Dynamic Programming: Arithmetic, Geometric, Harmonic, and Power Means

Caleb Princewill Nwokocha

Department of Computer Science, The University of Manitoba

Winnipeg, Manitoba, Canada

August 24, 2022

Acknowledgments

To my family and professors, I am very grateful for your support before and during this writing.

Abstract

Arithmetic, Geometric, and Harmonic means are the well-known Pythagorean means [1, 2, 4, 3]. Power mean, also known as Hölder mean, is a generalized mean for the Pythagorean means [5]. Formulas of these means have been used to derive brute-force algorithms that run at polynomial space and time $O(n)$, but these brute-force algorithms cost more time and space when compared to their dynamic counterpart. In this paper, I derive dynamic algorithms using technique of dynamic programming for computing Pythagorean and generalized power means. Derivations of the dynamic algorithms are written in the sections 2-5, and a weblink to Java implementation of the dynamic algorithms is provided at the end of section 2. Derivations of the dynamic and brute-force mean algorithms make use of formulas found in statistics resources like Moore et al, 1998; Stan, 2009; and Chou, 1969; although the dynamic mean algorithms have elements of dynamic programming. I show that dynamic mean algorithms require only constant time and space $O(c)$ to compute the mean of a variable. These dynamic algorithms are less costly than the brute-force algorithms because the latter require $O(n)$ time and space. Implementation of the dynamic algorithms may be found in applications that process instant mathematical mean of very large data with respect to the time each datum is acquired. An example of a simple application is provided in the introduction section, where Bob calculates the mean number of oranges that he receives each day from Alice.

Keywords: Dynamic mean algorithm, dynamic programming, arithmetic mean, geometric mean, harmonic mean, power mean.

1. Introduction

The objective of calculating the mean is finding an average of a variable. Consider a case where Alice and Bob are friends, and every day, Alice gives Bob x_t oranges. The symbol x_t is a variable that represents the number of oranges that Bob receives from Alice each day. Suppose that on day 1, Alice gave Bob 2 oranges; on day 2, Alice gave Bob 6 oranges; and on day 3, Alice gave Bob only 1 orange. If the t subscript of x_t represent the day that Bob received an orange, then this Alice and Bob case can be stated statistically as the following:

When $t = 1$, $x_t = 2$
When $t = 2$, $x_t = 6$
When $t = 3$, $x_t = 1$

The case can be stated even more succinctly as the following:

$$[x_{t=1} = 2, x_{t=2} = 6, x_{t=3} = 1] \rightarrow O(3) \text{ space}$$

If Bob aim to calculate the arithmetic mean of number of oranges he receives each day, then Bob will typically do his calculation by the following brute-force mean algorithm:

$$\text{On day } t = 1\text{: } \bar{x}_{t=1} = \frac{x_{t=1}}{t=1} = \frac{2}{1} = 2 \ \rightarrow O(c) \text{ time}$$

The symbol $\bar{x}_t$ represents the mean on day t.

$$\text{On day } t = 2\text{: } \bar{x}_{t=2} = \frac{x_{t=1}+x_{t=2}}{t=2} = \frac{2+6}{2} = 4 \rightarrow O(c) \text{ time}$$

$$\text{On day } t = 3\text{: } \bar{x}_{t=3} = \frac{x_{t=1}+x_{t=2}+x_{t=3}}{t=3} = \frac{2+6+1}{3} = 3 \rightarrow O(c) \text{ time}$$

Observe that on the second day, Bob must remember $x_{t=1}$ to calculate $\bar{x}_{t=2}$ on day 2. Also, Bob must sum $x_{t=1}$ and $x_{t=2}$ as part of his calculation of $\bar{x}_{t=2}$ on day 2. Again, On the third day, Bob must remember $x_{t=1}$ and $x_{t=2}$, and sum $x_{t=1}$, $x_{t=2}$ and $x_{t=3}$ during his calculation of $\bar{x}_{t=3}$ on day 3. In general, Bob must remember $x_{t=1}, \cdots, x_{t=n}$ and sum $x_{t=1}, \cdots, x_{t=n+1}$ in the process of calculating $\bar{x}_{t=n+1}$ on day $n + 1$. Here n represents the last day t before day $n + 1$ that Bob received oranges from Alice. For example, if $n = 3$ then day n is equal to day 3, and day $n + 1$ is equal to day 4.

