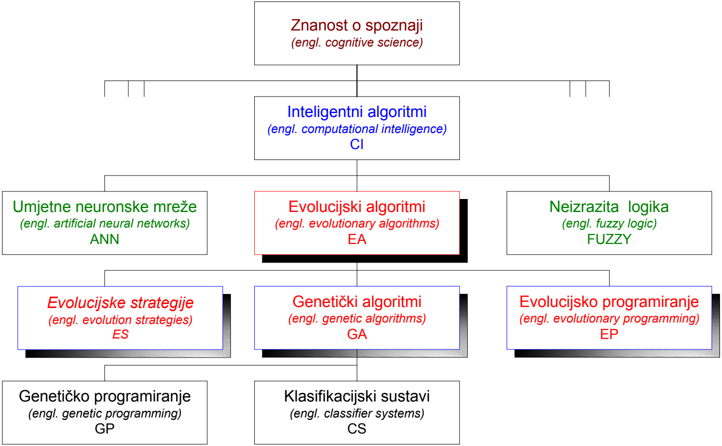
INTELLIGENT CONTROL SYSTEMS

Mato Baotić
Jadranko Matuško

EVOLUTIONARY ALGORITHMS

October, 1st, 2014

EVOLUTIONARY ALGORITHMS



Evolutionary algorithms

- Searching and optimization methods based/motivated by a biological and physical evolution.
- These algorithms reflect an engineer's approach to problem solving, i.e. starting from some initial solution, we are trying to find an improved, but not necessarily a theoretically optimal solution.

Classification of searching and optimization methods

- Principal goal is to find (parametric) vector $x \in S \subset \mathbb{R}^n$ for which the objective function $F(x)$ has a supremum, where S is a searching space.
- If we denote the optimal vector as $x^\star$ we can write:

$$x^\star = arg\,sup\,F(x), \;\; x \in S. \tag{1}$$

- Generally the supremum of a continuous objective function lies either in the interior of the set S or at its boundary.
- This is illustrated by the example shown in Fig.1

EXAMPLE 1

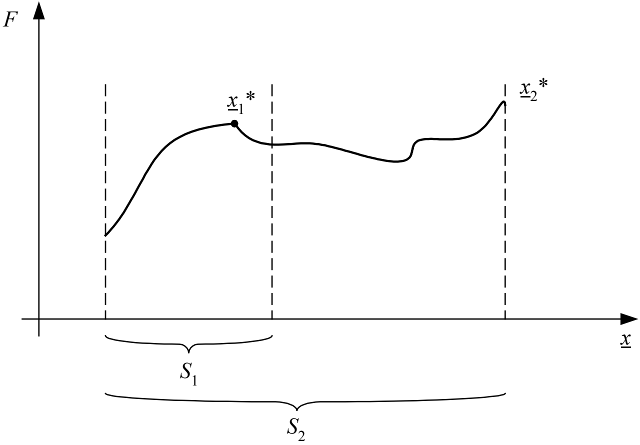


Figure : Continuous objective function $F(x)$ with different searching spaces S_i.

Classification of searching and optimization methods

- In the previous examples we can write:

$$\sup_{x \in S} F(x) = \max_{x \in S} F(x) \tag{2}$$

Searching and optimization methods can be generally classified into following three groups:

- Analytical methods
- Specification methods;
- Stochastic methods.

ANALYTICAL METHODS

- Analytical methods are based on necessary and/or sufficient conditions that a continuos function $F(x)$ need to satisfy at $x = x^\star$
- Analytical methods can be divided into:
  - direct methods,
  - indirect methods.

INDIRECT ANALYTICAL METHOD

- We analyze the gradient function $\nabla F(x)$ to obtain singular points of the function $F(x)$ that lies inside the searching space (local maximum, local minimum ili saddle point).
- Singular points are obtained as a solution to the nonlinear system of equations:

$$\nabla F(x) = 0. \tag{3}$$

- Among the candidates we chose the global optimum $x^{\star}$.

