Moreno, A., Viveros, A. y Ramírez, R. (2021). Statistical significance of the median of a set of points on the plane. *The College Mathematics Journal*, 52(3), 205-218

https://doi.org/10.1080/07468342.2021.1909969

PREPRINT

Statistical significance of the median of a set of points on the plane.

The word median appears in two apparently different mathematical contexts.

In statistics it appears as a measurement of central position and in geometry as one of the lines determined by the barycenter. Why do we use the same word if they are apparently different things?

We can intuit a first relationship. In geometry, a medial line in a triangle passes through a vertex and splits the triangle into two equal areas. This property could be related to the definition of the median on the real line, but the intersection point of the medians is the barycenter, which turns out to be the arithmetic mean of the three vertices, understanding the barycenter as the centre of gravity or a figure's centre of mass.

Different authors have provided interesting connections between the meanings of statistical concepts and their geometric interpretation (Herr, 1980, Margolis, 1979, Sakar et al., 2016). In this study, we intend to provide an interpretation of the median in the plane as a measurement of position from the analysis of its geometric properties.

1. Geometric properties of the median on the line of real numbers.

To begin, we will be analyzing some well-known number and algebraic properties. We will consider *n* points in the line of numbers.

We will see, then, what happens with the median when we have a set of points located on a line:

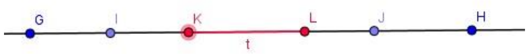
If *n* is odd, the median is located at the point that is at the position $\frac{n+1}{2}$

We can illustrate this in the following way:

Figure 1: Median when *n* is odd

The median is found at point M, coinciding with the position indicated above.

If *n* is even, the median tends to associate with the interval formed by the central points while considering the median as the average value of the extremes of this interval in many cases, as we can see in the following example:



Moreno, A., Viveros, A. y Ramírez, R. (2021). Statistical significance of the median of a set of points on the plane. *The College Mathematics Journal*, 52(3), 205-218

https://doi.org/10.1080/07468342.2021.1909969

PREPRINT

Figure 2: Median when n is even

If we generalize this idea, any point between K and L that we choose could be the point that divides the distribution of points into two equal parts, leaving three to the right and three to the left.

We could geometrically interpret that there are infinite points *in the middle*, that is, there are as many points of distribution to their left as to their right.

We observe a characteristic that all these points fulfill:

The Median minimizes the sum of all absolute deviations. In other words, if we represent the median by M, we get,

$$\min_k \sum_{i=1}^{n} |x_i - k| = \sum_{i=1}^{n} |x_i - M|.$$

When the constant, with respect to which the deviations are taken, is equal to the median, the sum of absolute deviations is minimal.

We can illustrate this, taken $f(k) = \sum_{i=1}^{n} |x_i - k| \; with \; x_i = 1,2,4,10$

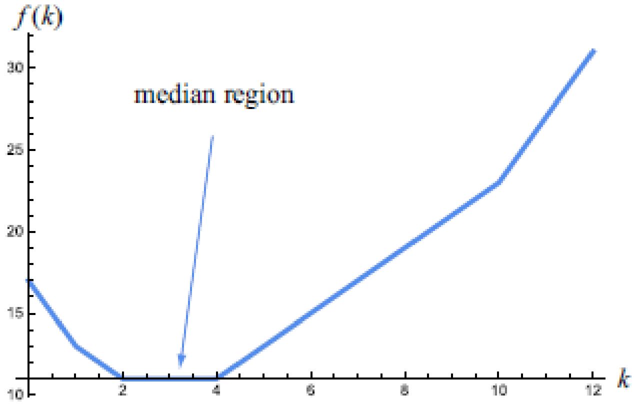


Figure 3: Median for the points 1,2,4,10

However, if we consider the uniqueness of the median, as in the case when n is odd, we refer to the midpoint of the interval formed by K and L.

The previous property admits a version in the Euclidean plane, substituting the absolute value for distance in the plane. In this case, we look for the point of the plane that minimizes the sum of distances to all points.

Moreno, A., Viveros, A. y Ramírez, R. (2021). Statistical significance of the median of a set of points on the plane. *The College Mathematics Journal*, 52(3), 205-218

https://doi.org/10.1080/07468342.2021.1909969

PREPRINT

For this, we pose two questions to guide our reflection:

- How do we find the point that minimizes the sum of the distances?
- Does the previous point provide any significance about the central position of the set of points?

2. Fermat-Torricelli

The answer to the first question was given for 3 points around the year 1659 and, subsequently, was generalized to n points in a distribution.

This point is called the Fermat-Toricelli.

The name *Fermat point* is given to the point of the plane for which the sum of the distances to the vertices of a given triangle, each of whose interior angles are less than 120º, is the minimum possible (García (2005) and Bernal and Tornel, sf.)

According to the questions that we have previously posed, the definition of median in the plane is optimizing a geometric problem relative to the sum of distances. In relation to question 2, we will analyze some of the most common procedures to calculate it and to obtain properties relative to the position (as a measure of statistical centralization).

