mathematics

Article

Minimal Impact One-Dimensional Arrays

Leo Egghe [1] **and Ronald Rousseau** [2,3,*]

1 University of Hasselt, 3500 Hasselt, Belgium; leo.egghe@uhasselt.be
2 Faculty of Social Sciences, University of Antwerp, 2020 Antwerpen, Belgium
3 Department MSI, KU Leuven and Centre for R&D Monitoring (ECOOM), 3000 Leuven, Belgium
* Correspondence: ronald.rousseau@uantwerpen.be or ronald.rousseau@kuleuven.be

Received: 15 April 2020; Accepted: 13 May 2020; Published: 17 May 2020

Abstract: In this contribution, we consider the problem of finding the minimal Euclidean distance between a given converging decreasing one-dimensional array X in $(\mathbf{R}^{+})^{\infty}$ and arrays of the form $A_a = \left(\underbrace{a, a, \ldots, a}_{a\ times}, 0, 0, \ldots\right)$, with a being a natural number. We find a complete, if not always unique, solution. Our contribution illustrates how a formalism derived in the context of research evaluation and informetrics can be used to solve a purely mathematical problem.

Keywords: generalized h-index; generalized g-index; minimization problem

---

1. Introduction

Let $(\mathbf{R}^{+})^{\infty}$ be the positive cone of all infinite sequences with non-negative real values. Elements in this cone will be referred to as one-dimensional arrays, in short, arrays. We recall that any finite sequence with non-negative values can be considered as an element in $(\mathbf{R}^{+})^{\infty}$ by adding infinitely many zeros. Let X = $(x_r)_{r=1,2\ldots}$ and Y = $(y_r)_{r=1,2\ldots}$ be elements of $(\mathbf{R}^{+})^{\infty}$, then X ≤ Y if for all r = 1, 2, …, $x_r \le y_r$. Equality only occurs if for all r, $x_r = y_r$. In this way, $(\mathbf{R}^{+})^{\infty}$ becomes a cone with a (natural) partial order ≤. An array X = $(x_r)_{r=1,2\ldots}$ in $(\mathbf{R}^{+})^{\infty}$ is said to be decreasing if for all r = 1, 2, …, $x_r \ge x_{r+1}$.

We recall the definition of the h-index as introduced by Hirsch [1]. Consider, $(c_r)_{r=1,\ldots,R}$, the list of received citations of the articles (co-) authored by scientist S, ranked according to the number of citations each of these articles has received. Articles with the same number of citations are given different rankings. Then, the h-index of scientist S is h if the first h articles each received at least h citations, while the article ranked h + 1 received strictly less than h + 1 citations. Stated otherwise, scientist S' h-index is h if h is the largest natural number such that the first h publications each received at least h citations.

This index, although having many disadvantages in practical use ([2,3]), has received a lot of attention. At this moment [1], it has received already more than 4300 citations in the Web of Science. Because of these disadvantages, many alternatives have been proposed, among which the most popular is the g-index, introduced and studied by Egghe [4]. This g-index is defined as follows: as with the calculation of the h-index, articles are ranked in decreasing order of the number of citations received; then, the g-index of this set of articles is defined as the highest rank, g, such that the first g articles together received at least g^2 citations. This can be reformulated as follows: the g-index of a set of articles is the highest-rank g such that the first g (>0) articles have an average number of citations equal to or higher than g. Indeed, $\sum_{j=1}^{g} c_j \ge g^2 \Leftrightarrow \frac{1}{g}\sum_{j=1}^{g} c_j \ge g$. For more information on the h-index and related indices, we refer to [5–7].

In [8], we defined the h- and the g-index for infinite sequences as follows:

Definition 1. *The h-index for infinite sequences:*

Let X =$(x_r)_{r=1,2\ldots}$ be a decreasing array in **$(\mathbf{R}^+)^\infty$**. *The h-index of X, denoted h(X), is the largest natural number h such that the first h coordinates each have at least a value h. If all components of a decreasing array X are strictly smaller than 1, then h(X) = 0. We will further consider only arrays X with at least one component larger than or equal to 1, hence with h(X) ≥ 1.*

Note that an h-index is defined here only for decreasing arrays (although a generalization exists, see [9]). The same remark is valid for the other indices used in this article.

Similarly, a g-index has been defined in [8] as follows:

Definition 2. *The g-index for infinite sequences:*

Let X =$(x_r)_{r=1,2\ldots}$ be a decreasing array in **$(\mathbf{R}^+)^\infty$**. *The g-index of X, denoted g_X, is defined as the highest natural number g such that the sum of the first g coordinates is at least equal to g^2 or, equivalently, if the average of the first g coordinates is at least equal to g.*

Notation. We denote by [[a,b]] for a, b natural numbers such that a ≤ b, the intersection of the real-valued interval [a,b] and **N**, *the set of natural numbers.*

2. Introducing the Research Problem

Definition 3. *For each natural number a > 0, we define the minimal impact array of level a, denoted as A_a, as follows:*

$$A_a = \left(\underbrace{a, a, \ldots, a}_{a\ times},\ 0, 0,\ \ldots \right)$$

It is easy to see that A_a is the smallest array X (for the partial order ≤) for which h(X) = g(X) = a. We note that the sequence $(A_n)_n$ is increasing for ≤.

