Part I Evolutionary Computation

Chapter 4
Evolution Strategies

Evolution Strategies

- **4-1 (1+1)-ES**
- **4-2 Generic ES Algorithm**
- **4-3 Implementation of ES**
- **4-4 Others**

ES

- **1960s, Rechenberg and Schwefel.**
- **Based on the concept of the evolution of evolution.**
- **Since biological processes have been optimized by evolution and evolution is a biological process itself, then it must be the case that evolution optimizes itself.**
- **The emphasis is toward the phenotypic bahavior of individuals.**

ES

- **Each individual is represented by its genetic building blocks and a set of strategy parameters that models the behavior of that individual in its enviroment.**
- **Evolution then consists of evolving both the genetic characteristics and the strategy parameters, where the evolution of the genetic characteristics is controlled by strategy parameters.**

ES

- Mutation are only accepted in the case of success---mutated individuals are only accepted if the mutation resulted in improving the fitness of the individual.
- Offspring can be produced from more than two parents.

4-1 (1+1)-ES

- **Do not make use of a population**
- **A single individual is used from which offspring is produced through application of a mutation operator.**
- **One of the first evolutionary algorithms that represents an individual as a tuple to consist of the decision vector x ,to be optimized and a vector of strategy parameters σ.**

$$\chi(t) = (\mathbf{x}(t), \sigma(t))$$

- **The strategy parameter vector represents the mutational step size for each dimension, which is adapted dynamically according to performance.**
- **According to the biological observation that offspring are similar to their parents and that smaller deviations from the parent occur more often than larger ones.**

- **the offspring is created**

$$\chi'(t) = (\mathbf{x}'(t), \sigma'(t))$$

$$\begin{aligned} x_j'(t) &= x_j(t) + N_j(0, \sigma_j(t)) \\ &= x_j(t) + \sigma_j(t) N_j(0, 1) \end{aligned}$$

- **Strategy parameters are adapted based on the 1/5 success rule(Rechenberg)**
- **Increase deviations σ, if the relative frequency of successful mutations over a certain period is larger than 1/5; otherwise deviations are decreased.**

$$\sigma(t+1)=\begin{cases}\sigma(t)/c & \varphi(k)>1/5\\ \sigma(t)*c & \varphi(k)<1/5\\ \sigma(t) & \varphi(k)=1/5\end{cases}$$

最后k代迭代中成功变异的比率，k>10

0.817<=c<=1

Selection

- **The selection operator selects the best between the parent and the offspring.**

$$\mathbf{x}(t+1) = \begin{cases} \mathbf{x}'(t) & \text{if } f(\mathbf{x}'(t)) < f(\mathbf{x}(t)) \\ \mathbf{x}(t) & \text{otherwise} \end{cases}$$

$$\sigma(t+1) = \begin{cases} \sigma'(t) & \text{if } f(\mathbf{x}'(t)) < f(\mathbf{x}(t)) \\ \sigma(t) & \text{otherwise} \end{cases}$$

(μ+1)-ES

- Rechenberg suggested that the (1+1)-ES can be extended to a multimembered ES, denoted as the (μ+1)-ES. This strategy uses a population of μ>1 parents. Two parents are randomly selected and recombined by discrete, multipoint crossover to produce one offspring.

$$x_j'(t) = \begin{cases} x_{1j}(t) & \text{if } r_j \leq 0.5 \\ x_{2j}(t) & \text{otherwise} \end{cases}$$

$$r_j \sim U(0,1), j = 1, \ldots, n_x.$$

$$\sigma_j(t) = \begin{cases} \sigma_{1j}(t) & \text{if } r_j \leq 0.5 \\ \sigma_{2j}(t) & \text{otherwise} \end{cases}$$

- The offspring is mutated as for (1+1)-ES.
- An elitist approach is followed to select the new population.
- The best μ individuals out of the $\mu+1$ survive to the next generation.