From an analytical perspective, and reducing to three points in the plane, the function to be minimized is:

$$f(x,y) = \sqrt{(a_1 - x)^2 + (a_2 - x)^2} + \sqrt{(b_1 - x)^2 + (b_2 - x)^2} + \sqrt{(c_1 - x)^2 + (c_2 - x)^2}$$

where (a_1, a_2), (b_1, b_2), (c_1, c_2) are the coordinates of the vertices in a plane, (x, y) is the point we are looking for.

Therefore, the first thing we should consider is an analytical view of the problem in order to evaluate what information it gives us.

2.1 Analytical view of the *Fermat Point*

We will illustrate this section by using the position of the triangle proposed in Sángari and Egües (2012).

Applying the appropriate rotations, reflections, and translations, we can assume that the three vertices of the triangle are, A = (0,0), B = (a, 0), C = (b, c):

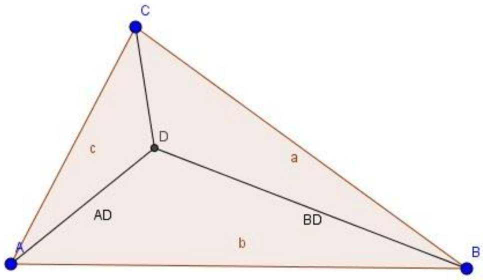


Figure 4: Initial position

Once we have established the starting situation, if we make the sum of the distances to the Fermat point (D = (x, y)), we can obtain the function that we have to minimize:

$$f(x,y) = \sqrt{x^2+y^2} + \sqrt{(x-a)^2+y^2} + \sqrt{(x-b)^2+(y-c)^2}\ .$$

To be more concise, henceforth $d_1 = \sqrt{x^2+y^2}$, $d_2 = \sqrt{(x-a)^2+y^2}$ and $d_3 = \sqrt{(x-b)^2+(y-c)^2}$.

Accordingly, to minimize the sum of distances, we set the gradient of f to the vector (0, 0):

$$\nabla f(x,y) = \frac{(x,y)}{d_1} + \frac{(x-a,y)}{d_2} + \frac{(x-b,y-c)}{d_3} = (0,0).$$

To continue, the gradient is a system of equations with two unknowns:

$$\frac{x}{d_1} + \frac{x-a}{d_2} + \frac{x-b}{d_3} = (0,0)$$

$$\frac{y}{d_1} + \frac{y}{d_2} + \frac{y-c}{d_3} = (0,0)$$

A solution to the above system is also a solution to the system:

$$\left(\frac{x}{d_1}\right)^2 + \left(\frac{x-a}{d_2}\right)^2 + 2\frac{(x-0)(x-a)}{d_1 d_2} = \left(\frac{x-b}{d_3}\right)^2$$

$$\left(\frac{y}{d_1}\right)^2 + \left(\frac{y}{d_2}\right)^2 + 2\frac{(y-0)(y-0)}{d_1 d_2} = \left(\frac{y-c}{d_3}\right)^2$$

Moreno, A., Viveros, A. y Ramírez, R. (2021). Statistical significance of the median of a set of points on the plane. *The College Mathematics Journal*, 52(3), 205-218

https://doi.org/10.1080/07468342.2021.1909969

PREPRINT

If we now add the two equations, we will obtain:

$$\left(\frac{x}{d_1}\right)^2+\left(\frac{x-a}{d_2}\right)^2+2\frac{(x-0)(x-a)}{d_1d_2}+\left(\frac{y}{d_1}\right)^2+\left(\frac{y}{d_2}\right)^2+2\frac{(y-0)(y-0)}{d_1d_2}$$
$$=\left(\frac{x-b}{d_3}\right)^2+\left(\frac{y-c}{d_3}\right)^2$$

Using the relations for d_1, d_2, d_3 in terms of x, y, a, b, c, we have the identity:

$$\left(\frac{x}{d_1}\right)^2+\left(\frac{x-a}{d_2}\right)^2+\left(\frac{y}{d_1}\right)^2+\left(\frac{y}{d_2}\right)^2-\left(\frac{x-b}{d_3}\right)^2\left(\frac{y-c}{d_3}\right)^2=1$$

Then, substituting in the previous equation, we obtain the following equation:

$$2\frac{(x-0)(x-a)+(y-0)(y-0)}{d_1d_2}+1=0$$

By using the cosine definition of the angle between two vectors, we can observe the following:

$$\frac{(x-0)(x-a)+(y-0)(y-0)}{d_1d_2}=\frac{(x,y)\cdot(x-a,y)}{d_1d_2}=\cos(\widehat{uv}),$$

Where $\widehat{uv}$ is the angle at D between AD and BD.

