

## AS 的重要步骤

AS 算法对 TSP 的求解流程主要有两大重要步骤：

- ① 路径构建
- ② 信息素更新

## 路径构建

定义：AS 中的随机比例规则

对每只蚂蚁  $k$ ，路径记忆向量  $\mathbf{R}^k$  按照访问顺序记录了所有  $k$  已经经过的城市序号。设蚂蚁  $k$  当前所在的城市为  $i$ ，则其选择城市  $j$  作为下一个访问对象的概率为：

$$p^k(i, j) = \begin{cases} \frac{[\tau(i, j)]^\alpha [\eta(i, j)]^\beta}{\sum_{u \in J_k(i)} [\tau(i, u)]^\alpha [\eta(i, u)]^\beta}, & j \in J_k(i) \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

其中； $J_k(i)$  表示从城市  $i$  可以直接到达的且又不在蚂蚁访问过的城市序列  $\mathbf{R}^k$  中的城市集合； $\eta(i, j)$  是一个启发式信息，通常由  $\eta(i, j) = 1/d_{ij}$  直接计算； $\tau(i, j)$  表示边  $(i, j)$  上的信息量。

## 信息素更新

- 初始化:  $\tau_0 = m/C^{nn}$
- 更新公式:

$$\tau(i, j) = (1 - \rho) \cdot \tau(i, j) + \sum_{k=1}^m \Delta\tau^k(i, j) \quad (2)$$

$$\Delta\tau^k(i, j) = \begin{cases} (C_k)^{-1}, & (i, j) \in \mathbf{R}^k \\ 0, & \text{otherwise} \end{cases}$$

其中,  $m$  是蚂蚁的个数;  $C^{nn}$  是由贪心算法构造的路径的长度;  $\rho \in (0, 1]$  是信息素的蒸发率, 通常设置为  $\rho = 0.5$ ;  $\Delta\tau^k(i, j)$  是第  $k$  只蚂蚁在它经过的边上释放的信息素量;  $C_k$  表示路径的长度, 它是  $\mathbf{R}^k$  中所有边的长度和.

# 智能优化技术

## 分布估计算法 (Estimation of Distribution Algorithms)

龚文引 (教授、博士生导师)

中国地质大学 (武汉) 计算机学院

May 14, 2020

# 1. 大纲

算法简介

几个简单示例

典型 EDA 简介

小结

## 2. 算法简介

### 智能优化中的学习与建模 (learning/modeling)

- 智能优化方法具有**自适应性** (adaptive)
  - 下一次的**搜索 (方向和步长)** 依赖于对以前搜索经验的估计 (如: 函数值、梯度等)
- 传统的演化算法采用杂交和变异产生新个体, 其**学习和建模**过程是**隐式的** (implicit)
- 有的优化算法采用**显式的** (explicit) 学习和建模产生新个体
  - Nelder-Mead 单纯形法
  - 基于响应面 (Response surfaces) 的方法: meta-modeling, surrogate function
  - 基于概率估计 (Probability estimation) 的方法: 分布估计算法 (Estimation of Distribution Algorithms, EDA)

## 2. 算法简介

### 分布估计算法

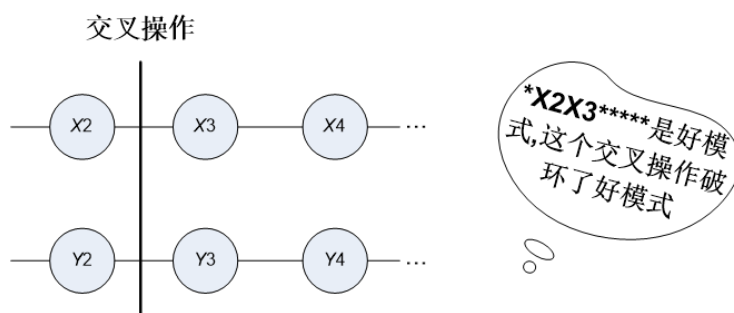
- 简称 EDA;
- 基于种群的新型演化算法;
- 思想起源于遗传算法;
- 采用（无监督的）密度估计或产生式统计模型.