DIRECT GRADIENT METHODS

- We calculate, analytically or numerically, a local gradient function $\nabla F(x_k)$ which is then used in a gradient based searching procedure:

$$x_{k+1} = x_k + \eta s_k(x_k), \quad k = 0, 1, 2, \ldots, \tag{4}$$

where $\eta > 0$ is a variable step size ($\eta = \eta_k$), and $s_k(x_k)$ is a search direction in $\mathbb{R}^n$.

DIRECT GRADIENT METHODS

- Among many variants of the gradient based methods, two of them are predominantly used:
  - steepest ascend method.
  - conjugate gradient method.
- Steepest ascend method typically start with the initial value x_0 (see (4)) and the associate search direction: $s_k(x_k) = \nabla F(x_k)$.
- Using this iterative procedure leads to finding a stationary point in the vicinity of x_0. This method converges relatively slowly and tends to become unstable near the point of maximum.

Conjugate gradient method, CGM

- Conjugate gradient method provides an improvement of the steepest ascend method in terms of its convergence and stability.
- The searching directions, from Eq. (4), are calculated as:

$$\begin{aligned} s_0 &= \nabla F(x_0) \\ s_k &= \nabla F(x_k) - \beta_k s_{k-1}, \quad k = 1, 2, 3, \ldots, \end{aligned} \tag{5}$$

with scaling factor β_k defined as:

$$\beta_k = \frac{\nabla F^T(x_k) \nabla F(x_k)}{\nabla F^T(x_{k-1}) \nabla F(x_{k-1})} \tag{6}$$

- The drawback of the CGM is that the objective function $F(x)$ has to be in an explicit form and at least once differentiable.

Conjugate gradient method, CGM

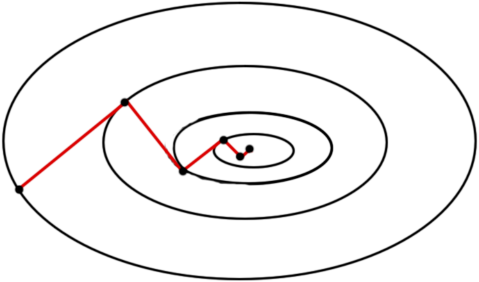


Figure : Illustration of conjugate gradient nethod

Non-gradient direct methods

Hook-Jeeves method (pattern search method)

- The idea of this method is to cyclically sample the parameter space according to some pattern, check the local ascend direction and continues with searching in that direction.
- If we assume that the component x_1 of the vector x increases for step λ, we obtain $F^{+}(x_1+\lambda, x_2, ..., x_n)$. If $F^{+}(x_1+\lambda, x_2, ..., x_n) > F(x)$ then the searching direction is successful and we set $x^1 = (x_1+\lambda, x_2, x_3, ..., x_n)$.
- Otherwise we check $F^{-} = (x_1-\lambda, x_2, ...,) > F(x)$. If it holds then we set $x^1 = (x_1-\lambda, x_2, x_3, ..., x_n)$
- If none of above conditions is met then we set $x^1 = x$
- Once x^1 is set we proceed to the next component of the vector x with the analogous procedure.

Hook-Jeves Algorithm

1. Choose an initial value x_0, step size λ and coefficient $0 < \alpha < 1$
2. Check vector x with step size λ and comparing value $F(x)$. If no improvement is achieved. i.e. $x = x^n$, then we set $\lambda := \alpha\lambda$ and go back to the step 2. Otherwise proceed with the step 3.
3. Check the parameter value $x^n + (x^n - x)$ with (old) λ and comparing value $F(x)$. If $x^n + (x^n - x) = [x^n + (x^n - x)]^n$ then we set $x := x^n$, $\lambda = \alpha\lambda$ and go back to the step 2. Otherwise proceed with the step 3.

Maximum is found if $x = x^n$ in step 2 and λ becomes sufficiently small.

SPECIFICATION METHODS

- The basic idea of this method is to calculate the objective function $F(x)$ for all $x \in S$ and then choose the one with the largest associated objective function $F(x)$ as the optimal one.
- This approach is appropriate only for small scale problems, since increasing the problems dimensionality leads to a combinatorial explosion of the problem complexity.
- The main advantage of the specification method, comparing to the analytical ones, stems from the fact that it doesn't require the knowledge of the objective function. The objective function can be given in the form of a so-called oracle.