Cooperative Combination of the Cross-Entropy Method and the Evolutionary Particle Swarm Optimization to Improve Search Domain Exploration and Exploitation

Leonel Carvalho, Vladimiro Miranda, Armando Leite da Silva, Carolina Marcelino, Elizabeth Wanner, Jean Sumaili

leonel.m.carvalho@inesctec.pt

17-07-2017
Methodological Approach

• **Combination** of two optimization methods
  – Cross-Entropy (CE) Method for **exploration**
  – Evolutionary Particle Swarm Optimization (EPSO) for **exploitation**

• **EPSO parameters were tuned** using an iterative optimization process based on a $2^2$ **factorial design**
  – Only the mutation rate $\tau$ and the probability of communication $P$ were optimized
CE Method for Optimization

• **Monte Carlo approach** to combinatorial and continuous non-linear optimization proposed by Reuven Rubinstein

• Starts by defining a *sampling distribution* for the *optimization variables* and gradually adjust the parameters (e.g. mean, standard deviation) according to the performance of part of the samples
  – Example of distributions: Bernoulli, Binomial, **Gaussian**
CE Method for Optimization

Select $\mu_0$ and $\sigma_0^2$, the number of samples per iteration $N$, the rarity parameter $\rho$, the smoothing parameter $\alpha$, $k := 0$

Do

$k := k + 1$

Generate a sample of $X_1$, ..., $X_N$ from the sampling distribution $N(\mu_{k-1}, \sigma_{k-1}^2)$

Compute $S(X_1)$, ..., $S(X_N)$ and order the samples from the worst to the best performing ones, i.e. $S(X_1) < S(X_2) < ... < S(X_N)$

Compute $\gamma_k$ as the $\rho$th quantile of the performance values and select $N_{\text{elite}} = \rho N$; let $\psi$ be the subset from the ordered set of samples that contains all the $N_{\text{elite}}$ samples, i.e., the samples $S(X) < \gamma_k$

For $j = 1$ to $n$

$$\mu_{kj} := \sum_{i \in \psi} \frac{X_{ij}}{N_{\text{elite}}}$$

$$\sigma_{kj}^2 := \sum_{i \in \psi} \frac{(X_{ij} - \mu_{kj})^2}{N_{\text{elite}}}$$

End For

Apply smoothing

$$\mu_k := \alpha \mu_k + (1 - \alpha) \mu_{k-1}$$

$$\sigma_k^2 := \alpha \sigma_k^2 + (1 - \alpha) \sigma_{k-1}^2$$

 Until $k < k^{\text{MAX}}$
CE Method for Optimization

- Test Bed OPF: Case 1
EPSO

- **Evolutionary Particle Swarm Optimization (EPSO)** is an hybrid between Evolutionary Strategies (ES) and Particle Swarm Optimization (PSO) proposed by Vladimiro Miranda
  - **Replication**: each individual is replicated $r$ times
  - **Mutation**: the $r$ clones have their weights $w$ mutated
  - **Recombination**: the $r+1$ individuals generate one offspring
  - **Evaluation**: each offspring has its fitness evaluated
  - **Selection**: the best particle out of the $r+1$ survives to be part of a new generation
EPSO

- Movement Rule

\[ X_{\text{new}} = X + V_{\text{new}} \]

\[ V_{\text{new}} = w_I^* V + w_M^*(X_M - X) + w_C^* P(X_G^* - X) \]

- **Inertia**: movement in the same direction
- **Memory**: attraction towards the individual best solution
- **Cooperation**: attraction to a region near the global best position
EPSO

• Weights are mutated and subjected to selection

\[ w^* = w[1 + \sigma N(0,1)] \]

• The individuals are attracted to a region near the best solution found

\[ X^*_G = X_G [1 + w^*_{GB} N(0,1)] \]

• Matrix P acts as a communication barrier

\[
V^{\text{new}} = w^*_I V + w^*_M (X_M - X) + w^*_C P (X^*_G - X)
\]

\[
P = \begin{bmatrix}
1 & 0 & 0 \\
0 & \ddots & 0 \\
0 & 0 & 0
\end{bmatrix}
\]

If \( i \neq j \) then \( P_{ij} = 0 \)
ElseIf \( U(0,1) > p \) then \( P_{ij} = 0 \)
Else \( P_{ij} = 1 \)
Example

- Test Bed OPF: Case 1 – 31 independent runs
EPSO Parameter Tuning

• **Iterative method** based on $2^2$ factorial design and the Two-way ANOVA to guarantee the best EPSO performance in every problem

  Define maximum allowable interval for $\tau$ and $P$ (e.g. [0.2, 0.8])

  Run a $2^2$ factorial design (**4 experiments**)

  Perform $4 \times 31$ runs of EPSO for the 4 combinations of the limit values of $\tau$ and $P$

