Integrate Local Replacement Strategy into Differential Evolution for the Solutions of Dynamic Optimization Problems

Shih-Jie PAN and **Daim-Yuang SUN***

Department of Chemical and Materials Engineering,
National Chin-Yi University of Technology,
Correspondent e-mail: dysun@mail.ncut.edu.tw

May 3, 2013

Abstract

In this study, a modified version of differential evolution (DE) is used to solve dynamic optimization problems (DOPs). To promote the efficiency of the basic DE, the local replacement (LR) strategy is introduced to intensify the search in the neighborhood of the current best solution. To solve by the proposed method, the DOP is converted a nonlinear programming (NLP) problem via control vector parametrization (CVP). The final results from two cases of DOP show that such modification is simple and effective in the promotion of convergent rate and numerical accuracy.

Keywords: Differential Evolution, Local Replacement, Dynamic Optimization Problem, Control Vector Parametrization

1. Introduction

In general, most engineering optimization problems can be formulated as nonlinear programming (NLP) problems as follows

$$\min_{\mathbf{\Phi}} \ \mathcal{J} \tag{1}$$

subject to

$$\begin{aligned} &\mathbf{L} \leq \mathbf{\Phi} \leq \mathbf{U} \\ &\mathbf{h}(\mathbf{\Phi}) = \mathbf{0} \\ &\mathbf{g}(\mathbf{\Phi}) \leq \mathbf{0} \end{aligned} \tag{2}$$

where $\mathcal{J}$ is the objective function; $\mathbf{\Phi}^{\mathsf{T}} \equiv [\varphi_1, \cdots, \varphi_n]$ denote decision vector, $\mathbf{L}$ and $\mathbf{U} \in \mathcal{R}^n$ respectively denote the decision vector, the lower bound and the upper bound. $\mathbf{h}$ and $\mathbf{g}$ respectively denote the equality and inequality constraints on $\mathbf{\Phi}$. Thus, the above NLP problem represented is cast as finding the global optimal solution $\mathbf{\Phi}^*$ that minimizes the objective function $\mathcal{J}$ and simultaneously satisfying Eq. (2).

To solve the above problems, gradient-base methods and direct search methods are commonly used. Using the information derived from the first or second derivatives, the gradient-based methods, such as successive quadratic programming (SQP), may converge the problem very rapidly, especially when the initial guess is closed to the optimum. However, for the systems with highly nonlinear and/or multimodal nature, to acquire a global optimum might become very difficult. Based on whether the objective function is improved, the direct search methods directly update the decision vectors. However, even the direct search methods may approach the global solution, they often require very large objective function evaluations. In recent, it is very popular to use meta-heuristics algorithms, such as simulated annealing (SA), genetic algorithm (GA) and differential evolution (DE), in the solution of engineering optimization problems. Even the convergent rate is unable to be proved via strict mathematical procedure, these methods still demon-

*Author to whom all correspondence should be addressed.

rithms, such as simulated annealing (SA), genetic algorithm (GA) and differential evolution (DE), in the solution of engineering optimization problems. Even the convergent rate is unable to be proved via strict mathematical procedure, these methods still demonstrate very nice convergence.

We now integrate local replacement (LR) strategy into DE to optimize dynamic chemical systems. The results show that the strategy of intensifying the search in the neighborhood of the current best solution is effective in promoting the convergent rate of DE. In the following sequel, Section 2 presents the principle of the basic DE. Section 3 introduces the local replacement strategy. The formulation and parametrization of dynamic optimization problems (DOPs) is shown in Section 4. The numerical examples and the related discussions are provided in Section 5. The conclusion is cast in the final section, Section 6.

2. The Basic DE

Derived from the concept of evolution in biology, the basic DE [1] is proposed in 1996 to tackle difficult optimization. Without using any information from the derivative of the objective function, the method is still effective in solving non-differentiable systems. In general, four primary operations are included in the basic DE such initialization, mutation, crossover and selection. The details of the operations are shown below.