P. Larranaga and J. A. Lozano (eds.). *Estimation of Distribution Algorithms: a new tool for evolutionary computation*. Kluwer Academic Publishers, 2002.

## 2. 算法简介

### 算法思想

- 改进遗传算法的交叉操作和变异操作，防止破坏积木块;
- 采用概率建模和抽样的显式形式产生新个体;
- 其主要思想是把优化问题转化成概率分布的搜索过程.



## 2. 算法简介

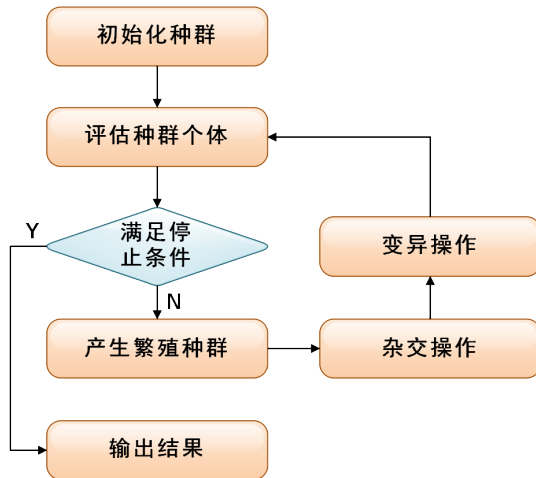


Figure: GA 流程图

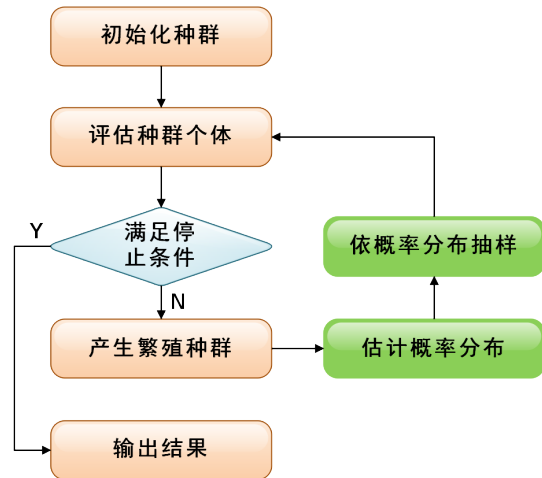


Figure: EDA 流程图

## 2. 算法简介

### 发展历史

- 开山始祖
  - Population-based incremental learning (PBIL)
  - Univariate marginal distribution algorithm (UMDA)
- 早期算法专注于二进制编码
  - Mutual information maximizing input clustering (MIMIC)
  - Bayesian optimization algorithm (BOA)
  - Bivariate marginal distribution algorithm (BMDA)
  - Extended compact genetic algorithm (eCGA)
- 逐渐扩展到连续 EDA
  - PBILc
  - UMDAc
  - ...

## 2. 算法简介

### EDA 方法分类

不同 EDA 算法的主要区别在于：**估计概率密度的方法不同，即建模的不同**

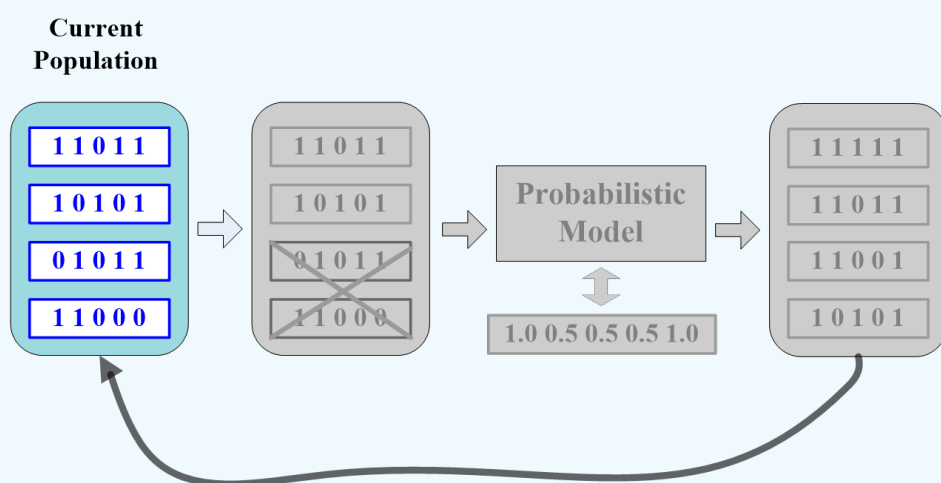
- 独立变量 EDA
  - Population-based incremental learning (PBIL)
  - Univariate marginal distribution algorithm (UMDA)
  - Compact genetic algorithm (cGA)
- 双变量依赖 EDA
  - Mutual information maximizing input clustering (MIMIC)
  - Bivariate marginal distribution algorithm (BMDA)
- 多变量依赖 EDA
  - Bayesian optimization algorithm (BOA)
  - Linkage-tree Genetic Algorithm (LTGA)