Analysis shows that if a computer was performing the same calculation as Bob, the computer would require $O(1) + O(1) + O(1) = O(3)$ space for storage of $x_{t=1}, x_{t=2}, x_{t=3}$. For Alice and Bob case, this $x_{t=1}, x_{t=2}, x_{t=3}$ is equal to $x_{t=1}, \cdots, x_{t=n}$ if $t = n = 3$. Hence, the brute-force mean algorithm require $O(n)$ for storage of $x_{t=1}, \cdots, x_{t=n}$. It also require $O(n + 1)$ time for summation of $x_1, \cdots, x_{n+1}$. If Bob knew an algorithm that require only $O(c)$ time and space, he would be better off using that algorithm for his calculation, as such algorithm would save Bob time and memory for use on other activities. In the case of a computer dealing with very large data $x_t, \cdots, x_{t+k}$, where each datum is available at certain times $t, \cdots, t + k$, such algorithm will not require stored data $x_t, \cdots, x_{t+k-1}$ of past times $t, \cdots, t + k - 1$ to calculate the current mean $\bar{x}_{t+k}$ at time $t + k$. It will only require the current datum x_{t+k} at current time $t + k$ and the stored mean of past data $\bar{x}_{t+k-1}$ for calculation of $\bar{x}_{t+k}$. Therefore, such algorithm runs at $O(c)$ space and time complexity.

In the proceeding sections, I derive dynamic mean algorithms that require only $O(c)$ space and time to compute arithmetic, geometric, harmonic, and power means.

2. Dynamic Arithmetic Mean

It is true that the arithmetic mean of x_t and x_{t+1}, where $t = 2$, is

$$\bar{x}_{t+1} = \frac{t\bar{x}_t + x_{t+1}}{t+1} = \frac{(2 \cdot x_2) + x_3}{3}$$

In section 3.1, I explain the method I used to derive $\frac{t\bar{x}_t + x_{t+1}}{t+1}$, and also show the derivation of the dynamic arithmetic mean.

2.1 Derivation

All derivations in this writing begins by considering statements that are contradictions. These derivations make use of a kind of proof known as proof-by-contradiction.
It is widely known that the standard formula for calculating arithmetic mean AM is

$$AM = \bar{x}_t = \frac{x_1 + \cdots + x_t}{t}$$

Where time i range from 1 to t, and x_i is an observation in the set of observations x. In the set x, there are t observations observed at the time interval 1 to t. That means, observation x_i is observed at time i. For example, if Bob received x_1 oranges at time 1, x_2 oranges at time 2, and x_3 oranges at time 3, then the group of oranges $x_i : 1 \leq \forall i \leq 3$ is in the set of x, therefore, x stores three groups of oranges (i.e., $\{x_1, x_2, x_3\} \subseteq x = \{x_i : 1 \leq \forall i \leq 3\} \subseteq x$) gotten three times: $t = 1$, $t = 2$, and $t = 3$. The symbol i can also be interpreted as the index of any arbitrary x_i in x, and t can be interpreted as the index of the last observation x_t in x.

Consider the question: If Bob know the arithmetic mean AM of $x_1, \cdots, x_t$ at time t, and later at time $t+1$, Alice gave Bob x_{t+1} oranges; is $\frac{AM + x_{t+1}}{t+1} = \frac{x_1 + \cdots + x_{t+1}}{t+1}$?

Suppose the statement $\frac{AM + x_{t+1}}{t+1} = \frac{x_1 + \cdots + x_{t+1}}{t+1}$ is a contradiction, then

$$\frac{AM + x_{t+1}}{t+1} \neq \frac{x_1 + \cdots + x_{t+1}}{t+1}$$

$$\frac{\frac{x_1 + \cdots + x_t}{t} + x_{t+1}}{t+1} \neq \frac{x_1 + \cdots + x_{t+1}}{t+1}$$

$$\frac{x_1 + \cdots + x_t}{t} + x_{t+1} \neq x_1 + \cdots + x_{t+1}$$

$$t\left(\frac{x_1 + \cdots + x_t}{t}\right) + x_{t+1} = x_1 + \cdots + x_{t+1}$$

