2016 International Symposium on Nonlinear Theory and Its Applications,
NOLTA2016, Yugawara, Japan, November 27th-30th, 2016

On a rotationally invariant of PSO

Yosuke Hariya, Takuya Shindo and Kenya Jin'no

Nippon Institute of Technology
Saitama 345-0826, Japan
Email: e1112350@estu.nit.ac.jp, t-shindo@nit.ac.jp, jinno@nit.ac.jp

Abstract—Particle swarm optimization (PSO) is a stochastic population-based algorithm that is designed for real-parameter optimization problems. PSO is a simple and powerful algorithm. However, the performance of PSO is degraded in the case of non-separable problems. In this article, we discuss rotationally invariant PSOs and its performance.

1. Intriduction

Optimization problem is an important issue in various fields. Single-objective continuous optimization problem is a problem of finding a real-valued vector that minimizes an objective function $f : \mathbb{R}^n \rightarrow \mathbb{R}$. In particular, an optimization problem that the analytic form is not known is called as black-box optimization problem. Many stochastic population-based algorithms have been proposed to solve the black-box optimization problem.

Particle swarm optimization (PSO)[1][2] is one of the stochastic population-based algorithms that is based on swarm intelligence. PSO is simple and powerful algorithm. However, its search performance is depended on the coordinate system of the objective function [3][4][5]. Such property is referred to as *rotation variance*, and this property is related to separability of the objective function [4]. In the black-box optimization problem, the algorithm whose search performance is affected by the property of the objective function is undesirable. To overcome this problem, several rotationally invariant PSO have been proposed [6][7][8][9][10][11]. In this article, we discuss the rotation variance of PSO and we introduce the typical rotationally invariant PSO.

Separability: If the function can be rewritten as $f(\boldsymbol{x}) = \sum_{i=1}^{n} f_i(x_i)$, the function f is said to be separable [4]. Namely, the function f is corresponds to each dimension is independent.

In general, if the number of dimensions increases linearly, the volume of the search space increases exponentially. However, since the separable function can be rewritten as the sum of the 1-dimensional functions, the complexity of the problem increases linearly. Thus, the separable functions is said to be an easier problem than the non-separable function.

In almost cases, a separable function can be transformed into the non-separable function by rotation of the coordinate system. From this fact, the performance of the algorithm is depended on the separability of the objective function is referred to as rotation variance.

2. Particle Swarm Optimization

PSO has been proposed by Kenedy and Clerc [1][2]. Each particle contains theree vectors: the position $\boldsymbol{x}_i^t$, the velocity $\boldsymbol{v}_i^t$ and the personal best position $\boldsymbol{p}_i^t$, where i denotes the number of particles and t denotes the iterations. The particle swarm has global best position $\boldsymbol{g}^t$, it is the best of personal best position. In each iteration, the position and the velocity are updated by the following equations.

$$\boldsymbol{v}_i^{t+1} = \omega \boldsymbol{v}_i^t + \phi_1 \boldsymbol{R}_1(\boldsymbol{p}_i^t - \boldsymbol{x}_i^t) + \phi_2 \boldsymbol{R}_2(\boldsymbol{g}^t - \boldsymbol{x}) \quad (1)$$

$$\boldsymbol{x}_i^{t+1} = \boldsymbol{x}^t + \boldsymbol{v}^{t+1} \quad (2)$$

where ω denotes an inertia weight coefficient and ϕ_1, ϕ_2 are acceleration coefficients. $\boldsymbol{R}_1$ and $\boldsymbol{R}_2$ are randomly generated diagonal matrices. Each element of these matrices is a uniform random number in interval $[0, 1]$.

Rotation variance of PSO: The reason of the rotation variance of PSO is the search direction biase [3]. Figure 1 shows the histogram of the search direction and the trajectory of particles on 2-dimensional sphere function. The angle of the velocity vector means the search direction. Sphere function is isotropic. However, the search direction is biased in parallel to the coordinate axes.