[https://en.wikipedia.org/wiki/Estimation\\_of\\_distribution\\_algorithm](https://en.wikipedia.org/wiki/Estimation_of_distribution_algorithm)

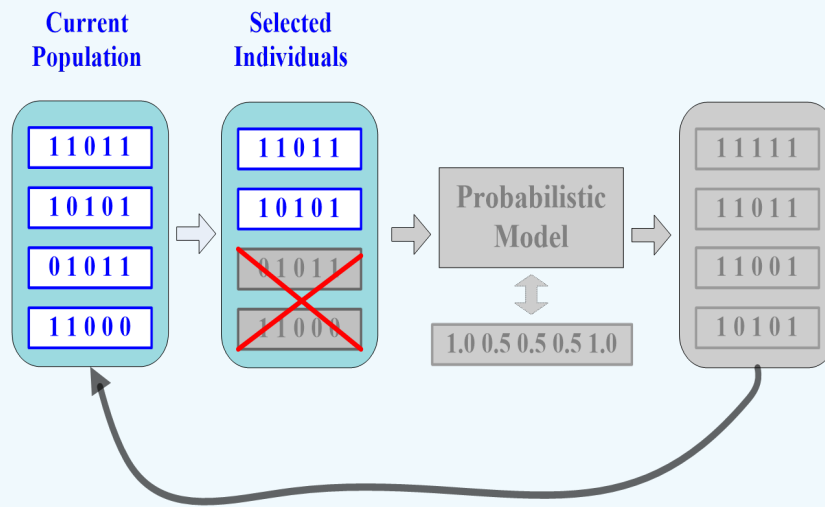
## 3. 几个简单示例

### 示例 1

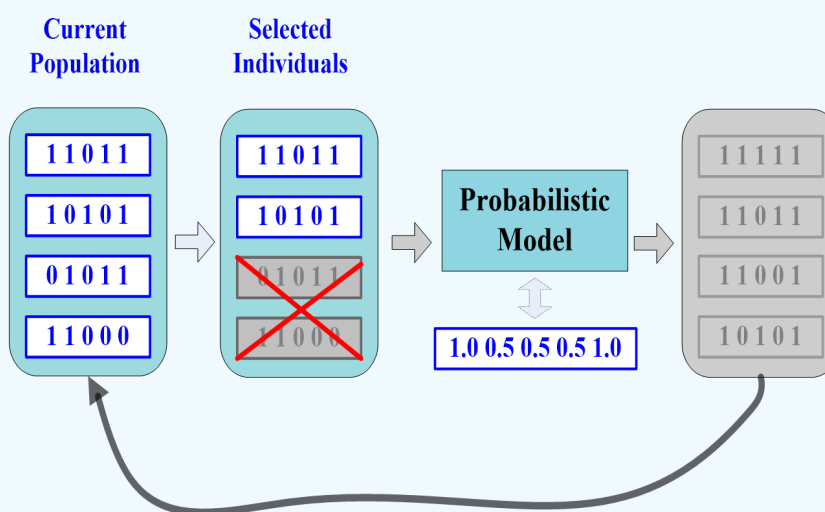
#### 初始群体



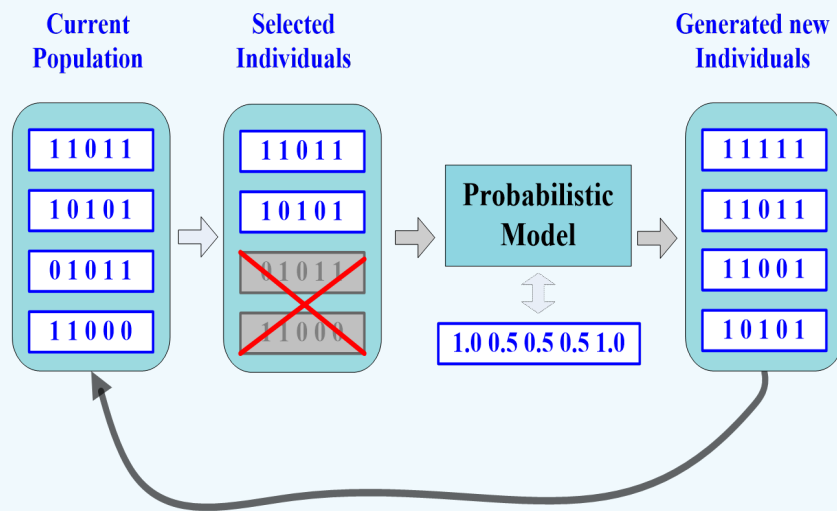
## 繁殖选择



## 概率密度估计 (modeling)



## 新种群产生 (sampling)



## 问题

假设有如下离散优化问题：

$$\min f(\mathbf{x}) = x_1 + x_2$$

其中,  $x_1 \in \{1, 2, 3, 4, 5\}$ ,  $x_2 \in \{6, 7, 8, 9, 10\}$ .



## 建立概率模型

建立一个概率模型。这里我们只需建立一个离散的概率模型（假设  $x_1, x_2$  相互独立），初始化如下：

$x_1$	1	2	3	4	5
$p_1$	0.2	0.2	0.2	0.2	0.2
$x_2$	6	7	8	9	10
$p_2$	0.2	0.2	0.2	0.2	0.2

## 采样与选择

- 1) 采样 (Sampling) :
  - 根据概率模型，采用若干个点（6 个）。假设采到的点为：  