Initialization

Randomly generate $N_p(\geq\ 4)$ decision vectors from the lower and upper bounds, the j-th component in $\mathbf{\Phi}_i$ is produced by

$$\varphi_{ij}^{g} = L_j + rand() \cdot (U_j - L_j), \qquad j = 1, \cdots, n$$

to form the first generation ($g = 1$) of target vectors set

$$\widetilde{\mathcal{T}}^{g} = \left\{ \mathbf{\Phi}_1, \mathbf{\Phi}_2, \cdots, \mathbf{\Phi}_{N_p} \right\} \tag{3}$$

where $rand() \in [0, 1]$ denotes an uniform random number.

Mutation

Generate the mutation vector set

$$\widetilde{\mathcal{M}}^{g} = \left\{ \mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_{N_p} \right\} \tag{4}$$

by the three different vectors randomly chosen from $\widetilde{\mathcal{T}}^{g}$ with the following relation

$$\mathbf{v}_i = \mathbf{\Phi}_{r1} + F \cdot (\mathbf{\Phi}_{r2} - \mathbf{\Phi}_{r3}), \qquad i = 1, \cdots, N_P \tag{5}$$

where F is the *mutation factor* with the value in the range of [0,2].

Crossover

For each $\mathbf{\Phi}_i$ and $\mathbf{v}_i$, use the following relation to execute the following operation

$$u_{ij} = \begin{cases} v_{ij} & \text{if } (randb(i) \leq CR) \\ & \text{or } i = rnbr(j) \\ \varphi_{ij} & \text{otherwise} \\ & j = 1, 2, \cdots n \end{cases}$$

for the formation of the trial vector set

$$\widetilde{\mathcal{C}}^{g} = \left\{ \mathbf{c}_1, \mathbf{c}_2, \cdots, \mathbf{c}_{N_p} \right\} \tag{6}$$

where $\mathbf{c}_i \equiv [c_{i1}, \cdots, c_{in}]$ denotes the i-th trial vector, $randb(i)$ is the uniform random number ranging from 0 to 1, $rnbr(j)$ is the randomly-chosen number from the index set of $\{1, 2, ..., n\}$, CR means the crossover constant ranging from 0 to 1.

Selection

Using the following relation, either $\mathbf{x}_i$ or $\mathbf{u}_i$ having the better objective value is selected to form $\widetilde{\mathcal{T}}^{g+1}$

$$\mathbf{x}_i = \begin{cases} \mathbf{u}_i & \text{if } \mathcal{J}(\mathbf{u}_i) \leq \mathcal{J}(\mathbf{x}) \\ \\ \mathbf{x}_i & \text{otherwise} \end{cases} \tag{7}$$

The following diagram illustrated the above procedure.

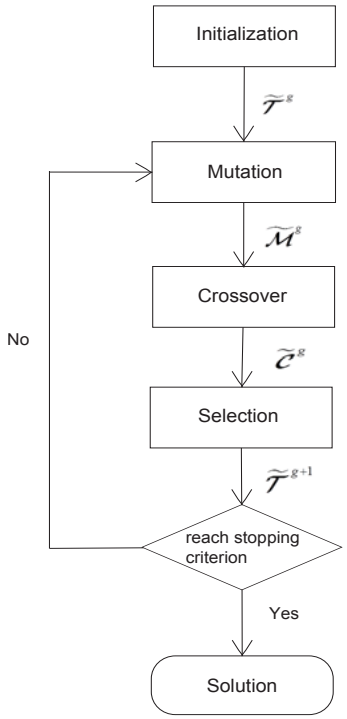

Figure 1: Schematic of DE

3. Local Replacement Strategy

In most evolutionary algorithm, to diverse the trial vectors and to intensify the search in the region the the current best solution are two primary policies for promoting the convergent efficiency. Although the basic DE may approach global optimum by the diversified trial vectors, it is well known that its convergent rate is very slow. In this study, we propose the use of integrating the local replacement strategy into the basic DE to improvement the convergence. The so-called LR strategy is to replace some worse target vectors in $\widetilde{\mathcal{T}}^{g+1}$ by the same numbers of the vectors from the neighborhood of the current best solution to intensify the refinement of the solutions [2]. The detail steps are shown as follows

Ordering

The target vectors in $\widetilde{\mathcal{T}}^{g+1}$ are sorted in ascending order for minimization problems according to the associated objective values.

