COMP 6660 Fall 2022 Assignment 1a

Joshua Boyd
jcb0216@auburn.edu

August 31, 2022

Abstract

Random search is a process by which solutions to a problem are randomly generated and compared against a test. The goal of this assignment was to implement a random search algorithm which generates bridges by uniform random placement of coordinates within a predefined rectangular area. Generated bridges were subjected to a simple physics engine test to estimate structural integrity. After an optimal bridge structure was found, statistical analysis was performed between the results and mystery algorithm data. In conclusion, the random search algorithm perform statistically significantly better than the mystery algorithm.

Introduction

Random search is a process by which solutions to a problem are randomly generated and compared against a test. The result of the solution when subjected to the defined test is measured as the fitness of that solution. Solutions with a higher fitness are kept, while the lower fitness solutions are disregarded. Termination of the search happens when either a predefined fitness solution or iteration limit is reached.

Problem Statement

In this experiment, random search was applied in searching for optimal bridge structure. Bridge structure is represented by a uniform random placement of coordinates within a predefined rectangular area. The coordinate pairs represent nodes (connection points), which are automatically connected by frame elements (edges) if within a set distance from one another. Pre-existing nodes at fixed locations have been provided, and serve as the foundation of the bridge and connection points along the roadbed.

Once a bridge structure solution has been generated, it is evaluated against a physics simulation. In this simulation, frame elements stretch, compress, and bend slightly under load until the forces in the bridge are at equilibrium, allowing their structural integrity to be determined. This is done by applying an increasingly-heavy load to points within the roadbed until the bridge fails. Elements of the bridge will fail if their tensile, compressive, or bending stresses exceed the maximum amounts supported by their material and geometry. The bridge is considered to have failed if any element fails. After simulation, a fitness score is given to the solution. It is this fitness score that the algorithm must attempt to maximize in order to design the strongest bridge possible under these conditions.

Results

Bridge Generation

The experiment was ran for a total of 30 iterations, with each iteration generating 5000 solutions each. Best fitness values per run iteration are given in table 1 on page 4. The best solution found was given a fitness score of **62,000,000**. This fitness was first hit in the **5th iteration** of the search, at **solution number 161**. The bridge created for this solution is shown in figure 1 on page 2. A plot of the fitness progress for the 5th iteration is shown in figure 2 on page 3.