Where u = AD and v = BD, therefore we can see that:

$$\cos(\widehat{uv}) = -\frac{1}{2}$$

This implies that $\widehat{uv} = 120º$, as we expected from the bibliographic references and in consequence each interior angle of ΔABC must be less than 120º. In addition, performing the corresponding calculations of the Hessian we can see that if f has a critical point in the points where it is infinitely differentiable, then that critical point will be a minimum.

In relation to the “central position” of the solution, an interesting characteristic is obtained: the angles at D between AD and CD and between BD and CD are also 120º. However, there are cases in which the median point does not meet this condition, as is the case in which a triangle has an angle greater than 120º.

We are going to study this difference between cases based on geometric constructions.

2.2 Geometric construction of Fermat's point.

The Torricelli algorithm, to whom the authorship of the first solution of the problem is attributed, consists of three stages:

- First, we form the triangle formed by A, B and C, the three vertices.

- By using each of the sides of the triangle as a base, an equilateral triangle is constructed with the new vertex outside ΔABC.

- The circumscribed circles are drawn to each of the initial equilateral triangles. The intersection of the three circles is a single point, which is called the *Fermat-Torricelli point*, and produces the minimum sum of the distances to the vertices.

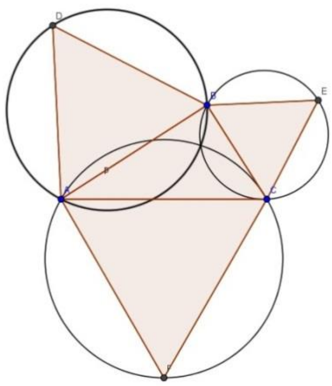


Figure 5: Torricelli's demonstration

The only difference between the algorithm devised by Torricelli and the one described later by Simpson is found in the last step, which is modified as follows:

- (Simpson) Through a segment, each of the outer vertices of the previous equilateral triangles is joined to the vertex of the triangle that does not belong to the equilateral triangle.

Moreno, A., Viveros, A. y Ramírez, R. (2021). Statistical significance of the median of a set of points on the plane. *The College Mathematics Journal*, 52(3), 205-218

https://doi.org/10.1080/07468342.2021.1909969

PREPRINT

- These lines are called Simpson lines and they intersect in the Fermat-Torricelli point.

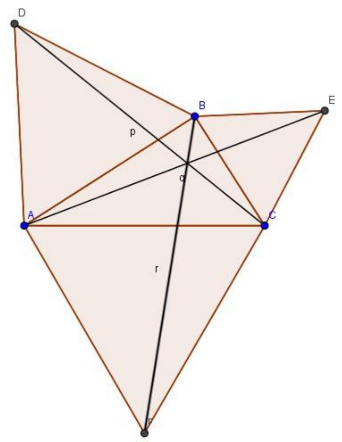


Figure 6: Simpson's demonstration

For more details of the demonstration by means of the congruence, see Cárdenas (2004).

However, an important exception to the previous results was discovered by Heinen in 1834, and it states that if one of the angles of the triangle is greater than or equal to 120º, neither of the two previous methods is valid, so the optimal point to minimize the sum of distances corresponds to the vertex of the triangle related to the angle greater than or equal to 120º (Krarup and Vajda, 1997).

Finally, Bernal and Tornel (S.F.), assure us that through the results we have obtained, in the conditions where the interior angles of the triangle are less than 120º, the solution to the problem is unique and it is the one given by Simpson's resolution.

Finally, we can offer a conclusion of the central position for 3 points:

- If the triangle has the angles less than 120º:

  The two construction procedures (the capable arch and the similarity of triangles) are based on locating the point that has the angle of *view* at any two of the vertices of 120º.

- If the triangle has an angle greater than 120º:

  The Fermat point is located at the vertex of the triangle that forms the largest angle.

In the limited case in which the obtuse angle of a triangle approaches 180º, the Fermat point will approach the median of the three vertices as aligned points (real case).

As a measurement of centralization, we can interpret that the median of the three points of the plane is a point from which the viewing angles to the rest of points (distinct from it) must be greater than or equal to 120º (which is the division of 360º into three equal parts).

We will study the case of four points to see if this characteristic is generalized.

3. Localization of Fermat Point for n points.

We will check if the previous feature characterizes the median for a set of points greater than three. Is the 120º angle of the view maintained?

In the case of n points in the plane, the search for 120º viewing angles is associated with the construction of Steiner trees.

The Steiner tree of a set of points is a graph that connects all points and has minimum length. Its solution highlights only the 120º property (Figure 5) and indeed the Fermat point determines the Steiner tree for three points.

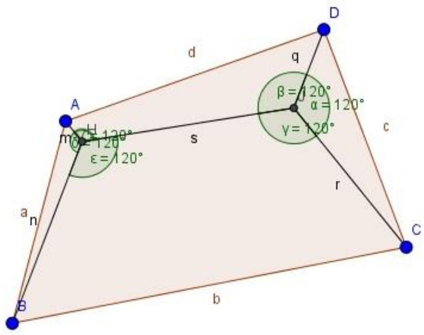


Figure 7: Steiner tree for a quadrilateral

The Steiner tree for a quadrilateral set is not unique and it does not determine a single point that minimizes the sum of the distances to the vertices.