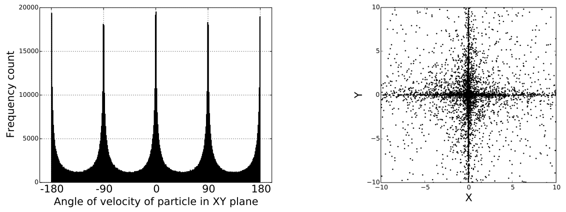
In order to clarify the reason generating the bias of the search direction, we consider the simple velocity update rule that the reference position is one and without the inertia weight coefficient, as $\boldsymbol{v}_i^{t+1} = \phi \boldsymbol{R}(\boldsymbol{b}_i^t - \boldsymbol{x}_i^t)$. Figure 2 shows the histogram of the 2-dimensional velocity vector when the angle of the reference position is fixed. Since the sign is not reversed. the distribution of the angle of the velocity vector is biased when the reference position vector is close to the coordinate axis.

3. Rotationally invariant PSOs

Several rotationally invariant PSOs have been proposed. In this section, we introduce the typical rotationally invariant PSOs.

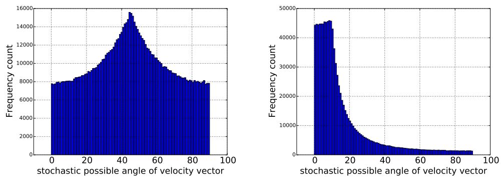
Linear PSO (LPSO) is the most simple rotationally invariant PSO [6][7]. The velocity update rule of LPSO is described by the following equation,

$$\boldsymbol{v}_i^{t+1} = \omega \boldsymbol{v}_i^t + \phi_1 r_1(\boldsymbol{p}_i^t - \boldsymbol{x}_i^t) + \phi_2 r_2(\boldsymbol{g}^t - \boldsymbol{x}), \quad (3)$$



(a) The histogram of angle of velocity vector (b) The trajectory of particles

Figure 1: The search direction and the trajectory of particles on 2-dimensional sphere function



(a) $\theta = 45°$ (b) $\theta = 10°$

Figure 2: PSO: the histogram of the 2-dimensional velocity vector when the angle to the reference position θ is fixed.

whre r_1 and r_2 are uniform random numbers in the interval $[0, 1]$. In LPSO, the search direction is always pointed to the reference position. Since the random number is a scalar, the diversity of LPSO is poor.

Rotation PSO (RPSO) was proposed by Wilke et al [6][7]. In RPSO, a rotation matrix is multiplied to the velocity vector.

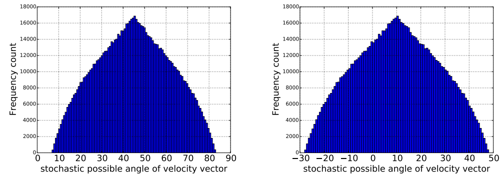
$$\boldsymbol{v}_i^{t+1} = \omega \boldsymbol{v}_i^t + \phi_1 \boldsymbol{M}_1(\boldsymbol{p}_i^t - \boldsymbol{x}_i^t) + \phi_2 \boldsymbol{M}_2(\boldsymbol{g}^t - \boldsymbol{x}) \quad (4)$$

$$\boldsymbol{M} = \boldsymbol{I} + \frac{\alpha\pi}{180}\left(\boldsymbol{E} - \boldsymbol{E}^\top\right) \quad (5)$$

where $\boldsymbol{M}$ is a random rotation matrix with the rotation angle α, and $\boldsymbol{E}$ is randomly generated matrix whose elements are uniform random numbers in the interval $[-0.5, 0.5]$. The dynamics of RPSO is closest to the dynamics of PSO. However, the calculation amount of the generating of the random rotation matrix is $O(n^2)$. Also, the calculation amount of the product of the velocity vetor and the rotation matrix is $O(n^2)$.

We consider the simple velocity update rule of RPSO as $\boldsymbol{v}_i^{t+1} = \phi \boldsymbol{M}(\boldsymbol{b}_i^t - \boldsymbol{x}_i^t)$. Figure 3 shows the histogram of the 2-dimensional velocity vector when the angle to the reference position is fixed. In RPSO, the biased histogram of the angle of the velocity vector is not obserbed.

Bonyadi et al. proposed rotationally invariant PSO using the exact rotation matrix [8]. In this article, we call this method as **Modified Rotation PSO** (MRPSO). In general, the calculation amount of the generating of the exact



(a) $\theta = 45°$ (b) $\theta = 10°$

Figure 3: RPSO: the histogram of the 2-dimensional velocity vector when the angle to the reference position θ is fixed.

rotation matrix is $O(n^5)$. However, almost all elements of the rotation matrix are zero. By using the advantage of this fact, the calculation amount can be reduced to $O(n^2)$. Based on this exact rotation matrix's advantage, Bonyadi et al. proposed the method changes the rotation angle adaptively in the search process.