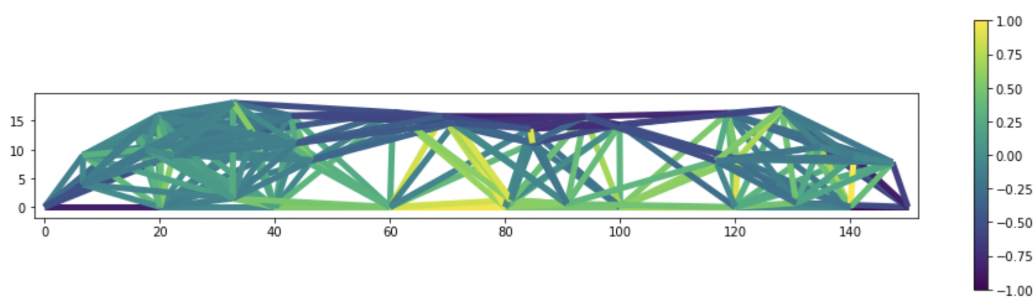

Figure 1: Best Solution Bridge

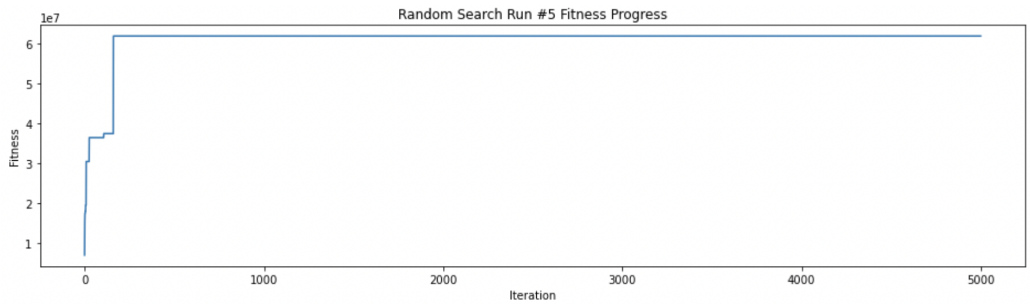

Figure 2: Run Number 5 Solution Fitness Progress

Statistical Analysis

The best fitness per run iteration data was compared against a mystery algorithm data sample provided. The values for this mystery data can be seen in table 2 on page 6. The distribution of the data is not known to be normal and the sample size is greater than 29. Therefore, a two sample F-Test was performed for variances. The table showing the results of this test can be found in figure 3 on page 5. The chosen value for α was 0.025. The sample size for each sample was 30. The variance for sample 1 was 18,822,988,505,747.1 and the variance for sample 2 was 88,466,954,022,988.4. The calculated test statistic F was valued at 0.212768583632437. The nearest critical value to F(F Critical one-tail) was 0.475964774310031. According to the test results, the value of F is smaller than one and is also smaller than F Critical one-tail. This means the null hypothesis of equal variances is rejected, and the two sets are not shown to be equal.

With the variances of the data sets not shown to be equal according to the F-Test, a Two-tailed two-sample t-test assuming unequal variances was used next for comparison. The table showing the results of this test can be found in figure 4 on page 7. The chosen value for α was 0.05. The sample size for each sample was 30. The sample mean for sample 1 was 51,933,333.33 and the sample mean for sample 2 was 47,083,333.33. The sample variance for sample 1 was 18,822,988,505,747.1 and the sample variance for sample 2 was 88,466,954,022,988.4. The calculated test statistic t was found to be 2.56461902490105. The upper critical value (t Critical two-tail) was 2.01954097044138.

Iteration Number	Best Solution Fitness Score
1	53000000
2	50000000
3	53000000
4	57000000
5	62000000
6	49500000
7	47000000
8	54500000
9	51500000
10	48500000
11	47500000
12	48500000
13	53000000
14	62000000
15	50000000
16	48000000
17	59000000
18	58000000
19	48000000
20	45500000
21	54000000
22	52500000
23	53500000
24	49000000
25	48500000
26	55000000
27	47000000
28	50000000
29	50500000
30	52500000

Table 1: Best Solution Fitness Score Per Run Iteration

Conclusion

The calculated test statistic t Stat is greater than 0 and greater than t Critical two-tail. Therefore, the two experiments produced significantly different mean fitness. The experiment with the highest sample mean can then be assumed to produce a significantly higher mean fitness than the other. Since experiment 1 produced a higher sample mean, we can conclude that experiment 1 (random search algorithm) is statistically significantly better than experiment 2 (mystery algorithm data).

F-Test Two-Sample for Variances		
	Variable 1	*Variable 2*
Mean	51933333.33	47083333.33
Variance	1.8823E+13	8.8467E+13
Observations	30	30
df	29	29
F	0.212768584	
P(F<=f) one-tail	3.89984E-05	
F Critical one-tail	0.475964774	

Figure 3: F-Test Two-Sample for Variance Excel Data

Row Number	Value
1	47500000
2	55500000
3	48500000
4	21500000
5	44000000
6	45500000
7	43000000
8	40500000
9	36000000
10	48500000
11	47500000
12	58500000
13	46000000
14	50000000
15	29500000
16	56000000
17	45500000
18	47000000
19	51000000
20	54500000
21	36500000
22	66000000
23	45500000
24	44000000
25	48500000
26	51500000
27	55500000
28	57000000
29	32500000
30	59500000

Table 2: Mystery Algorithm Sample Data

t-Test: Two-Sample Assuming Unequal Variances		
	Variable 1	*Variable 2*
Mean	51933333.33	47083333.33
Variance	1.8823E+13	8.8467E+13
Observations	30	30
Hypothesized Mean Difference	0	
df	41	
t Stat	2.564619025	
P(T<=t) one-tail	0.007043949	
t Critical one-tail	1.682878002	
P(T<=t) two-tail	0.014087897	
t Critical two-tail	2.01954097	

Figure 4: t-Test Two-Sample Assuming Unequal Variances Excel Data