We say that an array X is l^2-converging if $\sum_{i=1}^{\infty} x_i^2$ is finite. As we only use this form of convergence, we will further on omit the specification "l^2" and simply say converging.

Next, we formulate the research problem of this contribution.

Research Problem

Given a converging decreasing array X in **$(\mathbf{R}^+)^\infty$**, find the largest natural number a such that the Euclidean distance $d(X,A_a)$ is minimal.

We note that the analogous problem for differentiable functions Z(r) and a real number a has already been studied and solved in [10]. We further note that the requirements to be decreasing and convergent are independent. Indeed, if a decreasing array is convergent and we add its sum (or a larger number) to any term, except the first, then the resulting array is still convergent but not decreasing anymore. Further, the array with terms $\frac{1}{\sqrt{n}}$ is decreasing but not convergent.

Minimizing $d(X,A_a)$ is the same as finding a minimal value for the function

$$f_X : \mathbf{N_0} \to \mathbf{R} : a \to f_X(a)$$

where $\mathbf{N_0}$ denotes the set of natural number without zero and

$$f_X(a) = d^2(X, A_a) = \sum\nolimits_{i=1}^{a} (x_i - a)^2 + \sum\nolimits_{i=a+1}^{\infty} x_i^2 = \sum\nolimits_{i=1}^{\infty} x_i^2 - 2a\left(\sum\nolimits_{i=1}^{a} x_i\right) + a^3. \qquad (1)$$

Equation (1) shows why we need convergent arrays. Note also that a minimal value a depends on X. Hence, we write it as a_X. It is trivial to see that if X = A_b for some natural value b, then b = a_X (for this X) and f(b) = 0. It is clear that arrays X of the form A_b are the only ones for which the corresponding function f_X becomes zero.

This leads us to the following questions:

1. Does a_X exist for each X, converging and decreasing in $(\mathbf{R}^+)^\infty$?
2. Given X, converging and decreasing in $(\mathbf{R}^+)^\infty$, how do we find a_X (if it exists)?
3. If a_X exists, is it unique?

3. Results

3.1. Characterizing the Minimum of f_X

Taking into account that a_X is possibly not unique for some X, we want to characterize a_X—if it exists and is strictly larger than 1—as the largest natural number such that

$$f_X(a_X - 1) \geq f_X(a_X) \tag{2}$$

Note that if the minimum of f_X occurs in two (or more) natural numbers, we choose the largest one. We still have to show that inequality (2) actually characterizes the minimum we are searching for. Indeed, theoretically, it may happen that the function $f_X(a)$ decreases first to a (local) minimum b, then increases again, and then decreases to a lower minimum value than the one in b. This might, in theory, even occur infinitely many times. We will prove that this behavior does not occur. Moreover, if we want to use inequality (2), we first have to deal with the case $a_X = 1$, as this case is not covered by inequality (2).

Remark 1. *We first note that if $h_X \geq 2$, then certainly $a_X > 1$. Indeed, if $h_X \geq 2$, then $x_1 \geq x_2 \geq 2$. Then, $f_X(1) = \sum_{i=1}^{\infty} x_i^2 - 2x_1 + 1 < f_X(2) = \sum_{i=1}^{\infty} x_i^2 - 4(x_1 + x_2) + 8$ is equivalent to $1 < 8 - 2x_1 - 4x_2$ or $2x_1 + 4x_2 < 7$. This inequality never holds; hence, the minimum of f_X does not occur in 1. We conclude that $a_X = 1$ can only occur if $h_X = 1$.*

If $a_X = 1$, then $f_X(1) < f_X(2)$. This inequality is equivalent to $-2x_1 + 1 < -4(x_1 + x_2) + 8$ or $2x_1 + 4x_2 < 7$ or $2x_1 + 4x_2 - 7 < 0$.

Taken the constraints $x_1 \geq 1$ and $x_1 \geq x_2$ into account yields the following area (see Figure 1) in which $a_X = 1$. This is the area R situated within the polygon with vertices (1,0), (1,1), (7/6,7/6), (7/2,0), where points on the line $2x_1 + 4x_2 - 7 = 0$ are excluded. We note that for all points in this area, $h_X = 1$.

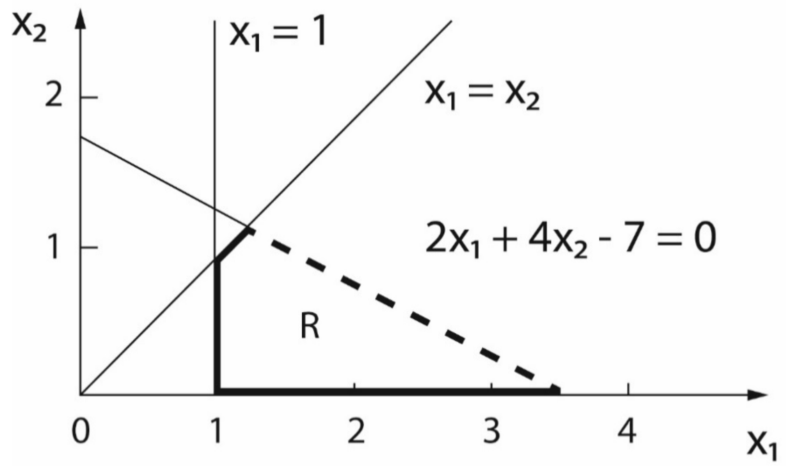

Figure 1. Zone R in (x_1,x_2)-plane where $a_X = 1$.

