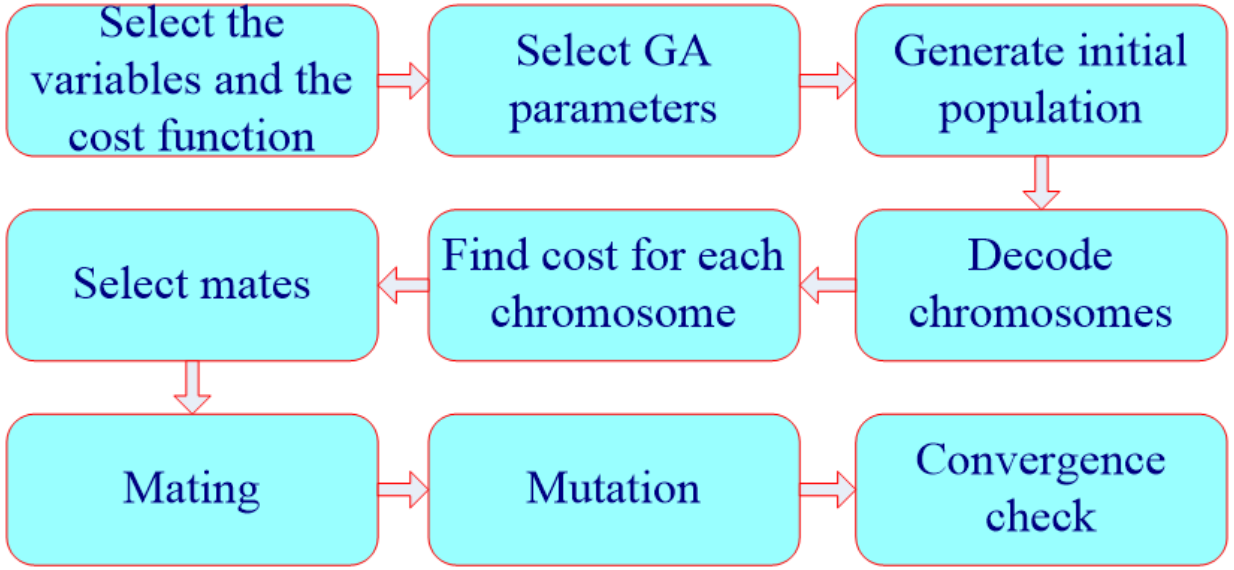
聚类——GAKFCM

作者：凯鲁嘎吉 - 博客园 http://www.cnblogs.com/kailugaji/

参考文献：黄白梅. 基于GA优化的核模糊C均值聚类算法的研究[D]. 武汉科技大学, 2013.

一、遗传算法



二、算法具体细节

1.Real Coding Mode

Each individual is represented by C×D real numbers, where C is the number of clusters and D is the dimension of the data.

2.Nonlinear Ranking Select Measurement

$$f(j) = q(1-q)^{j-1}$$
$$q \in (0,1),\ j = 1,2,\ldots,n$$

where q is the parameter, j is the sorting number and n is the number of individuals.

3.Adaptive Crossover Strategy

$$f_c(t) = p_{c_0}(1-\frac{t}{T})$$
$$p_{c_0} \in (0,1)$$

where p_{c0} is the initial crossover rate, t is the current evolution time of individuals and T is the maximum number of iterations.

4.Adaptive Mutation Strategy

$$f_m(t) = p_{m_0}(1-\frac{t}{T})$$
$$p_{m_0} \in (0,1)$$

where p_{m0} is the initial mutation rate, t is the current evolution time of individuals and T is the maximum number of iterations.

