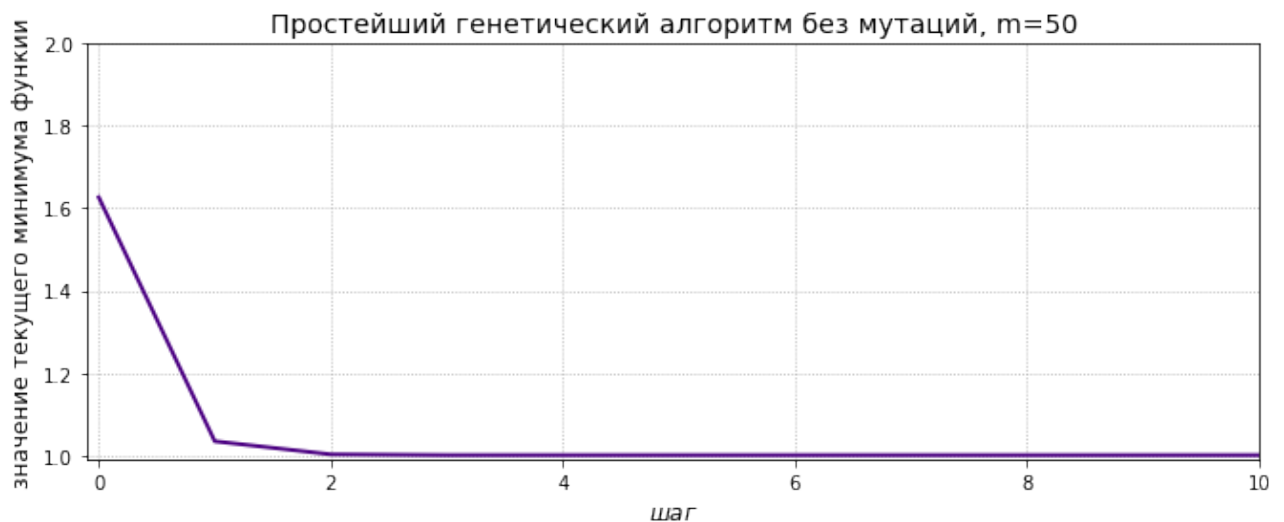
Начальный этап: случайно генерируем начальную популяцию и задаем фитнес-функцию.

Критерий остановки: пройденное количество шагов (100).

Функция скрещивания: поскольку мы работаем в двумерном пространстве, то в ходе скрещивания наследник со случайной вероятностью p получает первую координату от первого родителя, а вторую - от второго.

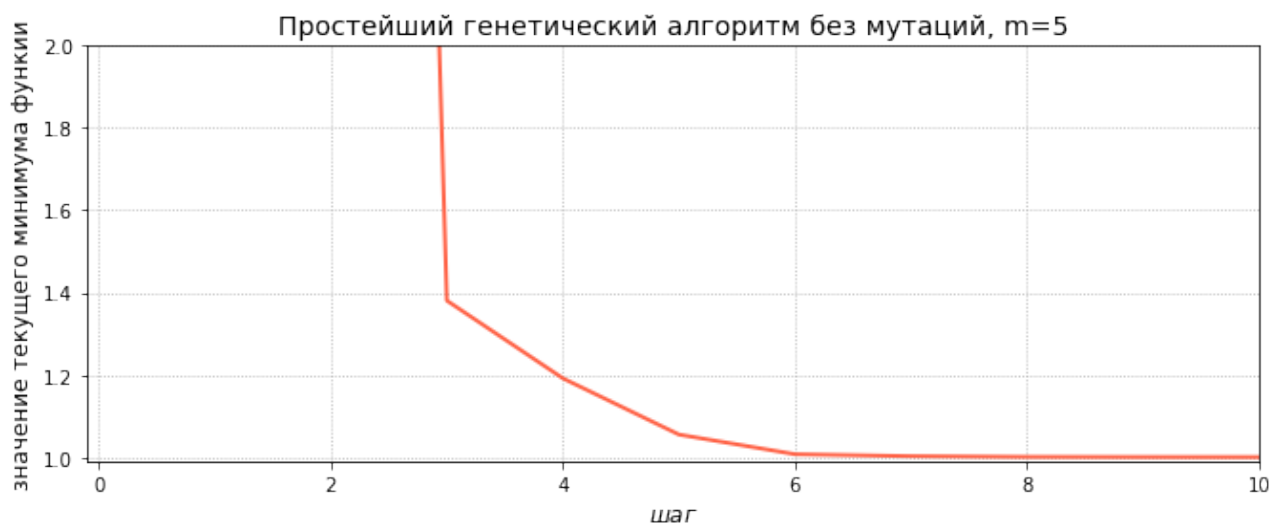
Создание новой популяции: рассматриваем наследников каждой пары точек, чтобы не упустить возможный оптимум.

