



机器学习及其MATLAB实现—从基础到实践 第9课

【声明】 本视频和幻灯片为炼数成金网络课程的教学资料，所有资料只能在课程内使用，不得在课程以外范围散播，违者将可能被追究法律和经济责任。

课程详情访问炼数成金培训网站

<http://edu.dataguru.cn>

- 第一课 MATLAB入门基础
- 第二课 MATLAB进阶与提高
- 第三课 BP神经网络
- 第四课 RBF、GRNN和PNN神经网络
- 第五课 竞争神经网络与SOM神经网络
- 第六课 支持向量机 (Support Vector Machine, SVM)
- 第七课 极限学习机 (Extreme Learning Machine, ELM)
- 第八课 决策树与随机森林
- **第九课 遗传算法 (Genetic Algorithm, GA)**
- 第十课 粒子群优化 (Particle Swarm Optimization, PSO) 算法
- 第十一课 蚁群算法 (Ant Colony Algorithm, ACA)
- 第十二课 模拟退火算法 (Simulated Annealing, SA)
- 第十三课 降维与特征选择

- 遗传算法 (Genetic Algorithm , GA) 是一种进化算法 , 其基本原理是仿效生物界中的 “物竞天择、适者生存” 的演化法则 , 它最初由美国Michigan大学的J. Holland教授于1967年提出。
- 遗传算法是从代表问题可能潜在的解集的一个**种群 (population)** 开始的 , 而一个种群则由经过基因 (gene) 编码的一定数目的**个体(individual)**组成。因此 , 第一步需要实现从表现型到基因型的映射即编码工作。初代种群产生之后 , 按照适者生存和优胜劣汰的原理 , 逐代 (generation) 演化产生出越来越好的近似解 , 在每一代 , 根据问题域中个体的**适应度 (fitness)** 大小选择个体 , 并借助于自然遗传学的遗传算子 (genetic operators) 进行组合交叉和变异 , 产生出代表新的解集的种群。这个过程将导致种群像自然进化一样 , 后世代种群比前代更加适应于环境 , 末代种群中的最优个体经过解码 (decoding) , 可以作为问题近似最优解。

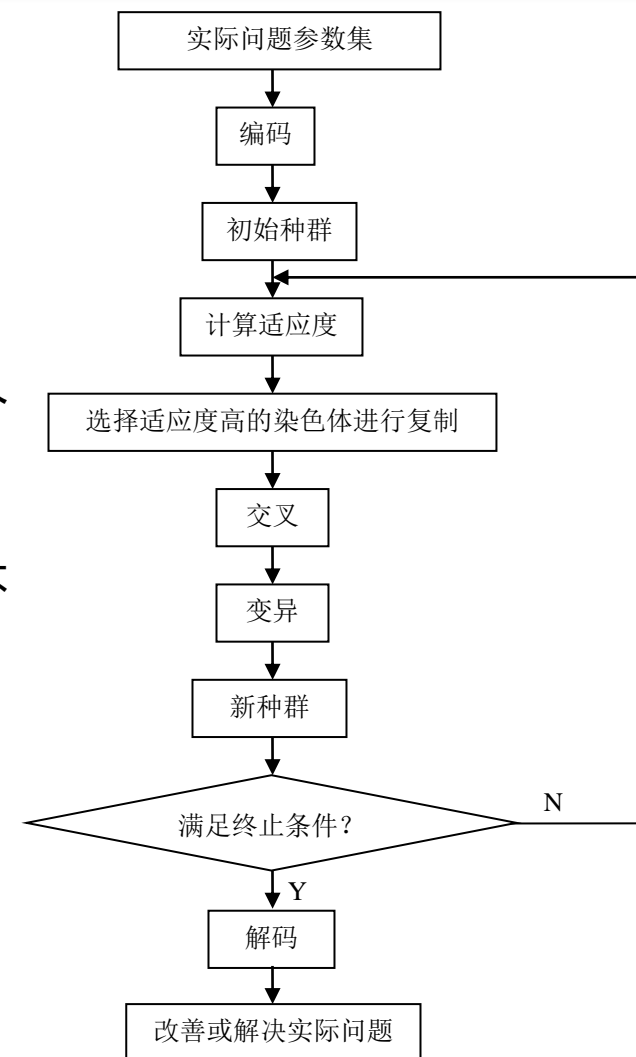
- 遗传算法有三个基本操作：**选择 (Selection)**、**交叉 (Crossover)**和**变异 (Mutation)**。
- (1) **选择**。选择的目的是为了从当前群体中选出优良的个体，使它们有机会作为父代为下一代繁衍子孙。根据各个个体的适应度值，按照一定的规则或方法从上一代群体中选出一些优良的个体遗传到下一代种群中。选择的依据是适应性强的个体为下一代贡献一个或多个后代的概率大。
- (2) **交叉**。通过交叉操作可以得到新一代个体，新个体组合了父辈个体的特性。将群体中的各个个体随机搭配成对，对每一个个体，以交叉概率交换它们之间的部分染色体。
- (3) **变异**。对种群中的每一个个体，以变异概率改变某一个或多个基因座上的基因值为其他的等位基因。同生物界中一样，变异发生的概率很低，变异为新个体的产生提供了机会。