 $\mathbf{y}_1 = (2, 7), \mathbf{y}_2 = (3, 9), \mathbf{y}_3 = (3, 10), \mathbf{y}_4 = (4, 8), \mathbf{y}_5 = (2, 9), \mathbf{y}_6 = (5, 6);$
- 2) 计算适应值 (Fitness evaluation) :
  - $f(\mathbf{y}_1) = 9, f(\mathbf{y}_2) = 12, f(\mathbf{y}_3) = 13, f(\mathbf{y}_4) = 12, f(\mathbf{y}_5) = 11, f(\mathbf{y}_6) = 11;$
- 3) 选择 (Selection) :
  - 选择最优的 3 个点:  $\mathbf{y}_1, \mathbf{y}_5, \mathbf{y}_6.$

## 更新概率模型 1

根据所选择的点更新概率模型:  $\mathbf{y}_1 = (2, 7)$ ,  $\mathbf{y}_5 = (2, 9)$ ,  $\mathbf{y}_6 = (5, 6)$

1) 更新变量  $x_1$  的分布:

- $p'_1(x_1 = 1) = 0$ ,  $p'_1(x_1 = 2) = 2/3$ ,  $p'_1(x_1 = 3) = 0$ ,  $p'_1(x_1 = 4) = 0$ ,  
 $p'_1(x_1 = 5) = 1/3$ .

2) 更新变量  $x_2$  的分布:

- $p'_2(x_2 = 6) = 1/3$ ,  $p'_2(x_2 = 7) = 1/3$ ,  $p'_2(x_2 = 8) = 0$ ,  $p'_2(x_2 = 9) = 1/3$ ,  
 $p'_2(x_2 = 10) = 0$ .