4-2 Generic ES Algorithm

Set the generation counter, $t = 0$;
Initialize the strategy parameters;
Create and initialize the population, $\mathcal{C}(0)$, of μ individuals;
for *each individual*, $\chi_i(t) \in \mathcal{C}(t)$ **do**
    Evaluate the fitness, $f(\mathbf{x}_i(t))$;
end
while *stopping condition(s) not true* **do**
    for $i = 1, \ldots, \lambda$ **do**
        Choose $\rho \geq 2$ parents at random;
        Create offspring through application of crossover operator on parent genotypes and strategy parameters;
        Mutate offspring strategy parameters and genotype;
        Evaluate the fitness of the offspring;
    end
    Select the new population, $\mathcal{C}(t + 1)$;
    $t = t + 1$;
end

Main components of Generic ES

- **Initialization**: for each individual, its genotype is initialized to fall within the problem boundary constraints. Also the strategy parameters.
- **Recombination**: offspring are produced through application of crossover operator on two or more parents.
- **Mutation**: offspring are mutated, where mutational step sizes are determined from self-adaptive strategy parameters.

main components

- **Evaluation: an absolute fitness function is used to determine the quality of the solution represented by the genotype of the individual.**
- **Selection:select parents for recombination and determine which individuals survive to the next generation.**

1. Representation

- 适合于连续变量优化

- **1**）二元表示：包括目标变量和标准差

$$(\mathrm{x},\sigma)=(x_1,x_2,\cdots x_n,\sigma_1,\sigma_2,\cdots\sigma_{n_\sigma})$$

n或者1

- **2**）三元表示：增加一个坐标旋转角度**α**，表示两个分量之间的旋转角度。

$$\alpha=(\alpha_1,\alpha_2,\cdots\alpha_s),s=\frac{n(n-1)}{2}$$

2. Mutation

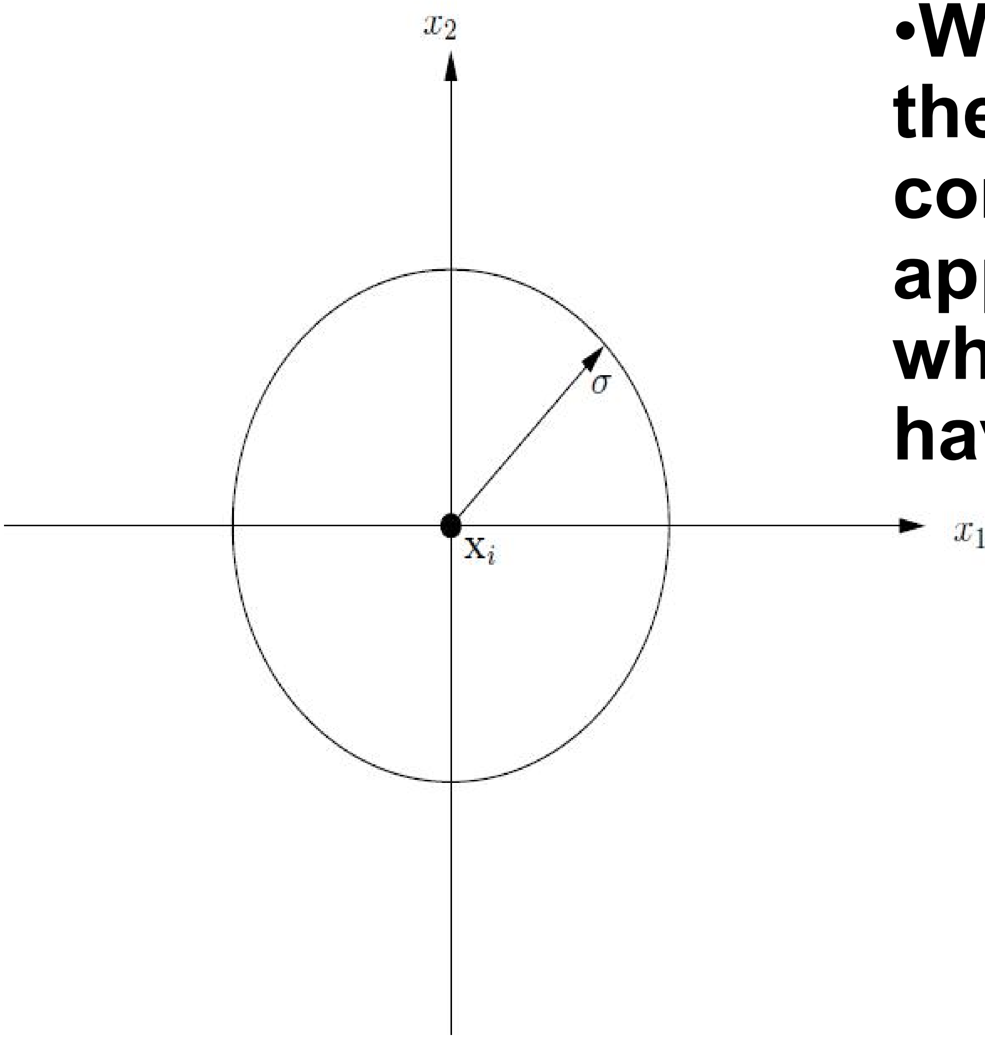
- **1) only one deviation parameter is used for all components of the genotype.**

