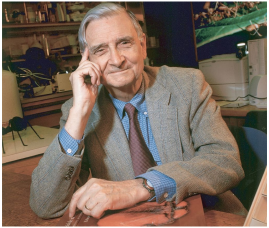
Particle Swarm Optimization

Jayanti Prasad

Inter-University Centre for Astronomy & Astrophysics
Pune, India (411007)

February 20, 2015

Swarms in Nature



E. O. Wilson (1929-)

"In theory at least, individual members of the school can profit from the discoveries and previous experience of all other members of the school during the search for food. This advantage can become decisive, outweighing the disadvantages of competition for food items, whenever the resource is unpedictably distributed in patches."

What is Optimization ?

- Find a vector $X_0 \in \mathbb{R}^n$ for which:

$$f(X_0) \leq f(X), X \in [a, b] \quad \text{and} \quad a, b \in \mathbb{R}^n \tag{1}$$

with or without constraints of the following kind:

$$\begin{aligned} g_j(X) & \leq 0 \quad \text{for} \quad j = 1, .., M \quad \text{(inequality constrains)} \\ h_j(X) & = 0 \quad \text{for} \quad j = 1, .., p \quad \text{(equality constrains)} \end{aligned} \tag{2}$$

- The function $f(X)$ is called an "optimization function" or the "cost function".

Ordinary χ^2

- Let $d \in \mathbb{R}^M$ is our data vector, $x \in \mathbb{R}^n$ is our parameter vector and $A \in \mathbb{R}^{M\times N}$ is a transformation kernel and they are related as :

$$d = Ax + n, \tag{3}$$

where $n \in \mathbb{R}^M$ is the noise vector.

- The above equation can be solved by finding the minimum of the following cost function:

$$f(x) = (d - Ax)^T C_n^{-1}(d - Ax), \tag{4}$$

where $C_n =< nn^T >$ is the noise covarainnce matrix.

- For the case when noise is un-correlation we can identify $f(X)$ with χ^2:

$$f(x) = \chi^2 = \frac{1}{M}\sum_{i=1}^{M}\frac{1}{\sigma_i^2}\left[d_i - \sum_{j=1}^{N} A_{ij}x_j\right]^2 \tag{5}$$

Common Optimization strategies

- Deterministic:
  - Grid based search : Computational cost grows as exponentially with the dimension N of the search.
  - Gradient based search : Newton-Raphson.
  - conjugate Gradient : when A is symmetric, positive definite.
  - Downhill-Simplex Methods.
  - Powell's method.
- Stochastic
  - Artificial Neural Networks (ANN).
  - Genetic Algorithm (GA).
  - Ant Colony Optimization (ACO), Particle Swarm Optimization (PSO), Bat Algorithm (BA), Firefly Algorithm (FA), Bacteria Foraging Optimization (BFO) etc.
  - Continuous Opinion Dynamics Optimizer (CODO) [Rishemjit Kaur & Kapur (2013)].

Particle Swarm Optimization (PSO)

- PSO was presented by Kennedy & Eberhart in [Kennedy & Eberhart (1995)] inspired by the work of E. O. Wilson [Wilson (1975)].
- There are many good books [Kennedy & Eberhart (2001); Engelbrecht (2002); Jun Sun & Wu (2012)] and review articles available about PSO and a dedicated website [http://www.swarmintelligence.org/].
- PSO is a stochastic method for optimization of continuous nonlinear functions motivated by the movement of organisms in a bird flock or fish school.
- In PSO, a "team of particles" (computational agents) search for the global maximum/minimum of a non-linear function (fitness/coast function) exhibiting random motions.
- Random motion of particles is governed by their "personal experience" (`Pbest`) and "social learning" (`Gbest`).

PSO Applications

- Applications of PSO in Engineering problems have been common [Boeringer & Werne (2004); Robinson & Rahmat-Samii (2004)], however, recently it has been used for Binary Inspiral parameter estimation in gravitational wave studies [Wang & Mohanty (2010)], Gravitational Lens Modeling [Rogers & Fiege (2011)], Semi-Analytic Models of Galaxy Formation [Ruiz et al. (2015)], Cosmological Parameter Estimation [Prasad & Souradeep (2012); Prasad (2014)], Pulsar Timing Arrays [Taylor et al. (2012); Lentati et al. (2013)].
- There are many variations of PSO some of them are Adaptive PSO (APSO), landscape adaptive PSO (LAPSO), Distribution Vector PSO (DVPSO), Constriction Factor PSO (CFPSO), Inertia Weight PSO (IWPSO) [Yisu (2008)].
- At present we will stick to the original version of PSO given in [Kennedy & Eberhart (1995)].

PSO Variables

1	$X^i(t)$	Position of i^{th} particle at "time" t (discrete)
2	$V^i(t)$	"Velocity" of i^{th} particle at "time" t
3	P^i	Personal best "Pbest " of i^{th} particle
4	G	Global best or "Gbest "
5	$[X_{\min}, X_{\max}]$	Search Range
6	$V_{\max}$	Maximum Velocity

- Note that $X, V, P, G \in \mathbb{R}^N$ where N is the dimensionality of the search space.
- If $f(X^i(t)) = f^i(t)$ then we can define P^i and G in the following way:
  - Pbest : $P^i = X^i(s)$ such that $f^i(s) \leq f^i(t)$ for $0 < s < t$.
  - Gbest : $G = P^k$ such $f(P^k) \leq f(P^i)$.