Случай 1: Положим, размер популяции равным 50. Это и даст необходимую точность, и не приведет к излишне долгим вычислениям. Алгоритм возвращает найденный минимум: 1.0, точность довольно большая. Строим график:



Как мы видим, значение действительно уменьшается с течением времени и оптимум уже с довольно хорошей точностью достигался на втором шаге.

Случай 2: Положим теперь размер популяции равным 5. Алгоритм возвращает найденный минимум: 1.0000078. Как мы можем заметить, точность результата уже стала заметно меньше. Построим график:



Однако если запустить этот алгоритм еще раз при другой сгенерированной начальной выборке, то можно получить ответ: 2.241927.

Итак, сравним случаи 1 и 2. В первом случае мы рассматривали размер популяции, равный

50, во втором же случае он был в 10 раз меньше и равнялся 5. Как можно отметить по графику, во втором случае алгоритм сходил за большее число шагов, он достиг более-менее точного результата только при 7–8 шагах, в то время как в первом случае нам хватило двух. Более того, полученный нами ответ оказался менее точным, чем в первом случае, хотя общее количество шагов и было одинаковым. Однако в то же время в первом случае размер популяции был 50, так что в ходе сортировки до того, как мы достигли приемлемого результата, мы перебрали $50 \cdot 2 = 100$ особей (пробных решений), где 2 - число шагов, 50 - размер популяции, в то время как во втором случае это число составило $5 \cdot 8 = 40$. И все же во втором случае алгоритм мог и не сойтись, поскольку размер популяции очень маленький, так что если бы на начальном шаге мы сгенерировали популяцию неудачно, то к минимуму могли бы и не прийти, что нам и демонстрируем второй случай, когда мы получили значение, равное 2.24.