$$n_\sigma = 1$$

$$\sigma_i'(t) = \sigma_i(t)e^{\tau N(0,1)} \qquad \tau = \frac{1}{\sqrt{n_x}}.$$

$$x_i' = x_i + \sigma' N_i(0,1)$$

标准差乘以服从对数正态分布的随机数



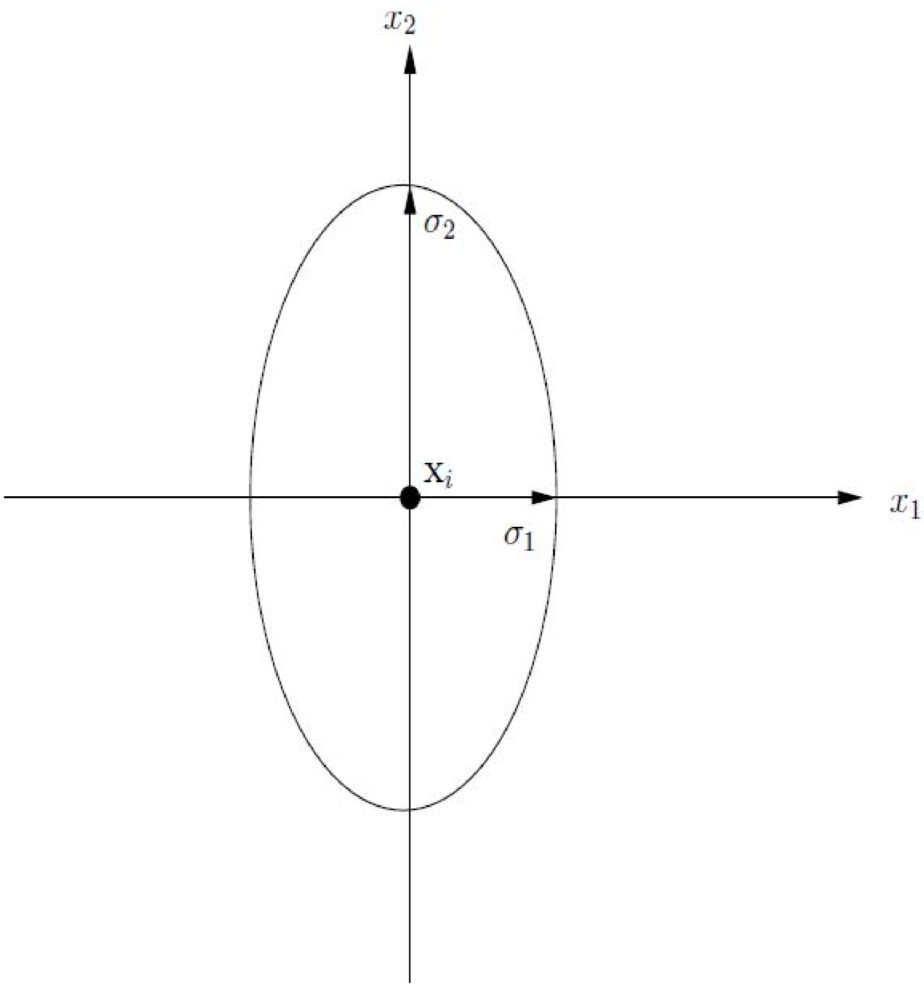
- While adjustment of the single parameter is computational fast, the approach is not flexible when the coordinates have different gradients.

- 2）each component has its own deviation parameter.