3.1 Generalization of Fermat's problem, Weber's problems

In the 20th century the generic problem of Fermat (median two-dimensional) and its solution were taken up again to be used in real number cases by Alfred Weber (1868-1958) (García, 2005).

Weber's problem consists of calculating the optimal location of a set of points so that the weighted sum of the distances from that point to a set of given points is minimal.

The Hungarian mathematician Endre Vaszonyi Weiszfeld (now known as Andrew Vaszonyi) was the first to develop an iterative algorithm to find the solution to the Weber problem and, therefore, for the spatial median.

The simplest and most recurrent technique to solve Weber's problem is the so-called Weiszfeld algorithm, which is still the most common method for solving it, although there may be a set of initial points for which the algorithm does not converge.

3.1.1 Weiszfeld solution

Weiszfeld considers the solution including weights, that is, that some points have some characteristics in such a way that they are more important than others.

Although we give a general solution, we will see examples assuming that all weights are equal, which is the case that we are studying.

We consider the function:

$$W(x,y) = \sum_{i=1}^{m} w_i \sqrt{(x-a_{i1})^2 + (y-a_{i2})^2}$$

where w_i are the weights of each of the points.

If the vertices are contained in a line, the problem is easily solved. Therefore, even if it is not specified, we will assume that the vertices are not aligned.

We try to get the minimum of the previous function by imposing the first-order optimization condition and by obtaining the following system of non-linear equations:

$$\frac{\partial W(x,y)}{\partial x} = \sum_{i=1}^{m} \frac{w_i(x-a_{i1})}{\sqrt{(x-a_{i1})^2+(y-a_{i2})^2}} = 0$$

$$\frac{\partial W(x,y)}{\partial y} = \sum_{i=1}^{m} \frac{w_i(y-a_{i2})}{\sqrt{(x-a_{i1})^2+(y-a_{i2})^2}} = 0$$

The function $W(x,y)$ is convex, as demonstrated by the mathematician Love in 1967. Therefore, these conditions define a minimum. However, it is evident that partial derivatives do not exist when (x, y) matches some of the given points $a_i = (a_{i1}, a_{i2})$.

However, we can verify at the beginning of the algorithm if the optimal point is one of the vertices (Cañavate et al, 2001).

Nevertheless, we are interested in observing the position of the general solution in order to obtain the solution of the equations system. To achieve this, we must first partially resolve the previous equations, x and y:

Moreno, A., Viveros, A. y Ramírez, R. (2021). Statistical significance of the median of a set of points on the plane. *The College Mathematics Journal*, 52(3), 205-218

https://doi.org/10.1080/07468342.2021.1909969

PREPRINT

$$x = \frac{\sum_{i=1}^{m} \frac{w_i a_{i1}}{\sqrt{(x-a_{i1})^2 + (y-a_{i2})^2}}}{\sum_{i=1}^{m} \frac{w_i}{\sqrt{(x-a_{i1})^2 + (y-a_{i2})^2}}}$$

$$y = \frac{\sum_{i=1}^{m} \frac{w_i a_{i2}}{\sqrt{(x-a_{i1})^2 + (y-a_{i2})^2}}}{\sum_{i=1}^{m} \frac{w_i}{\sqrt{(x-a_{i1})^2 + (y-a_{i2})^2}}}$$

In view of these results, a resolution method was proposed by Weiszfeld (1937), whose main idea is to generate a succession of points so that each new point is closer to verifying the previous equation.

The idea of looking for a fixed-point equation to obtain an iterative method that leads to the solution is well known in the field of numerical analysis and belongs to the class of procedures known as *one-point successive approximation methods*, since we only need the current point to determine the next.

Therefore, the solution of the system of nonlinear equations can be obtained using the following iterative method which, given the point corresponding to the iteration k-th, x_k will generate the following by means of the subsequent function:

$$x_{k+1} = \frac{\sum_{i=1}^{m} \frac{w_i a_{i1}}{\sqrt{(x_k-a_{i1})^2 + (y_k-a_{i2})^2}}}{\sum_{i=1}^{m} \frac{w_i}{\sqrt{(x_k-a_{i1})^2 + (y_k-a_{i2})^2}}}$$

$$y_{k+1} = \frac{\sum_{i=1}^{m} \frac{w_i a_{i2}}{\sqrt{(x_k-a_{i1})^2 + (y_k-a_{i2})^2}}}{\sum_{i=1}^{m} \frac{w_i}{\sqrt{(x_k-a_{i1})^2 + (y_k-a_{i2})^2}}}$$

Thus, once we have the point (x_k, y_k), we can build the following by substituting the corresponding values in the previous formula.

