

Занятие №4

Генетический алгоритм решения многокритериальных многопараметрических переборных оптимизационных задач

Теоретические сведения

В основу генетического алгоритма положен естественный биологический метод отбора.

Цель биологической эволюции состоит в наилучшем приспособлении к жизненным условиям. С технической точки зрения наиболее приспособленными оказываются оптимальные по некоторому критерию системы.

Для оценки систем используют:

- степень приспособленности (биологические системы);
- критерии качества (технические системы).

Данное направление, получившее название техническая генетика является усечённым понятием биологической генетики.

Основные положения:

1) Генетическая модель – способ представления информации о системе в генетической форме.

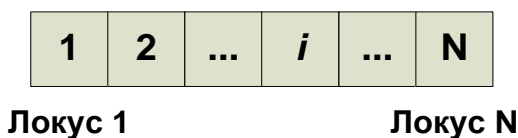
2) Генетический алгоритм – процесс приспособляемости, оптимизации системы.

3) Правила остановки - момент, на котором останавливается процесс оптимизации системы.

Основой построения генетической модели является двоичное кодирование. Носителем этого свойства системы является хромосома.

Свойств, которые отражаются в хромосоме, может быть много. Существует такое понятие как код признака, т.е. каждое свойство располагается в определённом месте хромосомы – в i -ом локусе ($i=1 \dots N$).

Хромосома:



В каждом локусе размещается один ген. Ген – это конкретный двоичный код признаков, выбранных из множества аллелей (аллелеформ).

Ген	000	} Аллелеформы (аллели)
	001	
	010	
	...	
	111	

Совокупность всех генов образует генофонд объекта.

Существуют следующие способы кодирования информации в генах:

- дискретное кодирование (перебор возможных различных вариантов; в соответствии с признаком задаётся код);
- непрерывное кодирование (например, параметр имеет значение от 0 до К с интервалом в единицу; берётся аллель, который соответствует в двоичном коде определённому числу).

Рассмотрим пример.

Пусть задан граф (рис. 4.1)

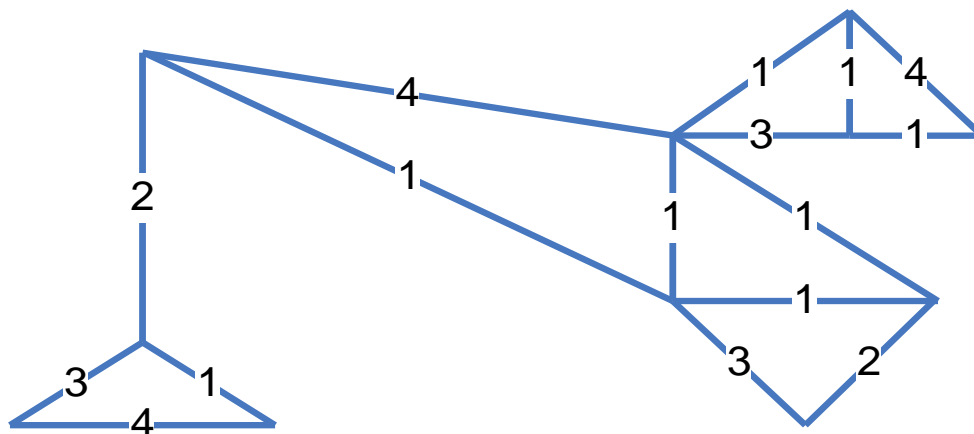


Рис 4.1 Неориентированный простой граф $G(X, V, W)$ порядка $n=12$ (цифрами указаны веса W соответствующих ребер).

Требуется найти оптимальное дихотомическое разбиение графа G для внешних параметров $n_1=5$ и $n_2=7$, которое обеспечивает минимальный вес разреза $Q(X_1, X_2)$.

Число допустимых решений равно $C_{12}^5 = 792$, два из которых приведены в Таблице 4.1.