•遗传算法的基本步骤：

1)编码：GA在进行搜索之前先将解空间的解数据表示成遗传空间的基因型串结构数据，这些串结构数据的不同组合便构成了不同的点。

2)初始群体的生成：随机产生N个初始串结构数据，每个串结构数据称为一个个体，N个个体构成了一个群体。GA以这N个串结构数据作为初始点开始进化。

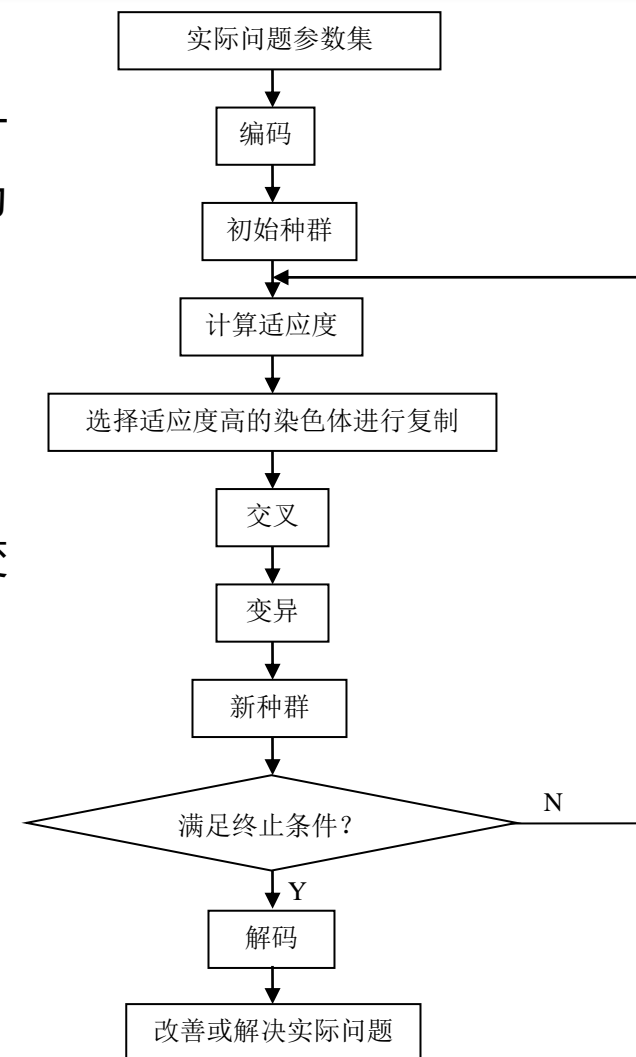
3)适应度评估：适应度表明个体或解的优劣性。不同的问题，适应性函数的定义方式也不同。



4)选择：选择的目的是为了从当前群体中选出优良的个体，使它们有机会作为父代为下一代繁殖子孙。遗传算法通过选择过程体现这一思想，进行选择的原则是适应性强的个体为下一代贡献一个或多个后代的概率大。选择体现了达尔文的适者生存原则。

5)交叉：交叉操作是遗传算法中最主要的遗传操作。通过交叉操作可以得到新一代个体，新个体组合了其父辈个体的特性。交叉体现了信息交换的思想。

6)变异：变异首先在群体中随机选择一个个体，对于选中的个体以一定的概率随机地改变串结构数据中某个串的值。同生物界一样，GA中变异发生的概率很低，通常取值很小。



- MATLAB内嵌遗传算法工具箱: gadst
- Sheffield大学遗传算法工具箱: gatbx
- 北卡罗来纳大学遗传算法工具箱: gaot

- **optimtool**
- **initializega**
- **ga**
- **normGeomSelect**
- **arithXover**
- **nonUnifMutation**

重点函数解读 —— initializega

```
pop = initializega(populationSize,variableBounds,evalFN,evalOps,options)
```

参数	意义
pop	随机生成的初始种群
populationSize	种群大小即种群中个体的数目
variableBounds	变量边界的矩阵
evalFN	适应度函数的名称
evalOps	适应度函数的参数
options	精度及编码形式，1为浮点编码，0为二进制编码

```
[x,endPop,bPop,traceInfo] = ga(bounds,evalFN,evalOps,startPop,opts,termFN,termOps,selectFN,selectOps,xOverFNs,xOverOps,mutFNs,mutOps)
```

参数	意义
bounds	变量上下界的矩阵
evalFN	适应度函数的名称
evalOps	适应度函数的参数
startPop	初始种群
opts	精度、编码形式及显示方式，1为浮点编码，0为二进制编码，默认为[1e-6 1 0]
termFN	终止函数的名称
termOps	终止函数的参数
selectFN	选择函数的名称
selectOps	选择函数的参数
xOverFNs	交叉函数的名称
xOverOps	交叉函数的参数
mutFNs	变异函数的名称
mutOps	变异函数的参数

