Dualcenter CMA-ES Experimentation

# System Description

This proposed system requires prior knowledge of the CMA-ES algorithm. It adds two more components to the regular CMAE-ES algorithm. Firstly, there is additional center that is calculated differently from the pulled-back center (original center). Secondly, there is an adaptive method used in determining how many solutions are to be generated from each of the centers.

1. This center vector is derived from an exponentially weighted average of the best offspring generated from the previous w-generations where w is the size of the history-window (usually about 10-30).
2. The adaptive evolution adds two variables to the model, which I call, orig-scale & best-scale. These scales are used to determine how much offspring will be generated from the two centers. These scales are calculated after each generation by summing the weights of all solutions generated from the original center that were selected, and for all the solutions generated from the best center that were selected. More specifically:
   * orig-scale = 1 + sum(weights of selected original offspring)
   * best-scale = 2 - orig-scale

One can infer that these two scales will always sum to 2. Now, these scales are each multiplied by ƛ/2. For example, if (ƛ = 12, orig-scale = 1.5, best-scale = 0.5) then the number of offspring to be generated from the original center is 1.5 x 12/2 = floor(9) = 9, and the number of offspring to be generated from the best center is ƛ - 9 = 3. Whatever types of offspring are selected all depend of the fitness values of these solutions. Lastly, if the scales become completed favored in one direction, for example, orig-scale = 2 & best-scale = 0 then this only happens for the current generation. The scales would then be reset to orig-scale = 1.5 & best-scale = 0.5 (or whatever default setting you chose).

## 

## Experiment 1

**Settings**

|  |  |
| --- | --- |
| orig-scale | 1.5 |
| best-scale | 0.5 |
| weights for finding scales |  |
| weights for finding bestcenter |  |
| window size for the best solution from previous generations | 10 |

**Hypothesis**

The scales are set to favor generating solutions from the original center, with the minority of solutions being generating from the best-center. The reason for this is to maintain the reliability of CMA-ES and to give it the option to generate more solutions from the best-center if those solutions become favorable. In the weight equations, the problem dimension is added to to produce a flatter, or more a more evenly distributed set of log-weights. This set of log-weights may be better at maintaining the stability or the change rate of the scale values. Therefore, less resetting of the scales should be observed (this happens when the scales become 2 & 0). A best chromosome window size of 10 is good and can help the algorithm to recover from stagnation if the past generations produced some bad solutions.

**Experiment 1 Results**

|  |  |  |
| --- | --- | --- |
| **levy** | | |
| dimension | normal wins | dual-center wins |
| 5 | 16 | 14 |
| 10 | 12 | 18 |
| 25 | 13 | 17 |
| 50 | 15 | 15 |
| **elliptical** |  |  |
| dimension | normal wins | dual-center wins |
| 5 | 15 | 15 |
| 10 | 6 | 24 |
| 25 | 11 | 19 |
| 50 | 15 | 15 |
| **ackley** | | |
| dimension | normal wins | dual-center wins |
| 5 | 9 | 21 |
| 10 | 8 | 22 |
| 25 | 8 | 22 |
| 50 | 11 | 19 |
| **griewank** | | |
| dimension | normal wins | dual-center wins |
| 5 | 15 | 15 |
| 10 | 10 | 20 |
| 25 | 5 | 25 |
| 50 | 7 | 23 |
| **rosenbrock** | | |
| dimension | normal wins | dual-center wins |
| 5 | 9 | 21 |
| 10 | 5 | 25 |
| 25 | 4 | 26 |
| 50 | 8 | 22 |

|  |  |  |
| --- | --- | --- |
| **dim** | **normal wins** | **dual-center wins** |
| 5 | 64 | 86 |
| 10 | 41 | 109 |
| 25 | 41 | 109 |
| 50 | 56 | 94 |

Which dimension was overall better for dual-center wins across these functions?

From the tally results here, it appears the most success was at dimensions 10 & 25. Since the dimension was used in calculating the weights, perhaps 10 & 25 may be good constants for calculating the scoring weights or the best-center. The new scoring weights will be tested with variants of these constants, followed by the best-center weights and then together.