When it comes to arrays X for which $a_X = 1$, this set consists of all decreasing convergent arrays with (x_1,x_2) in the area R.

3.2. The Generalized Discrete h- and g-Index

We next show that a_X exists for each converging and decreasing X in $(\mathbf{R}^+)^\infty$. For this, we recall the definitions of the generalized discrete h- and g-index [11].

Definition 4. *The generalized discrete h- and g-index [11]:*
Given X, a decreasing array in $(\mathbf{R}^+)^\infty$*. Let* $\theta > 0$*, then*

(1) $z = h_\theta(X)$, *in short* h_θ, *iff z is the largest index such that* $x_z \geq \theta z$*; if such an index does not exist, namely when* $x_1 < \theta$*, then we define* $z = h_\theta(X) = 0$*;*
(2) $z = g_\theta(X)$, *in short* g_θ, *iff z is the largest index such that* $\sum_{i=1}^{z} x_i \geq \theta z^2 \Leftrightarrow \frac{1}{z} \sum_{i=1}^{z} x_i \geq \theta z$ *; if such an index does not exist, e.g., when* $\sum_{i=1}^{\infty} x_i < \theta$*, then we define* $z = g_\theta(X) = 0$*.*

We note that if a and b are natural numbers and $X = (x_r)_{r=1,2...}$ is a decreasing array in $(\mathbf{R}^+)^\infty$, then the property for $r \leq a$: $\sum_{i=1}^{r} x_i \geq \theta a^2$ implies that $g_\theta(X) \geq a$; similarly, the property for $r > b$: $\sum_{i=1}^{r} x_i < \theta b^2$ implies $g_\theta(X) \leq b$.

We finally also define the discrete f-index, already introduced in [10], for the continuous case.

(3) $z = f_\theta(X)$, *in short* f_θ, *if z is the largest index such that* $\frac{1}{2}\left(x_z + \frac{1}{z} \sum_{i=1}^{z} x_i\right) \geq \theta z$*. Again, if such an index does not exist, we define* $z = f_\theta(X) = 0$*.*

In [10], we found that in the continuous case, the solution of our problem was obtained as $f_{(3/4)}(X)$ (where f is the continuous analog of the discrete f-index introduced above). We will show further on that this is not the case for the discrete case studied here.

Proposition 1. *The indicators* $h_\theta(X)$*,* $g_\theta(X)$*, and* $f_\theta(X)$ *are each decreasing in* θ*.*

Proof of Proposition 1. Let $\theta_1 > \theta_2$. If $z_1 = h_{\theta_1}(X)$ and $z_2 = h_{\theta_2}(X)$, then $x_{z_1} \geq \theta_1 z_1 > \theta_2 z_1$. As z_2 is the largest index such that $x_{z_2} > \theta_2 z_2$, it follows that $z_1 = h_{\theta_1}(X) \leq z_2 = h_{\theta_2}(X)$. Consequently, $\theta_1 > \theta_2$ implies $h_{\theta_1}(X) \leq h_{\theta_2}(X)$, showing that $h_\theta(X)$ is decreasing in θ.

Similarly, if $\theta_1 > \theta_2$, $z_1 = g_{\theta_1}(X)$, and $z_2 = g_{\theta_2}(X)$, then $\frac{1}{z_1} \sum_{i=1}^{z_1} x_i \geq \theta_1 z_1 > \theta_2 z_1$. As z_2 is the largest index such that $\frac{1}{z_2} \sum_{i=1}^{z_2} x_i > \theta_2 z_2$, it follows, like in the case for the generalized discrete h-index, that $g_\theta(X)$ is decreasing in θ. Finally, it also follows that $f_\theta(X)$ is decreasing in θ. □

Theorem 1. *For all X, decreasing in* $(\mathbf{R}^+)^\infty$ *and all* $\theta > 0$*,* $h_\theta(X) \leq f_\theta(X) \leq g_\theta(X)$*. Hence,* $f_\theta(X) \in [[h_\theta(X), g_\theta(X)]]$*.*

Proof of Theorem 1. Let $a = f_\theta(X)$, then, by the definition of $f_\theta(X)$,

$$\frac{1}{2}(x_{a+1} + x_{a+1}) \leq \frac{1}{2}\left(x_{a+1} + \frac{1}{a+1}\sum_{i=1}^{a+1} x_i\right) < \theta(a+1)$$

Hence, $x_{a+1} < \theta(a+1)$ and thus $a+1 > h_\theta(X)$, leading to $a = f_\theta(X) \geq h_\theta(X)$. Now, for $a = f_\theta(X)$, we further have $\frac{1}{a}\sum_{i=1}^{a} x_i \geq \frac{1}{2}\left(x_a + \frac{1}{a}\sum_{i=1}^{a} x_i\right) \geq \theta a$, hence $a \leq g_\theta(X)$. This proves Theorem 1. □

3.3. Excluding the Theoretical Case of Infinitely Many Minima

Next, we need two lemmas.

Lemma 1. *If X is decreasing, then* $\overline{\overline{X}}$ *with* $\left(\overline{\overline{X}}\right)_i = \frac{1}{i}\sum_{j=1}^{i} x_j$ *is also decreasing. This decrease is strict if* $x_1 > x_2$*.*

Proof of Lemma 1. The easy proof is left to the reader. □

As for given X and n > 0, $x_n = \theta n$, for $\theta = \frac{x_n}{n}$, it is clear that $\{h_\theta(X),\ \theta > 0\} = \mathbf{N}$ (where n = 0 is reached for $\theta > x_1$). Now, we prove a similar result for $g_\theta(X)$.

