

Homework #2

Q1. ES and steady-state GAs

Discuss why ES and steady-state GAs form two extremes regarding the population size and the number of offspring created.

Population Size

- **ES**
 - Uses relatively larger populations to widely explore the search space.
 - Generates diverse solutions to enhance global search.
 - Avoids premature convergence.
- **Steady-State GAs**
 - Generally works with smaller populations.
 - Replaces only a few individuals at a time.
 - Population remains constant throughout.
 - Leads to a more incremental and localized improvement of existing solutions.

Number of Offspring Created

- **ES**
 - Produces a large number of offspring per generation.
 - Common setting: seven times the parent count ($\lambda \approx 7\mu$).
 - Maximizes diversity and exploration.
 - Higher computational cost.
- **Steady-State GAs**
 - Generate only a few offspring per iteration and immediately evaluate and integrate them.
 - Allows for steady, computationally efficient convergence
 - But limits broad exploration.
 - Lower computational cost.

Summary

ES favors exploration with large populations and many offspring, ideal for broad solution spaces. Steady-state GAs focus on exploitation with small populations and fewer offspring, making them suited for refining solutions. This places them at opposite ends of the spectrum regarding population size and offspring creation.

Q2. Change in allele frequency

Given a population of μ individuals, which are bit-strings of length L . Let the frequency of allele 1 be 0.25 at position i , that is, 25% of all individuals contain a 1, and 75% a 0 at the i th position on the chromosome. How does this allele frequency change after performing k crossover operations with one-point crossover? How does it change if uniform crossover is performed?

The initial allele frequency of 25% at position i will not change after performing crossover operations, regardless of the type of crossover used. This is because in the absence of any selection or mutation, crossover alone does not alter the allele frequency in the population.

Both one-point and uniform crossover are mechanisms to shuffle genetic material between individuals. They don't inherently introduce new alleles or change the overall allele frequencies in the population.

Therefore, the allele frequency at position i will remain at 25%, even after multiple crossover operations.

Q3. Tables of (1,1)-ES and (1+1)-ES using correlated Gaussian mutation

(1,1)-ES – with correlated Gaussian mutation

	Number of Generations		
(1,1)-ES	$\sigma = 0.01$	$\sigma = 0.1$	$\sigma = 1.0$
Run #1	10000000	10000000	10000000
Run #2	10000000	10000000	10000000
Run #3	10000000	10000000	10000000
Run #4	10000000	10000000	10000000
Run #5	10000000	10000000	10000000
Run #6	10000000	10000000	10000000

Run #7	10000000	10000000	10000000
Run #8	10000000	10000000	10000000
Run #9	10000000	10000000	10000000
Run #10	10000000	10000000	10000000
Average:	10000000	10000000	10000000
STD:	0	0	0

(1+1)-ES – with correlated Gaussian mutation

	Number of Generations		
(1+1)-ES	$\sigma = 0.01$	$\sigma = 0.1$	$\sigma = 1.0$
Run #1	489	371891	10000000
Run #2	519	186349	10000000
Run #3	438	186528	10000000
Run #4	556	134240	10000000
Run #5	500	727107	10000000
Run #6	570	205467	10000000
Run #7	412	268314	10000000
Run #8	412	84586	10000000
Run #9	571	82344	10000000
Run #10	527	317246	10000000
Average:	499.4	256407.2	10000000
STD:	61.1	190283.2	0

Q4. Observe the running processes of problem 3

Compare and contrast the results you obtained in problem 3 and discuss what you think about the difference between (1,1)-ES and (1+1)-ES.

Impact of Step Size σ

- **Smaller σ :** The ES performs better with smaller step sizes, as we can tell by the result in (1+1)-ES. It also gives us a smaller standard deviation compared to using a larger σ . This is because smaller steps allow for more precise exploration of the search space without over shooting.
- **Larger σ :** Larger step sizes can lead to faster convergence, but can also lead to premature convergence to suboptimal solutions which results in poor performance. We can also tell this by the result in (1+1)-ES.