## Experiment 2

**Settings**

|  |  |
| --- | --- |
| orig-scale | 1.5 |
| best-scale | 0.5 |
| weights for finding scales |  |
| weights for finding best-center |  |
| window size for the best solution from previous generations | 10 |

**Experiment 2 Results**

|  |  |  |
| --- | --- | --- |
| **levy** | | |
| dimension | normal wins | dual-center wins |
| 5 | 7 | 23 |
| 10 | 11 | 19 |
| 25 | 17 | 13 |
| 50 | 14 | 16 |
| **elliptical** |  |  |
| dimension | normal wins | dual-center wins |
| 5 | 15 | 15 |
| 10 | 7 | 23 |
| 25 | 10 | 20 |
| 50 | 20 | 10 |
| **ackley** | | |
| dimension | normal wins | dual-center wins |
| 5 | 11 | 19 |
| 10 | 12 | 18 |
| 25 | 5 | 25 |
| 50 | 11 | 19 |
| **griewank** | | |
| dimension | normal wins | dual-center wins |
| 5 | 15 | 15 |
| 10 | 5 | 25 |
| 25 | 7 | 23 |
| 50 | 8 | 22 |
| **rosenbrock** | | |
| dimension | normal wins | dual-center wins |
| 5 | 8 | 22 |
| 10 | 8 | 22 |
| 25 | 5 | 25 |
| 50 | 8 | 22 |

**Observation**

|  |  |  |
| --- | --- | --- |
| **dim** | **normal wins** | **dual-center wins** |
| 5 | 56 | 94 |
| 10 | 43 | 107 |
| 25 | 44 | 106 |
| 50 | 61 | 89 |

Which dimension was overall better for dual-center wins across these functions?

From the tally results here, it appears the most success is **again**, at dimensions 10 & 25, with a slight increase in dim 5 wins and a slight decrease in dim 50 wins. In the next experiment I will test using this same weight distribution for calculating the best center, but not for the scoring system.

## Experiment 3

**Settings**

|  |  |
| --- | --- |
| orig-scale | 1.5 |
| best-scale | 0.5 |
| weights for finding scales |  |
| weights for finding best-center |  |
| window size for the best solution from previous generations | 10 |

**Experiment 3 Results**

|  |  |  |
| --- | --- | --- |
| **levy** | | |
| dimension | normal wins | dual-center wins |
| 5 | 12 | 18 |
| 10 | 11 | 19 |
| 25 | 11 | 19 |
| 50 | 9 | 21 |
| **elliptical** |  |  |
| dimension | normal wins | dual-center wins |
| 5 | 13 | 17 |
| 10 | 10 | 20 |
| 25 | 12 | 18 |
| 50 | 13 | 17 |
| **ackley** | | |
| dimension | normal wins | dual-center wins |
| 5 | 11 | 19 |
| 10 | 10 | 20 |
| 25 | 8 | 22 |
| 50 | 12 | 18 |
| **griewank** | | |
| dimension | normal wins | dual-center wins |
| 5 | 9 | 21 |
| 10 | 12 | 17 |
| 25 | 6 | 24 |
| 50 | 10 | 20 |
| **rosenbrock** | | |
| dimension | normal wins | dual-center wins |
| 5 | 9 | 21 |
| 10 | 7 | 23 |
| 25 | 6 | 24 |
| 50 | 6 | 24 |

**Observation**

|  |  |  |
| --- | --- | --- |
| **dim** | **normal wins** | **dual-center wins** |
| 5 | 54 | 96 |
| 10 | 50 | 100 |
| 25 | 43 | 107 |
| 50 | 50 | 100 |

Which dimension was overall better for dual-center wins across these functions?

With this new weight distribution for determining the best-center, it appears that the dual-center system has most wins across all functions. As a result, this new parameter is the front-runner so far and may hint at the fact that the way this center is calculated is very important.