Journal of Frame and Matrix Theory (JFMT) 2 (2025) No. 1, 36–59
Online ISSN: 3041-8526
Published by Faculty of Mathematics and Computer Science
Hakim Sabzevari University
http://dx.doi.org/10.22034/jfmt.2025.491401.1019

Non-uniformity measurement

Mohammad Zarghani[a,*], Alireza Ghodsi[b]

[a] *Farhangian University, Sabzevar, Iran.*

[b] *Faculty of Mathematics and Computer Science, Hakim Sabzevari University, Sabzevar, Iran.*

(Communicated by Reza Saadati)

Abstract

In this paper, we show that the coercive functions reach their minimum value on closed subsets of $\mathbb{R}^n$. We then introduce the concept of sequence variance, a statistical metric designed to quantify the degree of non-uniformity in a sequence of observations (the irregularity in the spatial arrangement of points). This metric is calculated as the average squared distance between ordered data points. Additionally, the paper introduces the sequence correlation coefficient and examines its properties. Finally, we present a method for detecting outliers in a sequence of data points within Euclidean spaces.

Keywords: Sequence variance, Non-uniformity, Sequence covariance, Sequence correlation, Outlier.
MSC 2020: Primary 62R20; Secondary 62A99, 54B99.

*Corresponding author
Email addresses: zarghanimohammad@gmail.com (Mohammad Zarghani), ghodsiir@gmail.com (Alireza Ghodsi)

Received: 29 November 2024 *Accepted:* 17 March 2025

1 Introduction

Anomaly detection, or outlier detection, involves identifying data points that significantly deviate from the norm, categorized into three types: global anomalies (individual outliers), contextual anomalies (deviations based on specific contexts), and collective anomalies (groups of data points with unusual behavior) [1, 7, 2]. Hawkins defined an outlier in [9] as follows: "An outlier is an observation which deviates so much from the other observations as to arouse suspicions that it was generated by a different mechanism." The process is crucial in various fields, as anomalies may indicate errors or noteworthy events, such as fraudulent transactions or unusual temperature readings. To analyze and quantify these deviations, various techniques are employed, including depth-based methods and statistical indicators like variance and standard deviation. This article focuses on the importance of a particular indicator in evaluating data non-uniformity. Non-uniformity (or irregularity) in spatial data refers to the concept of uneven or non-homogeneous distribution of points, events, or observations within a geographic or geometric space. In other words, this phenomenon occurs when data is not evenly spread across the space, and some areas have a higher density of points while others have fewer or even no points at all.

Now, briefly, we summarize the contents of this article as follows.

Section 2 introduces statistics and Euclidean spaces as preliminaries.

In section 3, we determine the global minimum of the coercive function $f(\mathbf{w}_1, \ldots, \mathbf{w}_k) = \sum_{i=1}^{k} \frac{1}{m_i} \|\mathbf{w}_i\|^2$ on the closed set $\prod_{i=1}^{k} D_i$ and we also give a global maximum of f on the compact set $\prod_{i=1}^{k} D_i^*$ (see Theorem 3.2). Next, we define the concept of sequence variance, which serves as a statistical metric for assessing the degree of non-uniformity within a sequence of observations. This indicator is computed as the arithmetic mean of the squared differences between adjacent observations in the sequence. Then, we expand the notion of sequence variance to d-dimensional Euclidean spaces. Furthermore, this section introduces specific indices, including the sequence covariance and the sequence correlation coefficient, along with an analysis of their properties (for instance, see Proposition 3.13). It is important to note that, unlike variance and co-variance, the newly defined concepts sequence variance, sequence covariance, and sequence correlation are not invariant under permutations of the data. These indicators are specifically designed for application to ordered data points.

Finally, in section 4, we generally deal with the detection of outlier data. Outlier detection is a fundamental aspect of data analysis, enabling us to uncover hidden insights, identify errors, and make better decisions. The choice of method depends on the nature of the data, the domain, and the specific goals of the analysis.

2 Preliminaries

In probability theory and statistics, "variance" is the squared deviation from the mean of a sample (or a population) [6]. Variance is an indicator of dispersion, meaning it is a metric of how far a set of numbers is spread out from their average value. The variance of data points $x_1, \ldots, x_n$, $n > 1$, is defined by

$$S_x^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 \ ,$$

where $\bar{x}$ is the mean of x_i's. The root of S_x^2 is denoted by S_x and called "standard deviation". The "co-variance" of data points $(x_1, y_1), \ldots, (x_n, y_n)$, $n > 1$, is defined by

$$S_{xy} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) \ ,$$

where $\bar{x}$ and $\bar{y}$ is the mean of x_i's and y_i's, respectively. The "correlation coefficient" of data points $(x_1, y_1), \ldots, (x_n, y_n)$ is defined by $r = \frac{S_{xy}}{S_x S_y}$, where S_x and S_y are standard deviation of x_i's and y_i's. [10].

A reflexive, weak or non-strict partial order, commonly referred to simply as a partial order, is a relation $\leq$ on a set P that is reflexive, antisymmetric, and transitive. That is, for all $a, b, c \in P$ it must satisfy:
Reflexivity: $a \leq a$, i.e. every element is related to itself.
Antisymmetry: if $a \leq b$ and $b \leq a$ then $a = b$, i.e. no two distinct elements precede each other.
Transitivity: if $a \leq b$ and $b \leq c$ then $a \leq c$.
A total order or linear order is a partial order in which any two elements are comparable, and a set equipped with a total order is called a totally ordered set.

It is known that the lexicographical (lexical) order between d-tuples in the Euclidean d-space $\mathbb{R}^d$ with $d \geq 2$ is defined by
$(a_1, \ldots, a_d) \leq^{lex} (b_1, \ldots, b_d)$ if $a_1 < b_1$ else, if $a_1 = b_1$ and $a_2 < b_2$ ... else, if $a_i = b_i$ for all $i < d$, and $a_d \leq b_d$; in other words, $(a_1, \ldots, a_d) <^{lex} (b_1, \ldots, b_d)$ if $a_i < b_i$ for the first i, and $(a_1, \ldots, a_d) = (b_1, \ldots, b_d)$ if $a_i = b_i$ for all i. Particularly, the lexicographical order between couples in $\mathbb{R}^2$ is $(x_1, y_1) \leq^{lex} (x_2, y_2)$ if $x_1 < x_2$, or ($x_1 = x_2$ and $y_1 \leq y_2$).

In statistics, the k'th order statistic of a statistical sample is equal to its k'th-smallest value [3]. For example, suppose that four numbers are observed or recorded, resulting in a sample of size 4. If the sample values are $6, 9, 3, 8$ the order statistics would be denoted by

$$x_{(1)} = 3, \quad x_{(2)} = 6, \quad x_{(3)} = 8, \quad x_{(4)} = 9,$$

where the subscript (i) enclosed in parentheses indicates the i'th order statistic of the sample. The first order statistic (or smallest order statistic) is always the minimum of the sample,

that is, $x_{(1)} = \min\{ x_1, \ldots, x_n \}$. Similarly, for a sample of size n, the n'th order statistic (or largest order statistic) is the maximum, that is, $x_{(n)} = \max\{ x_1, \ldots, x_n \}$. The sample range is the difference between the maximum and minimum. It is a function of the order statistics: $\text{Range}\{ x_1, \ldots, x_n \} = x_{(n)} - x_{(1)}$.

Let $\mathbb{R}^d$ be the Euclidean d-space. The length of the vector $\mathbf{v} = (v_1, \ldots, v_d)$ is captured by the formula [8], $\|\mathbf{v}\| := \sqrt{v_1^2 + \cdots + v_d^2}$. This is the Euclidean norm, which gives the ordinary distance from the origin to the point $\mathbf{v}$. Especially the absolute value $\|\mathbf{v}\| = |\mathbf{v}|$ is a norm on the one-dimensional vector spaces formed by the real numbers.

In Euclidean d-space $\mathbb{R}^d$, an open ball of radius r and center $\mathbf{x}$, denoted by $B(\mathbf{x}, r) = \{\mathbf{v} \in \mathbb{R}^d | \, \|\mathbf{v} - \mathbf{x}\| < r\}$, is the set of all points of distance less than r from $\mathbf{x}$. A closed ball of radius r and center $\mathbf{x}$, denoted by $\bar{B}(\mathbf{x}, r) = \{\mathbf{v} \in R^d | \, \|\mathbf{v} - \mathbf{x}\| \leq r\}$, is the set of all points of distance less than or equal to r away from $\mathbf{x}$. A unit ball (open or closed) is a ball of radius 1.