Comparison the Results

- **(1,1)-ES:** Performs very poorly, with all runs hitting the 10,000,000 generations threshold no matter which σ is used. This is most likely because (1,1)-ES always replaces the parent with the offspring, regardless of fitness. This leads to a loss of good solutions.
- **(1+1)-ES:** Performs relatively well, especially with smaller σ values, compared to that of (1,1)-ES. This is most likely because (1+1)-ES selects the fitter individual between the parent and offspring. This ensures that the population always improves or at least stays the same, leading to more efficient exploration of the search space.

Q5. Tables of (1,1)-ES and (1+1)-ES using uncorrelated Gaussian mutation

(1,1)-ES – with uncorrelated Gaussian mutation

	Number of Generations		
(1,1)-ES	$\sigma = 0.01$	$\sigma = 0.1$	$\sigma = 1.0$
Run #1	10000000	10000000	10000000
Run #2	10000000	10000000	10000000
Run #3	10000000	10000000	10000000

Run #4	10000000	10000000	10000000
Run #5	10000000	10000000	10000000
Run #6	10000000	10000000	10000000
Run #7	10000000	10000000	10000000
Run #8	10000000	10000000	10000000
Run #9	10000000	10000000	10000000
Run #10	10000000	10000000	10000000
Average:	10000000	10000000	10000000
STD:	0	0	0

(1+1)-ES – with uncorrelated Gaussian mutation

	Number of Generations		
(1+1)-ES	$\sigma = 0.01$	$\sigma = 0.1$	$\sigma = 1.0$
Run #1	2678	7637	113157
Run #2	2410	3794	4122
Run #3	4072	3318	309506
Run #4	2668	4098	8174
Run #5	21733	9772	28617
Run #6	1804	1854	5078
Run #7	9999	16046	148211
Run #8	7553	1341	8010
Run #9	3054	3816	122068
Run #10	8239	2887	154968
Average:	6421.0	5456.3	90191.1
STD:	6093.7	4520.4	99514.9

Q6. Compare and contrast the results you obtained in Q3 and Q5.

Discuss what you think about self-adaptation.

Impact of Step Size σ

- **Small Step Size ($\sigma = 0.01$)**
 - **Correlated Gaussian Mutation:** Benefits from the consistent, small steps, leading to more precise exploration of the search space.
 - **Uncorrelated Gaussian Mutation:** While initially less efficient, the adaptive step size mechanism allows it to gradually reduce the step size, ultimately achieving comparable performance.
- **Medium Step Size ($\sigma = 0.1$)**
 - **Correlated Gaussian Mutation:** Struggles with the fixed step size, potentially leading to premature convergence or overshooting the optimal solution.
 - **Uncorrelated Gaussian Mutation:** Benefits from the adaptive step size, allowing it to adjust the step size based on the search progress and avoid getting stuck in local optima.
- **Large Step Size ($\sigma = 1$)**
 - **Correlated Gaussian Mutation:** Struggles with the fixed step size, leading to premature convergence and overshooting the optimal solution. This leads to poor performance.
 - **Uncorrelated Gaussian Mutation:** Benefits from the adaptive step size, allowing it to adjust the step size based on the search progress and avoid getting stuck in local optima. Outperforms the usage of correlated Gaussian mutation.

Comparing the Results

Based on the results of Q3 and Q5, it seems that **uncorrelated Gaussian mutation** has performed better, especially for larger step sizes. This is likely due to its ability to adapt the step size for each dimension independently, allowing for a more targeted exploration of the search space.

Self-Adaptation

Self-adaptation is a powerful technique employed in evolutionary algorithms to dynamically adjust the mutation strength (step size in this case) during the optimization process. This allows the algorithm to adapt to the changing landscape of the problem and avoid getting stuck in local optima.

However, the parameters τ , τ' , and ϵ_0 play a crucial role in controlling the rate of step size adaptation. Knowing that minimizing the n-dimensional sphere model requires a smaller step

size (as seen in the result of Q3), I added another parameter, Ω , to prevent the step size from going too large. Proper tuning of these parameters may make or break the algorithm. After thorough tests, the final value of these parameters are:

$$\tau = \frac{1}{\sqrt{2\sqrt{n}}} \quad \tau' = \frac{1}{\sqrt{2n}} \quad \varepsilon_0 = 10^{-6} \quad \Omega = 1$$

Q7. Tables of (1,1)-ES and (1+1)-ES using correlated Gaussian mutation and the 1/5 Rule

