How to Gather Domain Knowledge (with CMA-ES)

by a semi-practitioner

Nikolaus Hansen
Inria & École polytechnique, France

February 2026

Context

- continuous search domain
- single objective
- mostly black-box
- mostly anytime scenario (budget is not input to the algorithm)
- the budget allows more than one single run

1. Understand and (re)shape the objective function (it should be smooth and well-conditioned)
   - variable encoding (small dimensionality, uniform sensitivity, small correlation, “high” smoothness)
   - using knowledge of mathematics (absolute values are not smooth, max(.) is not smooth,…)
   - using domain knowledge
   - modality can be tweaked, sometimes
   - scrutinizing single-run data, in particular / for example using `cma.plot()`
   - rather tune the function than the algorithm

2. On “new” problems: anticipate that the first found “optimal” solutions exploit some mishap in the formulation of the objective

3. A first attempt: explore the neighbourhood of the *best known solution* with a (very) small initial step-size, use SLSQP

   the initial step-size ***was*** small when it *increased by a factor of 3 or more* right from the start

4. Be aware: determining local optimality (on a non smooth function) is nontrivial

   continue a terminated run after increasing the step-size by a factor of 3
   check whether reruns end up in the same solution(s) up to a high precision
   ending up *in the same solution* twice is a good indicator of local optimality
   the opposite is false: not ending up in the same solution is *not* a good indicator of anything *specific*

5. Plot sections through the landscape

   1-D information is cheap, `cma.optimization_tools.Sections` may help

   informs about variable scaling order of magnitude (see point 1), ruggedness/noise,…

   look at any “cheaply” available information, without looking we won’t see what we miss

6. Decide on population size (and noise handling) and restart with (widely) different initial solutions and initial step-sizes (as far as budget allows)

   do we end up in the same optima?

   any locally optimal solution we did not know may be interesting domain knowledge

7. “Play” with larger population sizes (IPOP)

   and different initial solutions (and step-sizes)