Definition 2.1. [11]. A set $D \subseteq \mathbb{R}^d$ is bounded if there exists a constant $M > 0$ such that $\|\mathbf{v}\| < M$ for all $\mathbf{v} \in D$.

Proposition 2.2. In Euclidean spaces:

1. A continuous preimage of a closed set is closed.
2. A set $D \subseteq \mathbb{R}^d$ is said to be compact if it is closed and bounded.
3. A closed subset of a compact set is compact.
4. The product of finite collection of compact sets is compact.

Theorem 2.3. [11]. Let D be a compact (closed and bounded) subset of $\mathbb{R}^d$. If $f(\mathbf{v})$ is a continuous real function on D, then $f(\mathbf{v})$ has a global maximizer and a global minimizer on D.

Definition 2.4. [11]. A real continuous function $f(\mathbf{v})$ that is defined on $\mathbb{R}^d$ is coercive if $\lim_{\|\mathbf{v}\| \to \infty} f(\mathbf{v}) = +\infty$.

Theorem 2.5. [11] Let $f(\mathbf{v})$ be a continuous function defined on all of $\mathbb{R}^d$. If $f(\mathbf{v})$ is coercive, then $f(\mathbf{v})$ has a global minimizer. Furthermore, if the first partial derivatives of $f(\mathbf{v})$ exist on all of $\mathbb{R}^d$, then any global minimizers of $f(\mathbf{v})$ can be found among the critical points of $f(\mathbf{v})$.

3 Sequence variance

Given a set of linear equations and a norm function, the task is to find the minimizer of the norm function subject to the constraint that the given equations are satisfied. To

achieve this, we require a lemma akin to Theorem 2.5. We are providing a proof for this lemma, even though it may already be found in a reference book such as [4].

Lemma 3.1. Every coercive function on a closed subset of $\mathbb{R}^n$ has a global minimum.

Proof . Let choose any point in D, call it $\mathbf{v}_0$. Since $f(\mathbf{v})$ is coercive, $\lim_{\|\mathbf{v}\|\to\infty} f(\mathbf{v}) = +\infty$. This means that if $\|\mathbf{v}\|$ is large, then so is $f(\mathbf{v})$. Accordingly there is a number $r > 0$ such that if $\|\mathbf{v}\| > r$, then $f(\mathbf{v}) \geq 1 + f(\mathbf{v}_0)$. This guarantees that $f(\mathbf{v}) > f(\mathbf{v}_0)$. Let $\bar{B}(0,r) = \{\mathbf{v} \in \mathbb{R}^d | \, \|\mathbf{v}\| \leq r\}$. Also $\mathbf{v}_0 \in \bar{B}(0,r) \cap D$ else $\|\mathbf{v}_0\| > r$ implies $f(\mathbf{v}_0) \geq 1 + f(\mathbf{v}_0)$ which is a contradiction. The function $f(\mathbf{v})$ is continuous at each point of the set $\bar{B}(0,r) \cap D$ and the set $\bar{B}(0,r) \cap D$ is compact. From Theorem 2.3, it follows that f takes a minimum value on $\bar{B}(0,r) \cap D$ at a point $\mathbf{v}^*$ in $\bar{B}(0,r) \cap D$; i.e, $f(\mathbf{v}) \geq f(\mathbf{v}^*)$ for all $\mathbf{v} \in \bar{B}(0,r) \cap D$. In particular since $\mathbf{v}_0 \in \bar{B}(0,r) \cap D$, we see that $f(\mathbf{v}^*) \leq f(\mathbf{v}_0)$. On the other hands, if $\mathbf{v} \in \bar{B}^c(0,r) \cap D$, where $\bar{B}^c(0,r) = \mathbb{R}^n - B(0,r)$, then $f(\mathbf{v}) > f(\mathbf{v}_0) \geq f(\mathbf{v}^*)$. Summarizing, we have seen that $\mathbf{v} \in D = (\bar{B}(0,r) \cap D) \cup (\bar{B}^c(0,r) \cap D)$ implies $f(\mathbf{v}^*) \leq f(\mathbf{v})$. This shows $\mathbf{v}^*$ is a global minimizer of $f(\mathbf{v})$ on all D. This completes the proof. □

Notation 1. For indices i and s, we let $\delta_{is} = 1$ if $i = s$ and $\delta_{is} = 0$ if $i \neq s$.

Theorem 3.2. Assume that $f : \mathbb{R}^{m_1} \times \cdots \times \mathbb{R}^{m_k} \longrightarrow \mathbb{R}$ defined by

$$f(\mathbf{w}_1, \ldots, \mathbf{w}_k) = \sum_{i=1}^{k} \frac{1}{m_i} \|\mathbf{w}_i\|^2$$

be a real function and for $1 \leq i \leq k$, let

$$D_i := \{\mathbf{w}_i \in \mathbb{R}^{m_i} \mid \sum_{j=1}^{m_i} v_{ij} = b_i\} \text{ and } D_i^* := \{\mathbf{w}_i \in \mathbb{R}^{m_i} \mid \sum_{j=1}^{m_i} v_{ij} = b_i; v_{ij}, b_i \geq 0\},$$

where v_{ij} be the j'th component of $\mathbf{w}_i \in \mathbb{R}^{m_i}$ and b_i a constant. Then the function f on the closed set $\prod_{i=1}^{k} D_i$ has a minimum value if $v_{ij} = \frac{b_i}{m_i}$, for all $1 \leq i \leq k$ and $1 \leq j \leq m_i$. Furdermore, if b_i and v_{ij}, for all i, j, are non-negative, then the function f on $\prod_{i=1}^{k} D_i^*$ has also a global maximum.

Proof . The coercivity of f follows from

$$\|(\mathbf{w}_1, \ldots, \mathbf{w}_k)\| \to \infty \quad \text{if and only if} \quad \sum_{i=1}^{k} \frac{1}{m_i} \|\mathbf{w}_i\|^2 \to \infty \ .$$

Now, we show that the set $\prod_{i=1}^{k} D_i$ is closed. For this, let the function $s : \mathbb{R}^{m_1} \times \cdots \times \mathbb{R}^{m_k} \longrightarrow \mathbb{R}^k$ defined by

$$s(\mathbf{w}_1, \ldots, \mathbf{w}_k) = (\sum_{j=1}^{m_1} v_{1j}, \ldots, \sum_{j=1}^{m_k} v_{kj}).$$

Since s is continuous, and the preimage of a continuous function on a closed set is closed, the set $\prod_{i=1}^{k} D_i$ is closed. Hence, by Lemma 3.1, coercive function f on a closed subset $\prod_{i=1}^{k} D_i$ has a global minimum. Now, we let the Lagrangian function $\mathcal{L}$ defined by

$$\mathcal{L}(\mathbf{w}_1, \ldots, \mathbf{w}_k, \lambda_1, \ldots, \lambda_k) := f(\mathbf{w}_1, \ldots, \mathbf{w}_k) - \sum_{i=1}^{k} \lambda_i (\sum_{j=1}^{m_i} v_{ij} - b_i),$$

where λ_i are Lagrangian coefficients for $1 \leq i \leq k$. Then the first order derivatives of the function $\mathcal{L}$ are

$$\frac{\partial \mathcal{L}}{\partial v_{ij}} = \frac{2v_{ij}}{m_i} - \lambda_i \quad \text{and} \quad \frac{\partial \mathcal{L}}{\partial \lambda_i} = \sum_{j=1}^{m_i} v_{ij} - b_i$$

for all $1 \leq i \leq k$ and $1 \leq j \leq m_i$. By setting the partial derivatives equal to zero, we get $\lambda_i = \frac{2v_{ij}}{m_i}$, for all $1 \leq j \leq m_i$, and

$$\sum_{j=1}^{m_i} v_{ij} = b_i \tag{3.1}$$

where $1 \leq i \leq k$. Therefore for all $1 \leq i \leq k$, $v_{i1} = \cdots = v_{im_i}$ (and all are equal to $\frac{m_i \lambda_i}{2}$). Substituting this in Equation 3.1, we have $m_i v_{ij} = b_i$ or equivalently $v_{ij} = \frac{b_i}{m_i}$ for all $1 \leq i \leq k$ and $1 \leq j \leq m_i$, so that the critical point $(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ has been achieved, where for any $1 \leq i \leq k$,

$$\mathbf{v}_i = (v_{i1}, \ldots, v_{im_i}) = (\frac{b_i}{m_i}, \ldots, \frac{b_i}{m_i}) \ .$$