If $AM = t\left(\frac{x_1 + \cdots + x_t}{t}\right)$ then $AM + x_{t+1} = x_1 + \cdots + x_{t+1}$, which will correct the contradiction.

$$\therefore \frac{t\left(\frac{x_1 + \cdots + x_t}{t}\right) + x_{t+1}}{t+1} = \frac{x_1 + \cdots + x_{t+1}}{t+1}$$

$$\left(\frac{t}{t+1}\right)\left(\frac{x_1 + \cdots + x_t}{t}\right) + \frac{x_{t+1}}{t+1} = \frac{x_1 + \cdots + x_{t+1}}{t+1}$$

$$\left(\frac{t}{t+1}\right)\left(\frac{x_1+\cdots+x_t}{t}\right)+\frac{x_{t+1}}{t+1}=\left(\frac{t}{t+1}\right)\left(\frac{1}{t}\sum_{i=1}^{t}x_i\right)+\frac{x_{t+1}}{t+1}$$

$$\left(\frac{t}{t+1}\right)\left(\frac{1}{t}\sum_{i=1}^{t}x_i\right)+\frac{x_{t+1}}{t+1}=\frac{1}{t+1}\sum_{i=1}^{t+1}x_i$$

$$\frac{t}{t+1}\bar{x}_t+\frac{x_{t+1}}{t+1}=\bar{x}_{t+1}$$

$$\bar{x}_{t+1}=\frac{t\bar{x}_t+x_{t+1}}{t+1}$$

$$\bar{x}_{t+k}=\frac{1}{t+k}\left[\left(\sum_{i=t+1}^{t+k}x_i\right)+t\bar{x}_t\right]$$

2.2 Algorithm

In this subsection, I show algorithms and complexities of the brute-force and dynamic arithmetic means. The brute-force algorithm is the following:

(1) bruteForceArithmeticMean (data): $\longrightarrow$ $O(n)$ space
(2) sum = 0.0
(3) for each datum in data: sum += datum $\longrightarrow$ $O(n)$ time
(4) arithmeticMean = sum / length(data)
(5) return arithmeticMean

Below is the dynamic algorithm:

(1) dynamicArithmeticMean (time, datum): $\longrightarrow$ $O(c)$ space
(2) currentArithmeticMean = $\frac{((\text{time}-1) * \text{previousArithmeticMean}) + \text{datum}}{\text{time}}$ $O(c)$ time
(3) return currentArithmeticMean

There need to be a datum generator function generateDatum(t) that provides datum at time t. The generated datum will be used by the brute-force and dynamic mean algorithms. Below is an algorithm for the datum generator:

(1) generateDatum(t):
(2) data = $x_1 \cdots x_k$
(3) return t index of data

Finally, let there be a main algorithm that call the above algorithms:

(1) main ():
(2) initialize data
(3) for range $t = 1$ to $t \leq n$:
(4) add generateDatum(t) to index t of data
(5) bruteForceArithmeticMean (data)

$O(n^2)$ time and $O(n)$ space

```
(6)   for range t = 1 to t ≤ n:
(7)       datum = generateDatum(t);
(8)       dynamicArithmeticMean (t, datum)
```

$O(n)$ time and $O(c)$ space

It is shown that the dynamicArithmeticMean algorithm run more efficiently than the bruteForceArithmeticMean. Similar algorithms like the dynamicArithmeticMean can be adapted to compute the dynamic geometric, harmonic, and power means. To compute either the dynamic geometric, harmonic, or power mean, change the mean formula at line (2) of the dynamicArithmeticMean algorithm to the appropriate dynamic mean formulas shown in section 4, 5, and 6. The dynamic algorithms still remain more efficient for computing geometric, harmonic, and power means.

Java implementation of these algorithms can be found at *github.com/calebnwokocha/Dynamic-Means*.

3. Dynamic Geometric Mean

It is true that the geometric mean of x_t and x_{t+1}, where $t = 2$, is

$$\bar{x}_{t+1} = ((\bar{x}_t)^t x_{t+1})^{\frac{1}{t+1}} = ((x_2)^2 \cdot x_3)^{\frac{1}{3}}$$

3.1 Derivation

It is well known that the standard formula for calculating geometric mean GM is

$$GM = \bar{x}_t = (x_1 \cdots x_t)^{\frac{1}{t}}$$

Where $x_i; 1 \leq i \leq t$ represents quantity of oranges Bob receives.