Таблица 4.1

Допустимые решения для графа G

Разбиение (X_1, X_2)	Характеристики			
	f_1	f_2	F	Q
$X_1 = (x_1, x_7, x_8, x_{10}, x_{11});$ $X_2 = (x_2, x_3, x_4, x_5, x_6, x_9, x_{12})$	11	18	29	7
$X_1^* = (x_2, x_3, x_4, x_5, x_6);$ $X_2^* = (x_1, x_7, x_8, x_9, x_{10}, x_{11}, x_{12})$	14	20	34	2

Дихотомическое разбиение (X_1^*, X_2^*) является оптимальным разбиением графа, приведенного на рис. 4.1.

Обобщая вышесказанное, цель эволюции первоначально заданной популяции $P^0 = (a_1^0, \dots, a_v^0)$ в течение жизненного цикла T, можно сформулировать следующим образом.

Отношения между особями и внешней средой, приводящие к избирательной элиминации (“гибели”) менее приспособленных и выживанию более приспособленных особей, должны быть построены таким образом, чтобы в течение смены поколений в хромосомном наборе популяции накапливались такие новые качественные признаки (гены и генотипы), которые обеспечивают увеличение средней степени приспособленности особей по популяции в целом:

$$MAX_{P^t} \left\{ \frac{1}{v} \sum_{i=1}^v \mu(a_i^t) \right\}$$

При этом генотипы особей $a_k^t \in P^t$, $t=0, 1, 2, \dots, T$, заданные с помощью бинарных строк $E(a_k^t)$, в процессе своего преобразования в результате факторов эволюционного развития популяции по Дарвину в каждом поколении должны обладать следующими свойствами:

- наследственности, которая закрепляет у “потомков” лучшие признаки, полученные от “родителей” в результате их размножения;
- изменчивости, которая служит основой образования новых признаков за счет изменения генетического состава популяции в результате мутаций;
- соревновательности, которая определяет направление генетических изменений в популяции в результате естественного отбора по степени приспособленности особей к условиям внешней среды.

В дальнейшем под *генетическим алгоритмом* будем понимать алгоритмический подход к решению экстремальных задач однокритериального выбора, основанный на моделировании основных факторов эволюционного развития популяции.

Большую роль в развитии генетических алгоритмов сыграли I. Holland, D. Goldberg и L. Davis, которые заложили и развили теоретические основы генетического подхода к решению задач оптимизации. Не останавливаясь на обзоре этих работ, приведем обобщенную схему генетического алгоритма, структура которого является типичной для широкого круга публикаций по этому вопросу.

Базовая структура “Генетического алгоритма”:

1. Формирование начальной популяции P^0 из v особей (a_1^0, \dots, a_v^0) :
 - 1.1. Генерация хромосомного набора из v бинарных строк $E(a_k^t)$, удовлетворяющих требованиям, предъявляемым к символьной модели исходной экстремальной задачи.
 - 1.2. Преобразование бинарных строк $E(a_k^t)$ в соответствующие им векторы управляемых переменных $\bar{x}^k \in D$ и вычисление степени приспособленности $\mu(a_k^t)$ для каждой особи a_i^0 , обладающей генотипом $E(\bar{x}^i)$.
 - 1.3. Особи $a_k^0, k = \overline{1, v}$ образуют начальную популяцию P^t для поколения $t=0$.
2. Воспроизводство “потомков” с наследственными признаками “родителей”:
 - 2.1. Выбор конкретной “родительской” пары $(a_k^t, a_l^t) \in P^t$ для участия в процессе размножения.
 - 2.2. Выбор схемы размножения.