Therefore the global minimum f on $\prod_{i=1}^{k} D_i$ is equal to

$$f(\mathbf{v}_1, \ldots, \mathbf{v}_k) = \sum_{i=1}^{k} (\frac{1}{m_i} \sum_{j=1}^{m_i} v_{ij}^2) = \sum_{i=1}^{k} (\frac{1}{m_i} \sum_{j=1}^{m_i} (\frac{b_i}{m_i})^2) = \sum_{i=1}^{k} (\frac{b_i}{m_i})^2 \ .$$

In the case that $v_{ij} \geq 0$, for all i, j, we have

$$f(\mathbf{w}_1, \ldots, \mathbf{w}_k) = \sum_{i=1}^{k} \frac{1}{m_i} \sum_{j=1}^{m_i} v_{ij}^2 \leq \sum_{i=1}^{k} \frac{1}{m_i} (\sum_{j=1}^{m_i} v_{ij})^2 = \sum_{i=1}^{k} \frac{b_i^2}{m_i} \ . \tag{3.2}$$

Consequently, the set $\prod_{i=1}^{k} D_i^*$ is bounded and so, by Proposition 2.2, is compact. Using Theorem 2.3, f has also a global maximum on $\prod_{i=1}^{k} D_i^*$. For any $1 \leq i \leq k$, let choose an index $1 \leq s_i \leq m_i$, $1 \leq i \leq k$, such that $v_{is_i} = b_i$, for $i = s_i$, and $v_{is_i} = 0$, for all $i \neq s_i$. (Note that the global maximum value does not change by choosing any desired $1 \leq s_i \leq m_i, 1 \leq i \leq k$). Now, let

$$\mathbf{v}_i = (v_{i1}, \ldots, v_{im_i}) = (b_i \delta_{s_i 1}, \ldots, b_i \delta_{s_i m_i}) \ ,$$

where $1 \leq i \leq k$. Then, using non-equality in Equation 3.2, the maximum value of f on $\prod_{i=1}^{k} D_i^*$ is equal to

$$f(\mathbf{v}_1, \ldots, \mathbf{v}_k) = \sum_{i=1}^{k} \frac{1}{m_i} (b_i^2) \ .$$

□

In optimization problems involving weighted graphs, the weights assigned to edges can be considered as variables. These weights typically represent quantities such as distance, cost, or capacity. By defining the edge weights as the minima of norm functions, we can establish constraints or relationships between different edges.

Example 3.3. In Figure 1, the length of 31 springs are all positive. For $1 \leq i \leq 31$, if w_i is

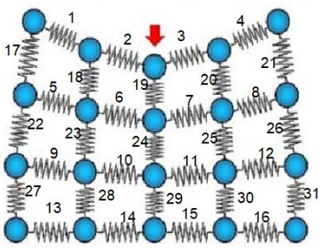

Figure 1: Stretched graph of springs

the length of i-th spring labeled in Figure 1, then minimizing the function $f : \mathbb{R}^{16} \times \mathbb{R}^{15} \longrightarrow \mathbb{R}$ defined by

$$f((w_1, \ldots, w_{16}), (w_{17} \cdots, w_{31})) = \frac{\sum_{i=1}^{16} w_i^2}{16} + \frac{\sum_{i=17}^{31} w_i^2}{15}$$

subject to $\sum_{i=1}^{16} w_i = 32$ and $\sum_{i=17}^{32} w_i = 15$ gives us the following piecewise homogeneous graph with the constant spring lengths $w_1 = \cdots = w_{16} = 2$ and $w_{17} = \cdots = w_{31} = 1$ (see Figure 2).

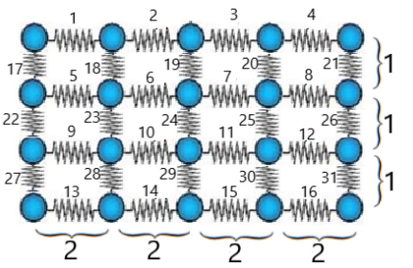

Figure 2: Piecewise homogeneous graph

The following remark is just the special case of Theorem 3.2.

Remark 3.4. Assume that $f : \mathbb{R}^m \longrightarrow \mathbb{R}$ defined by $f(\mathbf{w}) = \frac{1}{m} \|\mathbf{w}\|^2$ be a real function and let $D = \{\mathbf{w} = (w_1, \ldots, w_m) \in \mathbb{R}^m \mid \sum_{i=1}^m w_i = b\}$, where b is supposed as a constant. Then the function f on the closed set D has the global minimum value $(\frac{b}{m})^2$ at $(w_1, \ldots, w_m) = (\frac{b}{m}, \ldots, \frac{b}{m})$. Furdermore, if b and w_i's, for all $1 \leq i \leq m$, are non-negative, then, for any fixed index $1 \leq s \leq m$ the function f on $D^* = \{\mathbf{w} = (w_1, \ldots, w_m) \in \mathbb{R}^m \mid \sum_{i=1}^m w_i = b; w_i, b \geq 0\}$ has also the global maximum value $\frac{b^2}{m}$ at $(w_1, \ldots, w_m) = (b\delta_{1s}, \ldots, b\delta_{ms})$.

The subsequent corollary demonstrates that when the variables of a sequence are uniformly distributed within a specified domain, the associated sequence variance attains its minimum value. Conversely, when the data are concentrated along the edges of the domain, the resulting sequence variance reaches its maximum value.

Corollary 3.5. Let $\{x_i\}_{i=1}^n$, $n \geq 2$, be a finite sequence of variables in $\mathbb{R}$ with the norm function $W^2 = \frac{1}{n-1} \sum_{i=1}^{n-1} w_i^2$ (we call it **sequence variance** of x_i's), where $w_i := x_{i+1} - x_i$, $1 \leq i \leq n-1$, and for any constant R, let

$$D := \{(w_1, \ldots, w_{n-1}) \in \mathbb{R}^{n-1} \mid \sum_{i=1}^{n-1} w_i = R\},$$

and

$$D^* = \{(w_1, \ldots, w_{n-1}) \in \mathbb{R}^{n-1} \mid \sum_{i=1}^{n-1} w_i = R; w_i, b \geq 0\}.$$

Then, using Remark 3.4, the following statements are stablished.

1. W^2 on D has a global minimum $(\frac{R}{n-1})^2$ whenever $x_{i+1} = x_i + \frac{R}{n-1}$, for all $1 \leq i \leq n-1$. If $R > 0$ (respectively, $R < 0$), the uniformly distributed sequence $\{x_i\}_{i=1}^n$, $n \geq 2$, is strictly monotonically increasing (respectively, strictly monotonically decreasing).
2. If w_i's, for all $1 \leq i \leq n-1$, (and so R) are non-negative, then W^2 has a global maximum $\frac{R^2}{n-1}$ on D^* whenever $x_1 = \cdots = x_s$, $x_{s+1} = x_s + R$ and $x_{s+1} = \cdots = x_n$, for an arbitrary constant $0 \leq s \leq n-1$.

Now, we extend the concept of sequence variance to Euclidean spaces.

Definition 3.6. let $\{\mathbf{x}_i\}_{i=1}^n$, $n \geq 2$, be a finite sequence in $\mathbb{R}^d$. We define

$$W^2 = \mathrm{SeqVar}(\mathbf{x}) = \frac{1}{n-1} \sum_{i=1}^{n-1} \|\mathbf{x}_{i+1} - \mathbf{x}_i\|^2$$

and refer to it as the **sequence variance** of $\mathbf{x}_i$'s, which represents the variance of ordered data". Furthermore, the square root of W^2, denoted by W, is called **non-uniformity indicator**.

Note that in the special case that $\{x_i\}_{i=1}^n$, $n \geq 2$, is a finite sequence in $\mathbb{R}$, the norm $\|x_{i+1} - x_i\|$ is exactly the same as $|x_{i+1} - x_i|$ for all $1 \leq i \leq n-1$. So sequence variance on $\mathbb{R}^d$ is an extension of that on $\mathbb{R}$.