Clerc proposed **Standard PSO2011** (SPSO)[9] that realizes the rotation invariance by changing the shape of the seach area. The velocity update rule of SPSO is described by the following equations.

$$\boldsymbol{v}_i^{t+1} = \omega \boldsymbol{v}_i^t + H(\boldsymbol{C}_i^t, \|\boldsymbol{C}_i^t - \boldsymbol{x}_i^t\|) - \boldsymbol{x}_i^t \quad (6)$$

$$\boldsymbol{C}_i^t = \boldsymbol{x}_i^t + \frac{\phi_1(\boldsymbol{p}_i^t - \boldsymbol{x}_i^t) + \phi_2(\boldsymbol{g}^t - \boldsymbol{x}_i^t)}{3} \quad (7)$$

$H(\boldsymbol{a}, b)$ is a hypersphere function with the center $\boldsymbol{a}$ and the radius b. Since the shape of the search area of SPSO is spherical, the biased search direction is not observed.

Locally convergence rotationally invariant PSO (LcRiPSO) was proposed by Bonyadi et al [10]. LcRiPSO is the method combining the perturbed PSO [12] and LPSO. The random number of LcRiPSO is scalar as well as LPSO. However, since adding a normal random number to the reference position, the diversity of LcRiPSO is richer than LPSO.

$$\begin{aligned}\boldsymbol{v}_i^{t+1} = \omega \boldsymbol{v}_i^t &+ \phi_1 r_1 \left(\mathcal{N}(\boldsymbol{p}_i^t, (\sigma_1^t)^2 \boldsymbol{I}) - \boldsymbol{x}_i^t\right) \\ &+ \phi_2 r_2 \left(\mathcal{N}(\boldsymbol{g}^t, (\sigma_2^t)^2 \boldsymbol{I}) - \boldsymbol{x}_i^t\right) \quad (8)\end{aligned}$$

Bonyadi et al. proposed the method changes the variance σ_1^t, σ_2^t adaptively in the search process.

Norm Linked PSO (NLPSO) was proposed by us [11]. In NLPSO, the information of direction to the reference position is the sign only that is given by the sign function. Thus, the distribution of the angle of the velocity vector is not biased when the reference position vector is close to the coordinate axis.

$$\begin{aligned}\boldsymbol{v}_i^{t+1} = \omega \boldsymbol{v}_i^t &+ \phi_1 \boldsymbol{R}_1 \|\boldsymbol{p}_i^t - \boldsymbol{x}_i^t\|_2 \mathbf{sing}(\boldsymbol{p}_i^t - \boldsymbol{x}_i^t) \\ &+ \phi_2 \boldsymbol{R}_2 \|\boldsymbol{g}^t - \boldsymbol{x}_i^t\|_2 \mathbf{sing}(\boldsymbol{g}^t - \boldsymbol{x}_i^t) \quad (9)\end{aligned}$$

Figure 4 shows the performance of PSO and rotationally invariant PSOs on 2-dimensional ellipse function [3]. The

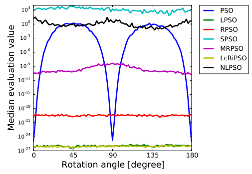
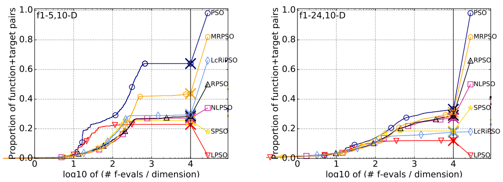


Figure 4: The rotation variance of the PSOs on ellipse function.



(a) Separable functions. (b) All functions.

Figure 5: The results of BBOB.

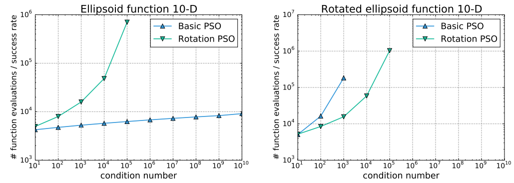
horizontal axis denotes the rotated angle of the coordinate system and the vertical axis denotes the median value in 25 trials. In each trial, the number of evaluations is 10000. From the result, the performance of PSO is degraded when the coordinate system is rotated. On the other hand, rotationally invariant PSOs are not depended on the angle of the coordinate system. Since these experiments use only 2-dimensional function, it is not possible to discuss the performance of these PSOs from the experimental results.