Lemma 2. *For given X, decreasing and convergent,* $\{g_\theta(X),\ \theta > 0\} = \boldsymbol{N}$

Proof of Lemma 2. It is clear that $\{g_\theta(X),\ \theta > 0\} \subset \mathbf{N}$ (recall that $g_\theta(X) = 0$ if $\theta > \sum_{j=1}^{\infty} x_j$). Next, we consider the opposite relation.

The value $n = 0$ results from all $\theta > \sum_{j=1}^{\infty} x_j$. If $n \neq 0$, we define $\theta = \frac{1}{n^2}\sum_{j=1}^{n} x_j > 0$. Then, we have, using Lemma 1, for all $i \leq n$, $\frac{1}{i}\sum_{j=1}^{i} x_j \geq \frac{1}{n}\sum_{j=1}^{n} x_j = \theta n$. Consequently,

$$\sum_{j=1}^{i} x_j \geq \theta n i \geq \theta i^2 \tag{3}$$

Now, for all $i > n$, using Lemma 1 again, $\frac{1}{i}\sum_{j=1}^{i} x_j \leq \frac{1}{n}\sum_{j=1}^{n} x_j = \theta\, n < \theta i$ and hence

$$\sum_{j=1}^{i} x_j < \theta\, i^2 \tag{4}$$

It follows from (3) and (4) and the definition of g_θ that $n = g_\theta$. This shows that $\{g_\theta(X),\ \theta > 0\} = \mathbf{N}$. □

Theorem 2. *Given X is decreasing and convergent and $a > g_{(0.5)}(X)$, then $f_X(x)$ is strictly increasing for $x > a$.*

Proof of Theorem 2. From Lemma 2, it follows that there exists $\theta_0 < 0.5$ such that $a = g_{\theta_0}(X)$. Indeed, $g_\theta(X)$ is a decreasing function of θ and $a > g_{(0.5)}(X)$. Hence

$\sum_{i=1}^{a} x_i \geq \theta_0 a^2$ and $\sum_{i=1}^{a+1} x_i < \theta_0 (a+1)^2$. From this inequality, we derive that $a^3 - 2a\sum_{i=1}^{a} x_i \leq a^3 - 2a\theta_0 a^2 = a^3(1 - 2\theta_0)$ and $(a+1)^3 - 2(a+1)\sum_{i=1}^{a+1} x_i > (a+1)^3 - 2(a+1)\theta_0(a+1)^2 = (a+1)^3(1 - 2\theta_0)$. Consequently, $(a+1)^3 - 2(a+1)\sum_{i=1}^{a+1} x_i > a^3 - 2a\sum_{i=1}^{a} x_i$, which shows that $f_X(x)$ is strictly increasing for $x > g_{(0.5)}(X)$. □

It follows from Theorem 2 that if a_X exists, it belongs to $[[1, g_{(0.5)}(X)]]$, which excludes the theoretical case of infinitely many minima.

3.4. Excluding the Case of More Than One Minimum

Next, to exclude the case of a local maximum, following a first local minimum, we continue as follows.

Theorem 3. *For all X, decreasing and convergent in $(\boldsymbol{R}^+)^\infty$ and for all $a \in \mathbf{N_0}$, we have the following property:*

$$f_X(a + 1) > f_X(a) \text{ implies that } f_X(a + 2) > f_X(a + 1).$$

Proof of Theorem 3.

$$
\begin{aligned}
&\mathrm{fx}(a+1) > \mathrm{fx}(a) \\
\Leftrightarrow &\sum_{i=1}^{\infty} x_i^2 - 2(a+1)\left(\sum_{i=1}^{a+1} x_i\right) + (a+1)^3 > \sum_{i=1}^{\infty} x_i^2 - 2a\left(\sum_{i=1}^{a} x_i\right) + a^3 \\
\Leftrightarrow &\, 2(a+1)\left(\sum_{i=1}^{a+1} x_i\right) - 2a\left(\sum_{i=1}^{a+1} x_i - x_{a+1}\right) < (a+1)^3 - a^3 = 3a^2 + 3a + 1 \\
\Leftrightarrow &\, 2\left(\sum_{i=1}^{a+1} x_i\right) + 2ax_{a+1} < 3a^2 + 3a + 1
\end{aligned}
\tag{5}
$$

From (5), we also note that $(2(a+1)+2a)x_{a+1} = (4a+2)\, x_{a+1} < 3a^2 + 3a + 1$, or $x_{a+1} < a+1$.
Now, we have to show that

$$
\begin{aligned}
&\mathrm{fx}(a+2) > \mathrm{fx}(a+1) \\
\Leftrightarrow &\sum_{i=1}^{\infty} x_i^2 - 2(a+2)\left(\sum_{i=1}^{a+2} x_i\right) + (a+2)^3 > \sum_{i=1}^{\infty} x_i^2 - 2(a+1)\left(\sum_{i=1}^{a+1} x_i\right) + (a+1)^3 \\
\Leftrightarrow &\, 2\left((a+2)\left(\sum_{i=1}^{a+2} x_i\right) - (a+1)\left(\sum_{i=1}^{a+1} x_i\right)\right) < (a+2)^3 - (a+1)^3 = 3a^2 + 9a + 7 \\
\Leftrightarrow &\, 2\left((a+2)\left(\sum_{i=1}^{a+1} x_i + x_{a+2}\right) - (a+1)\left(\sum_{i=1}^{a+1} x_i\right)\right) < 3a^2 + 9a + 7 \\
\Leftrightarrow &\, 2\left(\sum_{i=1}^{a+1} x_i\right) + 2(a+2)x_{a+2} < 3a^2 + 9a + 7
\end{aligned}
$$