Example 3.7. Figure 3 is a part of the electrocardiogram (ECG) designed and simulated with Geogebra software. The corresponding sequence variance is equal to

$$W^2 = \frac{1^2 + 1.18^2 + 1.03^2 + 0.9^2 + 0.71^2 + 6.39^2 + 7.07^2 + 1.09^2 + 1.16^2 + 1.57^2 + 1.26^2 + 1.2^2}{12} \approx 8.63$$

and so the non-uniformity indicator, W, is approximately 2.94.

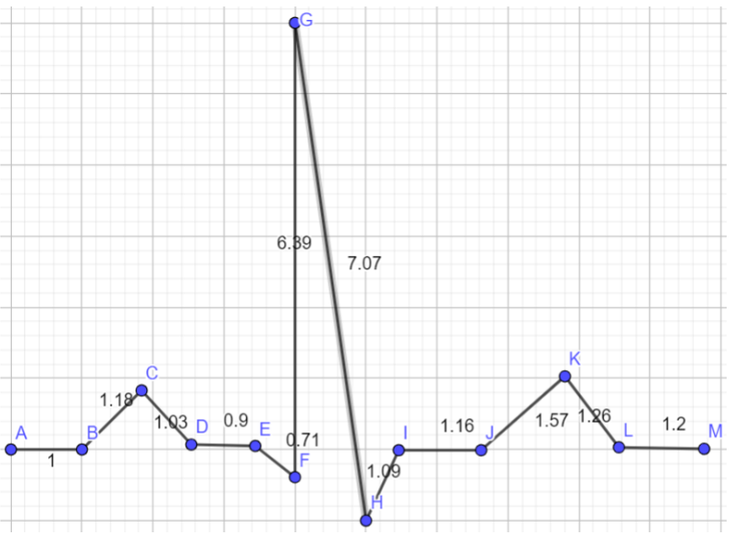

Figure 3: Electrocardiogram

The following proposition establishes basic properties of sequence variance such as:

Proposition 3.8. 1. Sequence variance of a sequence in $\mathbb{R}^d$ is non-negative because the squares are positive or zero; that is

$$W^2 \geq 0 \ .$$

2. The sequence variance of a constant sequence in $\mathbb{R}^d$ is zero. Conversely, if the sequence variance of a sequence in $\mathbb{R}^d$ is 0, then it is almost surely a constant sequence (that is, all terms of the sequence are constant); in other words, for a finite sequence $\{x_i\}_{i=1}^n$, $n > 1$,

$$W^2 = \frac{1}{n-1} \sum_{i=1}^{n-1} \|\mathbf{x}_{i+1} - \mathbf{x}_i\|^2 = 0 \ \text{ if and only if } \ \mathbf{x}_1 = \cdots = \mathbf{x}_n \ .$$

3. Sequence variance is invariant with respect to changes in a location parameter; more precisely, for any constant vector $\mathbf{b}$,
$$\mathrm{SeqVar}(\mathbf{x}+\mathbf{b}) = \mathrm{SeqVar}(\mathbf{x}) \ .$$
4. If all terms are scaled by a constant a, the sequence variance is scaled by the square of that constant; that is
$$\mathrm{SeqVar}(a\mathbf{x}) = a^2\mathrm{SeqVar}(\mathbf{x}) \ .$$

The following proposition asserts that the sequence variance of a sequence in the Euclidean d-space can be calculated by summing the sequence variances of its individual components. It also suggests that, under certain conditions, the maximum (or minimum) sequence variance of the sequence corresponds to the sum of the maxima (or minima) sequence variance of its i-th components.

Proposition 3.9. Let $\{\mathbf{x}_i\}_{i=1}^n$, $n > 1$, be a finite sequence of variables in $\mathbb{R}^d$, $d \geq 2$, which, for $1 \leq j \leq d$, the j-th component of the vector $\mathbf{x}_i$, $1 \leq i \leq n-1$, is denoted by x_{ij}. Suppose that $\{R_i\}_{i=1}^d$ are non-negative constants and let $R := \sqrt{R_1^2 + \cdots + R_d^2}$. For $1 \leq j \leq d$, let $W_j^2 := \frac{1}{n-1}\sum_{i=1}^{n-1} w_{ij}^2$ be the sequence variance of the sequence $\{x_{ij}\}_{i=1}^n$ in $\mathbb{R}$, where $w_{ij} := |x_{i+1,j} - x_{ij}|$, $1 \leq i \leq n-1$, and set
$$D_j := \{(w_{1j}, \ldots, w_{n-1,j}) \in \mathbb{R}^{n-1} \mid \sum_{i=1}^{n-1} w_{ij} = R_j\} \ .$$
Then the following statements are stablished.

1. $W^2 = \sum_{j=1}^d W_j^2$, where W^2 is the sequence variance of $\{\mathbf{x}_i\}_{i=1}^n$.
2. For variables $d_i := \|\mathbf{x}_{i+1} - \mathbf{x}_i\|$, $1 \leq i \leq n-1$, the function $W^2 = \frac{1}{n-1}\sum_{i=1}^{n-1} d_i^2$ has a global minimum (respectively global maximum) on the set
$$D := \{(d_1, \ldots, d_{n-1}) \in \mathbb{R}^{n-1} \mid \sum_{i=1}^{n-1} d_i = R\} \ ,$$
which is equal to the sum of the global minima (respectively global maxima) of W_j^2's on D_j, $1 \leq j \leq d$.

Proof .

1. Using definition, we have
$$\begin{aligned} W^2 &= \tfrac{1}{n-1}\textstyle\sum_{i=1}^{n-1} \|\mathbf{x}_{i+1} - \mathbf{x}_i\|^2 \\ &= \tfrac{1}{n-1}\textstyle\sum_{i=1}^{n-1}\sum_{j=1}^{d} |x_{i+1,j} - x_{ij}|^2 \\ &= \textstyle\sum_{j=1}^{d} \tfrac{1}{n-1}\sum_{i=1}^{n-1} |x_{i+1,j} - x_{ij}|^2 \\ &= \textstyle\sum_{j=1}^{d} W_j^2. \end{aligned}$$

2. By Remark 3.4, W^2 has a global minimum on D whenever $d_i = \frac{R}{n-1}$, for all $1 \leq i \leq n-1$ and it is equal to $(\frac{R}{n-1})^2$. On the other hand, by Remark 3.4, for any $1 \leq j \leq d$, W_j^2 has the global minimum value $(\frac{R_j}{n-1})^2$ on D_j whenever $(w_{1j}, \ldots, w_{n-1,j}) = (\frac{R_j}{n-1}, \ldots, \frac{R_j}{n-1})$. So, in this case, we have

$$\begin{aligned}
\sum_{j=1}^{d} \min W_j^2 \,(\text{on } D_j) &= \sum_{j=1}^{d} \tfrac{1}{n-1} \sum_{i=1}^{n-1} (\tfrac{R_j}{n-1})^2 \\
&= \sum_{j=1}^{d} (\tfrac{R_j}{n-1})^2 \\
&= \tfrac{\sum_{j=1}^{d} R_j^2}{(n-1)^2} \\
&= (\tfrac{R}{n-1})^2 \\
&= \min W^2(\text{on } D).
\end{aligned}$$

Furthermore, by Remark 3.4, W^2 has the global maximum $\frac{R^2}{n-1}$ on D whenever

$$(d_1, \ldots, d_{n-1}) = (\delta_{1s_0} R, \ldots, \delta_{n-1,s_0} R)$$

for a constant $1 \leq s_0 \leq n-1$. Also, by Remark 3.4, for $1 \leq j \leq d$, W_j^2 has the global maximum $\frac{R_j^2}{n-1}$ on bounded and closed set D_j whenever

$$(w_{1j}, \ldots, w_{n-1,j}) = (\delta_{1s_0} R_j, \ldots, \delta_{n-1,s_0} R_j)\,.$$

Then, in this case, we have

$$\begin{aligned}
\sum_{j=1}^{d} \max W_j^2 \,(\text{on } D_j) &= \sum_{j=1}^{d} \tfrac{1}{n-1} \sum_{i=1}^{n-1} (\delta_{is_0} R_j)^2 \\
&= \tfrac{1}{n-1} \sum_{i=1}^{n-1} \sum_{j=1}^{d} (\delta_{is_0} R_j)^2 \\
&= \tfrac{1}{n-1} \sum_{i=1}^{n-1} \delta_{is_0} \sum_{j=1}^{d} R_j^2 \\
&= \tfrac{1}{n-1} \sum_{i=1}^{n-1} \delta_{is_0} R^2 \\
&= \tfrac{1}{n-1} (R^2) \\
&= \max W^2(\text{on } D)
\end{aligned}$$

This completes the proof.