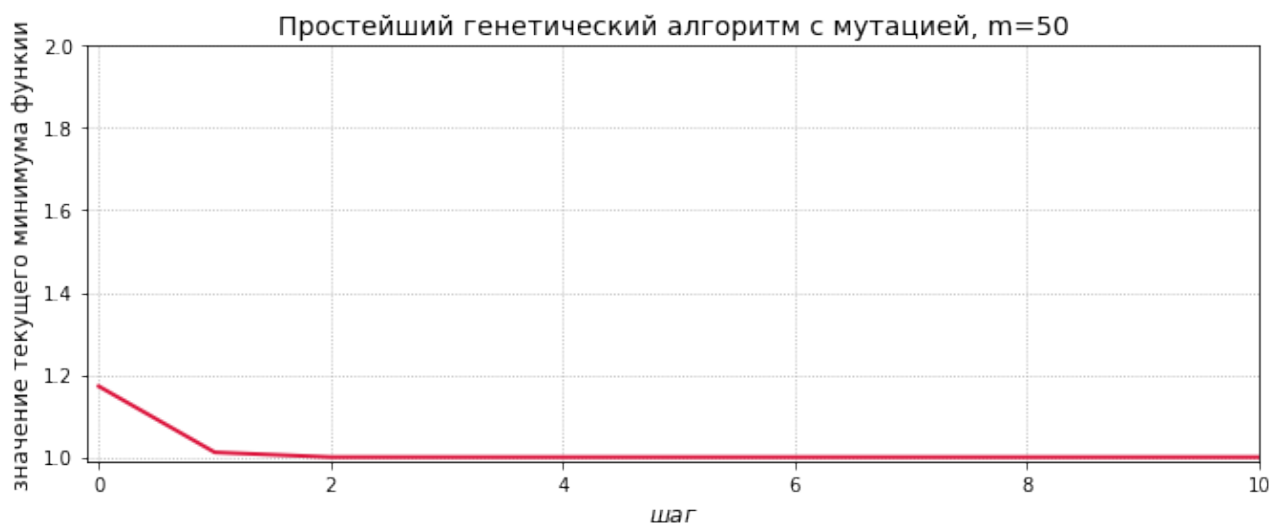
Таким образом, с одной стороны уменьшение размера начальной популяции уменьшило и точность, однако с другой стороны сильное увеличение размера популяции могло привести к большему времени работы, ведь нам бы пришлось скрещивать всех представителей, а при большом размере начальной популяции это заняло бы немало времени. Подводя итог, стоит заметить, что важно подобрать хороший размер начальной популяции, который предоставлял бы и необходимую точность, и в то же время не сильно увеличивал бы время работы. Данный размер может зависеть от задачи, поэтому в большинстве случаев программисту приходится задавать его самостоятельно.

Классический генетический алгоритм с использованием мутаций.

Во второй части нашего исследования мы реализуем классический генетический алгоритм с использованием мутаций. Критерий останова и функция скрещивания остаются прежними.

Мутация: в данном случае нам надо отметить две вещи: кто мутирует и как мутирует. Подвергать мутации мы будем только родителей. Таким образом, считаем, что вначале произошло скрещивание (в ходе которого характеристики родителей оставались неизменными), а затем произошла мутация, которая воздействовала только на старое поколение. Мы рассматриваем точки в двумерном пространстве, поэтому применим мутацию, которая будет случайным образом менять, например, вторую координату выбранной точки.

Рассматривать будем также функцию $func_1$, и после завершения алгоритма получим результат 1.0, опять же с довольно большой точностью.