4. The performance of rotationally invariant PSOs

We investigate the performance of these PSOs by BBOB [13]. In order to evaluate the performane, we use the empirical cumulative distribution functions that are generated by COCO [14]. We set the recomended parameters. Figure 5 shows the results. In Fig 5, the horizontal axis denotes the log of the number of evaluation divided by the number of dimensions, and the vertical denotes the success rate in instances of each function.

Rotationally invariant PSO is not the method to improve performance but the method which resolved rotation variance. Thus, RPSO, MRPSO and NLPSO shows the similar performance as PSO. However, the results of all functions indicate that the performance of PSO is the best. Namely, resolving the search direction bias is the factor of deteriorating the performance. Because, in separable function, the biased search of PSO is advantageous. Thus, in Fig. 5a, the success rate of PSO is higher than rotationally invariant PSOs.

In a particularly high-conditined and separable function, the biased search is advantageous. Figure 6 shows the performance of PSO and RPSO on the separable and non-separable convex fucntions [5]. In these experiments, the



(a) separable (b) non-separable

Figure 6: Performance of PSO and RPSO on the separable and non-separable convex function.

maximum number of evaluations is 10^6, and if the evaluation value reaches 10^{-4} untill the maximum number of evaluations, this trial is regarded as a success. From the results, if the condition number is increased, the performance of RPSO is deteriorated. On the other hand, In the case of separable function, if the condition number is increased, the performance of PSO is hardly changed. However, in the case of non-separable function, if the condition number is increased, the performance of PSO is rapidly deteriorated.

5. A Novel PSO for high-conditioned and non-separable functions

In order to solve the high-conditioned and non-sepable functions, we proposed new rotationally invariant PSO, it is described by the folloing equations [15].

$$\boldsymbol{v}_i^{t+1} = \omega \boldsymbol{v}_i^t + \phi_1 r_1 \left(\tilde{\boldsymbol{p}}_i^t - \boldsymbol{x}_i^t\right) + \phi_2 r_2 \left(\tilde{\boldsymbol{g}}^t - \boldsymbol{x}_i^t\right) \quad (10)$$

$$\tilde{\boldsymbol{p}}_i^t = \boldsymbol{p}_i^t + c_d \left(\boldsymbol{p}_{j_1}^t - \boldsymbol{p}_{j_2}^t\right) \quad (11)$$

$$\tilde{\boldsymbol{g}}^t = \boldsymbol{g}^t + c_d \left(\boldsymbol{p}_{j_3}^t - \boldsymbol{p}_{j_4}^t\right) \quad (12)$$

c_d is a constant number in the interval $[0, 1]$. j_1, j_2, j_3 and j_4 are random particle numbers, where $i \neq j_1 \neq j_2 \neq j_3 \neq j_4$. From the central limit theorem, the difference vector of the personal best follows the normal distribution. Thus, the proposed method is similar to LcRiPSO. However, the covariance matrix is different. In LcRiPSO, the covariance matrix of the perturbation becomes a diagonal matrix. On the other hand, the covariance matrix of the difference vector of the personal best is estimated as the inverse Hessian matrix of the objective function [15]. The estimation of the inverse Hessian matrix of the objective function is essential to solve the high-conditioned and non-separable functions [16]. Also, in order to improve the local search ability, we applied the selection mechanisim [15].

To confirm the performance of proposed method, we carry out experiments. Table 1 shows the test functions. For each function, 25 trials are conducted. The parameter settings are refer to [15]. The time evolution of the best evaluation value in each trial is shown in Fig. 7.

From the results, the performance of the proposed method is better than the conventional PSO in the high-conditioned and non-separable functions.

Table 1: Test functions, where $\boldsymbol{y} := \boldsymbol{Ax}$ and $\boldsymbol{A}$ is a rotation matrix.