□

For notations mentioned in Proposition 3.9, an immediate result is that, whenever $R := \sqrt{R_1^2 + \cdots + R_d^2}$, the global maximum value (respectively the global minimum value) of W^2 on D is equal to the global maximum value (respectively the global minimum value) of W^2 on $\prod_{j=1}^{d} D_j$.

Corollary 3.10. Let $\{(x_i, y_i)\}_{i=1}^n$, $n > 1$, be a finite sequence in $\mathbb{R}^2$ with sequence variance W^2. Let W_x^2 and W_y^2 be the sequence variance of the sequences $\{x_i\}_{i=1}^n$ and $\{y_i\}_{i=1}^n$, respectively. Then, we have $W^2 = W_x^2 + W_y^2$. Furthermore, suppose R_x and R_y are non-negative constants such that

$$R_x = \sum_{i=1}^{n-1} |x_{i+1} - x_i| \text{ and } R_y = \sum_{i=1}^{n-1} |y_{i+1} - y_i| .$$

Let $R := \sqrt{R_x^2 + R_y^2}$, and set $D := \{(d_1, \ldots, d_{n-1}) \in \mathbb{R}^{n-1} \mid \sum_{i=1}^{n-1} d_i = R\}$. Then the following statements are established.

1. W^2 has the global minimum $\frac{R^2}{(n-1)^2}$ on D, whenever $|x_{i+1} - x_i| = \frac{R_x}{n-1}$ and $|y_{i+1} - y_i| = \frac{R_y}{n-1}$, for all $0 \leq i \leq n-1$.
2. W^2 has the global maximum $\frac{R^2}{n-1}$ on D, whenever $|x_{i+1} - x_i| = \delta_{is} R_x$ and $|y_{i+1} - y_i| = \delta_{is} R_y$, where $1 \leq s \leq n-1$ is a constant index.

Proof . It follows from Proposition 3.9. □

Different ordering methods can lead to different interpretations of how "irregular" or "non-uniform" the spatial arrangement of points is. Here, we discuss the importance of defining and utilizing lexical order. To illustrate this, consider 8 points A, F, D, E, B, H, C, G arranged on a circle as shown in Figure 4.

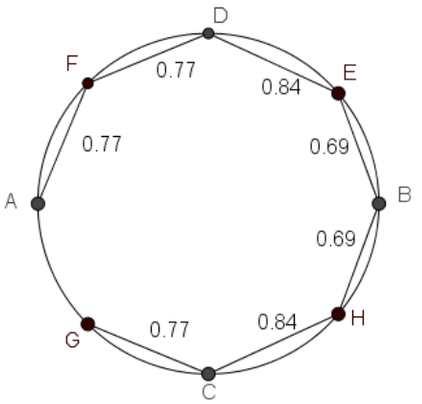

Figure 4: Data points ordered with a non-lexical order.

First, we define an order $\leq$ for the positions of the points as follows:

$$A^{\mathbf{p}} \leq F^{\mathbf{p}} \leq D^{\mathbf{p}} \leq E^{\mathbf{p}} \leq B^{\mathbf{p}} \leq H^{\mathbf{p}} \leq C^{\mathbf{p}} \leq G^{\mathbf{p}}.$$

where, for instance, $A^{\mathbf{p}}$ represents the coordinate of the point $A(x, y)$. Under this ordering, the corresponding sequence variance is calculated as:

$$W^2 = \frac{0.77^2 + 0.77^2 + 0.84^2 + 0.69^2 + 0.69^2 + 0.84^2 + 0.77^2}{7} \approx 0.59,$$

which gives the non-uniformity indicator $W \approx 0.77$. This result indicates that the average distance between the points on the circle is approximately 0.77 and so the data points are uniformly distributed along the boundary of the disk.

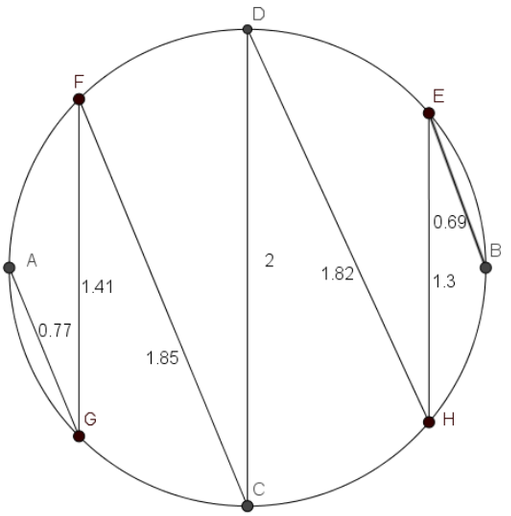

Figure 5: Data points ordered with a lexical order.

Next, we consider the lexical order $\leq^{lex}$ for the same set of points

$$A, F, D, E, B, H, C, G$$

as shown in Figure 5, defined as:

$$A^{\mathbf{p}} \leq^{lex} G^{\mathbf{p}} \leq^{lex} F^{\mathbf{p}} \leq^{lex} C^{\mathbf{p}} \leq^{lex} D^{\mathbf{p}} \leq^{lex} H^{\mathbf{p}} \leq^{lex} E^{\mathbf{p}} \leq^{lex} B^{\mathbf{p}}.$$

Under this lexical ordering, the corresponding sequence variance becomes:

$$W^2 = \frac{0.77^2 + 1.41^2 + 1.85^2 + 2^2 + 1.82^2 + 1.3^2 + 0.69^2}{7} \approx 2.21,$$

yielding a non-uniformity indicator $W \approx 1.49$. This suggests that the average distance between the points on the disk (inside and boundary of circle) is approximately 1.49, which is significantly larger than the value obtained under the non-lexical order. This is closer to the truth because, as a general rule, when data accumulates at the disk boundary, the

non-uniformity of the data on the disk tends to increase. This difference demonstrates that the choice of ordering method is significant, and that lexicographical order is particularly important for certain analyses on the plane.

Now, we introduce a new indicator called sequence covariance.

Definition 3.11. Let $\{(x_i, y_i)\}_{i=1}^n$, $n > 1$, be a finite sequence in $\mathbb{R}^2$. The **sequence covariance** of $\{(x_i, y_i)\}_{i=1}^n$, denoted by W_{xy} or $SeqCov(x, y)$, is given by

$$W_{xy} = SeqCov(x, y) = \frac{1}{n-1} \sum_{i=1}^{n-1} (x_{i+1} - x_i)(y_{i+1} - y_i) \ .$$

sequence covariance of linear combinations: For constants a, b, c and d, if (x, y) represents $\{(x_i, y_i)\}_{i=1}^n$, then

$$SeqCov(x, x) = SeqVar(x), \tag{3.3}$$

$$SeqCov(ax, by) = abSeqCov(x, y), \quad SeqCov(x + c, y + d) = SeqCov(x, y). \tag{3.4}$$

Definition 3.12. Let $\{(x_i, y_i)\}_{i=1}^n$, $n > 1$, be a finite sequence in $\mathbb{R}^2$. Let W_x and W_y be the non-uniformity indicator of the sequences $\{x_i\}_{i=1}^n$ and $\{y_i\}_{i=1}^n$, respectively, and W_{xy} the sequence covariance of $\{(x_i, y_i)\}_{i=1}^n$. The **sequence correlation coefficient** of $\{(x_i, y_i)\}_{i=1}^n$, denoted by r_{xy}, is given by

$$r_{xy} = \frac{W_{xy}}{W_x W_y} \ .$$

Proposition 3.13. Let $\{(x_i, y_i)\}_{i=1}^n$, $n > 1$, is a sequence in $\mathbb{R}^2$. Then

1. r_{xy} is a number between -1 and 1.
2. $r_{xy} = \pm 1$ if and only if x and y are linearly dependent and lie on a line with a non-zero slope.