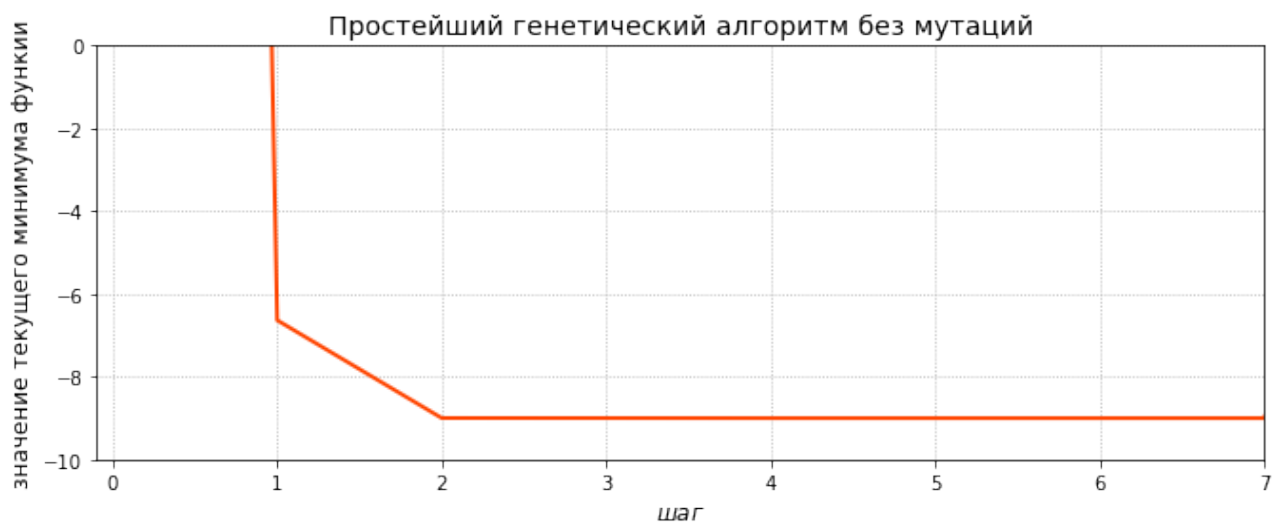


Как мы видим, значение действительно уменьшается с течением времени, причем быстрее, чем в предыдущем случае (без использования мутации). Почему это происходит? Когда мы включаем в наш алгоритм мутацию, мы увеличиваем множество принимаемых функцией значений, а затем выбираем из них все то же число наилучших, что и в первом случае. Однако теперь выбор

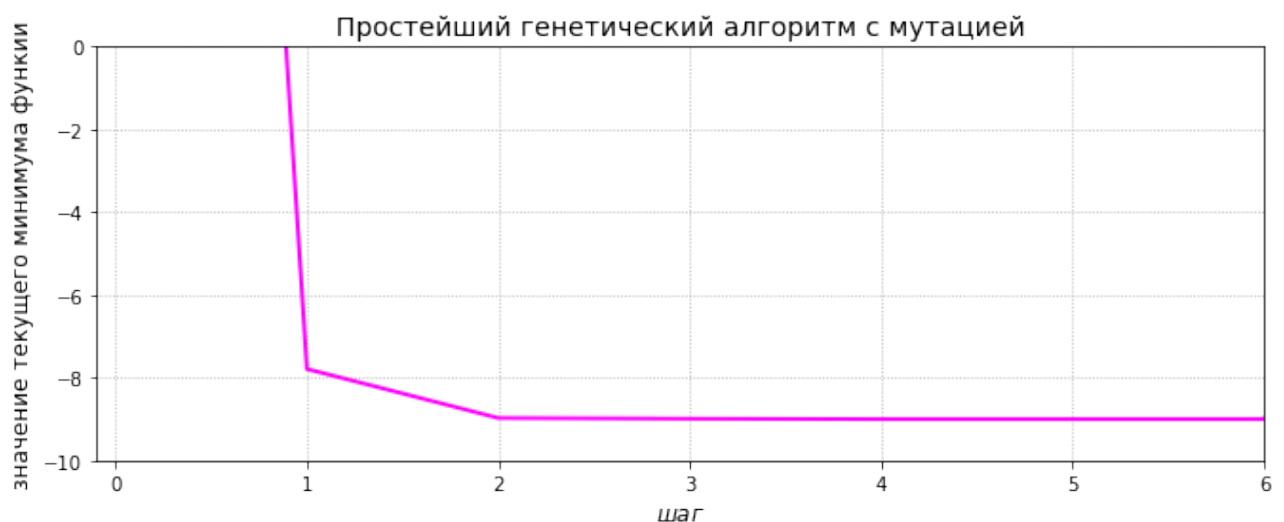
значений у нас больше, так что и сойтись мы можем быстрее. Но с другой стороны, мутация хотя и позволяет алгоритму сойтись за меньшее число шагов до нужной точности, она также влияет на время работы: при использовании мутации алгоритм будет работать дольше, ведь нам нужно вычислить еще одну дополнительную функцию от всей начальной популяции, а потом еще добавить множество ее значений к числу тех, из которых надо выбрать лучших.