$x_1$	1	2	3	4	5
$p'_1$	0	2/3	0	0	1/3
$x_2$	6	7	8	9	10
$p'_2$	1/3	1/3	0	1/3	0

## 更新概率模型 2

对分布函数进行平滑处理（防止有一些点的概率为 0，自变量的值相差较近则概率应该相差不大）

- 以  $x_1$  为例:

$$p''_1(x_1 = i) = \sum_{j=1}^5 \exp(-|i - j|) p'_1(x_1 = j)$$

并归一化. 同理求  $p''_2$ , 得到:

$x_1$	1	2	3	4	5
$p''_1$	0.139861	0.380183	0.16124	0.117459	0.201247
$x_2$	6	7	8	9	10
$p''_2$	0.277140	0.279550	0.125736	0.108491	0.209084

## 更新概率模型 3

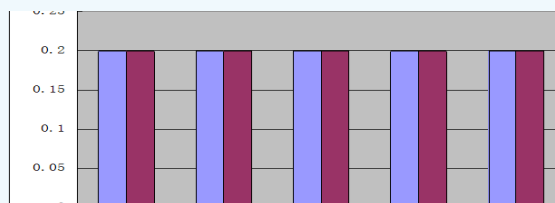
按照增量 (Incremental) 方式更新概率分布:

- $p_1 = \gamma \cdot p_1'' + (1 - \gamma) \cdot p_1$
- $p_2 = \gamma \cdot p_2'' + (1 - \gamma) \cdot p_2$
- 其中  $\gamma \in [0, 1]$  是学习因子, 此处设  $\gamma = 0.5$ .

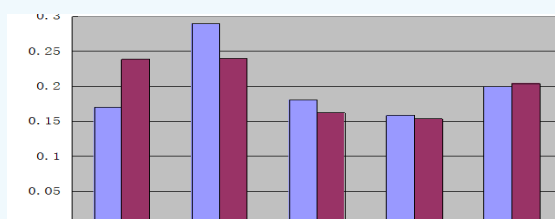
$x_1$	1	2	3	4	5
$p_1$	0.169931	0.290091	0.180625	0.158729	0.200624
$x_2$	6	7	8	9	10
$p_2$	0.238570	0.239775	0.162868	0.154245	0.204542

