Jingyun_thesis_problems

Current result

strategy	Decrease factor (emergency)	GP length scale	linear	quadratic	cubic	schwefel	quartic
Baseline (Arash's)		8	**503**	**214**	**198**	**1503**	**1236**
Emergency CSA	0.68	8	299	144	178	949	720
Emergency CSA	0.72	8	304	142	184	916.5	717.5
Emergency CSA	0.72	32	350	134	192	977	682

Problems

Since what I did is a combination of (1+1)-ES and $(\mu/\mu, \lambda)$-ES, specifically, $((\mu/\mu, \lambda) + (\mu/\mu, \lambda))$-ES ((mml+mml)-ES). I'm not sure it falls more into (1+1)-ES or mml-ES. The most confusing question is should I do the Section 3 Analysis on mml-ES and propose the emergency situation in step size adaptation and therefore use (mml+mml)-ES in Section 4 or just propose (mml+mml)-ES in Section 3 (I think the analysis would be different) referred to the question in 3 a iii.

Some clarifications:

1. In introduction, I'd focus on the effect of using a population in offspring generation that could potentially avoid poorer step size and render the bias due to the surrogate model.
2. In related work, I want to include
   a. Surrogate models
   In comparison, should I pays more attention to surrogate-assisted mml-ES and their result or (1+1)-ES.
   b. Step size adaptation
   In the context that we use mml-ES in each iteration, I'd the usually used step size adaptation mechanism CSA and CMA. Should I mention 1/5-rule since the idea of emergency is borrowed from that.
3. In analysis, the most confusing part. I'll try put the equation in your paper first and derive that if time permits. Some detailed questions
   a. Plot expected progress rate over normalized step size for different noise-to-signal ratio
      - i. for the solid line ($n \to \infty$), i could simply use the equation in your paper

$$\eta = E[\Delta_R^*] = \frac{\sigma^* c_{\mu/\mu,\lambda}}{\sqrt{1+\vartheta^2}} - \frac{(\sigma^*)^2}{2\mu}$$

- ii. for the dotted lines (n=10,100), I would use the following equation in your paper

$$\eta \approx \frac{c_{\mu/\mu,\lambda}\sigma^*(1+\sigma^{*2}/2\mu N)}{\sqrt{1+\sigma^{*2}/2\mu N)}\sqrt{1+\vartheta^2+\sigma^{*2}/2N}} - N\left[\sqrt{1+\frac{\sigma^{*2}}{\mu N}} - 1\right]$$

- iii. for circle (experimental result for n=100) and cross (experimental result for n=10) For experimental result should I use the mml-ES or (mml+mml)-ES (using normalized step size and Gaussian distributed noise to model GP)
- iv. dashed line (performance of surrogated assisted (1+1)-ES) for comparision
- iv. the other problem is how to make the plot clear. I think possibly I should include i-iii for different lambda but that will makes the figure very messy, wonder if you have any suggestions on that

4. I wrote a pseudo code and wonder if the surrogate training part should be included. If so I need to add another if condition right after $y_i \leftarrow x + \sigma z_i$ and that makes the algorithm ugly.

Algorithm 1 A Surrogate Assisted $((\mu/\mu, \lambda) + (\mu/\mu, \lambda))$-ES

$c \leftarrow \frac{\mu+2}{n+\mu+5}$
$d \leftarrow 1 + 2\max(0, \sqrt{\frac{\mu-1}{n+1}} - 1)$
$p \leftarrow 0$
$D \leftarrow 0.68$
while not terminate() **do**
  for $i = 1, 2, ..., \lambda$ **do**
    Generate standard normally distributed $z_i \in \mathbb{R}^N$
    $y_i \leftarrow x + \sigma z_i$
    Evaluate y_i using the surrogate model, yieding $\hat{f}(y_i)$
  end for
  $z = \frac{1}{\mu} \sum_{i=1}^{\mu} z_{i;\lambda}$
  $y = x + \sigma x$
  Evaluate y using true objective function, yieding $f(y)$
  Update surrogate modle
  if $f(x) < f(y)$ (Emergency) **then**
    $\sigma \leftarrow \sigma D$
  else
    $s \leftarrow (1-c)s + \sqrt{c(2-c)\mu z}$
    $\sigma \leftarrow \sigma \times \exp\left(\frac{c}{d} \frac{\|X\|}{E\|N(0,I)\|} - 1\right)$
  end if
end while