I consider the question: If Bob know the geometric mean GM of $x_1, \cdots, x_t$ at time t, and later at time $t + 1$, Alice gave Bob x_{t+1} oranges; is $(GM \cdot x_{t+1})^{\frac{1}{t+1}} = (x_1 \cdots x_{t+1})^{\frac{1}{t+1}}$?

Suppose the statement $(GM \cdot x_{t+1})^{\frac{1}{t+1}} = (x_1 \cdots x_{t+1})^{\frac{1}{t+1}}$ is a contradiction, then

$$(GM \cdot x_{t+1})^{\frac{1}{t+1}} \neq (x_1 \cdots x_{t+1})^{\frac{1}{t+1}}$$

$$\left((x_1 \cdots x_t)^{\frac{1}{t}} \cdot x_{t+1}\right)^{\frac{1}{t+1}} \neq (x_1 \cdots x_{t+1})^{\frac{1}{t+1}}$$

$$(x_1 \cdots x_t)^{\frac{1}{t}} \cdot x_{t+1} \neq x_1 \cdots x_{t+1}$$

$$\left[(x_1 \cdots x_t)^{\frac{1}{t}}\right]^t \cdot x_{t+1} = x_1 \cdots x_{t+1}$$

If $GM = \left[(x_1 \cdots x_t)^{\frac{1}{t}}\right]^t$ then $GM \cdot x_{t+1} = x_1 \cdots x_{t+1}$, which will correct the contradiction.

$$\therefore \left(\left[(x_1 \cdots x_t)^{\frac{1}{t}}\right]^t \cdot x_{t+1}\right)^{\frac{1}{t+1}} = (x_1 \cdots x_{t+1})^{\frac{1}{t+1}}$$

$$\left(\left[\left(\prod_{i=1}^{t} x_i\right)^{\frac{1}{t}}\right]^t \cdot x_{t+1}\right)^{\frac{1}{t+1}} = \left(\prod_{i=1}^{t+1} x_i\right)^{\frac{1}{t+1}}$$

$$((\bar{x}_t)^t \cdot x_{t+1})^{\frac{1}{t+1}} = \bar{x}_{t+1}$$

$$\bar{x}_{t+1} = ((\bar{x}_t)^t x_{t+1})^{\frac{1}{t+1}}$$

$$\bar{x}_{t+k} = \left((\bar{x}_t)^t \prod_{i=t+1}^{t+k} x_i\right)^{\frac{1}{t+k}}$$

4. Dynamic Harmonic Mean

It is true that the harmonic mean of x_t and x_{t+1}, where $t = 2$, is

$$\bar{x}_{t+1} = t + 1\left(\frac{t}{\bar{x}_t} + \frac{1}{x_{t+1}}\right)^{-1} = 3\left(\frac{2}{x_2} + \frac{1}{x_3}\right)^{-1}$$

4.1 Derivation

The standard formula of harmonic mean HM is the following:

$$HM = \bar{x}_t = \frac{t}{\frac{1}{x_1} + \cdots + \frac{1}{x_t}}$$

Where $x_i; 1 \leq i \leq t$ represents quantity of oranges Bob receives.

I consider the question: If Bob know the harmonic mean HM of $x_1, \cdots, x_t$ at time t, and later at time $t + 1$, Alice gave Bob x_{t+1} oranges; is $\frac{t+1}{HM+\frac{1}{x_{t+1}}} = \frac{t+1}{\frac{1}{x_1}+\cdots+\frac{1}{x_{t+1}}}$?