- 2.3. Построение по выбранной схеме из генотипов “родителей” $E(a_k^t), E(a_i^t)$ генотипов их “потомков” $E(b_i^t), i \geq 1$, сохраняющих наследственные признаки “родителей”.
- 2.4. Преобразование бинарных строк $E(b_i^t)$ в соответствующие векторы управляемых переменных $\bar{x}^i \in D$ и вычисление степени приспособленности “потомков”, обладающих генотипами $E(\bar{x}^i)$.
- 2.5. Вычисления повторяются с шага 2.1. до тех пор, пока не будет воспроизведено заданное число “потомков”.
3. Мутагенез, приводящий к генетическим изменениям “родительских” признаков:
- 3.1. Выбор типа мутации.
- 3.2. Построение по генотипу $E(a_k^t)$ одной из особей $a_k^t \in P^t$ генотипа $E(m_i^t)$ особи-”мутанта” m_i^t с помощью конкретного типа мутации.
- 3.3. Преобразование бинарной строки $E(m_i^t)$ в соответствующий вектор управляемых переменных $\bar{x}^i \in D$ и вычисление степени приспособленности особи-”мутанта” m_i^t , обладающей генотипом $E(\bar{x}^i)$.
- 3.4. Вычисления повторяются с шага 3.1. до тех пор, пока не будет создано заданное число “мутантов”.
4. Естественный отбор:
- 4.1. Определение среди “родителей”, “потомков” и “мутантов” особей, образующих репродуктивную группу, которая примет участие в естественном отборе.
- 4.2. Выбор схемы естественного отбора.
- 4.3. Формирование по выбранной схеме хромосомного набора популяции следующего поколения $P^{t+1} = (a_1^{t+1}, \dots, a_v^{t+1})$ из особей, принадлежащих репродукционной группе.
5. Проверка условий окончания процесса эволюции популяции P .
- Если условия окончания процесса эволюции не выполнены, то происходит смена поколений и все вычисления для популяции следующего $(t+1)$ - го поколения P^{t+1} повторяются с шага 2.
- В качестве условий окончания процесса эволюции популяции может использоваться одно из следующих неравенств:

$$t > T$$

или

$$D_B(t) = 0.$$

Выполнение неравенства $t > T$ означает, что эволюция популяции закончена в связи с тем, что она исчерпала свой жизненный цикл; окончание эволюции популяции при равенстве побитового разнообразия текущей популяции P^t нулю означает, что все генотипы в хромосомном наборе популяции P^t совпадают между собой.

В заключение приведем отличия генетических алгоритмов от поисковых методов оптимизации.

1. Генетические алгоритмы осуществляют *прямое манипулирование бинарными строками* $E(\vec{x})$, с помощью которых закодированы в двоичном коде допустимые решения $\vec{x} \in D$, а не самими внутренними параметрами $x_i, i = \overline{1, n}$, заданными действительными числами.
2. Генетические алгоритмы в каждом t -ом поколении *оперируют одновременно со всей совокупностью из v допустимых решений* $\vec{x}^k \in D$, образующих популяцию P^t , с целью получения хромосомного набора популяции следующего поколения P^{t+1} .

Таким образом генетические алгоритмы на каждой итерации, совпадающей с текущим поколением, *позволяют определять v новых допустимых решений*, в то время как классические методы поиска на каждой итерации определяют единственное новое допустимое решение. Например, градиентный метод минимизации реализуется с помощью рекуррентного выражения:

$$\vec{x}^{k+1} = \vec{k}^k - \lambda \nabla Q(\vec{k}^k), k=0,1,2,\dots$$

где \vec{x}^0 - начальное приближение;

$\nabla Q(\vec{k}^k)$ - градиент минимизируемой функции;

λ - шаг вдоль градиента.

3. Генетические алгоритмы основаны на вероятностных схемах преобразования бинарных строк $E(\vec{x})$, составляющих хромосомный набор популяции P^t , которые *моделируют биологические механизмы популяционной генетики*.

4. Генетические алгоритмы - это *методы нулевого порядка*, стратегия поиска, в которых построена только на вычислении значений критерия оптимальности Q и не требует знания дополнительной информации о производных, константе Липшица и т.д., что характерно для градиентных и квази-ньютоновских методов.
5. Генетические алгоритмы являются *робастными методами по отношению к виду минимизируемой функции*, т.к. при их применении не требуется, чтобы критерий оптимальности был непрерывным, дифференцируемым, унимодальным и т.д. Они осуществляют поиск оптимального решения по одной и той же стратегии как для унимодальных, так и для многоэкстремальных функций.