Proof . Let W_x^2 and W_y^2 be the sequence variance of the sequences $\{x_i\}_{i=1}^n$ and $\{y_i\}_{i=1}^n$, respectively. It is easily to see that

$$W_{x+y}^2 = W_x^2 + W_y^2 + 2W_{xy} \ . \tag{3.5}$$

Now, let $z_i := \frac{y_i}{W_y} - \frac{x_i}{W_x}$. Then, using Equation 3.5, we have

$$\begin{aligned}
SeqVar(\tfrac{y_i}{W_y} - \tfrac{x_i}{W_x}) &= SeqVar(\tfrac{y_i}{W_y}) + SeqVar(-\tfrac{x_i}{W_x}) + 2SeqCov(\tfrac{y_i}{W_y}, \tfrac{-x_i}{W_x}) \\
&= SeqVar(\tfrac{y_i}{W_y}) + SeqVar(\tfrac{x_i}{W_x}) - 2SeqCov(\tfrac{y_i}{W_y}, \tfrac{x_i}{W_x}) \\
&= 1 + 1 - 2\tfrac{SeqCov(x,y)}{W_x W_y} \qquad \text{(by Proposition 4.2)} \\
&= 2 - 2r_{xy}.
\end{aligned}$$

Since $W_z^2 \geq 0$, we have $r_{xy} \leq 1$. The assertion $r_{xy} \geq -1$ is obtained by interchanging $z_i := \frac{y_i}{W_y} - \frac{x_i}{W_x}$ with $z_i := \frac{y_i}{W_y} + \frac{x_i}{W_x}$ and using Equation 3.5.

To prove the second assertion, first, consider $r_{xy} = 1$. Then, by placing it in the relation

$$SeqVar(\frac{y_i}{W_y} - \frac{x_i}{W_x}) = 2 - 2r_{xy},$$

we have $SeqVar(\frac{y_i}{W_y} - \frac{x_i}{W_x}) = 0$ and hence $\frac{y_{i+1}}{W_y} - \frac{x_{i+1}}{W_x} = \frac{y_i}{W_y} - \frac{x_i}{W_x}$, for all $1 \leq i \leq n-1$. Equivalently,

$$\frac{y_{i+1} - y_i}{W_y} = \frac{x_{i+1} - x_i}{W_x} \tag{3.6}$$

for all $1 \leq i \leq n-1$. Thus all of (x_i, y_i)'s lie on a line with a positive slope $\frac{W_y}{W_x}$, and the equation of this line is equal to $y - \overline{y} = \frac{W_y}{W_x}(x - \overline{x})$. Now, let $r_{xy} = -1$. Using equation $SeqVar(\frac{y_i}{W_y} + \frac{x_i}{W_x}) = 2 + 2r_{xy}$, we have

$$\frac{y_{i+1} - y_i}{W_y} = \frac{x_{i+1} - x_i}{-W_x} \tag{3.7}$$

for all $1 \leq i \leq n-1$. So that all of (x_i, y_i)'s lie on the line $y - \overline{y} = -\frac{W_y}{W_x}(x - \overline{x})$ whose slope is negative. For the reverse, suppose there is a linear relationship $y = ax + b$ between x_i's and y_i's, where a and b are constants. Then

$$\begin{aligned}
r_{xy} &= \frac{W_{xy}}{W_x W_y} \\
&= \frac{SeqCov(x, ax+b)}{\sqrt{SeqVar(x) SeqVar(ax+b)}} \\
&= \frac{a SeqCov(x,x)}{\sqrt{a^2 SeqVar(x) SeqVar(x)}} && \text{(by Equation 3.4 and Proposition 3.8)} \\
&= \frac{a SeqVar(x)}{|a| SeqVar(x)} && \text{(by Equation 3.3)} \\
&= \frac{a}{|a|} \ ,
\end{aligned}$$

So $r_{xy} = 1$ if $a > 0$, and $r_{xy} = -1$ if $a < 0$. This completes the proof. □

Proposition 3.14. Let $\{(x_i, y_i, z_i)\}_{i=1}^n$, $n > 1$, is a sequence in $\mathbb{R}^3$. Then the following statements are equivalent.

1. $r_{xy} = \pm 1$ and $r_{yz} = \pm 1$.
2. (x_i, y_i, z_i)'s lie on a line not perpendicular to X, Y and Z axes.

Proof . First, suppose that (1) holds. To show that (2) holds, first assume that $r_{xy} = r_{yz} = 1$, then, by the proof of Proposition 3.13, we have $\frac{y_{i+1} - y_i}{W_y} = \frac{x_{i+1} - x_i}{W_x}$ and $\frac{z_{i+1} - z_i}{W_z} = \frac{y_{i+1} - y_i}{W_y}$. So (x_i, y_i, z_i)'s lie on the line

$$\frac{x - \overline{x}}{W_x} = \frac{y - \overline{y}}{W_y} = \frac{z - \overline{z}}{W_z} \ .$$

If $r_{xy} = 1$ and $r_{yz} = -1$, again by the proof of Proposition 3.13, we have $\frac{y_{i+1}-y_i}{W_y} = \frac{x_{i+1}-x_i}{W_x}$ and $\frac{z_{i+1}-z_i}{-W_z} = \frac{y_{i+1}-y_i}{W_y}$. So that (x_i, y_i, z_i)'s lie on the line

$$\frac{x-\overline{x}}{W_x} = \frac{y-\overline{y}}{W_y} = \frac{z-\overline{z}}{-W_z} \ .$$

Similarly, if $r_{xy} = -1$ and $r_{yz} = 1$, (x_i, y_i, z_i)'s lie on the line

$$\frac{x-\overline{x}}{-W_x} = \frac{y-\overline{y}}{W_y} = \frac{z-\overline{z}}{W_z} \ ,$$

and, finally if $r_{xy} = r_{yz} = -1$, (x_i, y_i, z_i)'s lie on the line

$$\frac{x-\overline{x}}{W_x} = \frac{y-\overline{y}}{-W_y} = \frac{z-\overline{z}}{W_z} \ .$$

Now, we show that (2) yields (1). According to the assumption, (x_i, y_i, z_i)'s lie on the line

$$\frac{x-x_0}{a} = \frac{y-y_0}{b} = \frac{z-z_0}{c} \ ,$$

where a, b and c are non-zero constants and (x_0, y_0, z_0) is a fixed point in $\mathbb{R}^3$. Therefore, $y = \frac{b}{a}x - \frac{b}{a}x_0 + y_0$ and $z = \frac{c}{b}y - \frac{c}{b}y_0 + z_0$, and so, by the proof of Proposition 3.13,

$$r_{xy} = \frac{W_{xy}}{W_x W_y} = \frac{SeqCov(x, \frac{b}{a}x - \frac{b}{a}x_0 + y_0)}{\sqrt{SeqVar(x)SeqVar(\frac{b}{a}x - \frac{b}{a}x_0 + y_0)}} = \frac{\frac{b}{a}}{|\frac{b}{a}|} \ .$$

If a and b are with the same sign, then $r_{xy} = 1$, else $r_{xy} = -1$. Similarly, we can show $r_{yz} = \pm 1$. □

Table 1: The height data obtained from www.stampscreeningtool.org

hight	average hight for i'th years, i=1,...,19	sequence correlation
boys	87.1, 96.1, 102,5, 109.6, 115.9, 121.9, 127.9, 133.3, 138.4, 143.4,148.4,154.8,162.4,168.9,173.4,175.9,177	0.95
girls	85.7, 95.0, 101.5, 108.9, 115.3, 121.3, 127.3, 132.8, 138.4, 144.1, 149.8, 155.3, 159.6, 162.2, 163.2, 163.5, 163.5	0.89

Now, we provide an example that illustrates the use of the sequence correlation coefficient.

Example 3.15. Table 1 illustrates the average height of boys and girls aged between 2 and 18 years in UK.

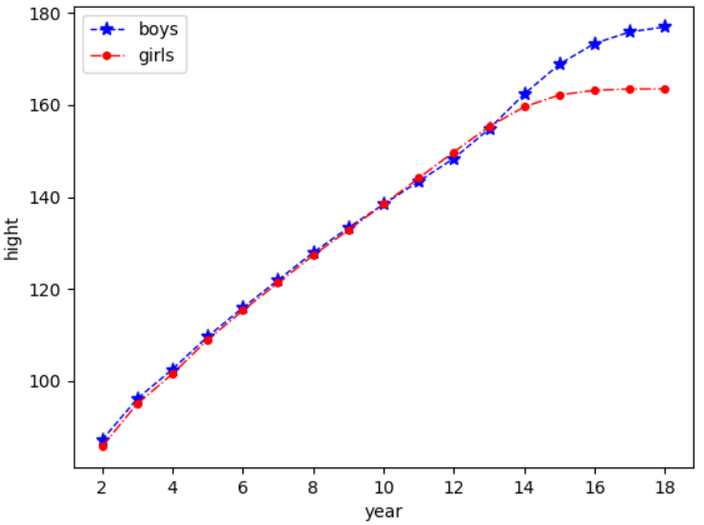

Figure 6: Time series curves of height in Figure 1

As shown in Figure 6, girls height increases at a slower rate starting from age 14, resulting in a decrease in the slope of their growth curve. Additionally, the sequence correlations presented in table 1 indicate that the relationship between age and height is stronger for boys compared to girls during the ages of 2 to 18 years.