Suppose the statement $\frac{t+1}{HM+\frac{1}{x_{t+1}}} = \frac{t+1}{\frac{1}{x_1}+\cdots+\frac{1}{x_{t+1}}}$ is a contradiction, then

$$\frac{t+1}{HM + \frac{1}{x_{t+1}}} \neq \frac{t+1}{\frac{1}{x_1} + \cdots + \frac{1}{x_{t+1}}}$$

$$\frac{t+1}{\frac{t}{\frac{1}{x_1} + \cdots + \frac{1}{x_t}} + \frac{1}{x_{t+1}}} \neq \frac{t+1}{\frac{1}{x_1} + \cdots + \frac{1}{x_{t+1}}}$$

$$\frac{t}{\frac{1}{x_1} + \cdots + \frac{1}{x_t}} + \frac{1}{x_{t+1}} \neq \frac{1}{x_1} + \cdots + \frac{1}{x_{t+1}}$$

$$\left[\left(\frac{t}{\frac{1}{x_1}+\cdots+\frac{1}{x_t}}\right)\left(\frac{1}{t}\right)\right]^{-1}+\frac{1}{x_{t+1}}=\frac{1}{x_1}+\cdots+\frac{1}{x_{t+1}}$$

If $HM=\left[\left(\frac{t}{\frac{1}{x_1}+\cdots+\frac{1}{x_t}}\right)\left(\frac{1}{t}\right)\right]^{-1}$ then $HM+\frac{1}{x_{t+1}}=\frac{1}{x_1}+\cdots+\frac{1}{x_{t+1}}$, which will correct the contradiction.

$$\therefore \frac{t+1}{\left[\left(\frac{t}{\frac{1}{x_1}+\cdots+\frac{1}{x_t}}\right)\left(\frac{1}{t}\right)\right]^{-1}+\frac{1}{x_{t+1}}}=\frac{t+1}{\frac{1}{x_1}+\cdots+\frac{1}{x_{t+1}}}$$

$$t+1\left(\left[t\left(\sum_{i=1}^{t}\frac{1}{x_i}\right)^{-1}\left(\frac{1}{t}\right)\right]^{-1}+\frac{1}{x_{t+1}}\right)^{-1}=t+1\left(\sum_{i=1}^{t+1}\frac{1}{x_i}\right)^{-1}$$

$$t+1\left(\left[\bar{x}_t\left(\frac{1}{t}\right)\right]^{-1}+\frac{1}{x_{t+1}}\right)^{-1}=\bar{x}_{t+1}$$

$$\bar{x}_{t+1}=t+1\left(\frac{t}{\bar{x}_t}+\frac{1}{x_{t+1}}\right)^{-1}$$

$$\bar{x}_{t+k}=t+k\left(\frac{t}{\bar{x}_t}+\sum_{i=t+1}^{t+k}\frac{1}{x_i}\right)^{-1}$$

5. Dynamic Power Mean

It is true that the power mean of x_t and x_{t+1}, where $t=2$ and $m=1$, is

$$\bar{x}_{t+1}=\left(\frac{\bar{x}_t^m t+x_{t+1}^m}{t+1}\right)^{\frac{1}{m}}=\left(\frac{(x_2^1\cdot 2)+x_3^1}{3}\right)^{\frac{1}{1}}$$

5.1 Derivation

The standard formula of power mean PM is the following:

$$PM=\bar{x}_t=\left(\frac{x_1^m+\cdots+x_t^m}{t}\right)^{\frac{1}{m}}$$

Where $x_i; 1\le i\le t$ represents quantity of oranges Bob receives.

I consider the question: If Bob know the power mean PM of $x_1,\cdots,x_t$ at time t, and later at time $t+1$, Alice gave Bob x_{t+1} oranges; is $\left(\frac{PM+x_{t+1}^m}{t+1}\right)^{\frac{1}{m}}=\left(\frac{x_1^m+\cdots+x_{t+1}^m}{t+1}\right)^{\frac{1}{m}}$?

Suppose the statement $\left(\frac{PM+x_{t+1}^m}{t+1}\right)^{\frac{1}{m}} = \left(\frac{x_1^m+\cdots+x_{t+1}^m}{t+1}\right)^{\frac{1}{m}}$ is a contradiction, then

$$\left(\frac{PM + x_{t+1}^m}{t+1}\right)^{\frac{1}{m}} \neq \left(\frac{x_1^m + \cdots + x_{t+1}^m}{t+1}\right)^{\frac{1}{m}}$$

$$\left(\frac{\left(\frac{x_1^m + \cdots + x_t^m}{t}\right)^{\frac{1}{m}} + x_{t+1}^m}{t+1}\right)^{\frac{1}{m}} \neq \left(\frac{x_1^m + \cdots + x_{t+1}^m}{t+1}\right)^{\frac{1}{m}}$$

$$\left(\frac{x_1^m + \cdots + x_t^m}{t}\right)^{\frac{1}{m}} + x_{t+1}^m \neq x_1^m + \cdots + x_{t+1}^m$$

$$t\left(\left[\left(\frac{x_1^m + \cdots + x_t^m}{t}\right)^{\frac{1}{m}}\right]^m\right) + x_{t+1}^m = x_1^m + \cdots + x_{t+1}^m$$