(1,1)-ES – with correlated Gaussian mutation and the 1/5 rule

	Number of Generations		
(1,1)-ES	$\sigma = 0.01$	$\sigma = 0.1$	$\sigma = 1.0$
Run #1	10000000	10000000	10000000
Run #2	10000000	10000000	10000000
Run #3	10000000	10000000	10000000
Run #4	10000000	10000000	10000000
Run #5	10000000	10000000	10000000
Run #6	10000000	10000000	10000000
Run #7	10000000	10000000	10000000
Run #8	10000000	10000000	10000000
Run #9	10000000	10000000	10000000
Run #10	10000000	10000000	10000000
Average:	10000000	10000000	10000000
STD:	0	0	0

(1+1)-ES – with correlated Gaussian mutation and the 1/5 rule

	Number of Generations
--	-----------------------

(1+1)-ES	$\sigma = 0.01$	$\sigma = 0.1$	$\sigma = 1.0$
Run #1	326	322	467
Run #2	341	262	548
Run #3	408	284	487
Run #4	418	316	530
Run #5	319	299	433
Run #6	416	328	481
Run #7	310	316	500
Run #8	327	303	481
Run #9	456	348	466
Run #10	370	201	514
Average:	369.1	297.9	490.7
STD:	51.7	41.5	33.6

Q8. Compare and contrast the results you obtained in problems 3, 5, and 7

Discuss what you think about the 1/5-rule for the self-adaptation of strategic parameters.

Comparing the Results

The results from Q7 demonstrate that the 1/5 rule can significantly enhance the performance of correlated Gaussian mutation, particularly when using larger initial step sizes. In these cases, it often outperforms uncorrelated Gaussian mutation.

A notable observation from the experiments is the convergence of the step size to approximately 0.2, regardless of the initial value. This suggests that for the specific problem of minimizing the n-dimensional sphere function, a step size of 0.2 is optimal.

This finding highlights the importance of adaptive step size mechanisms in evolutionary algorithms. By dynamically adjusting the step size based on the optimization progress, these algorithms can efficiently explore the search space and find optimal solutions.

The 1/5 Rule for Self-Adaptation

The 1/5 rule is a simple yet effective strategy for adapting the step size in evolutionary algorithms. It involves monitoring the success rate of mutations and adjusting the step size accordingly. This rule helps to balance the exploration of new regions of the search space with the exploitation of promising areas, resulting in great performance.

The parameters G and a play an important role in the adaptation of the step size. After careful consideration, the final value of these parameters are:

$$G = 50 \quad a = 1.5$$

Conclusion

Comparing the Three Approaches

1. **Fixed Step-Size Mutation:** This approach resulted in limited adaptability, particularly for (1,1)-ES. Without self-adjustment, both (1,1)-ES and (1+1)-ES struggle to maintain efficiency across different step-sizes. (1+1)-ES showed better performance with smaller step sizes ($\sigma = 0.01$), where it could explore the search space more precisely, while (1,1)-ES was unable to achieve the termination criteria in any run, indicating significant limitations.
2. **Uncorrelated Gaussian Mutation:** This method improved performance by allowing separate step-sizes for each dimension, which helped in adapting to the search landscape. (1+1)-ES performed well across a range of initial step-sizes, dynamically fine-tuning its search in different dimensions. However, (1,1)-ES still struggled, especially with larger step-sizes, due to the lack of selection pressure to retain good solutions, leading to slower convergence.
3. **1/5-Rule for Self-Adaptation:** This approach had the best overall performance, particularly for (1+1)-ES. By adjusting the step-size based on the success rate, (1+1)-ES could dynamically balance exploration and exploitation, leading to efficient convergence across all initial step-sizes. For (1,1)-ES, however, while the 1/5-rule brought some improvements, it could not fully compensate for the algorithm's inherent lack of selective pressure.

Summary

Overall, **(1+1)-ES consistently outperformed (1,1)-ES**, especially with the 1/5-rule and uncorrelated Gaussian mutation. The selective pressure in (1+1)-ES, which retains better solutions, allows it to adapt more effectively to the search landscape. In contrast, (1,1)-ES lacks this selective mechanism, leading to frequent replacement of good solutions and poorer convergence. These results highlight the advantages of adaptive step-sizes and selection in evolutionary strategies, especially for complex optimization tasks.