Dynamics of PSO particles

- Equation of Motion :

$$X^i(t+1) = X^i(t) + V^i(t+1) \tag{6}$$

- Velocity :

$$V^i(t+1) = wV^i(t) + c_1\xi_1[X^i_{\texttt{Pbest}} \ - X^i(t)] + c^2\xi_2[X_{\texttt{Gbest}} \ X^i(t)] \tag{7}$$

where w (inertia weight), c_1, c_2 (acceleration coefficients) are PSO design parameters. and ξ_1 and ξ_2 are uniform random numbers (URN) in range $[0-1]$.

Other Considerations

1. Fitness function $f(X)$ and search space $(X_{\min}, X_{\max})$.
2. Initial conditions:

$$
\begin{aligned}
X^i(t=1) &= X_{\min} + \xi_3 (X_{\max} - X_{\min}) \\
V^i(t=1) &= \xi_4 V_{\max}
\end{aligned}
\tag{8}
$$

   where

$$V_{max} = c_v (X_{\max} - X_{\min}) \tag{9}$$

   and ξ_3 and ξ_4 are URN in range $[0, 1]$.
3. Boundary conditions (reflecting boundary):

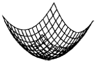
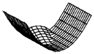
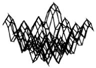
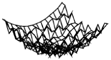
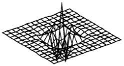
$$V^i(t) = -V^i(t) \quad \text{and} \quad X^i(t) = \begin{cases} X_{\max}, & \text{if } X^i(t) > X_{\max} \\ X_{\min}, & \text{if } X^i(t) < X_{\min} \end{cases} \tag{10}$$

4. Stopping criteria:

$$|f(X_{\mathtt{Gbest}}\ (t+1) - f(X_{\mathtt{Gbest}}\ (t))| < \epsilon \quad \text{for} \quad t = 1, n_{\text{stop}} \tag{11}$$

Some test problems (functions)

Table 1
Optimization test functions

Name	Formula	Dim. n	Range $[x_{\min}, x_{\max}]$	Optimal f	Goal for f	Sketch in 2D
Sphere	$f_0(\vec{x}) = \sum_{i=1}^{n} x_i^2$	30	$[-100, 100]^n$	0	0.01	
Rosenbrock	$f_1(\vec{x}) = \sum_{i=1}^{n-1} \left(100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2\right)$	30	$[-30, 30]^n$	0	100	
Rastrigin	$f_2(\vec{x}) = \sum_{i=1}^{n} \left(x_i^2 - 10\cos(2\pi x_i) + 10\right)$	30	$[-5.12, 5.12]^n$	0	100	
Griewank	$f_3(\vec{x}) = \frac{1}{4000} \sum_{i=1}^{n} x_i^2 - \prod_{i=1}^{n} \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$	30	$[-600, 600]^n$	0	0.1	
Schaffer's f6	$f_6(\vec{x}) = 0.5 - \frac{(\sin\sqrt{x_1^2+x_2^2})^2 - 0.5}{(1+0.001(x_1^2+x_2^2))^2}$	2	$[-100, 100]^2$	0	10^{-5}	

References

Boeringer, D. W., & Werne, D. H. 2004, IEEE Transc. on Antennas and propagation, 52, 771

Engelbrecht, A. P. 2002, Computational Intelligence, An Introduction (John Wiley & Son)

Jun Sun, C.-H. L., & Wu, X.-J. 2012, Particle Swarm Optimization:Classical and Quantum Perspectives (CRC)

Kennedy, J., & Eberhart, R. C. 1995, IEEE International Conference on Neural Networks, 4, 1992

—. 2001, Swarm Intelligence (Morgan Kufmann)

Lentati, L., Alexander, P., Hobson, M. P., Taylor, S., Gair, J., Balan, S. T., & van Haasteren, R. 2013, Phys. Rev. D , 87, 104021

Prasad, J. 2014, ArXiv e-prints

Prasad, J., & Souradeep, T. 2012, Phys. Rev. D , 85, 123008

Rishemjit Kaur, Ritesh Kumar, A. P. B., & Kapur, P. 2013, Scientific Reports, 3, 3008

Robinson, J., & Rahmat-Samii, Y. 2004, IEEE Trans. on Ant. and Prop., 727, 397

Rogers, A., & Fiege, J. D. 2011, Astrophys. J. , 727, 80

Ruiz, A. N., et al. 2015, Astrophys. J. , 801, 139

Taylor, S. R., Gair, J. R., & Lentati, L. 2012, ArXiv e-prints

Wang, Y., & Mohanty, S. D. 2010, Phys. Rev. D , 81, 063002

Wilson, E. O. 1975, Sociobiology (Harvard University Press)

Yisu, J. 2008, Applied Soft Computing, 8, 295