We rewrite the left-hand side of this inequality as

$$\left(2\sum_{i=1}^{a+1} x_i + 2ax_{a+1}\right) - 2ax_{a+1} + 4x_{a+2} + 2ax_{a+2}$$

Because of (5), we know that this expression is smaller than

$$\left(3a^2 + 3a + 1\right) + (2ax_{a+2} - 2ax_{a+1}) + 4x_{a+2} \tag{6}$$

As X is decreasing and hence $2ax_{a+1} \geq 2ax_{a+2}$, the expression (6) is smaller than or equal to

$$\left(3a^2 + 3a + 1\right) + 4x_{a+1} \tag{7}$$

Finally, because the note after inequality (5), expression (7) is smaller than

$$\left(3a^2 + 3a + 1\right) + 4(a+1) = 3a^2 + 7a + 5 \tag{8}$$

Finally, we see that

$$3a^2 + 7a + 5 < 3a^2 + 9a + 7 \tag{9}$$

which proves this theorem. □

Theorem 3 shows that a minimum for $f_X(x)$ exists and that a_X is uniquely defined. We note, however, that the minimum of f_X is not always unique. Indeed, the following example gives a case where there are two minima.

Let X = (5, 4.5). Then, $f_X(1) = 45.25 - 10 + 1 = 36.25$; $f_X(2) = 45.25 - 38 + 8 = 15.25$; $f_X(3) = 45.25 - 57 + 27 = 15.25$; $f_X(4) = 45.25 - 76 + 64 = 33.25$. Hence, $a_X = 3$.

From the previous proof, we know that $f_X(a+1) > f_X(a)$ implies $x_{a+1} < a + 1$. Hence, $a + 1 > h(X)$. Hence, $f_X(a)$ is decreasing on [[1, ..., h(X)]]. The next proposition shows that this is actually a strict decrease.

Proposition 2. *If h(X) > 1, then $f_X(a)$ is strictly decreasing for a in [[1, ..., h(X)]].*

Note that the requirement h(X) > 1 implies that $a_X > 1$.

Proof of Proposition 2. For any natural number a, such that a+1 ≤ h(X), we have $x_1 \geq \ldots x_a \geq x_{a+1} \geq a + 1$. Consequently, for all j = 1, ..., a + 1, $x_j - a - 1 \geq 0$.

Hence, $(x_j - a) > (x_j - a - 1) \geq 0$ and thus $(x_j - a)^2 > (x_j - a - 1)^2$.

Now, $f_X(a+1) = \sum_{i=1}^{a+1}(x_i - a - 1)^2 + \sum_{i=a+2}^{\infty} x_i^2 < \sum_{i=1}^{a+1}(x_i - a)^2 + \sum_{i=a+2}^{\infty} x_i^2 = \sum_{i=1}^{a}(x_i - a)^2 + (x_{a+1} - a)^2 + \sum_{i=a+1}^{\infty} x_i^2 - x_{a+1}^2 < \sum_{i=1}^{a}(x_i - a)^2 + \sum_{i=a+1}^{\infty} x_i^2 = f_X(a)$. Indeed, $(x_{a+1} - a)^2 - (x_{a+1})^2 = -2ax_{a+1} + a^2 = a(a - 2x_{a+1}) < 0$, as $2x_{a+1} \geq x_{a+1} \geq a + 1 > a$. □

As $h(X) \leq g(X) \leq g_{(0.5)}(X)$, this result shows that $a_X \in [\![h(X),\ g_{(0.5)}(X)]\!]$.

We next reformulate inequality (2), leading to a refinement of the previous observation.

Theorem 4. *Given an array X, converging and decreasing in* $\boldsymbol{(R^+)^\infty}$, *then a_X (≠1) is characterized as the largest natural number that satisfies the following inequality:*

$$x_{a_X} + \frac{1}{a_X - 1}\sum_{i=1}^{a_X} x_i \geq \frac{3a_X}{2} + \frac{1}{2(a_X - 1)}$$

Proof of Theorem 4. From Equations (1) and (2), we have

$$\begin{aligned} -2(a_X - 1)\left(\sum_{i=1}^{a_X-1} x_i\right) + (a_X - 1)^3 &\geq -2a_X\left(\sum_{i=1}^{a_X} x_i\right)a_X^3 \\ <=> -2(a_X - 1)\left(\sum_{i=1}^{a_X-1} x_i\right) + (a_X - 1)^3 &\geq -2(a_X - 1)\left(\sum_{i=1}^{a_X-1} x_i\right) - 2\left(\sum_{i=1}^{a_X} x_i\right) - 2(a_X - 1)a_X + a_X^3 \\ <=> (a_X - 1)^3) &\geq -2\left(\sum_{i=1}^{a_X} x_i\right) - 2(a_X - 1)a_X + a_X^3 \\ <=> 2\left(\sum_{i=1}^{a_X} x_i\right) + 2(a_X - 1)X_a &\geq 3a_X(a_X - 1) + 1 \\ <=> x_{a_X} + \frac{1}{a_X - 1}\sum_{i=1}^{a_X} x_i &\geq \frac{3a_X}{2} + \frac{1}{2(a_X - 1)} \end{aligned}$$