5.Fitness Function

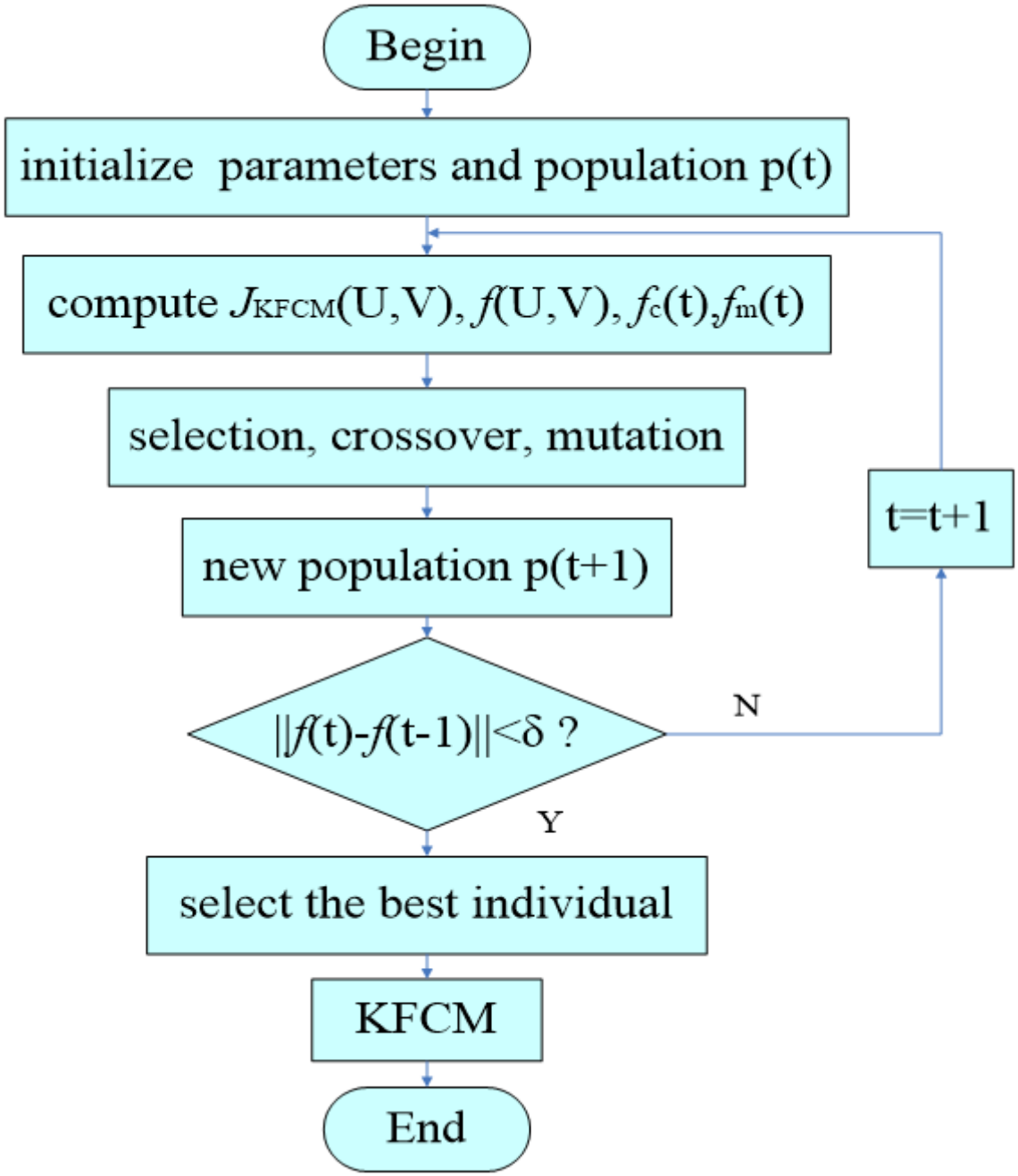
$$f(U,V)=\frac{1}{1+J_{KFCM}(U,V)}$$

其中J_{KFCM}见聚类——KFCM。

三、算法流程

- GAKFCM algorithm

1. Set the parameters of GAKFCM algorithm; Set the maximum number of iterations T and threshold $\varepsilon>0$; Initial population $p(t)$;
2. Compute $J_{KFCM}(U,V), f(U,V), f(j), f_c(t), f_m(t)$ of each individual in the population;
3. Compute the t population $p(t)$, selection-reproduction, crossover p_{ct}, as well as mutation p_{mt} of each individual in the t+1 population $p(t+1)$;
4. If $\|f^{(t+1)}(U,V)-f^{(t)}(U,V)\|>\varepsilon$ or $t<T$, t=t+1, go to step 2;
5. Select the best individual of the last generation as the algorithmi s final results; Clustering with KFCM.


Begin
initialize parameters and population p(t)
compute $J_{KFCM}(U,V)$, $f(U,V)$, $f_c(t)$,$f_m(t)$
selection, crossover, mutation
new population p(t+1)
$\|f(t)-f(t-1)\|<\delta$?
N
t=t+1
Y
select the best individual
KFCM
End

四、理解

GAKFCM是指用GA进行初始化KFCM的参数（聚类中心）。每个个体的大小与聚类中心的大小一致。