However, Weiszfeld's algorithm has its drawbacks: the method stops working if any iteration matches one of the vertices $a_i = (a_{i1}, a_{i2})$. Accordingly, the set of points for which the Weiszfeld algorithm stops working is a countable set.

By the end of the 1960s many algorithms for the optimization of nonlinear functions applicable to the function $W(x, y)$ had been developed.

However, the Weiszfeld algorithm is still used for its simplicity and elegance, and for being the closest to the solution at each step but the system $\frac{\partial W(x,y)}{\partial x} = 0$ and $\frac{\partial W(x,y)}{\partial y} = 0$ could be solved using by other algorithms such as the Newton´s method.

To comment on this algorithm, it can be affirmed:

- This algorithm needs an initial point, as it has been observed. And although there are few references in the scientific literature on the best choice of this point, Cañavate et al (2001) provide us with an initial point:

$$x_0 = \frac{\sum_{i=1}^{m} w_i a_i}{\sum_{j=1}^{m} w_j}$$

  However, for the case we are dealing with, the choice of the initial point is not relevant.

- Cañavate et al (2001) also propose a stop criterion for the iterative algorithm, based on an error which is previously established at the beginning.

To visualize these properties we will use the example given by these points:

A = (7,0), B = (3,4), C = (1,3), D = (0,1)
Next, by performing the algorithm, we obtain a fairly close position. We will use the point (0,0) as a start.

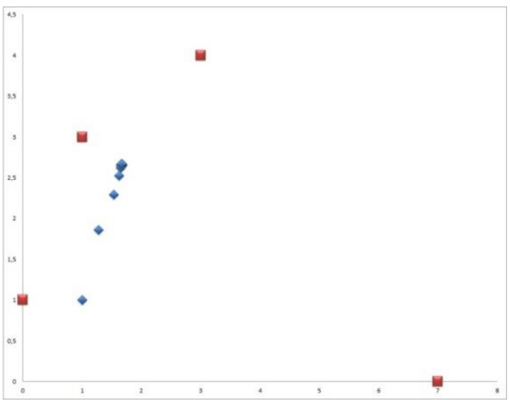


Figure 8: Representation of the algorithm

We can see the representation of the generated sequence as well as the final point to which they converge are in Figure 8 by the Weiszfeld algorithm.

Moreno, A., Viveros, A. y Ramírez, R. (2021). Statistical significance of the median of a set of points on the plane. *The College Mathematics Journal*, 52(3), 205-218

https://doi.org/10.1080/07468342.2021.1909969

PREPRINT

As we can see, this succession depends on the initial value that we choose and it is directed towards the midpoint of the distribution. Thus, the sequence generated by Weiszfeld's algorithm does not determine anything about the position of the median in the plane, because, depending on the chosen starting point, it will vary in its position in the plane and it will not allow us to evaluate anything about how it is approaching the midpoint.

Let us see a geometric interpretation of the obtained point.

4. Location of the median in the plane for quadrilaterals

We can observe the angles formed by the median with the vertices in the previous example:

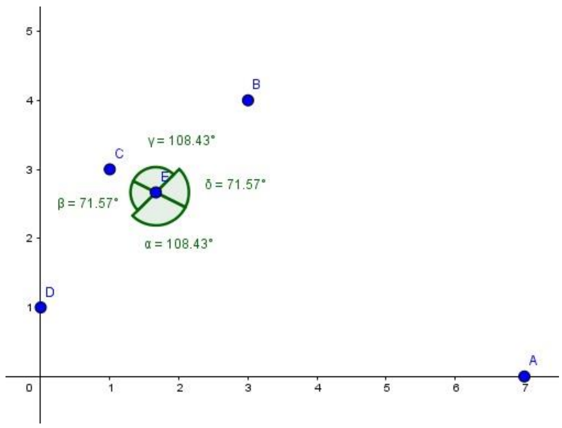


Figure 9: Previous quadrilateral and its median.

One characteristic suggested by Figure 9 is that for the four angles from the median to contiguous vertices of the quadrilateral, adjacent angles are supplementary, and non-adjacent angles are vertical angles.

One characteristic that can be deduced from the previous figure, is that the “viewing angles” of pairs of contiguous vertices with respect to the median are equal two to two.

In this way, if we add the angles α and β, we would obtain 180º, since, being the sum of the different angles, they must sum up to 180º.

Therefore, it is noted that C and A are on the same line when forming a 180º angle between C and A. With the same deduction, it is verified that B and D are on the same line.
Thence, an interesting feature is displayed: the median is the intersection of the diagonals of the convex quadrilateral.

To find the location of the median in quadrilaterals as the point of the intersection of the diagonals, let us check if this point complies with the Weiszfeld algorithm. We will focus on analysing the presented case to interpret the solution obtained geometrically rather than advancing towards a general presentation.

Except for the case of movements, turns and symmetries, we can assume that the quadrilateral is located in the first quadrant.