Альтернативный критерий остановки

В третьей части нашей работы мы рассмотрим другой критерий остановки: сходимость по значению функции. Мы заранее задаем необходимую точность ($eps = 0.0000001$), и когда разность значений функции на двух соседних шагах будет меньше заданного значения, тогда мы и остановимся. Рассмотрим функцию $func_2 = x^4 + y^4 - 2x^2 + 4xy - 2y^2 - 1$. К ней применим для начала алгоритм без использования мутаций. Полученный минимум: -8.9999 , что согласовывается с теоретически вычисленным результатом (-9). Потребовавшееся число итераций равно 7.



Теперь запустим для этой же функции алгоритм с мутацией и с этим же критерием остановки. Полученный минимум: -8.9999 . Как мы видим, ответ снова получился очень близкий к правильному. Потребовавшееся число итераций равно 6.



Данный случай наглядно демонстрирует нам, что при использовании мутаций мы можем сойтись быстрее, нежели без использования мутаций. Действительно, в первом случае мы пришли к решению за 7 шагов, а во втором - уже за 6. Кроме того, второй алгоритм работал также довольно быстро, не дольше первого. Поэтому выбор между решением с мутацией и решением

без использования мутаций - это еще одна задача, которая ложится на плечи программиста, и в зависимости от программы нам нужно решить, что именно мы хотим: более быструю сходимость или же более быстрое время работы. Кроме того, мутацию также нужно хорошо подобрать, потому что если ее выбрать неэффективно, то время работы увеличится, а вот результат лучше не станет.

Сравним эти два критерия останова

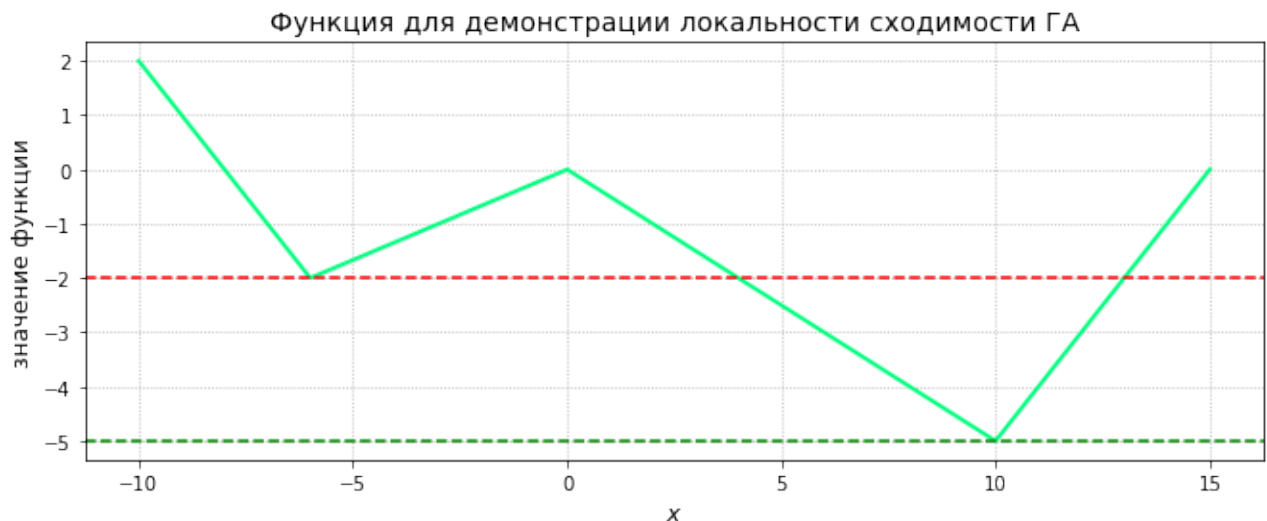
Когда мы заранее задаем количество итераций, то не можем предсказать, насколько точное решение получим. Например, с нашим числом итераций получилась очень большая точность, что с одной стороны очень хорошо, потому что ответ получился точный, но с другой стороны алгоритм работал относительно долго (в 6-8 раз дольше, чем во втором случае, когда критерием был модуль разности значений полученных на двух соседних итерациях оптимумов), да и не факт, что в реальных задачах такая большая точность действительно нужна. Во втором случае, мы можем сами регулировать точность, да и алгоритм работает быстрее, но если, например, минимумом функции будет $-\infty$, то во втором случае может получиться так, что мы будем работать без остановки. Например, если будет точность в районе 0.01, а значения минимумов на каждом из шагов будут убывать с большей скоростью и стремиться к $-\infty$, то ситуация получится не очень приятная. В случае же первого алгоритма мы гарантированно остановимся, когда совершим заданное число итераций. И в итоге в ответе получим, возможно, большое отрицательное число.