4 Outlier detection

In this section, we will begin by introducing a test for identifying outliers in $\mathbb{R}$, along with a screening test for detecting suspicious outliers in $\mathbb{R}^d$ when $d > 1$. Before presenting these tests, we will define outlier data.

Definition 4.1. Let $\mathbf{x}_1 \leq \cdots \leq \mathbf{x}_n$ be a sequence in $\mathbb{R}^d$ with sequence variance W. For $1 \leq i \leq n$, we call $\mathbf{x}_i$ an **outlier**, if the open unit ball centered at $\frac{\mathbf{x}_i}{W}$ contains one member of the set $\{\frac{\mathbf{x}_1}{W}, \ldots, \frac{\mathbf{x}_n}{W}\}$. Let $\mathbf{z}_i := \frac{\mathbf{x}_i}{W_{\mathbf{x}}}$, for all $1 \leq i \leq n$ and call them **standard data**.

Proposition 4.2. Let $\{\mathbf{x}_i\}_{i=1}^n$, $n > 1$, be a sequence in $\mathbb{R}^d$ with the sequence variance $W_{\mathbf{x}}^2$. Then the sequence of standard data $\{\mathbf{z}_i\}_{i=1}^n$ has the sequence variance $W_{\mathbf{z}}^2$ being equal to 1.

Proof . It follows from definition. □

Remark 4.3. (**Outlier Detection Test**) For a sequence $\{\mathbf{x}_i\}_{i=1}^n$, $n > 1$, in $\mathbb{R}^d$, preferably arranged with natural order on real numbers $\mathbb{R}$ if $d = 1$ and lexical order on $\mathbb{R}^d$ if $d > 1$, we let distances $d_i := \|\mathbf{x}_{i+1} - \mathbf{x}_i\|$, $i = 1, \ldots, n-1$, and define the standard distances $D_i := \frac{d_i}{W_{\mathbf{x}}}$, $1 \leq i \leq n-1$. Then

1. if D_1 (respectively D_{n-1}) is greater than 1, then $\mathbf{x}_1$ (respectively $\mathbf{x}_n$) is an outlier when $d = 1$ and $\mathbf{x}_1$ (respectively $\mathbf{x}_n$) is likely an outlier when $d > 1$.
2. For $1 \leq i \leq n-2$, if D_i and D_{i+1}, are greater than 1, then $\mathbf{x}_{i+1}$ is an outlier when $d = 1$ and $\mathbf{x}_1$ (respectively $\mathbf{x}_n$) is likely an outlier when $d > 1$.

It offers a method for detecting outliers in $\mathbb{R}$. Furthermore, when a point $\mathbf{x}$ in $\mathbb{R}^d$, $d > 1$, is suspected to be an outlier, we calculate the distance of $\frac{\mathbf{x}}{W_{\mathbf{x}}}$ from all the standard data of a cluster from which this data has been separated. If all distances are greater than 1, we confirm that $\mathbf{x}$ is indeed an outlier. This is a fast way to identify outliers in a sequence in $\mathbb{R}^d$. That is, when calculating the standard distance between ordered data points, we focus only on those points that are further away from their preceding and succeeding points, as these are considered potential outliers rather than all data points. For example, consider the time series

$$(1,1), (2,2), (3,3), (4,40), (5,5), (6,60), (7,7), (8,8), (9,9).$$

Using outlier detection test, only the data points $(4, 40)$, $(5, 5)$, and $(6, 60)$ are identified as being further away from their neighboring points and are thus considered potential outliers. However, since $(4, 40)$ and $(6, 60)$ are further away from all other data points, they are confirmed as outliers, whereas $(5, 5)$ is not.

The main similarities between variance and sequence variance as the indicators of the dispersion of data were given in Proposition 3.8. Now, we will examine the differences between variance and sequence variance. There are a number of key distinctions between these two concepts. The main differences pertain to their scope and application; variance indicators the dispersion of data points around the mean in 1-dimensional space $\mathbb{R}$, while sequence variance can be computed for any sequence in multidimensional spaces. Additionally, regardless of whether a dataset is ordered, variance yields a single numerical value, whereas this is not true for sequence variance. Below, we highlight further differences.

Remark 4.4. 1. **Sequence variance as an indicator for detecting outliers**: For the sequence $x_{(1)} \leq \cdots \leq x_{(n)}$ of variables with a constant range R, larger gaps between the data points are likely to result in a greater sequence variance (see Remark 3.4), although this does not necessarily apply to variance. For instance, if we examine two sequences with the same range in Table 2, we find that the sequence variance values differ significantly between the two sequences, while the variance for each sequence

remains relatively similar. This indicates the presence of an outlier in the first-row sequence

Table 2: non-uniformity measurement of 2 sequences

sequence	variance	sequence variance	non-uniformity indicator
$(1, 2, 3, 10)$	12.50	17.0	4.12
$(1, 4, 8, 10)$	12.19	9.67	3.11

For the sequence of the first row, using Part 1 of Remark 4.3, we have

$$D_1 = D_2 = \frac{1}{4.12} \approx 0.24 \quad \text{and} \quad D_3 = \frac{7}{4.12} \approx 1.7 \ .$$

Since D_3 is significantly greater than 1, then 10 is an outlier.
Now, we give another example to find outlier. For this, consider the following time series of values over consecutive time-stamps

$$3, 2, 3, 2, 3, 87, 86, 85, 87, 89, 86, 3, 84, 91, 86, 91, 88.$$

The time series is illustrated in Figure 7. It is evident that there is a sudden change

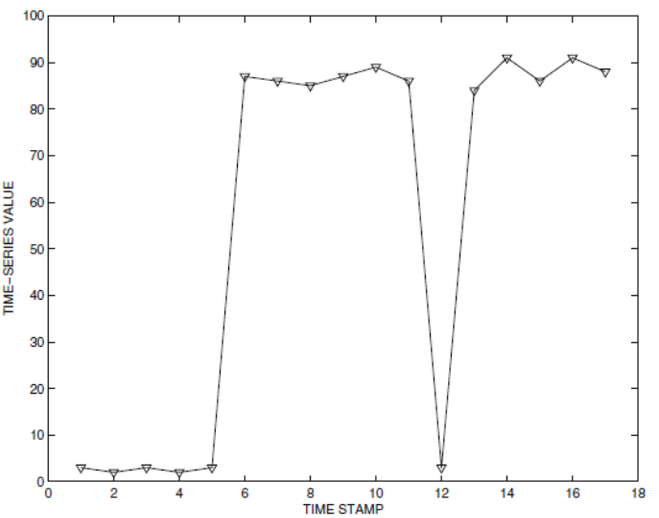

Figure 7: Example of Time Series. source: [2],Page22

the data values at time-stamp 12 from 86 to 3 and from 3 to 84. We express the series in the following format:
$(1, 3), (2, 2), (3, 3), (4, 2), (5, 3), (6, 87), (7, 86), (8, 85), (9, 87), (10, 89),$

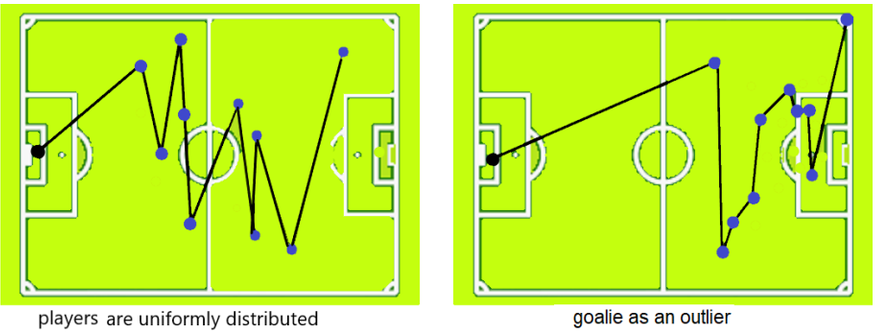

Figure 8: Non-uniformity of Team A's player in the right field

$(11, 86), (12, 3), (13, 84), (14, 91), (15, 86), (16, 91), (17, 88)$. Now, using Part 2 of Remark 4.3, we have sequence variance of bivariant data is 1290.8125. The distance between $(11, 86)$ and $(12, 3)$ is 83.006 and distance between $(12, 3)$ and $(13, 84)$ is 81.006.

$$D_{11} = \frac{83.006}{35.93} \approx 2.31 \quad \text{and} \quad D_{12} = \frac{81.006}{35.93} \approx 2.25 \ .$$

Since D_{11} and D_{12} are significantly greater than 1, then 3 corresponds to an outlier. The second method can also help identify outliers in a football match. For example, if a team's players, represented as blue dots in Figure 8, are predominantly active near the opposing goal, the goalkeeper may see himself/herself farther away from their teammates due to this increased distance. As a result, the goalkeeper's position is deemed an outlier.