With this we will demonstrate that the point that is at the intersection of the diagonals is the one that minimizes distances, leaving two pairs of equal opposite angles.

To visualize the required feature we will draw the right triangles shown in figure 10:

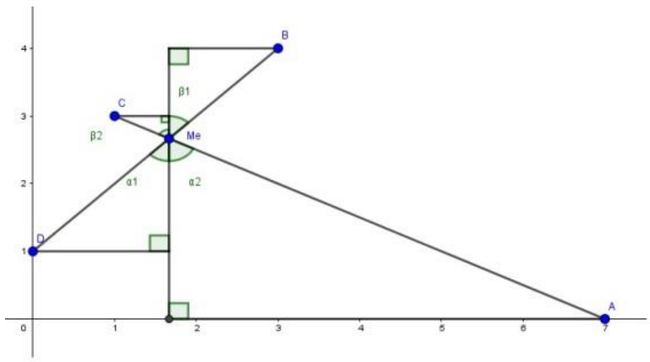


Figure 10: Construction using rectangle triangles

It is noted that the hypotenuses of the triangles are the segments that join the median point with each vertex.

In the triangle containing α_1, the opposite side to the angle would be $(x - a_{i1})$, and the hypotenuse of the triangle would be none other than $\sqrt{(x - a_{i1})^2 + (y - a_{i2})^2}$, where, in your notation, (a_{i1}, a_{i2}) are the coordinates of vertex i.

Then, from Weiszfeld's algorithm, we obtain that:

$$\frac{\partial W(x,y)}{\partial x} = \sum_{i=1}^{m} \frac{w_i(x-a_{i1})}{\sqrt{(x-a_{i1})^2+(y-a_{i2})^2}} = 0$$

$$\frac{\partial W(x,y)}{\partial y} = \sum_{i=1}^{m} \frac{w_i(y-a_{i2})}{\sqrt{(x-a_{i1})^2+(y-a_{i2})^2}} = 0$$

From the first equality we can obtain the following:

$$\frac{\partial W(x,y)}{\partial x} = \sum_{i=1}^{m} \frac{w_i(x-a_{i1})}{\sqrt{(x-a_{i1})^2+(y-a_{i2})^2}} = 0$$

Moreno, A., Viveros, A. y Ramírez, R. (2021). Statistical significance of the median of a set of points on the plane. *The College Mathematics Journal*, 52(3), 205-218

https://doi.org/10.1080/07468342.2021.1909969

PREPRINT

$$\frac{(x-a_1)}{\sqrt{(x-a_1)^2+(y-a_2)^2}}+\frac{(x-b_1)}{\sqrt{(x-b_1)^2+(y-b_2)^2}}+$$

$$\frac{(x-c_1)}{\sqrt{(x-c)^2+(y-c_2)^2}}+\frac{(x-d_1)}{\sqrt{(x-d_1)^2+(y-d_2)^2}}=0$$

Using figure 10, the Weiszfeld algorithm is reformulated in function of the sines, obtaining:

$$sin(\alpha_1)-sin(\alpha_2)-sin(\beta_1)+\ sin(\beta_2)=0$$

$$sin(\alpha_1)+\ sin(\beta_2)=sin(\beta_1)+\ sin(\alpha_2)$$

In this particular case, we have two angles that make the sine with a positive value and two others to which we have changed the negative sign to comply with the shape of the drawing. There are cases in which the position of the vertices results in having more positive values than negative ones, or more negative values than positive ones, preventing an equality like the one above.

However, to solve these cases it is necessary to pay attention to the second equation of the Weiszfeld algorithm, that is, the one given by the Y coordinate, thus making a demonstration in function of sines and cosines totally analogous to it.

Next, we are going to apply the trigonometric identity of the sum of sines:

$$2sin\left(\frac{\alpha_1+\beta_2}{2}\right)cos\left(\frac{\alpha_1-\beta_2}{2}\right)=$$
$$2sin\left(\frac{\beta_1+\alpha_2}{2}\right)cos\left(\frac{\beta_1-\alpha_2}{2}\right)$$

Finally, the result is:

$$\alpha_1+\beta_2=\beta_1+\alpha_2$$

$$\alpha_1-\beta_2=\beta_1-\alpha_2$$

From which we achieve:

$$\alpha_1=\beta_1$$

$$\beta_2=\alpha_2$$

Then $\alpha_1+\alpha_2=\beta_1+\beta_2$, which is indeed a condition that diagonals fulfill, since according to Thales’ Theorem, the angles opposite to the vertex which are formed when a line is cut are equal.

In this way we have come up with a particular solution of identity. But how do we know if it is the only one?

Because, indeed, it is the only solution: since, according to Cañavate et al (2001), Weiszfeld's algorithm guarantees the uniqueness of the median for any set of points.

Therefore, by using the Weiszfeld algorithm we can obtain the position of the median as the intersection of a quadrilateral's diagonals.