  Do

  Run Two-way ANOVA and select the variable with the highest F-test statistic

  Compute the main effect to determine which limit should be updated

  Update the limit to the central value of the interval (e.g. if the output decreases with the variable increase, then set the lower limit to the central value)

  Run a $2^2$ factorial design with the updated limits (**+2 experiments**)

  While there is evidence that $\tau$ and $P$ affect the output (check $p$-values) or if the difference between the limits is greater than a threshold (e.g. 0.1)
## EPSO Parameter Tuning

### Two-way ANOVA results

<table>
<thead>
<tr>
<th>Factor</th>
<th>F-test Statistic</th>
<th>P-value</th>
<th>Main Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1st Iteration</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\tau$</td>
<td>3.474</td>
<td>0.064</td>
<td>-350.73</td>
</tr>
<tr>
<td>$P$</td>
<td><strong>6.675</strong></td>
<td>0.011</td>
<td><strong>486.14</strong></td>
</tr>
<tr>
<td>$\tau \times P$</td>
<td>0.225</td>
<td>0.636</td>
<td>89.17</td>
</tr>
<tr>
<td></td>
<td>2nd Iteration</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\tau$</td>
<td>5.469</td>
<td>0.021</td>
<td>-548.35</td>
</tr>
<tr>
<td>$P$</td>
<td><strong>6.022</strong></td>
<td>0.015</td>
<td><strong>575.42</strong></td>
</tr>
<tr>
<td>$\tau \times P$</td>
<td>1.496</td>
<td>0.223</td>
<td>286.79</td>
</tr>
<tr>
<td></td>
<td>3rd Iteration</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\tau$</td>
<td>5.508</td>
<td>0.020</td>
<td>-448.7</td>
</tr>
<tr>
<td>$P$</td>
<td>1.448</td>
<td>0.231</td>
<td>230.01</td>
</tr>
<tr>
<td>$\tau \times P$</td>
<td>0.958</td>
<td>0.329</td>
<td>187.14</td>
</tr>
<tr>
<td></td>
<td>4th Iteration</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\tau$</td>
<td>2.849</td>
<td>0.093</td>
<td>-306.78</td>
</tr>
<tr>
<td>$P$</td>
<td>2.497</td>
<td>0.116</td>
<td>287.21</td>
</tr>
<tr>
<td>$\tau \times P$</td>
<td>1.807</td>
<td>0.181</td>
<td>244.34</td>
</tr>
</tbody>
</table>

### Setting values for $\tau$ and $P$

<table>
<thead>
<tr>
<th>Factor</th>
<th>Iteration</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau$ (low)</td>
<td>1st</td>
</tr>
<tr>
<td></td>
<td>0.2</td>
</tr>
<tr>
<td>$\tau$ (high)</td>
<td>0.8</td>
</tr>
<tr>
<td>$P$ (low)</td>
<td>0.2</td>
</tr>
<tr>
<td>$P$ (high)</td>
<td>0.8</td>
</tr>
</tbody>
</table>

Greater than the threshold level of 0.05

Box-plots for the final range of $\tau$ and $P$
Results

- Test Bed OPF: 31 independent runs

<table>
<thead>
<tr>
<th>Fitness</th>
<th>Case 1</th>
<th>Case 2</th>
<th>Case 3</th>
<th>Case 4</th>
<th>Case 5</th>
<th>Case 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Best</td>
<td>72682.29</td>
<td>72044.58</td>
<td>60282.36</td>
<td>71391.65</td>
<td>70566.15</td>
<td>60799.88</td>
</tr>
<tr>
<td>Mean</td>
<td>72686.53</td>
<td>72049.53</td>
<td>60286.37</td>
<td>71396.14</td>
<td>70572.44</td>
<td>60805.33</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>2.95</td>
<td>4.08</td>
<td>3.45</td>
<td>4.33</td>
<td>4.36</td>
<td>4.26</td>
</tr>
</tbody>
</table>
Results

• Test Bed OSDER: 31 independent runs

<table>
<thead>
<tr>
<th>Fitness</th>
<th>Case 33</th>
<th>Case 180</th>
</tr>
</thead>
<tbody>
<tr>
<td>Best</td>
<td>-5216.82</td>
<td>-2566.78</td>
</tr>
<tr>
<td>Mean</td>
<td>-5183.73</td>
<td>-2549.76</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>20.82</td>
<td>17.36</td>
</tr>
</tbody>
</table>
Final Remarks

• **Combination of methods** to address different stages of the search process can greatly **improve accuracy** and **robustness**
  – Need to establish intelligent mechanisms to switch from methods

• **Systematic parameter tuning** is essential to reduce the information required from the user
Thank you for your attention!

Leonel Carvalho
leonel.m.carvalho@inesctec.pt