## 更新概率模型

- 学习之前

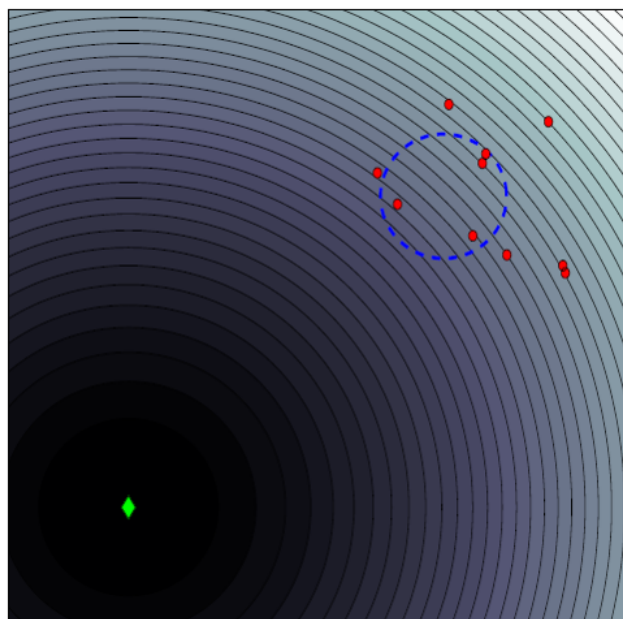


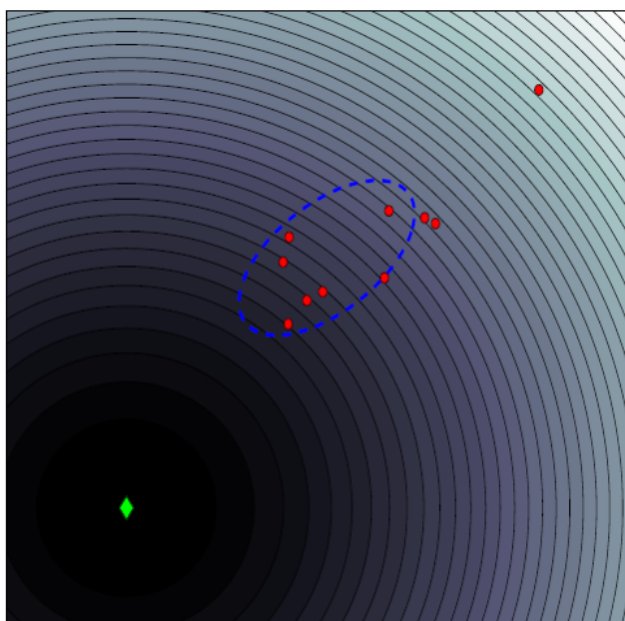
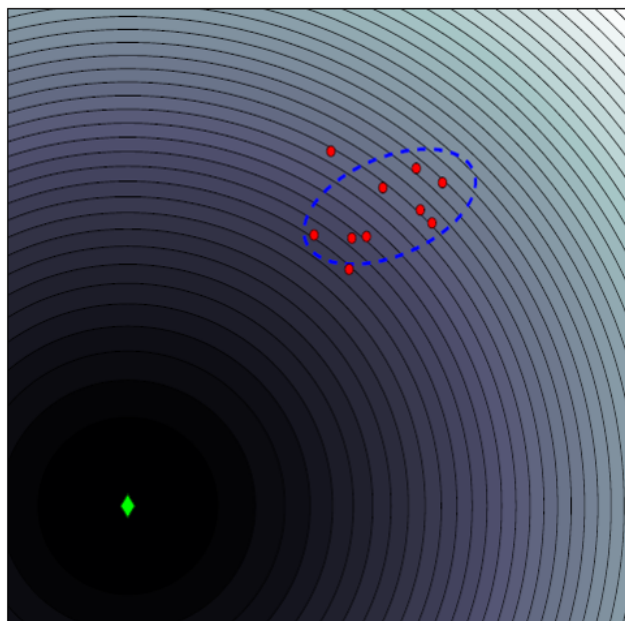
- 学习之后

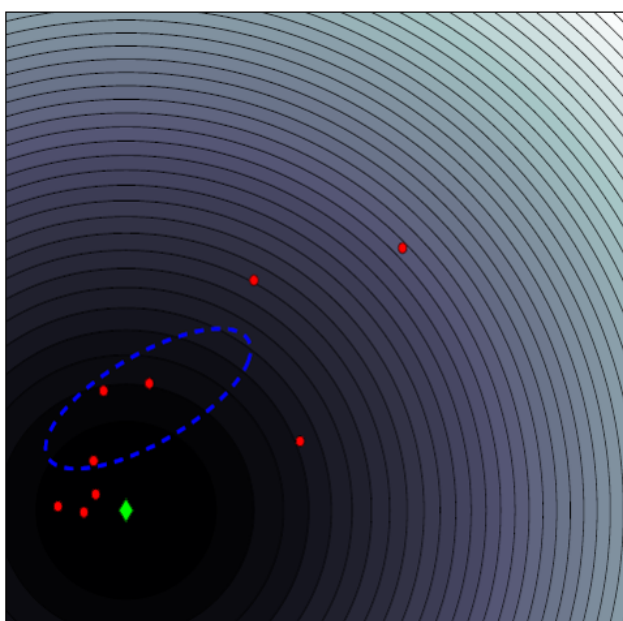
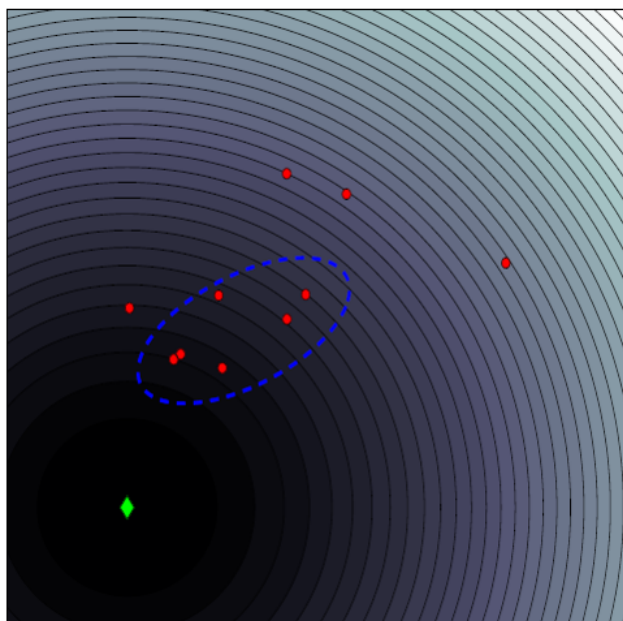