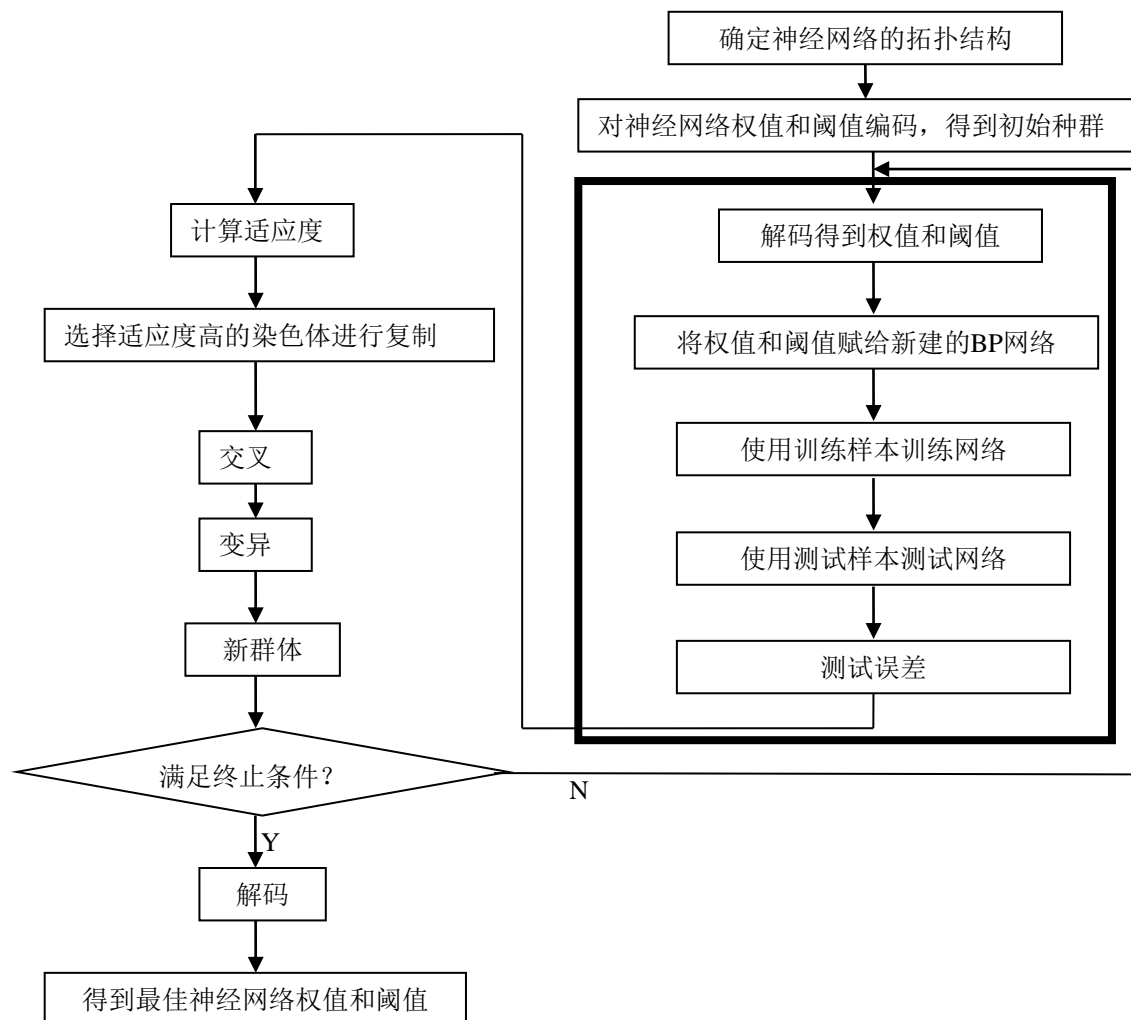
```
[x,endPop,bPop,traceInfo] = ga(bounds,evalFN,evalOps,startPop,opts,termFN,termOps,selectFN,selectOps,xOverFNs,xOverOps,mutFNs,mutOps)
```

参数	意义
x	优化计算得到的最优个体
endPop	优化终止时的最终种群
bPop	最优种群的进化轨迹
traceInfo	每代的最优适应度函数值和平均适应度函数值矩阵

一元函数优化

BP神经网络初始权值和阈值优化

遗传算法优化BP神经网络初始权值与阈值



- Dataguru (炼数成金) 是专业数据分析网站，提供教育，媒体，内容，社区，出版，数据分析业务等服务。我们的课程采用新兴的互联网教育形式，独创地发展了逆向收费式网络培训课程模式。既继承传统教育重学习氛围，重竞争压力的特点，同时又发挥互联网的威力打破时空限制，把天南地北志同道合的朋友组织在一起交流学习，使到原先孤立的学习个体组合成有组织的探索力量。并且把原先动辄成千上万的学习成本，直线下降至百元范围，造福大众。我们的目标是：低成本传播高价值知识，构架中国第一的网上知识流转阵地。
- 关于逆向收费式网络的详情，请看我们的培训网站 <http://edu.dataguru.cn>

Thanks

FAQ时间