□

Theorem 5. *If $f_{(3/4)}(X) > 1$, then, $f_{(3/4)}(X) \leq a_X$ and hence $a_X \in [\![f_{(\frac{3}{4})}(X),\ g_{(0.5)}(X)]\!]$.*

Proof of Theorem 5. If $a = f_{(3/4)}(X)$, then

$$\frac{1}{2}\left(x_a + \frac{1}{a-1}\sum_{i=1}^{a} x_i\right) = \frac{x_a}{2} + \frac{1}{2a}\sum_{i=1}^{a} x_i + \left(\frac{1}{2(a-1)} - \frac{1}{2a}\right)\sum_{i=1}^{a} x_i \geq \frac{3a}{4} + \frac{1}{2}\frac{1}{a(a-1)}\sum_{i=1}^{a} x_i. \quad (10)$$

As $a = f_{(3/4)}(X) \leq g_{(3/4)}(X)$, we have $\frac{1}{a}\sum_{i=1}^{a} x_i \geq \frac{3a}{4}$.

Consequently, by Theorem 4, we find that

$(10) \geq \frac{3a}{4} + \frac{1}{2(a-1)}\frac{3a}{4} \geq \frac{3a}{4} + \frac{3}{8(a-1)} > \frac{3a}{4} + \frac{1}{4(a-1)}$.

As a_X is the largest natural number with this property, this ends the proof of Theorem 5. □

3.5. Examples

Example 1. *We provide an example such that the following strict inequalities hold: $h_{(3/4)}(X) < f_{(3/4)}(X) < a_X < g_{(0.5)}(X)$.*

Let X = (6,1,1). Then, $h_{(3/4)}(X) = 1 < f_{(3/4)}(X) = 2 < a_X = 3 < g_{(0.5)}(X) = 4$.

This example shows that, contrary to the continuous case, $f_{(3/4)}(X)$ is not always the solution of the minimization problem. Stated otherwise, in general, $f_{(3/4)}(X) \neq a_X$. Yet, one may say that $f_{(3/4)}(X)$ is a (close) under limit.

As $g_{(3/4)}(X) \geq f_{(3/4)}(X)$, it was an upper limit for a_X in the continuous case. One may wonder if, in the discrete case, $g_{(3/4)}(X)$ is either an under or an upper limit for a_X. Yet, none of these two alternatives are correct. In the case of X = (6,1,1), $g_{(3/4)}(X) = 3 = a_X$. However, for X = (8,1), $h_{(3/4)}(X) = 1 < f_{(3/4)}(X) = 2 = a_X < g_{(3/4)}(X) = 3 < g_{(0.5)}(X) = 4$, while for X = (2, 0.9), $h_{(3/4)}(X) = 1 = f_{(3/4)}(X) = g_{(3/4)}(X) < a_X = g_{(0.5)}(X) = 2$.

We already observed that $g_{(3/4)}(X)$ can be smaller than, equal to, and larger than a_X. We next show that $a_X \leq g_{(3/4)}(X) + 1$.

Proposition 3. *Given an array X, converging and decreasing in* $(\mathbf{R}^+)^\infty$*, then $a_X \leq g_{(3/4)}(X) + 1$.*

Proof of Proposition 3. We show that if a = $g_{(3/4)}(X) + 1$, then $f_X(a+1) > f_X(a)$. This inequality is equivalent to

$$-2(a+1)\left(\sum_{i=1}^{a+1} x_i\right) + (a+1)^3 > -2a\left(\sum_{i=1}^{a} x_i\right) + a^3$$

$$\Leftrightarrow -2ax_{a+1} - 2\left(\sum_{i=1}^{a+1} x_i\right) + (a+1)^3 > a^3$$

$$\Leftrightarrow ax_{a+1} + \left(\sum_{i=1}^{a+1} x_i\right) < \frac{1}{2}\left(3a^2 + 3a + 1\right) = \frac{3}{2}a(a+1) + \frac{1}{2}$$

Now, a = $g_{(3/4)}(X) + 1 > g_{(3/4)}(X) \geq h_{(3/4)}(X)$ and hence $x_{a+1} \leq x_a < (3/4)a$.

Moreover, $\sum_{i=1}^{a+1} x_i = \sum_{i=1}^{a} x_i + x_{a+1} \leq \sum_{i=1}^{a} x_i + x_a < \frac{3}{4}a^2 + \frac{3}{4}a$.

Consequently, $ax_{a+1} + \left(\sum_{i=1}^{a+1} x_i\right) < \frac{3}{4}a^2 + \frac{3}{4}a^2 + \frac{3}{4}a = \frac{3}{2}a^2 + \frac{3}{4}a < \frac{3}{2}a(a+1) + \frac{1}{2}$, which proves Proposition 3. □

Example 2. *If X = (2,2,2), then $a_X = 3$ and $g_{(3/4)}(X) = 2$, providing an example where there is an equality for the expression $a_X \leq g_{(3/4)}(X) + 1$.*

3.6. An Upper Bound for a_X

We already know that $g_{(3/4)}(X)$ is not an upper bound for a_X and that $g_{(0.5)}(X)$ is. Hence, we wonder if there a number strictly between 0.5 and 0.75 that leads to an upper bound for all X.