Name	Function
Ellipsoid	$f_{\text{Ellipsoid}}(\boldsymbol{x}) = \sum_{i=1}^{n} 10^{6\frac{i-1}{n-1}} y_i^2$
Rosenbrock	$f_{\text{Rosenbrock}}(\boldsymbol{x}) = \sum_{i=1}^{n-1} 100(y_i^2 - y_{i+1})^2 + (y_i - 1)^2$
Schaffer	$f_{\text{Schaffer}}(\boldsymbol{x}) = \sum_{i=1}^{n-1} (y_i^2 + y_{i+1}^2)^{0.25} \left(\sin^2\left(50(y_i^2 + y_{i+1}^2)^{0.1}\right) + 1\right)$

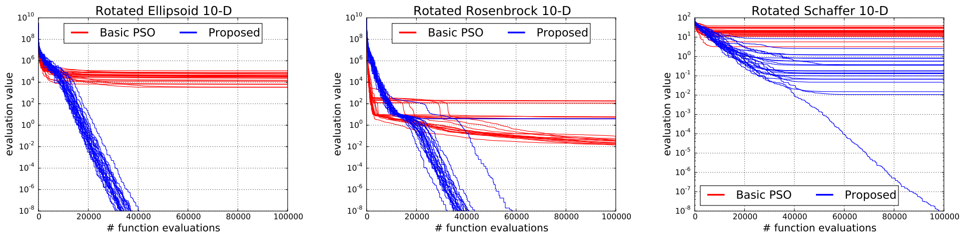


Figure 7: Time evolution of the best evaluation value for 10-dimensional test functions.

6. Conclusions

In this article, we clarify that the factor of the rotation variance of PSO. We introduced typical rotationally invariant PSOs. Futhmore, in order to evaluate the performance of these PSOs, we carried out experiments using BBOB. From the result, the general performance of PSO is better than rotationally invariant PSOs. The reason is that rotationally invariant PSOs do not solve the separable/non-separable and high-conditioned functions and PSO can solve the separable and high-conditioned functions.

To overcome this problem, we proposed new PSO for non-separable and high-conditioned functions. Also, in order to investigate the performance of proposed method, we carried out experiments. From the results, we clarified that our proposed method can solve the high-conditioned and non-separable functions.

Acknowledgment

This work was supported by JSPS KAKENHI Grant-in-Aid for Challenging Exploratory Research Number: 16K14271, Grant-in-Aid for Scientific Research(C) Number: 15K06077.

References

[1] J. Kennedy and R. Eberhart, in Proc. of IEEE ICNN, vol.4, pp.1942–1948 vol.4, 1995.

[2] Y. Shi and R. Eberhart, in Proc. of IEEE CEC, pp.69–73, 1998.

[3] W. M. Spears, D. T. Green, and D. F. Spears, Swarm Intelligence, vol.1, no.2, pp.34–57, 2010.

[4] N. Hansen et al., Applied Soft Computing Journal, vol.11, no.8, pp.5755–5769, 2011.

[5] A. Auger et al., 8th International Symposium SEA 2009, pp.3–15, 2009.

[6] D. N. Wilke et al., International Journal for Numerical Methods in Engineering, vol.70, no.8, pp.985–1008, 2007.

[7] D. N. Wilke et al., International Journal for Numerical Methods in Engineering, vol. 70, no. 8 pp.962–984, 2007.

[8] M. R. Bonyadi et al., Journal of Heuristics, pp.2–8, 2014.

[9] M. Clerc, "Standard Particle Swarm Optimisation," HAL open access archive, 2012.

[10] M. R. Bonyadi et al., Swarm Intelligence, vol. 8, no. 3, pp.159–198, 2014.

[11] Y. Hariya, T. Kurihara, T. Shindo, and K. Jin'no, "A study of robustness of PSO for non-separable evaluation functions," in Proc. NOLTA, pp.724–727, 2015.

[12] Z. Xinchao, "A perturbed particle swarm algorithm for numerical optimization," Applied Soft Computing, vol.10, no.1, pp.119–124, 2010.

[13] N. Hansen et al. , Report, pp.1–17, 2014.

[14] `http://coco.gforge.inria.fr/`

[15] Y. Hariya, T. Shindo and K. Jin'no "A Novel Particle Swarm Optimization Algorithm for Non-Separable and Ill-Conditioned Problems," IEEE SMC, accepted, 2016.

[16] N. Hansen, "The CMA evolution strategy: A tutorial," Vu le, vol.102, no.2006, pp.1–34, 2011.