Introduction

The Pearson correlation coefficient

What is pearsons r for

What is its use

Statement of the problem / research questions

* Yung research questions sa baba

Methodology

* Discuss how the C program is implemented
* Discuss how the runtime is timed. Why does this work? Program uses uh timespec, clock\_gettime, <https://stackoverflow.com/a/16276032>, https://linux.die.net/man/3/clock\_gettime
* Discuss how the big O notation will be analyzed – different methods
* Discuss how the theoretical running time is calculated – different methods

Results and Discussion

a. complexity

- what is the complexity something something? Explain here why

b. huge numbers

- say cant because too big, give the theoretical running time too or expected how many seconds

c. do the actual running time meet expected theoretical running time?

- say only at the start but it diverges afterwards. why? Might be because of additional overhead by the computations like multiply, addition, etc -> might take longer because the data gets bigger

Conclusions

* The complexity is this
* Because of the complexity cant run big
* Additionally there is added from the theoretical time because additional overhead

Research Questions:

1. What is the complexity of solving the Pearson Correlation Coefficient vector of an n x n square matrix X with a n x 1 vector y?

2. Is the program capable of running values of n > 10,000,000? If not, is it possible to run higher than 100,000,000? If not, then why, and what else is there to do to make it so?

3. Do the actual running time meet the expected theoretical running time?

4. Are there other ways to lower the average runtime without using extra processors or cores?