Theorem 6. *An upper bound for a_X is provided by $g_{(7/12)}(X)$.*

Proof of Theorem 6. Take a $\geq g_s(X)$, with s being any real number strictly smaller than 0.75.

Hence, a+1 $> g_s(X) \geq h_s(X)$. From these inequalities, we derive

$$\begin{cases} x_{a+1} < s(a+1) \\ \sum_{i=1}^{a+1} x_i < s(a+1)^2 \end{cases}$$

Multiplying the first inequality by a and adding the two resulting inequalities yields

$$ax_{a+1} + \sum_{i=1}^{a+1} x_i < sa(a+1) + s(a+1)^2$$

Now, from Proposition 3, we know that $f_X(a+1) > f_X(a)$, hence $a \geq a_X$ if

$$ax_{a+1} + \left(\sum_{i=1}^{a+1} x_i\right) \leq \frac{3}{2}a(a+1) + \frac{1}{2}$$

From this inequality, we find that s must satisfy the following inequality:

$$sa(a+1) + s(a+1)^2 = s\left(2a^2 + 3a + 1\right) \leq \frac{1}{2}\left(3a^2 + 3a + 1\right)$$

leading to

$$s \leq \frac{3a^2 + 3a + 1}{2(2a^2 + 3a + 1)} \tag{11}$$

This inequality must hold for any natural number a different from zero. As the right-hand side is increasing in a, we consider the inequality for $a = 1$, leading to $s \leq 7/12$. □

Corollary 1. *Given an array X, converging and decreasing in* $(\mathbf{R}^+)^\infty$*, then* $a_X \in [\![f_{(\frac{3}{4})}(X),\ g_{(\frac{7}{12})}(X)]\!]$.

We already observed that if $a_X = 1$, then $h(X) = 1$. What about the converse? The next proposition answers this question.

Proposition 4. *If $h(X) = 1$, then a_x can be larger than any natural number b.*

Proof of Proposition 4. Consider $X = (z, 1, 0, \ldots)$ and we want to find z such that $a_X \geq b$. If $a_X \geq b$, then $f_X(b) \geq f_X(b+1)$, or (with $b > 2$): $2(z+1) \geq 3b^2 + 3b + 1$. Hence, it suffices to take $z > (3b^2 + 3b - 1)/2$. □

An example: Suppose that we want $a_X \geq 18$. Taking $z = 512$ leads to $X = (512, 1, 0, \ldots)$ and $a_X = 18$. If, however, we want $a_X \geq 19$, then $z = 570$, leading to $X = (570, 1, 0, \ldots)$ with $a_X = 20$.

4. Applications

First, we give a new characterization of the classical h-index [1], i.e., the case $\theta = 1$.

Proposition 5. *Given X decreasing and convergent, then* $h(X) = \max\{a \in \mathbf{N};\ A_a \leq X\}$.

Proof of Proposition 5. Writing h(X) simply as h, we see that $A_h \leq X$ because for A_h and $j \leq h$, $x_j \geq h$, while for all $j > h$, $x_j \geq 0$. This shows that $h \leq \max\{a \in \mathbf{N};\ A_a \leq X\}$.

Now, let $a_m = \max\{a \in \mathbf{N};\ A_a \leq X\}$. Then, we see that for all $j \leq a_m$, $x_j \geq a_m$, while for all $j > a_m$, $x_j \geq 0$. As h is defined as the largest number with this property, we see that $h \geq a_m = \max\{a \in \mathbf{N};\ A_a \leq X\}$. This proves this proposition. □

Before continuing with the next proposition, we recall the definition of the majorization partial order for finite sequences.

Definition 5. *The majorization order [12]:*

Let X, Y ∈ $(\mathbf{R^+})^k$*, where k is any finite number in* $\mathbf{N_0}$ *= {1, 2, 3, … }. The array X is majorized by Y, or X is smaller than or equal to Y in the majorization order, denoted as X -< Y if for all i = 1, … ,N:*

$$\begin{cases} \sum_{i=1}^{N} x_i = \sum_{i=1}^{N} y_i \text{ and} \\ \sum_{j=1}^{i} x_j \leq \sum_{j=1}^{i} y_j;\ \forall\, i = 1, \ldots, N. \end{cases}$$

Proposition 6. *If X is finite with length N and* $\frac{1}{N}\sum_{j=1}^{N} x_j = \overline{x}$ *is a natural number, then* A_a *-< X ⇔ a =* $\overline{x}$*, where -< denotes the majorization partial order.*

Proof of Proposition 6. If A_a -< X, then, for all j = 1, … , N, $ja \leq \sum_{k=1}^{j} x_k$ and Na = $\sum_{k=1}^{N} x_k$. Consequently, a = $\overline{x}$.

Conversely, if a = $\overline{x}$ (and hence $\overline{x}$ must be a natural number), we have

$$A_a = \left(\underbrace{\overline{x}, \overline{x}, \ldots, \overline{x}}_{\overline{x}\ times},\ 0, 0,\ \ldots\right)$$

and hence for all j ≤ N, $j\overline{x} \leq \sum_{k=1}^{j} x_k$ and for j = N, $N\overline{x} = \sum_{k=1}^{N} x_k$. This shows that A_a -< X. □

Finally, we show that a_X is increasing in X.