## 更新概率模型

- 经过学习之后，优秀个体对应的坐标的概率会变高;
- 平滑处理可以保证:
  - 若一个个体位置和优秀个体距离较近，那么该位置会受惠于该优秀个体，概率也会比较高（基于这样一个假设：若个体相差不大，则其适应度应该也相差不大）







## Univariate marginal distribution algorithm (UMDA)

UMDA 是一种简单的 EDA，它通过估计所选择个体的边界概率（marginal probabilities）来建立概率模型。

- 从当前群体中选择  $\lambda$  个个体组成繁殖群体  $S(t)$ ;
- 依据  $S(t)$  建立如下概率模型:

$$p_{t+1}(x_i) = \frac{1}{\lambda} \sum_{\mathbf{x} \in S(t)} x_i, i = 1, \dots, n$$

## Population-based incremental learning (PBIL)

PBIL 是 UMDA 的一种改进，它通过增量的方式来建立概率模型。

- 从当前群体中选择  $\lambda$  个个体组成繁殖群体  $S(t)$ ;
- 依据  $S(t)$  建立如下概率模型:

$$p_{t+1}(x_i) = (1 - \gamma) \cdot p_t(x_i) + \gamma \cdot \frac{1}{\lambda} \sum_{\mathbf{x} \in S(t)} x_i$$

$$i = 1, \dots, n$$

- 其中,  $\gamma \in (0, 1]$  是学习因子 (learning rate) .

## 基于高斯模型的 EDA：算法流程

- ① 随机生成初始种群  $P(0)$ , 初始化高斯模型的均值  $\mu$  与方差  $\delta$ . 假设每维变量对应的高斯模型为:  $\{N(\mu_1, \delta_1), \dots, N(\mu_n, \delta_n)\}$ ,  $n$  为自变量维数;
- ② 根据各维变量对应的高斯模型, 抽样产生新种群  $T(t)$ ; // 抽样
- ③ 从新种群中选择优秀个体集合  $S(t)$ ; // 选择
- ④ 计算  $S(t)$  在各维变量上的均值与方差, 对原有的高斯模型进行更新; // 建模
- ⑤ 若未达到停止条件则返回步骤2; 否则算法退出.

## 基于高斯模型的 EDA：更新概率模型参数

- 更新均值:

$$\mu_j^{t+1} = (1 - \alpha) \cdot \mu_j^t + \alpha \cdot (x_j^{\text{best},1} + x_j^{\text{best},2} - x_j^{\text{worst}})$$

- 更新方差:

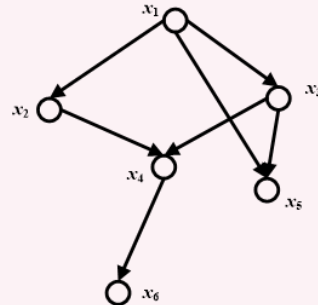
$$\delta_j^{t+1} = (1 - \alpha) \cdot \delta_j^t + \alpha \cdot \sqrt{\left( \sum_{k=1}^K (x_j^k - \bar{x}_j) \right) / K}$$

- 其中,  $\alpha \in [0, 1]$  是学习因子,  $K$  为所选择的优秀个体数目,  $\bar{x}$  是所选择的优秀个体的平均值向量,  $x^{\text{best},1}, x^{\text{best},2}$  是当前群里中最好的两个个体,  $x^{\text{worst}}$  是当前群体最差的个体;
- 更新的方式多种多样, 也可以直接用优秀个体的均值和方差替代原来的均值与方差.