Generate new trial vectors

Using the best target vector $\mathbf{\Phi}^*$ as the center, N_L new vectors $\tilde{\mathbf{\Phi}}$ are randomly chosen

$$\widetilde{\mathbf{\Phi}}_\ell = \widetilde{\mathbf{L}} + \mathbf{R}(\widetilde{\mathbf{U}} - \widetilde{\mathbf{L}}), \qquad \ell = 1, \cdots, N_L$$

from the neighborhood with the new lower and upper bounds ,$\widetilde{\mathbf{L}}$ and $\widetilde{\mathbf{U}}$

$$\widetilde{\mathbf{U}} = \widetilde{\mathbf{\Phi}}^* + \frac{\mathbf{U}-\mathbf{L}}{2} \times \frac{1}{N_p}$$
$$\widetilde{\mathbf{L}} = \widetilde{\mathbf{\Phi}}^* - \frac{\mathbf{U}-\mathbf{L}}{2} \times \frac{1}{N_p}$$

where $\mathbf{R}$ is a diagonal matrix whose diagonal entries with value ranging from 0 to 1.

Replacement

The newly generated vectors with the smaller objective values replace the last N_L target vectors (the worse solutions)

$$\mathbf{x}_{N_p-N_L+\ell} = \begin{cases} \tilde{\mathbf{x}}_\ell & \text{if } \mathcal{J}(\tilde{\mathbf{x}}_\ell) \le \mathcal{J}(\mathbf{\Phi}_{N_p-N_L+\ell}) \\ \mathbf{\Phi}_{N_p-N_L+\ell} & \text{otherwise} \end{cases}$$
$$\ell = 1, \cdots, N_L$$

4. The Formulation of DOPs

The dynamic optimization problems are in general described as

$$\min_{u(t)} \mathcal{J}(t_f) \tag{8}$$

subject to

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}, u(t), t) \tag{9}$$
$$\mathbf{x}(0) = \mathbf{x}_0 \tag{10}$$
$$h_k(\mathbf{x}, u, t) = 0 \quad (k = 1, \cdots, K) \tag{11}$$
$$L \le u(t) \le U \tag{12}$$
$$t \in [0, t_f] \tag{13}$$

where $\mathbf{x}$ denote an $(n \times 1)$ state vector. In chemical system, Eq. (9) acquired by the mass and/or energy balance are a set of ordinary differential equations with the initial conditions, Eq. (10). $u(t)$ means continuous control input from $t = 0$ to $t = t_f$ restricted by the lower bounds L and upper bounds U. Eq. (11) represents the K physical or chemical constraints.

Parametrization of DOP

In order to apply the proposed modification of DE to optimize dynamic systems. The above DOP is needed to be discretized into P sequential subproblems along the time horizon from $t = 0$ to $t = t_f$, $i.e.$

$$\min_{\mathbf{u}_i, \mathbf{t}_i, i=1,\cdots,P} \mathcal{J}(t_f) \tag{14}$$

subject to

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(x, u_i, t) \tag{15}$$
$$\mathbf{x}(0) = \mathbf{x}_0 \tag{16}$$
$$\mathbf{h}_k(\mathbf{x}(t), \mathbf{u}_i(t), t) = 0 \ \ (k = 1, \cdots, K) \tag{17}$$
$$L \leq u_i(t) \leq U \tag{18}$$
$$\forall t \in [t_i, t_{i+1}] \qquad i = 0, \cdots, P-1 \tag{19}$$

In the duration $[t_i \ t_{i+1}]$, the control input is interpolated by linear function

$$u_i(t) = \frac{u_{i+1} - u_i}{t_{i+1} - t_i}(t - t_i) + u_i \tag{20}$$

Besides $t_0 = 0$ and t_f, the $P - 2$ time grides are also considered as the decision variables. Thus, the original DOP with infinite dimension is parametrized as NLP problem with $2P$ dimensions.