Theorem 7. *If X < Y, then* $a_X \leq a_Y$*.*

Proof of Theorem 7. We know that a_X is the largest index such that

$$\sum_{i=1}^{\infty} x_i^2 - 2(a_X - 1)\left(\sum_{i=1}^{a_X-1} x_i\right) + (a_X - 1)^3 \geq \sum_{i=1}^{\infty} x_i^2 - 2a_X\left(\sum_{i=1}^{a_X} x_i\right) + (a_X)^3$$
$$\Leftrightarrow 2\left(\sum_{i=1}^{a_X-1} x_i\right) + (a_X - 1)^3 \geq -2a_X x_{a_X} + (a_X)^3$$

Now, we also know that for all $i \geq 1$, $y_i \geq x_i$. Hence, $\left(\sum_{i=1}^{a_X-1} y_i\right) \geq \left(\sum_{i=1}^{a_X-1} x_i\right)$ and $-2a_X y_{a_X} \leq -2a_X x_{a_X}$. This leads to

$$2\left(\sum_{i=1}^{a_X-1} y_i\right) + (a_X - 1)^3 \geq -2a_X y_{a_X} + (a_X)^3$$

Hence, also

$$\sum_{i=1}^{\infty} y_i^2 - 2a_X\left(\sum_{i=1}^{a_X-1} y_i\right) + 2\left(\sum_{i=1}^{a_X-1} y_i\right) + (a_X - 1)^3 \geq \sum_{i=1}^{\infty} y_i^2 - 2a_X\left(\sum_{i=1}^{a_X} y_i\right) + (a_X)^3$$

This can be written as

$$\sum_{i=1}^{\infty} y_i^2 - 2(a_X - 1)\left(\sum_{i=1}^{a_X-1} y_i\right) + (a_X - 1)^3 \geq \sum_{i=1}^{\infty} y_i^2 - 2a_X\left(\sum_{i=1}^{a_X} y_i\right) + (a_X)^3 \tag{12}$$

As a_Y is the largest index with property (12), this shows that $a_X \leq a_Y$. □

Remark 2. *If $X < Y$ (strict), then it is possible that $a_X = a_Y$. An example is given by $X = (6,1,1) < Y = (6,2,1)$, for which $a_X = a_Y = 3$.*

5. Conclusions

In this article, we studied the following problem:

Given a converging decreasing array X in $(\mathbf{R}^+)^\infty$, find the largest natural number a such that the Euclidean distance $d(X,A_a)$ is minimal.

We have shown that this problem has a solution, which is always situated in the interval $[\![h_{(\frac{3}{4})}(X),\ g_{(\frac{7}{12})}(X)]\!]$. Yet, the solution is not necessarily unique. It was shown that a discrete and an analogous continuous problem have related but not the same solutions. Our contribution illustrates how a formalism derived in the context of research evaluation and informetrics [1] can be used to solve a purely mathematical problem.

Author Contributions: Conceptualization, L.E.; Formal analysis, L.E. and R.R.; Writing—original draft, L.E.; Writing—review & editing, L.E. and R.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hirsch, J.E. An index to quantify an individual's scientific research output. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 16569–16572. [CrossRef] [PubMed]
2. Bouyssou, D.; Marchant, T. Ranking scientists and departments in a consistent manner. *J. Am. Soc. Inf. Sci. Technol.* **2011**, *62*, 1761–1769. [CrossRef]
3. Waltman, L.; van Eck, N.J. The inconsistency of the h-index. *J. Am. Soc. Inf. Sci. Technol.* **2012**, *63*, 406–415. [CrossRef]
4. Egghe, L. Theory and practise of the g-index. *Scientometrics* **2006**, *69*, 131–152. [CrossRef]
5. Alonso, S.; Cabrerizo, F.J.; Herrera-Viedma, E.; Herrera, F. H-index: A review focused in its variants, computation and standardization for different scientific fields. *J. Informetr.* **2009**, *3*, 273–289. [CrossRef]
6. Egghe, L. The Hirsch-index and related impact measures. *Annu. Rev. Inf. Sci. Technol.* **2010**, *44*, 65–114. [CrossRef]
7. Rousseau, R.; Egghe, L.; Guns, R. *Becoming Metric-Wise: A Bibliometric Guide for Researchers*; Chandos: Oxford, UK, 2018.
8. Egghe, L.; Rousseau, R. Polar coordinates and generalized h-type indices. *J. Informetr.* **2020**, *14*, 101024. [CrossRef]
9. Egghe, L.; Rousseau, R. Infinite sequences and their h-type indices. *J. Informetr.* **2019**, *13*, 291–298. [CrossRef]
10. Egghe, L.; Rousseau, R. Solution by step functions of a minimum problem in L^2 [0,T], using generalized h- and g-indices. *J. Informetr.* **2019**, *13*, 785–792. [CrossRef]
11. van Eck, N.J.; Waltman, L. Generalizing the h-and g-indices. *J. Informetr.* **2008**, *2*, 263–271. [CrossRef]
12. Hardy, G.H.; Littlewood, G.E.; Pólya, G. *Inequalities*; Cambridge University Press: Cambridge, UK, 1934.