## Bayesian optimization algorithm (BOA)

- BOA 采用贝叶斯网络 (Bayesian networks) 来描述变量之间概率依赖关系



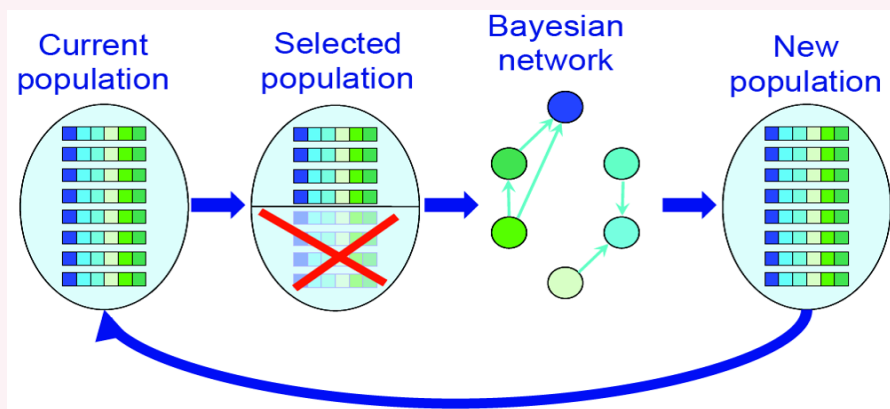
其中节点代表变量，边表示变量之间的概率依赖关系；

- 基于贝叶斯网络的联合概率密度为：

$$p(\mathbf{x}) = p(x_n | x_{n-1}, \dots, x_1) \cdot p(x_{n-1} | x_{n-2}, \dots, x_1) \cdots p(x_2 | x_1) \cdot p(x_1)$$

## Bayesian optimization algorithm (BOA)

- BOA 的关键问题：
  - 贝叶斯网络的结构如何确定？



### 课外作业

在网上下载 PBIL 源程序，读懂，并运行。

- ① 对每行代码添加中文注释。

## 5. 小结

### 本章小结

- ① EDA 简介
- ② 简单示例
- ③ 典型 EDA 简介

## 5. 小结

### 思考

在网上下载几种 EDA 源程序，读懂，并运行。思考以下问题：

- ① 不同 EDA 建立概率模型的方法有何异同？
- ② 能否把传统演化算法与 EDA 方法结合？

## 5. 小结

### 网络资源

- <http://www.iba.t.u-tokyo.ac.jp/english/EDA.htm>
- <http://martinpelikan.net/presentations.html>
- <http://medal.cs.umsl.edu/>
- <http://www-illigal.ge.uiuc.edu/>
- <http://www.evolution.re.kr/>

## 5. 小结

### 进一步阅读资料

- 周树德, 孙曾圻. “分布估计算法综述,” [自动化学报](#), vol. 33, no. 2, pp. 112 - 124, 2007.
- M Pelikan, D.E Goldberg & E Cantu-paz: “Linkage Problem, Distribution Estimation and Bayesian Networks,” [Evolutionary Computation](#), vol. 8, no. 3, pp. 311 - 340, 2000.
- W. Dong, T. Chen, P. Tino and X. Yao, “Scaling Up Estimation of Distribution Algorithms for Continuous Optimization,” [IEEE Transactions on Evolutionary Computation](#), vol. 17, no. 6, pp. 797-822, Dec. 2013.

## 6. 致谢

*Thank you!*

AUTHOR: GONG, Wenyin  
ADDRESS: School of Computer Science,  
China University of Geosciences,  
Wuhan, 430074, China  
E-MAIL: [wygong@cug.edu.cn](mailto:wygong@cug.edu.cn)  
HOMEPAGE: <http://www.escience.cn/people/wygong>