$$n_{\sigma} = n$$

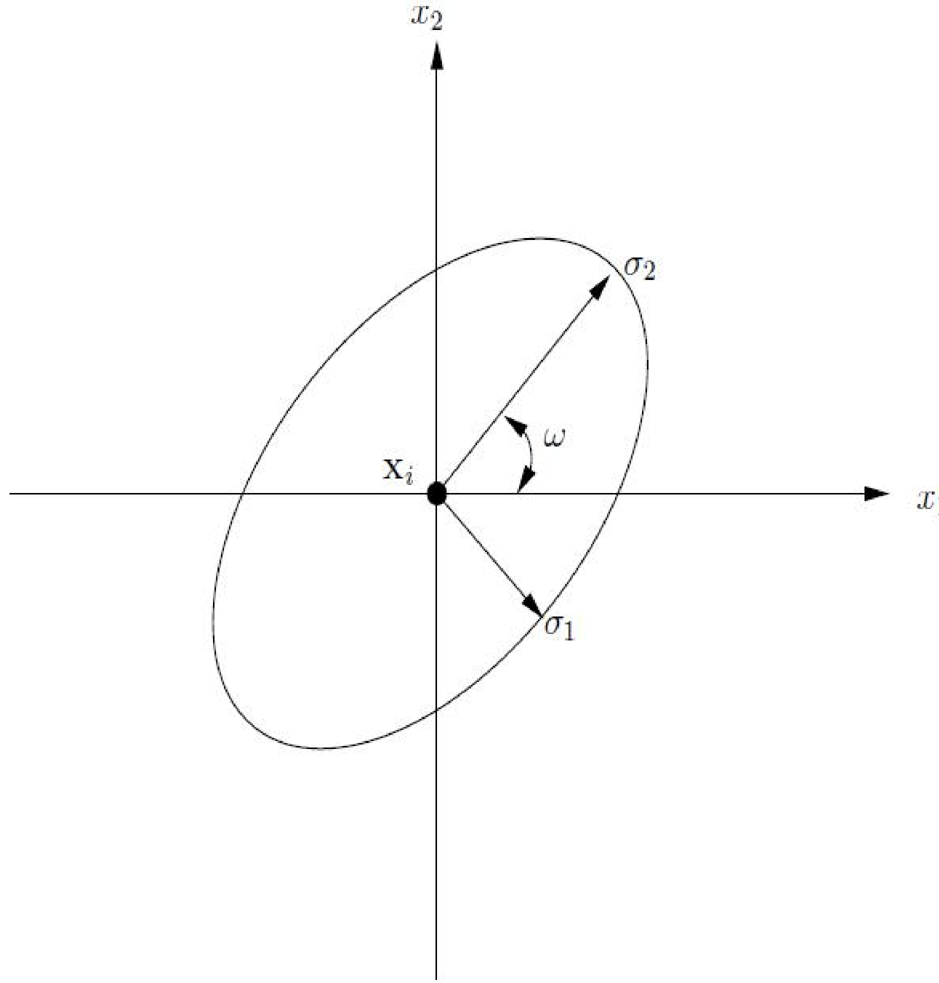
$$\sigma_{ij}^{'}(t) = \sigma_{ij}(t) e^{\tau^{'} N(0,1) + \tau N_j(0,1)}$$

$$\text{where } \tau^{'} = \frac{1}{\sqrt{2n_x}} \text{ and } \tau = \frac{1}{\sqrt{2\sqrt{n_x}}}$$



- **increased number of parameters causes a linear increase in computational complexity, but the added degrees of freedom provide for better flexibility.**
- **different gradients along the coordinate axes can now be taken into consideration.**

- **3) rotational angles are used which allow better approximation of the contours of the search space.**



3. Selection

- For each generation, λ offspring are generated from μ parents and mutated.

- Two main strategies
- (μ+λ): plus strategies
- (μ,λ):comma strategies

(μ+λ)-ES

- the next generation consists of the μ best individuals selected from μ parents and λ offspring.
- elitism ensure the fittest parents survive to the next generation.

(μ,λ)-ES

- The next generation consists of the μ best individuals selected from λ offspring.
- Elitism is not used and therefore this approach exhibits a lower selective pressure than the plus one.
- Diversity is therefore larger which results in better exploration.
- $\lambda >= \mu$ ($\lambda/\mu = 7$)
- more applicable

4. Crossover

- Crossover operators differ in the number of parents used to produce a single offspring and in the way that the genetic material and strategy parameters of the parents are combined to form the offspring.
- Local crossover: one offspring is generaed from two randomly selected parents.
- Glabal crossover: more than two parents.

- **Discrete recombination**: for each component of the genotype or strategy parameter vectors, the corresponding component of a randomly selected parent is used. （目标变量）
- **Intermediate recombination**: the offspring is a weighted average of the parents. （策略参数）

小结

- **(1+1):** 只是单个个体在演化，本质上是局部搜索。
- **(μ+1):** 交叉变异产生一个后代，如果比种群最差个体好，就替换。
- **(μ+λ):** 在原有的**μ**个体和新生成的**λ**个体中选择。
- **(μ,λ):** 从新生成的**λ**个体中选择。
- 变异是主要的遗传算子
- 交叉是辅助