If we want to know a comparative between mean and median location, then we show you a quadrilateral, its diagonals, contour plots of the system $\frac{\partial f(x,y)}{\partial x} = 0$ and $\frac{\partial f(x,y)}{\partial y} = 0$ (where $f(x, y)$ is the sum of the distances from (x, y) to the vertices).

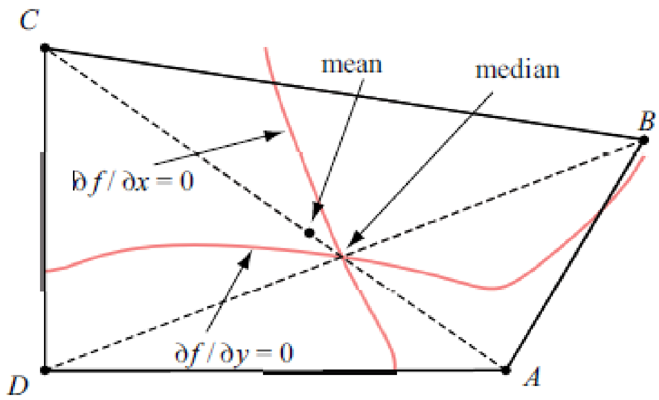


Figure 11: Median versus mean for the vertices of a convex quadrilateral.

Although we are aware that this extension is investigated in several complete studies in various articles, among them we will highlight that of Plastria (2006), which deals with the problem by using the classification of Kupitz and Martini (1997), describing the properties in the following way from the beginning:

The sum of the Euclidean distances to four fixed points p, q, r, s in the plane is minimized in the point of intersection of the diagonals, as long as the fixed points form a quadrilateral convex.

The function that we use in the Weiszfeld algorithm is a differentiable function, except for the vertices, where the function is not convex. If this is transferred to a quadrilateral when a quadrilateral is not convex, meaning that one of the interior angles measures more than 180º, the algorithm will result in the median being in one of the vertices of the quadrilateral.

This feature could be interpreted as analogous to that of 120º angles in the case of triangles.

In both cases, a property could be induced that is related in some way to that of the real line of "distributing half of the points on both sides".

The median of a quadrilateral is obtained by the intersection of the lines that meet the following criteria:

They are lines that pass through two of the points and divide the plane into two half-planes so that each one of them contains one of the remaining points.

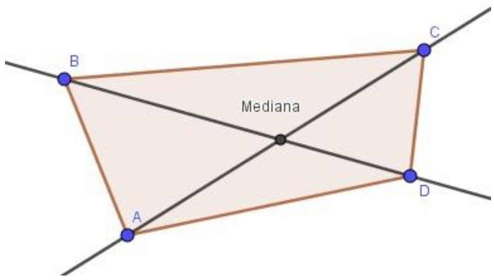


Figure 12: Quadrilateral convex

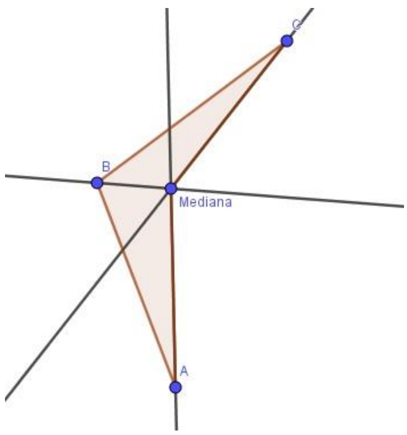


Figure 13: Quadrilateral concave

As we can see, in the case of the quadrilateral convex we have two lines that fulfill the cited property, those lines are the diagonals of the quadrilateral, and they cut in the median.

Finally, note that in the case of the quadrilateral concave, we have three lines that fulfill this criterion and they cut in the median.

In this case, a characteristic of angles of view of pairs of angles is not perceived, but it could be considered in angles of view of 180º for three points, although it could not be generalized in the case of the concave.

However, we do not consider that the characteristic of angles of view is generalized naturally as in the case of three points.

Conclusion

In the case of convex quadrilaterals, the median is obtained as the cut-off point of two lines (in this case the diagonals) that can be interpreted as lines that divide the distribution into two parts with an equal number of points.

In the case of quadrilaterals, the *median line* could be defined in that sense, being the line formed by joining two of its points and dividing the distribution into two parts with an equal number of points. When it corresponds to the diagonals, there are only two and they cut at one point.

In the case of parallelograms, it would even maintain the feature (similar in the triangle) of dividing the polygon into two parts of the same area.
However, these definitions and properties are not fulfilled for sets of n points.

Approximately four centuries ago, Fermat set out to find a point that minimized the sum of distances to three fixed points in the (Euclidean) plane, initiating, without knowing it, the family of *location problems*.

Although Fermat's original question found a complete answer by many scholars in his century, including such illustrious names as Torricelli, Ricci, Cavalieri, Viviani or Renieri, today it is still being studied vigorously from many points of view and for different extensions.