If $PM = t\left(\left[\left(\frac{x_1^m+\cdots+x_t^m}{t}\right)^{\frac{1}{m}}\right]^m\right)$ then $PM + x_{t+1}^m = x_1^m + \cdots + x_{t+1}^m$, which will correct the contradiction.

$$\therefore \left(\frac{t\left(\left[\left(\frac{x_1^m + \cdots + x_t^m}{t}\right)^{\frac{1}{m}}\right]^m\right) + x_{t+1}^m}{t+1}\right)^{\frac{1}{m}} = \left(\frac{x_1^m + \cdots + x_{t+1}^m}{t+1}\right)^{\frac{1}{m}}$$

$$\left(\left[t\left(\left[\left(\frac{1}{t}\sum_{i=1}^{t} x_i^m\right)^{\frac{1}{m}}\right]^m\right) + x_{t+1}^m\right] \cdot (t+1)^{-1}\right)^{\frac{1}{m}} = \left(\frac{1}{t+1}\sum_{i=1}^{t+1} x_i^m\right)^{\frac{1}{m}}$$

$$([\bar{x}_t^m t + x_{t+1}^m] \cdot (t+1)^{-1})^{\frac{1}{m}} = \bar{x}_{t+1}$$

$$\bar{x}_{t+1} = \left(\frac{\bar{x}_t^m t + x_{t+1}^m}{t+1}\right)^{\frac{1}{m}}$$

$$\bar{x}_{t+k} = \left(\frac{1}{t+k}\left[\left(\sum_{i=t+1}^{t+k} x_i^m\right) + \bar{x}_t^m t\right]\right)^{\frac{1}{m}}$$

Note that $x_1, \cdots, x_t$ is data Bob has about quantities of oranges he received from Alice, and x_i is a datum in x.

6. Conclusion

I have shown derivations of Pythagorean and generalized power mean algorithms designed by dynamic programming. The dynamic arithmetic mean algorithm implement the formula:

$$\bar{x}_{t+k} = \frac{1}{t+k}\left[\left(\sum_{i=t+1}^{t+k} x_i\right) + t\bar{x}_t\right]$$

The dynamic geometric mean algorithm implement the formula:

$$\bar{x}_{t+k} = \left((\bar{x}_t)^t \prod_{i=t+1}^{t+k} x_i\right)^{\frac{1}{t+k}}$$

The dynamic harmonic mean algorithm implement the formula:

$$\bar{x}_{t+k} = t + k\left(\frac{t}{\bar{x}_t} + \sum_{i=t+1}^{t+k} \frac{1}{x_i}\right)^{-1}$$

And finally, the dynamic power mean algorithm implement the formula:

$$\bar{x}_{t+k} = \left(\frac{1}{t+k}\left[\left(\sum_{i=t+1}^{t+k} x_i^m\right) + \bar{x}_t^m t\right]\right)^{\frac{1}{m}}$$

Dynamic mean algorithms run more efficiently than brute-force mean algorithms when computing Pythagorean and generalized power means of increasing data. Data is said to be increasing if its datum acquisition depends on change in t.

References

[1] Moore, D. S., McCabe, G. P., & Craig, B. A. (2009). Looking at Data – Distributions. In *Introduction to the practice of statistics.* Seventh edition. New York: W.H. Freeman.

[2] Geometric Mean. (2021, July 28). https://stats.libretexts.org/@go/page/4551

[3] *Statistical Analysis*, Ya-lun Chou, Holt International, 1969, ISBN 0030730953.

[4] http://www.scientificlib.com/en/Mathematics/Analysis/HarmonicMean.html

[5] Sykora, Stan. (2009). Mathematical Means and Averages: Basic Properties. 10.3247/SL3Math09.001.