5. Numerical Examples

In this study, the local replacement strategy is integrated with the basic DE and the modified DE [3]. Two typical examples are respectively solved by the basic DE, MDE, DE plus LR and MDE plus LR to compare the numerical efficiency. All these methods are programmed by MATLAB with absolute error are 1×10^{-6}. The values for F, CR and P are set as 0.3, 0.99, 5. Each example is iteratively solved 100 runs by these methods. In each run, the average objective vale (Aobj), standard deviation (SD) are recorded after 100 generations.

5.1. A batch reactor with consecutive chemical reaction

The model is described by the differential equations

$$\frac{dC_A}{dt} = -k_1 C_A^2 \tag{21}$$
$$\frac{dC_B}{dt} = k_1 C_A^2 - k_2 C_B \tag{22}$$

with the initial conditions

$$\begin{bmatrix} C_A & C_B \end{bmatrix} = \begin{bmatrix} 1 & 0 \end{bmatrix}$$

where

$$k_1 = 4000 exp(-2500/T) \tag{23}$$
$$k_2 = 620000 exp(-5000/T) \tag{24}$$

The input variable T is bounded by

$$298K \leq T \leq 398K \tag{25}$$

The objective of this problem is to determine the optimal temperature input to maximize the intermediate concentration C_B in the operating interval $0 \leq t \leq 1$,

$$\max_{T(t_f=1)} C_B$$

Logsdon and Biegler [4] report the solution of this problem to be 0.610767. Dadebo and Mcauley [5] and Babu and Angira [3] both acquire the maximal C_B to be 0.610070 and 0.610079, respectively. We solve this problem by DE plus LR and MDE plus LR and set $N_P = 180$ and $N_L = 20$. Compared with MDE+LR, DE plus LR as shown in **Table 1** seems to have better and consistent results.

Table 1: The maximum value of C_B for Case1 with $P = 5$

N_P	N_L	DE	MDE
200	0	0.610768	0.610727
190	10	0.610768	0.610756
180	20	0.610768	0.610762
150	50	0.610764	0.610747

The change of convergence for DE+LR with different N_L are shown in **Figure 2**. It is obvious that the convergent rate $N_L = 20$ is the fastest. With the increase of N_L, the convergent rate slows down.

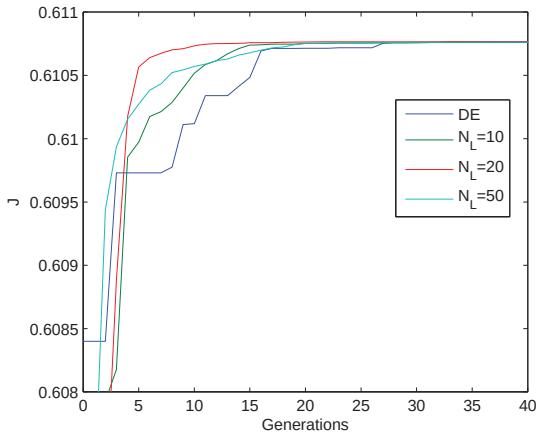

Figure 2: The comparison of convergent rate of DE and DE plus LR with different N_L

Table 2 illustrates the change of Aobj and SD with the increase of N_L. It seems to imply that for DE plus LR the average objective value and the standard deviation in 100 runs shows statistically better quality of convergence.

Table 2: Aobj and SD for Case1 with $P = 5$

N_L	DE	MDE
	Aobj (SD)	
0	0.610762 (8.38E-6)	0.610695 (3.08E-5)
10	0.610706 (5.99E-5)	0.610680 (5.22E-5)
20	0.610661 (8.12E-5)	0.610645 (7.99E-5)
50	0.610620 (1.25E-4)	0.610629 (8.65E-5)