2. **Sequence variance as a metric to determine dense data localities**: We begin by examining this property for one-dimensional data. Consider a collection of sequences with n terms, $x_{(1)} \leq \cdots \leq x_{(n)}$, and equal range R. For symmetric data, the variance is usually lower than that of right-skewed or left-skewed data, as variance measures the dispersion of data around the center. However, this does not hold true for sequence variance.

   To illustrate, consider Table 3, which contains three sequences with a fixed range of 50. The second row represents symmetric data, while the first and third rows represent asymmetrically distributed data. Using a Python code, we observe that the sequence variances of all three datasets are equal 46.67, whereas their variances differ. This indicates that sequence variance has approximately the same value for all sequences $\{x_i\}_{i=1}^{n}$, $n > 1$, that exhibit equal kurtosis at the center or edges of a fixed range.

   Algorithm for determining dense data localities:

   To accurately identify dense data localities within the three sequences in Table 3, we divide the domain into two equal intervals: $[0, 25]$ and $(25, 50]$. We then calculate

Table 3: Symmetric Data and Skewed Data

Sequence	Variance	sequence variance
$(0, 10, 12, 14, 16, 18, 20, 30, 40, 50)$	201.0	46.67
$(0, 10, 20, 22, 24, 26, 28, 30, 40, 50)$	177.0	46.67
$(0, 10, 20, 30, 32, 34, 36, 38, 40, 50)$	201.0	46.67

the sequence variance for the data points in each interval. A lower sequence variance generally indicates a higher density of points in that interval. If the sequence variances of the two intervals are equal, we further subdivide them into smaller intervals and repeat the process until the region with the highest data density is identified.

This algorithm works effectively in most cases, though exceptions may arise in special scenarios. For instance, applying this algorithm to the first sequence in Table 3, $(0, 10, 12, 14, 16, 18, 20, 30, 40, 50)$, we compute the sequence variance for the first segment $(0, 10, 12, 14, 16, 18, 20)$ as 20 and for the second segment $(30, 40, 50)$ as 100. Since the sequence variance of the first segment is lower than that of the second segment, and both segments have equal domain lengths, this indicates a higher data density in the first interval compared to the second."

Extension to higher dimensions:

This method can also be applied to determine dense regions in d-dimensional Euclidean space. For instance, consider a dataset in $\mathbb{R}^2$. First, we divide the plane containing the two-dimensional data into four equal parts, as shown in Figure 9. We calculate the sequence variance for the data points in each part, and iteratively subdivide the regions until the part with the lowest sequence variance and thus the highest data density is identified.

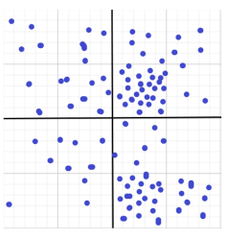

Figure 9: Determination of dense point locations through successive divisions.

In summary, sequence variance serves as an effective tool for identifying dense data localities in both one-dimensional and higher-dimensional spaces. Its ability to remain consistent across symmetric and skewed data distributions makes it particularly useful for analyzing sequential and spatial relationships.

3. **Sequence variance as an indicator of disorder**: Consider a dataset organized as

$\mathbf{x}_1 \leq \cdots \leq \mathbf{x}_n$. Whenever the values in the dataset change, we calculate the sequence variance while keeping the original order intact. By preserving the initial arrangement of the data, any alterations to the dataset contents in changes to the sequence variance, which in most instances leads to an increase in its value. For example, Table 4 presents

Table 4: Average monthly air temperature in a city

year	time series points	W^2
2021	(1,2),(2,9),(3,1),(4,19),(5,23),(6,25)	92.4
2022	(1,1),(2,7),(3,12),(4,14),(5,20),(6,22)	22.0

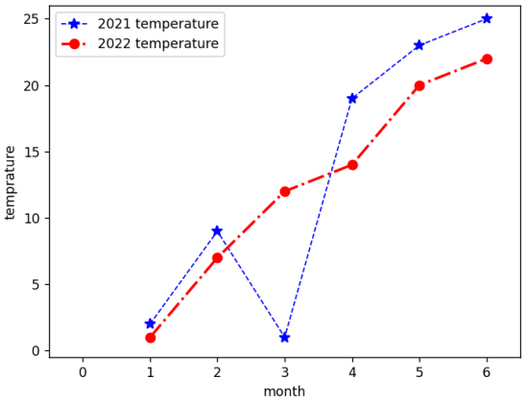

Figure 10: Time series curves of Table 4

two lists of temperatures recorded in a hypothetical city for the first six months of 2 years. Each i-th pair includes the i-th Gregorian month as the first element and the corresponding temperature as the second element. It is noteworthy that the time series is implicitly structured in lexicographical order. In March 2021, for instance, the air temperature dropped below the normal range, resulting in an extended length of the blue broken line depicted in Figure 10 due to the outlier. Consequently, the associated sequence variance is anticipated to increase (refer to the third column of Table 4). An increase in sequence variance indicates a higher degree of disorder.

Finally to make comparison clearer, we summarize the differences between variance an sequence variance in Table 5 .

Acknowledgements

We would like to express our sincere gratitude to Professor Ali Akbar Estaji and the anonymous referee(s) for their valuable assistance and insightful contributions to the math-

Table 5: A Comparative Analysis of Variance and sequence variance

property	variance, covariance	sequence variance, sequence covariance
definition	deviation from mean	distance between consecutive observations
invariant	invariant under permutations	no invariant under permutations
data type	suitable for unorder data	suitable for order data
applications	general statistical analysis	time-series analysis, sequence analysis
advantages	robust, widely applicable	focus on sequential relationships

ematical and statistical aspects of this paper.

References

[1] A. A. Estaji, A. Mahmoudi Darghadam, *Rings of real measurable functions vanishing at infinity on a measurable space*, J. Frame Matrix Theor., **1** (2024) 1–21. doi

[2] Ch. C. Aggarwal, *Outlier Analysis*, 2rd ed., Springer, Cham, New York, USA, 2017. zbl MR doi

[3] H. A. David, H. N. Nagaraja, *Order Statistics*, 3rd ed., Series, John Wiley and Sons, Hoboken, New Jersey, Canada, 2003. zbl MR doi

[4] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear programming. Theory and algorithms*, 3rd ed., Hoboken, NJ: John Wiley and Sons, New York, Amsterdam, 2006. zbl MR doi

[5] A. Bondy, U. S. R. Murty, *Graph Theory with Applications*, 5rd ed., Elsevier Science Publishing Co., Inc., the University of Michigan, USA, 1982. pdf

[6] J. E. Freund, *Mathematical Statistics*, 5th ed., Englewood Cliffs, NJ: Prentice-Hall International, Inc., 1992. zbl

[7] T. Haghdadi, *c_L-regular-open and c_L-regular-closed elements by clouser function in pointfree topology*, J. Frame Matrix Theor., **1** (2024) 22–36. doi

[8] K. Hoffman, R. Kunze, *Linear Algebra*, 2th ed., Frentice-Hall, inc., New Jersey, 1971. zbl MR

[9] D. M. Hawkins, *Identification of Outliers*, 2th ed., Chapman and Hall, London, 1980. zbl MR

[10] S. Y. T. Lin, Y. F. Lin, *Set Theory with Applications*, 2th ed., Mariner Publishing

Company, the University of Michigan, USA, 1981. link

[11] P. P. C. Alzate, J. R. G. Granada, *On the Coercive Functions and Minimizers*, Adv. stud. theor. phys., **11** (2017) 709–715. pdf