Локальность генетического алгоритма

Покажем, что наш алгоритм может сходиться к локальному минимуму, а не к глобальному. Рассмотрим кусочно-заданную функцию:

$$f(x) = \begin{cases} -x - 8, & x \leq -6 \\ \frac{x}{3}, & -6 < x \leq 0 \\ \frac{-x}{2}, & 0 < x \leq 10 \\ x - 15, & x > 10 \end{cases} \quad (1)$$

Поскольку ранее мы рассматривали функцию двух переменных, то и здесь также рассматриваем функцию двух переменных, однако вторую переменную считаем "фиктивной", то есть значение функции от нее зависеть не будет. В данном случае именно на примере такой функции довольно явно можно продемонстрировать локальную сходимость реализованного нами генетического алгоритма. Ее график выглядит так:



Как мы видим, у данной функции имеется в точке 10 глобальный минимум $y_{min} = -5$, а в точке -6 локальный минимум $y_{loc} = -2$. Генетический алгоритм основан на случайности. Мы случайно генерируем начальную популяцию, затем потомки наследуют случайное количество генов каждого из родителей, мутация также происходит случайно. Поэтому вполне может произойти такая ситуация, что начальная выборка получилась неудачная - все точки скопились около локального минимума. Тогда генетический алгоритм вполне может сойтись к нему, а не к глобальному оптимуму.

Например, если мы создадим начальную популяцию из 4 точек: $(-8, -2)$, $(-5.9, 0.3)$, $(-6.4, 14)$, $(-10, 18)$, то получим результат -1.9999 за 7 итераций (использовали второй критерий сходимости и алгоритм без мутации). Как мы видим, мы сошлись к локальному (-2) , а не к глобальному (-5) минимуму, что подтверждает один из недостатков генетического алгоритма: мы не можем исключать ситуацию, в которой сойдемся к локальному оптимуму.

Достоинства и недостатки генетических алгоритмов

Достоинства:

- Довольно часто генетические алгоритмы применяются в том случае, когда у нас очень большое пространство поиска.
- Есть возможность варьировать параметры поиска (увеличивая затраты на время, можно выиграть в точности, и наоборот)
- Большой выбор функций мутации/скрещивания/отбора, что позволяет оптимизировать работу той или иной задачи в зависимости от необходимости.
- Работает заведомо не хуже абсолютно случайного поиска

Недостатки:

- Как мы выяснили, генетический алгоритм может сходиться не к глобальному, а к локальному оптимуму, а иногда и вообще к произвольной точке. Более того, сходимость генетических алгоритмов не доказана.
- Необходимо самостоятельно подбирать более эффективную функцию мутации/скрещивания для каждой задачи, поэтому работа с алгоритмом может превратиться в попытки подобрать наиболее удачные параметры.
- Иногда в генетическом алгоритме отсутствует разнообразие в особях. Достаточно быстро находится локальный минимум, и в дальнейшем все остальные элементы популяции проигрывают ему отбор.

Способ борьбы с этим: сделать так, чтобы в отборе участвовали не только самые приспособленные (лучшие) особи, но вообще все, что мы и сделали в нашей реализации. В таком случае вероятность того, что результаты получатся разнообразными, возрастает.

- В простых целевых функциях генетические алгоритмы проигрывают по скорости простым алгоритмам поиска.