In this article, we think that the reflection on the generalization of the property of the real line median to the Euclidean plane adds some richness of meaning. The study of geometric properties of the Fermat point in simple cases, such as triangles and quadrilaterals, provides the elements that enrich the meaning of the related statistical concepts, especially with the interpretation of median as a measure of central position.

Moreno, A., Viveros, A. y Ramírez, R. (2021). Statistical significance of the median of a set of points on the plane. *The College Mathematics Journal*, 52(3), 205-218

https://doi.org/10.1080/07468342.2021.1909969

PREPRINT

The connection between Geometry and Statistics helps to give greater meaning to the contents that support versions from two perspectives. In the case of the median, this connection is reinforced by the statistical definition itself in the plane that is derived from a geometric property. These connections help to enrich the teaching process of elementary statistical elements both in secondary and in elementary university courses.

According to Bryant (1984), “The main benefit of this approach is an appreciation of the surprising power of a small number of underlying principles. The approach emphasizes the equivalence of the notions, expressed in different “languages”, rather than any expression by itself.”

It is an interesting study for us to obtain a statistical interpretation of the characteristic of the median given a set of n points in the plane and in superior dimensions.

Since the medians of a triangle intersect at the barycenter, which is the mean (the arithmetic average of the vertices), and the three Simpson lines meet at the Fermat point which is also the median of the vertices, then—with respect to the common meaning of the word median—the geometric median point and the statistical median are different notions.

Moreno, A., Viveros, A. y Ramírez, R. (2021). Statistical significance of the median of a set of points on the plane. *The College Mathematics Journal*, 52(3), 205-218

https://doi.org/10.1080/07468342.2021.1909969

PREPRINT

Bibliography

Bernal, R. J. C., Tornel, M. B. C. (S.F.). *Una revisión histórica de los métodos clásicos de resolución del problema de Fermat-Weber.*

Bryant, P. (1984). *Geometry, statistics, probability: variations on a common theme.* The American Statistician, 38(1), 38-48.

Cárdenas, S. (2004). *El punto de Fermat.* Misc. Matem, 40, 77-85.

Cañavate, R. J., Cobacho, B. & Rodríguez J. M. (2001). El algoritmo de Weiszfeld para la resolución del problema económico de Weber. *En ASEPELT. XV Reunión ASEPELT-España, p. 13. Vigo: España. ASEPELT. Retrived of* http://repositorio.upct.es/xmlui/bitstream/handle/10317/1507/awr.pdf?sequence=1&isAllowed=y

García, C. G. (2005). *Un modelo multicriterio de localización de centros no deseados con pesos* (Doctoral dissertation, University of Málaga).

Herr, D. G. (1980).*"On the History of the Use of Geometry in the General Linear Model"*, The American Statistician, 34, 43-4

Love, R. F. (1967). *A Note on the Convexity of Siting Depots*. International Journal of Production Research, 6:153-154.

Margolis, M. S. (1979). *Perpendicular projections and elementary statistics.* The American Statistician, 33(3), 131-135.

Krarup, J. & Vajda, S. (1997). *On Torricelli's geometrical solution to a problem of Fermat.* IMA Journal of Management Mathematics,8(3),215-224.

Kupitz, Y. S. & Martini, H. (1997). *Geometric aspects of the generalized Fermat-Torricelli problem.* Intuitive geometry, vol. 6. Bolyai Society, Mathematical Studies, pp. 55-127.

Sángari, A &, Egües, C. (2012). *Una síntesis del cálculo y la geometría: Localización del Punto de Fermat.* Revista de Educación Matemática, 27.

Plastria, F. (2006). *Four-point Fermat location problems revisited. New proofs and extensions of old results.* IMA Journal of Management Mathematics, 17(4), 387-396.

Sarkar, J. & Rashid, M. (2016). A geometric view of the mean of a set of numbers. *Teaching Statistics*, 38(3), 77-82.

Moreno, A., Viveros, A. y Ramírez, R. (2021). Statistical significance of the median of a set of points on the plane. *The College Mathematics Journal*, 52(3), 205-218

https://doi.org/10.1080/07468342.2021.1909969

PREPRINT

Saville, D. J. & Wood, G. R. (1986). A method for teaching statistics using N-dimensional geometry. *The American Statistician*, 40(3), 205-214.

Weiszfeld, E.V.(1937). Sur le point pour lequel la somme des distances de n points donnes est minimum. *The Tohoku Mathematical Journal*, 43, 335-386.

Summary

This article proposes a reflection for the statistics professor about the concept of the median, by delving into the geometric meaning of some of its properties. In the number line, the definition of median establishes a central measure relative to dividing the points of the distribution into two equal parts.

The definition in the plane is generalized from the property relative to minimizing distances of the Fermat point, approached from a geometric perspective and with a complex procedure. For any set of points, the statistical significance of the median is lost as a central measure of the distribution. The cases of three and four points are addressed in this work by interpreting the geometric solutions from a statistical perspective.