5.2. Nonlinear Continuous Stirred Tank Reactor Problem

The model for this case is described by:

$$\frac{dx_1}{dt} = -2(2+u)(x_1+0.25) + (x_2+0.5)exp(\frac{25x_1}{x_1+2}) \quad (26)$$

$$\frac{dx_2}{dt} = 0.5 - x_2 - (x_2+0.5)exp(\frac{25x_1}{x_1+2}) \quad (27)$$

$$\frac{dx_3}{dt} = x_1^2 + x_2^2 + 0.1u^2 \quad (28)$$

where the x_1 and x_2 are the dimensionless steady-state temperature and concentration respectively. The initial conditions are

$$\mathbf{x}(0) = \begin{bmatrix} 0.09 & 0.09 & 0 \end{bmatrix} \quad (29)$$

This objective of this problem is to minimize $x_3(t_f)$ at $t_f = 0.78$,

$$\min_u x_3(t_f)$$

Notably, the control input u in this case is unrestricted. Luus [6] have proved that this problem has one local solution 0.24425 and one global solution 0.13309.

Although the best value in **Table 3**, 0.13315, comes from the basic DE, the others from DE plus LR are still very closed. On the other hand, the objective values from MDE plus LR are higher. Meanwhile, DE plus LR also shows better quality in statistics in **Table 4**.

Table 3: The best value of $x_2(t_f)$ for Case2 with $P = 5$

N_P	N_L	DE	MDE
200	0	0.13315	0.13325
190	10	0.13316	0.13321
180	20	0.13316	0.13318
150	50	0.13318	0.13317

Table 4: Aobj and SD for Case2 with $P = 5$

N_L	DE	MDE
	Aobj (SD)	
0	0.13317 (1.1E-5)	0.13350(1.4E-4)
10	0.13321 (5.0E-5)	0.13349(2.0E-4)
20	0.13325 (1.1E-4)	0.13345(2.0E-4)
50	0.13344 (3.2E-4)	0.13341(2.1E-4)

Figure 3 illustrate an important property that DE with local replacement have better convergent rate than the basic DE.

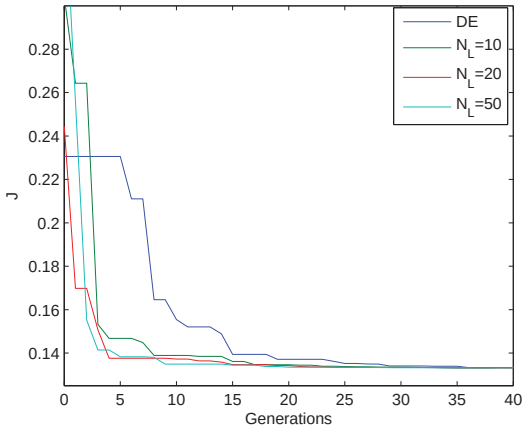

Figure 3: Convergence curves of DE for Case2

6. Conclusion

In this study, the basic DE and the modified DE are respectively used to local replacement strategy for the solution of dynamic optimization problems. To solve by these methods, the DOPs with infinite dimension are converted into NLP problems with finite dimensions via control vector parametrization. The results show that DE plus LR has better convergence than the basic DE and MDE plus LR. Because the choice of N_L might change with different problems, the setting of N_L is required to further investigate.

7. References

1. R. Storn and K. Price, "Differential evolution-a simple and efficient heuristic for global optimization over continuous spaces," *J. Global Optim.*, Vol. 11, No. 4, pp. 341-359, 1997

2. M. Y. Lai. "An improvement of differential evolution algorithm and applications in chemical engineering," Master's thesis, *National Chin-Yi University of Technology*, 2012

3. B. V. Babu and R. Angira, "Modified differential evolution (mde) for optimization of non-linear chemical processes," *Computers & Chemical Engineering*, Vol. 30, No. 6-7, pp. 989-1002 , 2006

4. J. S. Logsdon and L. T. Biegler, "Accurate solution of differential-algebraic optimization problems," *Industrial & Engineering Chemistry Research*, Vol. 28, No. 11, pp. 1628-1639 ,1989

5. S. A. Dadebo and K. B. Mcauley, "Dynamic optimization of constrained chemical engineering problems using dynamic programming," *Computers & Chemical Engineering*, Vol. 19, No. 5, pp. 513-525 ,1995

6. R. Luus, "Iterative Dynamic Programming," *Champman & Hall/ CRC,US*, 2000